

Convergent evolution of high elevation plant growth forms and geographically structured variation in Andean *Lupinus* (Fabaceae)

NATALIA CONTRERAS-ORTIZ^{1*}, GUY W. ATCHISON², COLIN E. HUGHES² and SANTIAGO MADRIÑÁN^{1,3}

¹Laboratorio de Botánica y Sistemática, Departamento de Ciencias Biológicas, Universidad de los Andes, Apartado Aéreo 4976, Bogotá, Colombia

²Department of Systematic and Evolutionary Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland

³Jardín Botánico de Cartagena “Guillermo Piñeres”, Turbaco, Bolívar, Colombia

Received 5 December 2016; revised 19 May 2017; accepted for publication 18 November 2017

The Andean *Lupinus* radiation presents one of the highest rates of net species diversification in plants and includes species with a wide variety of growth forms, but the detailed geographical and adaptive trajectories of diversification in the Andes remain unresolved. One of the most distinctive of these growth forms is the fistulose-inflorescence rosette, FIR, typical of tropical high-elevation habitats. To gain insights into the evolutionary origins of this growth form and the biogeography of Andean *Lupinus*, genome-scale nextRADseq data were generated for 124 individuals from the northern and central Andes. The RADseq data provide unprecedented phylogenetic resolution among the recently diverged Andean *Lupinus* spp. and suggest multiple independent evolutionary origins of the rosette growth form and marked geographical structure across the northern Andes at different scales. This suggests that both adaptive (ecological) and geographically driven non-adaptive diversification have contributed to rapid recent species diversification. Among the north-Andean FIR *Lupinus* spp., our results strongly support recognition of *L. alopecuroides* (Central Cordillera and northern Ecuador), *L. trianaanus* (southern portion of the Eastern Cordillera) and a putative new species from the ‘Pisba-Cocuy’ and ‘Ocetá’ páramos (northern portion of the Eastern Cordillera).

ADDITIONAL KEYWORDS: adaptive plant radiation – Andes – Colombia – Leguminosae – *Lupinus alopecuroides* – *Lupinus trianaanus* – nextRADseq – páramos – species delimitation.

INTRODUCTION

Evolutionary radiations characterized by rapid diversification of lineages and/or phenotypes (morphological, physiological, ecological divergence) are a universal feature of evolutionary diversification in many different geographical and ecological settings (Losos, 2010; Hughes, Nyffeler & Linder, 2015). In most cases, radiations involve elements of adaptive (phenotypic and ecological) diversification (sometimes termed disparification; e.g. Ackerley, 2009) and geographically driven non-adaptive species or lineage diversification (Sanderson, 1998; Losos & Mahler, 2010), but the relative importance of these two axes of evolutionary

diversification across organismal groups and ecological settings remains poorly understood. Nevertheless, well-studied radiations appear to vary greatly, ranging from relatively species-poor but spectacularly phenotypically diverse radiations, in which the adaptive axis is dominant, to species-rich clades, in which phenotypic diversity is less prominent and geographically driven species diversification predominates, with the latter referred to simply as ‘explosive species diversification’ (Givnish, 2015). Classical examples of adaptive radiation have focused on island and island-like systems where geographical isolation and ecological opportunity allowed groups of plants and animals to expand, diversify rapidly and occupy diverse ecological niches (Hughes & Eastwood, 2006; Seehausen, 2006; Losos & Ricklefs, 2009; Givnish *et al.*, 2009).

*Corresponding author. E-mail: na.contreras142@uniandes.edu.co

In recent years, the high elevation Andean grasslands have emerged as harbouring one of the greatest concentrations of large plant radiations on the planet (Madriñán, Cortés & Richardson, 2013; Hughes & Atchison, 2015). These high-elevation grassland ecosystems, which extend from Venezuela and Colombia to northern Argentina and Chile, began to emerge during the late Neogene, when the Andes reached their modern elevation (Gregory-Wodzicki, 2000; Hay *et al.*, 2002). They are found above the treeline and below the permanent snow line (c. 3200–5000 m), forming an archipelago of island-like land areas above c. 3200 m, and are locally referred to as páramo in the northern Andes and puna further south in the central Andes (Luteyn, 1999; Hughes & Eastwood, 2006).

This island-like isolation and new ecological opportunities following recent Andean uplift are hypothesized to have prompted the rapid and recent diversification of a suite of plant clades including Espeletiinae (Madriñán *et al.*, 2013; Diazgranados & Barber, 2017), *Gentianella* Moench (von Hagen & Kadereit, 2001), *Calceolaria* L. (Cosacov *et al.*, 2009), *Puya* Molina (Jabaily & Sytsma, 2013), *Valeriana* L. (Bell & Donoghue, 2005), *Oreobolus* R.Br. (Chacón *et al.*, 2006), *Lupinus* L. (Hughes & Eastwood, 2006), *Astragalus* L. (Scherson *et al.*, 2008), *Hypericum* L. (Nürk, Scheriau & Madriñán, 2013), *Bartsia* L. (Uribe-Convers & Tank, 2015) and *Diplostephium* Kunth (Vargas, Ortiz & Simpson, 2017). All these high-elevation Andean radiations are recent and are characterized by high species-diversification rates. However, the detailed geotemporal trajectories of diversification within these Andean radiations remain largely unknown because of the lack of resolution and incomplete taxon sampling in current phylogenetic studies, with the exceptions of *Bartsia* and *Diplostephium*, for which large targeted multi-gene DNA sequence and genome skimming data sets have provided insights into the evolutionary relationships at the species level in these Andean radiations (Uribe-Convers & Tank, 2015; Uribe-Convers, Settles & Tank, 2016; Vargas *et al.*, 2017). This means that little is known about the relative importance of adaptive vs. purely geographical components of diversification for this large cohort of young Andean plant radiations.

The archipelagic conformation of the Andean high elevation grasslands and the extreme physiographic heterogeneity of the Andes (Särkinen *et al.*, 2012) present many potential dispersal barriers and opportunities for geographical isolation suggesting a strong geographical component of diversification (Givnish, 2015). Maybe the most prominent such barrier for dispersal of high elevation grassland species is the Huancabamba Depression in northern Peru (Simpson, 1975; Parker *et al.*, 1985; Ayers, 1999), which is also

a potential centre of diversity for many taxa (Berry, 1982; Weigend, 2002; 2004). Similarly, the division of the Andes into three separate cordilleras in Colombia, and an outlying high-elevation Andean offshoot in Venezuela present additional obvious dispersal barriers and strong geographical isolation between them for montane species (Cadena, Klicka & Ricklefs, 2007; Londoño, Cleef & Madriñán, 2014).

At the same time, there is also no doubt that many of these natural dispersal barriers for high elevation species in the Andes have been transient, fluctuating and highly dynamic through the Pleistocene, with large elevational shifts in vegetation zones associated with successive glacials and interglacials (Simpson, 1974; Luteyn, 1999; Hooghiemstra & van der Hammen, 2004; Torres *et al.*, 2013) precipitating dramatic changes in area, connectivity and fragmentation of high-elevation grassland and montane forest habitats (Flantua *et al.*, 2014). However, this complexity remains poorly understood in terms of the impacts on species diversification and genetic diversity of populations across the Andes (e.g. Jabaily & Sytsma, 2013; Vásquez *et al.*, 2016; Kolář, Dušková & Sklenář, 2016). At the same time, the extremely steep and extended environmental gradients associated with elevation in the tropical Andes, present strong adaptive axes for diversification conducive to ecological divergence and are reflected in the striking disparification of growth forms adapted to different elevation zones for many high elevation Andean plant groups (Rosser *et al.*, 2012; Hughes & Atchison, 2015; Lagomarsino *et al.*, 2016).

The Andean radiation of *Lupinus* (Leguminosae) comprises a clade of c. 85 species with an estimated crown age of 1.19–3.50 My (Drummond *et al.*, 2012) and presents one of the highest documented rates of net species diversification in plants (Hughes & Eastwood, 2006; Drummond *et al.*, 2012; Hughes & Atchison, 2015). This clade is nested in a larger western New World clade characterized by a uniform chromosome number of $2n = 48$ (Conterato & Schifino-Wittmann, 2006) spanning the mountains of western North America, Mexico, Central America and the Andes, referred to as a ‘super-radiation’ by Drummond *et al.* (2012). The switch from annual to perennial life history has been suggested as a key adaptation facilitating colonization of montane habitats and enabling the subsequent diversification of growth forms that characterizes the Andean clade (Drummond *et al.*, 2012; Hughes & Atchison, 2015).

The Andean *Lupinus* clade includes species with a wide variety of growth forms occupying elevation zones from sea level to 5000 m at the upper limit for plant growth, with the greatest diversity of species above 3000 m. These species include dwarf prostrate mat-forming perennials <5 cm tall (3500–4800 m

elevation), erect woody shrubs and scrambling subshrubs (2700–3800 m elevation), small trees to 5–7 m height (2500–3200 m elevation), acaulescent clump-forming rosettes including one stem rosette species, *L. weberbaueri* Ulbr. (3800–4900 m elevation) and a small number of apparently evolutionarily derived herbaceous annuals (0–3000 m elevation) (Hughes & Eastwood, 2006; Hughes & Atchison, 2015).

One of the most distinctive of these growth forms is the weakly caulescent, or more frequently acaulescent rosette, characterized by long-petiolate clump-forming leaves, and including *L. weberbaueri*, which can form a stem rosette to 2 m tall (Fig. 1A–C). These plants are closely reminiscent of similar tropical high elevation rosettes such as the Espeletiinae, *Puya* and *Nicotiana thyrsiflora* Goodsp. in the Andes, giant lobelias in Hawaii, Africa and Asia (Givnish, 2010; Chen, Wang & Renner, 2016), members of the Hawaiian silversword alliance and the African giant senecios (*Dendrosenecio* B.Nord.), among others (Carlquist, 1974; Hedberg & Hedberg, 1979; Diazgranados & Barber, 2017). Typically, these rosette lupins have markedly fistulose (i.e. cylindrical and hollow or pithy) inflorescences (Fig. 1F, G) ranging in length from <1 cm up to 120 cm, sometimes held below the long-petiolate leaves or extending far above. These different arrangements (Fig. 1) combined with usually dense pubescence provide a suite of adaptations probably facilitating protection of flowers from light ground-layer frosts that can occur on virtually any night of the year in the high-elevation daily freeze-thaw zone in tropical mountains (Hedberg & Hedberg, 1979). For *Lupinus*, we refer to this growth form as the fistulose-inflorescence rosette (FIR).

There are c. 20 FIR *Lupinus* spp. scattered across the Andes from Venezuela to Bolivia, all of them restricted to the highest elevation super-páramo and puna between (3500–) 3800 m and 4900 m, close to the upper elevational limit for plant growth in the Andes (Luteyn, 1999). Lack of resolution among the Andean species in previous phylogenetic analyses (Käss & Wink 1997; Ainouche & Bayer, 1999; Ainouche, Bayer & Misset, 2004; Hughes & Eastwood 2006; Drummond et al., 2012) means that it is not clear whether these distinctive FIR species form a closely related group, or are independently derived from closely related mid-elevation shrubby, tree or herbaceous growth forms. In this study, we examine this question focusing on the FIR species of the northern Andes, especially Colombia.

The taxonomy of Andean *Lupinus* is the focus of on-going work to assemble a new taxonomic account (Hughes et al., in prep). Many questions remain about species delimitation in particular groups and the taxonomy is aggravated by the proliferation of names

perpetrated in the fundamentally flawed *Species Lupinorum* of C.P. Smith (1930–1953). Although most of the Andean FIR species are clearly distinct and readily delimited (e.g. *L. chrysanthus* Ulbr., *L. weberbaueri*), doubts remain concerning FIR species boundaries for the northern Andes.

Four north-Andean FIR species are generally recognized: (1) *L. carrikeri* C.P.Sm. (Fig. 1D), endemic to the Sierra Nevada de Santa Marta, in northern Colombia; (2) *L. trianaanus* C.P.Sm. (Fig. 1A, F), a name usually applied to populations in the Eastern Colombian Cordillera; (3) *L. alopecuroides* Desr. (Fig. 1B), usually assigned to populations spanning the Central and Eastern Cordilleras in Colombia and extending into Ecuador (Vásquez et al., 2016) and (4) *L. nubigenus* Kunth (Fig. 1K), a dwarf rosette mainly distributed in Ecuador but extending into northern Peru, which has been viewed by some as simply a diminutive version of *L. alopecuroides* and previously treated as a variety of the latter (Weddel, 1861), but a close relationship with this species has not been clearly established. Thus, the boundaries between these species remain poorly defined especially in the overlap zone of *L. trianaanus* and *L. alopecuroides* as traditionally circumscribed. Furthermore, FIR populations from páramos in the northern portion of the Colombian Eastern Cordillera that have generally been referred to *L. alopecuroides*, do not appear to match any of these species.

Previous phylogenetic studies of *Lupinus* were based on a small set of plastid, ITS and low-copy nuclear gene (*LEGCYC1* and *GPAT*) sequences and yielded phylogenetic trees with essentially no supported resolution of species relationships in the large Andean clade (Käss & Wink 1997; Ainouche & Bayer, 1999; Ainouche et al., 2004; Hughes & Eastwood, 2006; Drummond et al., 2012). The short time scale associated with such rapid episodes of species diversification, as seen in the Andean *Lupinus* radiation, mean that data from traditional sequencing of a handful of plastid and nuclear DNA sequence loci alone do not provide sufficient phylogenetic signal to build a well-resolved and robustly supported phylogenetic tree. In contrast, genome-wide data using next generation sequencing (NGS) approaches have recently been used to provide unprecedented phylogenetic resolution among Andean *Lupinus* spp. (Atchison et al., 2016; Nevado et al., 2016). In this study, we use a derivative method of RADseq, nextRAD (Nextera-tagmented, reductively-amplified DNA) (Emerson et al., 2015; Russello et al., 2015), which was successfully used to infer the origin of domestication of *L. mutabilis* Sweet, the Andean crop lupin (Atchison et al., 2016). The potential to generate genome-wide data sets quickly and cost-effectively makes RADseq and variants thereof powerful methods particularly for the investigation of relationships

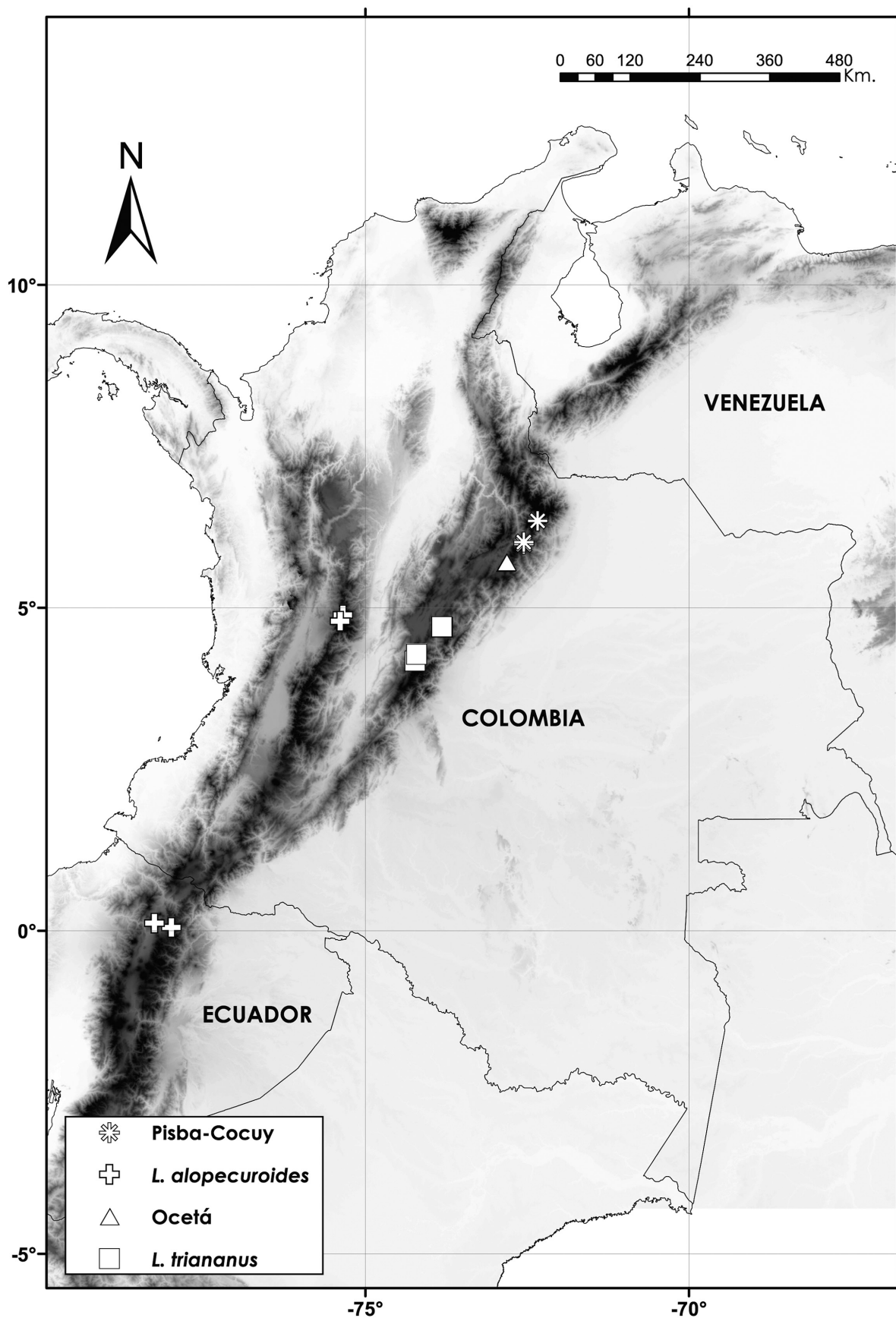


Figure 1. Fistulose-inflorescence rosette (FIR) *Lupinus* spp. in tropical high-elevation habitats across the Andes: A, *L. trianaanus*, at 3738 m, Páramo de Chingaza, Colombia. B, *L. alopecuroides*, Páramo de Puracé, Colombia. C, *L. weberbaueri*, stem rosette to 2 m height, at 4750 m, Cordillera Blanca, Ancash, Peru. D, *L. carrikeri*, acaulescent rosette, at 3900 m, Sierra Nevada de Santa Marta, Colombia. E, *L. puroserriceus*, weakly caulescent rosette, at 4200 m, Junin, Peru. F, cross section of fistulose inflorescence *L. trianaanus*. Inset, detail of inflorescence. G, cross section of stem of *L. trianaanus*. H, *L. tominensis*, at 3850 m, Quime, La Paz, Bolivia. I, *L. huaronensis*, at 4550 m, Cordillera Blanca, Ancash, Peru. J, *L. chrysanthus*, at 4220 m, Chalhuanca, Apurimac, Peru. K, *L. nubigenus*, at 4950 m, Cordillera Blanca, Ancash, Peru. Photographs A, Guy Atchison; B, Daniel Rojas; F and G, Natalia Contreras-Ortiz; F inset, Santiago Madriñán; C–E and H–K, Colin Hughes.

within and between closely related species (Davey & Blaxter, 2010; Andrews *et al.*, 2016).

In this study, we focus on the delimitation, relationships, biogeography and evolutionary trajectories of

the FIR species from the northern Andes, and particularly the Colombian species. We aim to test species boundaries and morphology-based species hypotheses among northern Andean FIRs, gain insights into



the evolutionary origin(s) of the FIR growth form across the Andean radiation and the biogeography of *Lupinus* in the Colombian Cordilleras. To accomplish this, we infer the phylogenetic relationships, reconstruct the evolution of the FIR growth form using ancestral state reconstruction (ASR) and estimate genetic structure among FIR populations of *Lupinus* from the northern Andes in the context of the Andean radiation as a whole.

MATERIAL AND METHODS

One hundred and twenty-four individuals of FIR and non-FIR (i.e. annual herbs, prostrate perennials, subshrubs, erect shrubs and treelets) species distributed from the northern (Colombia and Ecuador) and central Andes (Peru and Bolivia) (Supporting Information, Table S1) were included in this study. Only one accession (*L. purosericeus* C.P.Sm., MW2000142) was sampled from herbarium material and the remaining DNAs were extracted from silica-dried leaf samples collected in the field.

Sixty-one accessions represent FIR individuals from populations across the Central and Eastern Cordillera in Colombia and two localities in Ecuador (Fig. 2). These populations were densely sampled for phylogenetic and genetic structure analyses. Of the Colombian FIR species, only *L. carrikeri* (Fig. 1D), the endemic species from the remote Sierra Nevada de Santa Marta could not be included in the molecular analysis because available herbarium specimens did not yield DNA of sufficient quality and quantity (>50 ng high molecular weight DNA >20 kb, at >5 ng/μL). In addition, 16 FIR accessions representing eight species from the central Andes and 47 non-FIR accessions distributed across the central and northern Andes were included in the molecular phylogenetic analysis in order to ascertain the relationships of the northern FIR species in the context of the Andean *Lupinus* radiation as a whole. Although not all the estimated c. 85 Andean species (Hughes & Eastwood, 2006) were sampled, the accessions included in this study span the full range of growth forms and the whole distribution along the Andes, with denser sampling in the northern Andes. This sample was based on preliminary evidence suggesting two geographically structured clades, northern and south-central Andean, in the Andean radiation as a whole (Figs 4 and S1 from Atchison *et al.*, 2016). Herbarium specimens, silica-dried leaf samples, field notes and photographs were assembled in the field with voucher specimens deposited in the

herbaria of Universidad de los Andes (ANDES) in Bogotá, Colombia and the University of Zürich (Z), Switzerland (Supporting Information, Table S1).

DNA EXTRACTION AND NEXTRAD SEQUENCING

Genomic DNA was extracted using a CTAB DNA isolation protocol (modified from Doyle & Doyle, 1987). DNA quality was evaluated via 1% agarose gel electrophoresis and quantified with a Qubit Fluorometer (ThermoFisher, Basel, Switzerland). The extracted DNA concentrations from the samples included in this study ranged from 5.7–200.0 ng μL⁻¹ (with 38.0 ng μL⁻¹ for the MW2000142 sample derived from herbarium material).

We used a derivative method of RADseq, nextRAD (Nextera-tagmented reductively-amplified DNA) (Emerson *et al.*, 2015; Russello *et al.*, 2015) that has the potential to resolve species-level relationships in Andean *Lupinus* (Atchison *et al.*, 2016) by providing a large SNP dataset for phylogenetic and population genetic analyses. The nextRAD method differs from RADseq by using Nextera library preparation and selective PCR primers to amplify genome-wide loci consistently between samples, overcoming the restriction fragment length bias of RADseq (Davey *et al.*, 2013).

Genomic DNA was converted into nextRAD genotyping-by-sequencing libraries (Emerson *et al.*, 2015; Russello *et al.*, 2015) at SNPsaurus (SNPsaurus LLC, Oregon, USA), following the protocol described in Atchison *et al.* (2016). Different DNA concentrations (10–15 ng, 2–3× the normal concentration) were used for the fragmentation reaction, with greater input DNA for more degraded extracts, as determined by gel band quality in agarose gel electrophoresis. Sequencing of the libraries was carried out at the University of Oregon Sequencing Center, on an Illumina NextSeq 500 to generate 150-bp single-end reads.

RADSEQ DATA ASSEMBLY

Raw reads were processed using Trimmomatic v0.33 (Bolger, Lohse & Usadel, 2014) to remove bases at the ends of reads with a quality score <20. Reads were assembled de-novo into loci using pyRAD v3.0.5 (Eaton, 2014). Quality filtering of reads in pyRAD (step 2) converted bases with a Phred quality score < 20 into Ns and discarded reads with > 20 Ns. Within-sample clustering (step 3) was performed at a similarity threshold of 0.85. Clusters with a minimum read coverage of <5 were discarded. The consensus

Figure 2. Map of north-western South America showing the locations of sampled populations of fistulose-inflorescence rosettes from the Central and Eastern Cordillera of the northern Andes.



sequences were assembled for each cluster using error-rate and heterozygosity estimates calculated from the base counts at each site across all clusters. Consensus loci were clustered between samples at the same threshold as used in step 3. The other parameters were left at pyRAD default values.

For phylogenetic inference, we generated a Phylip formatted concatenated matrix of loci from pyRAD (step 7) with a minimum sample coverage of 5 (the minimum number of samples that have data for a given locus to be included in the final matrix). The assembled matrix contained 124 accessions (hereafter 124acc).

For inference of genetic structure among the FIR populations from the 'northern Andes' clade (Fig. S1 from Atchison *et al.*, 2016), we generated two separate matrixes of 30 and 31 FIR accessions (hereafter 30acc and 31acc). These matrixes correspond to the two monophyletic clades (Fig. 3A) of FIR populations, i.e. the *L. trianaus*/*L. alopecuroides* clade and the 'Océta'-Pisba-Cocuy' clade, together spanning the Central and Eastern Cordilleras and localities in Ecuador (Fig. 2; Supporting Information, Table S1). The 30acc and 31acc matrixes were generated with a minimum sample coverage of 10 and assembled using the unlinked SNPs approach with one randomly sampled SNP per locus with the output formatted for Structure (.str).

DNA samples that generated fewer than c. 200 000 reads, mostly from DNA sampled from herbarium material (results not shown), were removed because the previously described filtering steps progressively decreased read and loci numbers and because the high proportion of missing data for those samples could result in loss of resolution and support across the phylogenetic tree. Only one DNA sample derived from herbarium material generated enough reads to be included in the analyses: *L. purosericeus*, MW2000142 with 296 068 reads (Supporting Information, Table S1).

PHYLOGENETIC ANALYSIS

The 124acc alignment of concatenated loci obtained from PyRAD was filtered to remove invariant sites using the PyCogent version 1.5.3 toolkit (Knight *et al.*, 2007) to yield a matrix of 2 151 174 SNPs for phylogenetic analysis. For phylogenetic inference under maximum likelihood (ML) we used RAxML version 8.1.13 (Stamatakis, 2014) with a general-time-reversible

model of sequence evolution with gamma-distributed rate of variation (GTR + Γ) and conducted 100 replicates of rapid bootstrap (BS) algorithm in RAxML to describe uncertainty in the estimation of the topology. *Lupinus polyphyllus* Lindl., a North American perennial and a member of the $2n = 48$ chromosome western New World clade that includes the Andean radiation (see Drummond *et al.*, 2012) was used as an outgroup for rooting. Additionally, to further explore and corroborate the ML phylogenetic tree a species tree consistent with the multi-species coalescent was inferred with the software *tetrad* v.0.6.1 (<http://github.com/dereneaton/ipyrad>, accessed 9 May 2017; Eaton *et al.*, 2016), using the full SNP alignment of a subset of 32 samples chosen to represent the northern Andes clade.

ANCESTRAL STATE RECONSTRUCTION

The ASR of the FIR growth form was performed using the maximum likelihood method under the Markov k-state-1-parameter model of evolution (any change is equally probable) implemented in Mesquite v.3.2 (Maddison & Maddison, 2017). To account for topological uncertainty the reconstruction was performed across the 100 topologies obtained from the rapid bootstrap analysis performed in RAxML (see previous) using the option 'trace character over trees'. The trees were pruned leaving one accession per monophyletic species or morphospecies scored as either 1 or 0 for FIR or non-FIR growth form. Because of delimitation issues associated with cryptic species more than one accession was retained for some non-monophyletic species/morphospecies (i.e. *L. purosericeus*, *L. tominensis* Wedd., *L. lindleyanus* J. Agardh, *L. tauris* Benth., *L. microphyllus* Desr. and *L. sp. J*). Reconstruction results were mapped on the best ML tree and summarized using the option 'average frequencies across trees' that estimates the average likelihood of each state at each node across all the trees that have that node.

GENETIC STRUCTURE ANALYSIS

The 30acc and 31acc matrixes with 6273 and 6390 unlinked SNPs, respectively, after processing with PyRAD were used for genetic structure analysis of northern Andean FIR. To determine how well the genetic structure in the molecular data corresponds to

Figure 3. Phylogenetic and genetic structure analyses of Andean lupins. A, ML phylogenetic tree generated using RAxML based on the 124acc data matrix with a total of 2 151 174 SNPs. The tree was rooted using *L. polyphyllus* (Supporting Information Fig. S1) a North American member of the western New World clade (Drummond *et al.* 2012). Bootstrap support values (100 replicates) are shown at nodes for the main clades. Northern Andes: Ecuador and Colombia; Central Andes: Peru and Bolivia; non-FIR species: annual herbs, prostrate perennials, subshrubs, erect shrubs and treelets; FIR: fistulose-inflorescence rosette species/morphotypes names indicated and highlighted in grey/black. B, Genetic structure within and between FIR *Lupinus* spp. from the northern Andes, calculated from the 30acc (bottom) and 31acc (top) matrices of concatenated loci with a total of 13 914 unlinked SNPs (see text).

current species designations and morphology-based entities, we used the program STRUCTURE v.2.3.4 (Pritchard, Stephens & Donnelly, 2000) that implements a Bayesian clustering approach to assign individuals to populations or clusters, given a prior cluster number (K). In STRUCTURE, we used the admixture model and ran ten replicates at each K value between 1 and 7 for 50 000 generations of burn-in and 100 000 generations of sampling for each matrix. To identify the appropriate number of clusters (K), ΔK was calculated following the method of Evanno, Regnaut & Goudet (2005) with the online software STRUCTURE HARVESTER (Earl & vonHoldt, 2012). The largest value of ΔK was used as an indicator of the most probable K (Evanno et al., 2005).

RESULTS

NEXRAD DATA ASSEMBLY

An average of 2 739 653 sequence reads per sample were generated. The numbers of loci recovered after filtering were 76 432 for the 124acc matrix, 6826 for the 30acc matrix and 6876 for the 31acc. The numbers of variable sites and potentially parsimony-informative sites were 1 192 162 and 351 784 in the 124acc, 57 538 and 17 763 in the 30acc and 60 901 and 20 072 in the 31acc. The levels of missing data were c. 90% for the 124acc and c. 60% for both the 30acc and 31acc matrixes. Sequence reads are archived in the NCBI sequence read archive under BioProject PRJNA327946 (Supporting Information, Table S1).

PHYLOGENETIC ANALYSIS

The 124acc nextRADseq phylogenetic tree shows two major clades in the Andes, a robustly supported (100% BS) northern Andean clade (species from Colombia, Ecuador and one from northern Peru) sister to a weakly supported (67% BS) clade of central Andean species (all Peru and Bolivia except for Ecuadorian accessions of one species, *L. nubigenus*) (Fig. 3A; Supporting Information, Fig. S1). The 'northern Andes' clade comprises seven well-supported sub-clades and is consistent with the multi-species coalescent analysis (Supporting Information, Fig. S2): non-FIR species from the Central Cordillera (100% BS), a clade comprising accessions of *L. alopecuroides* also from the Central Cordillera (92% BS) and *L. trianaus* from the Eastern Cordillera (89% BS), non-FIR species from the Eastern Cordillera (100% BS), *L. interruptus* Benth., a treelet (also non-FIR) in the Eastern Cordillera (100% BS) and each of the FIR populations ('Ocetá' and 'Pisba-Cocuy') from the northern part of the Eastern Cordillera (100% BS) (Fig. 3A; Supporting Information, Fig. S1). Species denoted as non-FIR correspond to annual herbs, prostrate perennials, subshrubs, woody shrubs and treelets.

The first branching sub-clade in the 'northern Andes' clade corresponds to non-FIR species distributed mainly in the Central Cordillera from Colombia to northern Ecuador. This clade comprises the species *L. humifusus* Benth., *L. tauris*, *L. microphyllus*, *L. caespitosus* Torr. & A. Gray and three morphologically distinct entities designated *L. sp. K*, *L. sp. J* and *L. sp. H*. Sister to this sub-clade of non-FIR shrubs is a large clade made up of species all of them restricted to the Eastern Cordillera except for *L. alopecuroides*, a FIR species restricted to the Central Cordillera and northern Ecuador. *Lupinus alopecuroides* is robustly supported as sister to another FIR species, *L. trianaus* (100% BS) found in the southern portion of the Eastern Cordillera, on the páramos of Chingaza and Sumapaz (Fig. 3A).

The next two sub-clades of non-FIR species make up a paraphyletic grade of species of woody shrubs and treelets, distributed along the Eastern Cordillera: *L. boyacensis* C.P.Sm., *L. monserratensis* C.P.Sm., *L. chocontensis* C.P.Sm. and six morphologically distinct entities (*L. sp. D*, *L. sp. F*, *L. sp. C*, *L. sp. G*, *L. sp. B* and *L. sp. E*), plus a sub-clade comprising accessions of the treelet *L. interruptus*, which is restricted to the northern part of the Eastern Cordillera (Supporting Information, Fig. S1). This sub-clade of treelets is in turn sister to a second FIR clade comprising two distinct morphotypes both from the northern portion of the Eastern Cordillera. One of these morphotypes corresponds to FIR accessions restricted to the páramo of Ocetá, and the other to the páramos of Pisba and Cocuy ('Pisba-Cocuy'). Thus, the 'Ocetá'-'Pisba-Cocuy' clade of FIR species is nested in a paraphyletic grade made up of woody shrubs and treelets.

The 'central Andes' clade is here represented by a subset of FIR and non-FIR species from Peru and Bolivia. Although, the backbone of this clade is generally weakly supported compared to the more densely sampled 'northern Andes' clade, all the FIR species from Peru and Bolivia (*L. weberbaueri*, *L. huaronensis* J.F. Macbr., *L. chrysanthus*, *L. nubigenus*, *L. tominensis*, *L. purosericeus* and *L. pinguis* Ulbr.) are placed in this clade quite separate from the northern Andean FIR species (Fig. 3A). *Lupinus nubigenus* is the only species from the 'central Andes' clade that extends to the northern Andes in Ecuador and its placement with the central Andean species confirms that *L. nubigenus* is indeed a species distinct from *L. alopecuroides*.

Although several species/morphospecies with more than one accession formed monophyletic groups (e.g. *L. huaronensis*, *L. nubigenus*, *L. humifusus*, *L. monserratensis*, *L. interruptus*), accessions of other species were scattered within a clade (e.g. *L. tauris*, *L. sp. J*, *L. lindleyanus*, *L. purosericeus*) or even across distinct clades (e.g. *L. microphyllus*). This situation,

mostly observed among the non-FIR species and some central Andean FIR species (*L. purosericeus* and *L. tominensis*), may be a consequence of incomplete lineage sorting or mistaken species delimitation probably because of the existence of cryptic species.

ANCESTRAL STATE RECONSTRUCTION

The ASR of growth forms presented here supports at least three independent derivations of the FIR

growth form: one in the central and two in the northern Andes (Fig. 4): once in the *L. alopecuroides*-*L. trianaus* FIR clade in the Central Cordillera/southern portion of the Eastern Cordillera and a second in the 'Pisba-Cocuy'-'Oceta' FIR clade from the northern portion of the Eastern Cordillera. Growth form reconstruction in the central Andes remains uncertain as denoted by the proportion of node absence (weak BS support) on internal branches and as a result of incomplete taxon sampling.

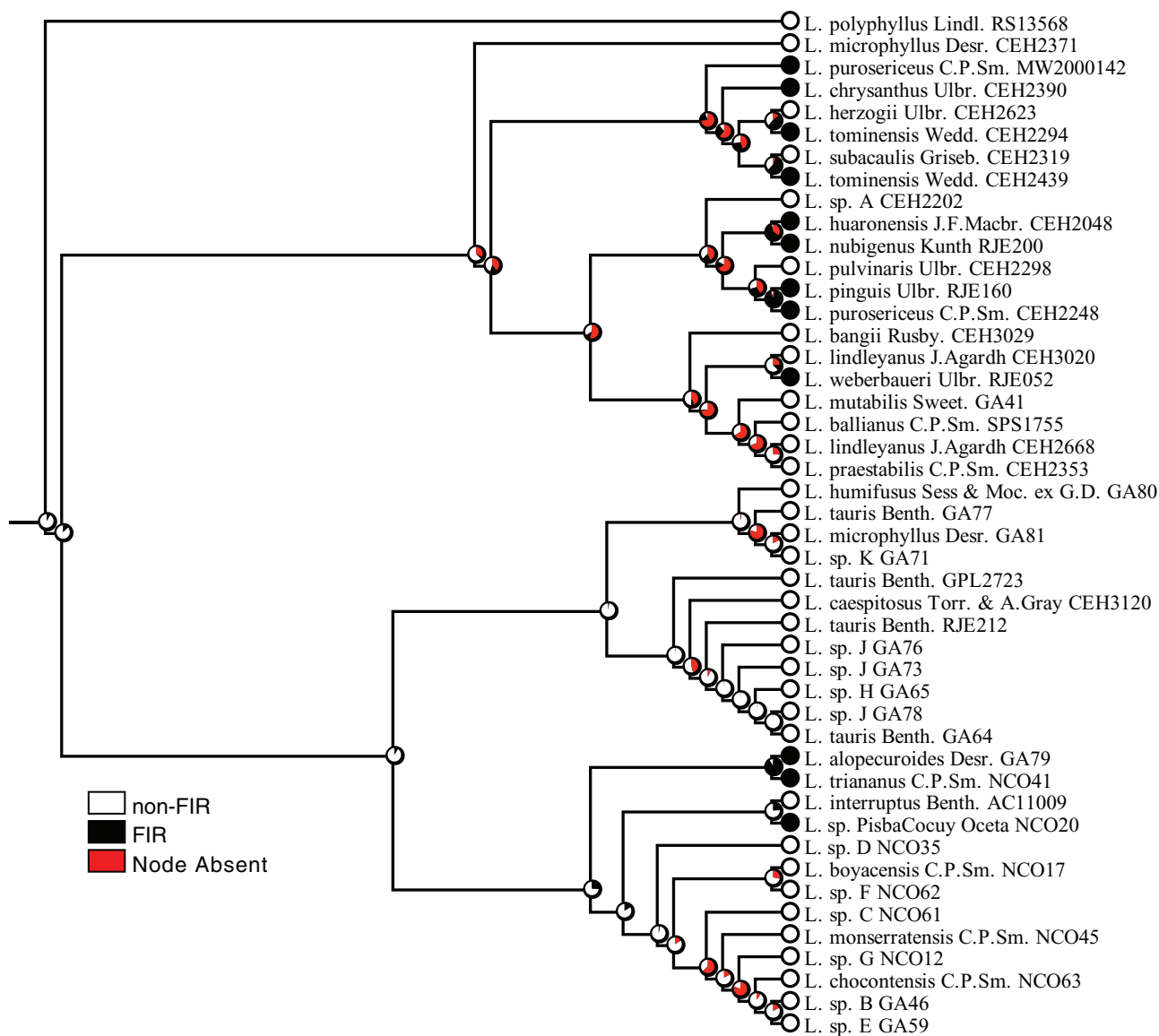


Figure 4. Maximum likelihood ASR of FIR and non-FIR growth forms using a Markov k-state-1-parameter model. The analysis was performed across the 100 topologies obtained from the rapid bootstrap analysis performed in RAxML and plotted on the best ML tree. Pie charts at internal nodes show the average likelihood received by each ancestral state (black and white) across all the trees that have that node present and the percentage of trees where the node is absent (red). FIR: fistulose-inflorescence rosette species/morphospecies; non-FIR: annual herbs, prostrate perennials, subshrubs, erect shrubs and treelets.

GENETIC STRUCTURE ANALYSIS

For both the 30acc and 31acc matrixes, STRUCTURE HARVESTER suggests that northern Andean FIR genetic structure is best explained by $K = 2$ (Fig. 3B), based on the largest value of ΔK . The four genetic clusters correspond to four putative morphologically and geographically distinct entities, *L. alopecuroides* and *L. trianaus* as sister species and ‘Ocetá’ and ‘Pisba-Cocuy’ also as sister groups (Fig. 3A). Geneflow and genetic admixture between these two sub-clades is probably infrequent or absent. The genetic differentiation between the geographically adjacent ‘Pisba-Cocuy’ and ‘Ocetá’ populations suggests almost no admixture between them, except for the one accession (NCO30) from Cocuy with c. 15% of its genetic material shared with the ‘Ocetá’ population and the other 85% assigned to the Pisba cluster. This strongly supports the hypothesis that the sampled populations from the immediately adjacent Pisba and Cocuy populations are conspecific, but leaves open the taxonomic status of the genetically distinct clades from ‘Pisba-Cocuy’ and ‘Ocetá’. In the clusters formed by *L. trianaus* and *L. alopecuroides*, a low percentage of genetic material is shared between a small number of samples of *L. alopecuroides* from Ecuador (PS10720 and PS10747) and *L. trianaus* from Sumapaz (NCO_PSum1, NCO_PSum2, NCO_PSum3 and NCO41), but in each case significantly < 50% admixture is present.

SPECIES DELIMITATION

The comparative morphology of the northern Andean FIR entities (Table 1) reveals a set of traits that vary among the four putative entities resolved by the genetic data. These include banner petal colour patterns, the disposition, length and persistence of the floral bracts during anthesis, leaflet shape and number, stipule length and form, indumentum and overall plant and inflorescence size (Fig. 5). Based on combined assessment of evidence from morphology, herbarium and field observations, phylogenetics, multi-species coalescent (Supplementary Information, Fig. S2) and genetic structure (Fig. 3), four operational taxonomic units (OTUs) can be defined: the two named species *L. alopecuroides* and *L. trianaus*; and the ‘Pisba-Cocuy’ and ‘Ocetá’ genotypes separated by only minor morphological differences that have not been assigned to any named species (Fig. 5).

DISCUSSION

The Andean *Lupinus* radiation is one of the most rapidly diversifying plant clades documented to date with an exceptionally high and apparently still accelerating rate of net species diversification (Drummond *et al.*, 2012; Hughes & Atchison, 2015). This is in line with

other páramo plant radiations that are predominantly young and fast (Madrinán *et al.*, 2013) with comparable rates of diversification to those documented for cichlid fish radiations in African lakes (Hughes & Eastwood, 2006; Seehausen, 2015) and with the somewhat older but still rapid radiations found in mid-elevation montane forests in the Andes (Lagomarsino *et al.*, 2016; Hughes, 2016). Andean *Lupinus* spp. also encompass spectacular variation in plant habit and plant height, with minute prostrate mat-forming plants, acaulescent cushion-like perennial herbs, large stem rosettes, woody shrubs and treelets to 5–6 m height and 10 cm stem diameter. Variation in growth forms is matched by similar disparification in inflorescence size, disposition and form. Furthermore, the Andean clade spans almost 5000 m of elevation with a few species at low elevation and cohorts of different species at mid, high and very high elevations, including a suite of species adapted to the extreme daily freeze-thaw conditions found between 4000 and 4900 m, close to the upper elevation limits for plant growth. This large phenotypic and ecological span suggests that the Andean *Lupinus* radiation is notable not only in terms of species diversification rates, but also shows rapid trait and ecological disparification strongly suggestive of a pronounced adaptive axis of evolutionary diversification (Hughes & Atchison, 2015) in line with other radiations in the high elevation Andean grasslands and montane forests (Nürk *et al.*, 2013; Hughes, 2016; Lagomarsino *et al.*, 2016; Diazgrandos & Barber, 2017).

Recent evidence from analysis of transcriptome data for 55 New World *Lupinus* spp., including 26 Andean accessions demonstrated unprecedented levels of accelerated natural selection acting on coding and regulatory regions genome wide and affecting up to 40% of genes among species in the rapidly diversifying Andean clade (Nevado *et al.*, 2016) also suggesting a strong adaptive component associated with the diversification of Andean lupins. In contrast, Givnish (2015) questioned the adaptive significance of the Andean *Lupinus* radiation, speculating instead that it is an example of geographically driven explosive species diversification driven primarily by dispersal limitation and geographical isolation and lacking a strong adaptive component of diversification.

In this study, we show both strong geographical phylogenetic structure indicative of dispersal limitation as a driver of diversification and the repeated occurrence of suites of locally occurring species with a rosette growth form adapted to high elevation zones within individual cordilleras in the northern Andes. The RADseq data generated here provide unprecedented resolution of relationships among the recently diverged Andean *Lupinus* spp. as demonstrated in the recent parallel study of Atchison *et al.* (2016) and in line with similar findings from other rapidly evolving

Table 1. Comparative morphology of northern fistulose-inflorescence rosette *Lupinus* spp. Elevation range and geographical distributions are based on herbarium records and field observations of sampled populations. Leaflet shapes follow Hickey (1973)

	<i>L. alopecuroides</i>	<i>L. trianaanus</i>	'Pisba-Cocuy'	'Ocetá'
Geographical distribution	Central Cordillera from Nevado del Ruiz (Colombia) to Chimborazo (Ecuador)	Chingaza and Sumapaz (East Cordillera, Colombia)	Pisba and Cocuy (East Cordillera, Colombia)	Ocetá (East Cordillera, Colombia)
Elevational range (m)	4000–4600	3700–3800	3600–4000	3500–3700
Habitat	Rocky-sand páramo, upper limits with superpáramo	Grass páramo, zones with episodic runoff of water	Grass páramo, zones with episodic runoff of water always found in association with <i>Senecio</i> spp.	Grass páramo, zones with episodic runoff of water or flooding, always found in association with <i>Senecio niveoaurus</i> Cuatrec.
General indumentum	Very dense	Dense	Sparse	Sparse
Leaflet shape	Oblanceolate	Narrow-ob lanceolate	Narrow-elliptic	Narrow-elliptic
Number of leaflets	12–14	13–16	13–17	16–18
Free-stipule length (cm)	2.0–3.4	2.0–4.0	2.3–4.6	0.6–1.8
Free-stipule form	Lanceolate	Lanceolate	Lanceolate	Triangular
Inflorescence length (cm)	40–60	50–60	70	100–110
Floral bract state during anthesis	Persistent, upcurved always concealing the flowers	Sometimes deciduous, upcurved only when flowers are in bud	Sometimes deciduous, upcurved only when flowers are in bud	Sometimes deciduous, upcurved only when flowers are in bud
Flowers	Sometimes visible, banner petal with white central blotch	Visible, banner petal with a conspicuous white central blotch	Visible, banner petal with a conspicuous white central blotch	Visible, banner petal with an incipient white central blotch restricted to the tip of the petal

clades, e.g. the African cichlids in Lake Victoria (Wagner *et al.*, 2013). Our results suggest that such data, especially when sampled densely across multiple populations and individuals of species, can generate robust hypotheses of relationships among even recently diverged species and thereby also provide evidence to test species boundaries and assist with species delimitation problems (Eaton & Ree, 2013; Wagner *et al.*, 2013; Pante *et al.*, 2015; Cavender-Bares *et al.*, 2015; Gohli *et al.*, 2015).

The proviso of dense sampling to achieve such resolution remains to be properly tested, but our results and those of Atchison *et al.*, (2016) show that support is higher in more densely sampled clades as seen for the northern Andean FIR species that are the focus of this study. By using an order of magnitude larger genome-wide DNA sequence data sets than those

employed previously by Drummond *et al.* (2012), well-resolved phylogenetic trees for Andean *Lupinus* generated using RADseq (nextRAD) technology provide the first opportunities to shed light on the trajectories of evolutionary diversification in this large radiation.

BIOGEOGRAPHY

The phylogenetic analysis shows a pattern of strong geographical structure at different scales (Fig. 3A): across the Andes as a whole; regionally among different cordilleras in the northern Andes and more locally between geographically adjacent páramos in the northern part of the Eastern Cordillera in Colombia. First, there is a marked biogeographic division between a clade of species mainly from Peru and Bolivia (central Andes clade) and a clade

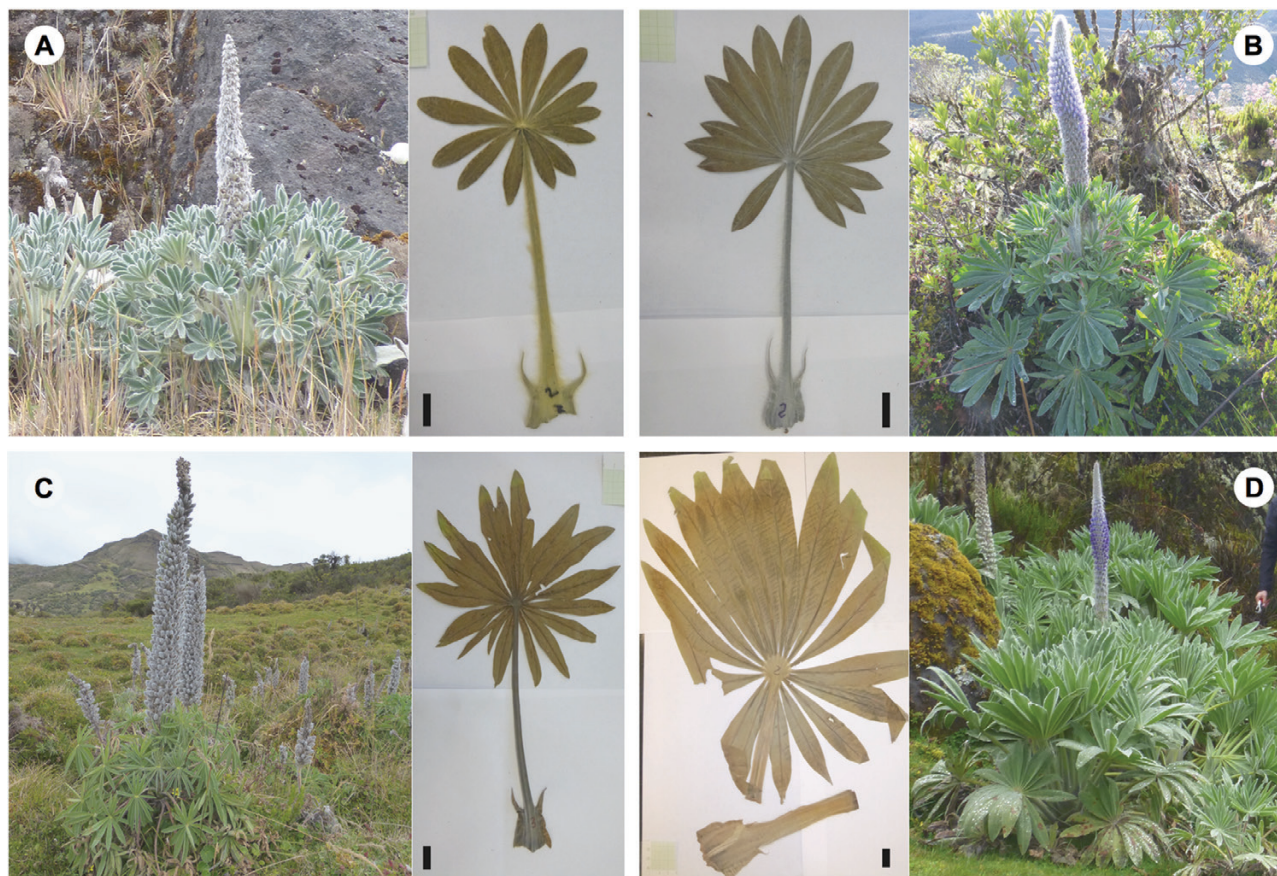


Figure 5. Colombian fistulose-inflorescence rosette *Lupinus* spp. A, *L. alopecuroides*, at 4460 m, Nevado de Santa Isabel, Caldas. B, *L. trianaus*, at 3600 m, Páramo de Sumapaz, Bogotá. C, Morphotype 'Pisba-Cocuy', at 3600 m, Páramo de Pisba, Boyacá. D, Morphotype 'Ocetá', at 3700 m, Páramo de Ocetá, Boyacá. Insets, typical leaves of each species/morphotype, scale bars to 2 cm. Photographs A, B, C and inset leaf in D, Natalia Contreras Ortiz; D habitat photograph, Gabriel S. Ortiz.

of species found from northern Peru to Colombia (northern Andes clade) suggesting a possible dispersal barrier between these areas (Fig. 3A; see also [Atchison et al., 2016](#)). Similar patterns of northern and central Andes phylogenetic structure have been documented in birds ([Bonaccorso, 2009](#); [Chaves & Smith, 2011](#); [Gutiérrez-Pinto et al., 2012](#)) and other high elevation Andean plant groups ([Ayers, 1999](#); [Cosacov et al., 2009](#); [Jabaily & Sytsma, 2013](#); [Vargas et al., 2017](#)). This break coincides with the lower elevation zone in the Andes between southern Ecuador and northern Peru around latitude *c.* 5° S referred to as the Huancabamba Depression. This well-known feature is thought to have acted as a possible dispersal barrier for montane lineages ([Simpson, 1975](#); [Parker et al., 1985](#)) and a centre of diversification for many taxa ([Berry, 1982](#); [Weigend, 2002, 2004](#)). Only one of the sampled *Lupinus* spp. spans this putative barrier, the FIR species *L. nubigenus* (Fig. 1K), which is the only member of the 'central Andes' clade

that is also found to the north of the Huancabamba depression in Ecuador. This gives at least one possible dispersal event for the sampled species across this barrier.

Second, the Andean *Lupinus* phylogenetic tree also shows a striking geographical structure in the northern Andes with distinct subclades occupying the Eastern and Central Cordillera of Colombia (Fig. 3A). Similar biogeographical patterns between taxa from the Central and Eastern Cordillera have been documented for other groups of organisms (e.g. [Cadena et al., 2007](#); [Gutiérrez-Pinto et al., 2012](#)) and suggest strong isolation and dispersal limitation between these cordilleras separated by the Río Magdalena valley. This pattern contrasts with the lack of differentiation among cordilleras found in populations of catfish in the genus *Loricaria* in the northern Andes, indicative of the extensive gene flow among populations and frequent long-distance dispersal for that particular group ([Kolar et al., 2016](#)).

Previous studies, which have sampled widely across the genus *Lupinus*, have shown that the Andean clade has a Northern Hemisphere origin (Hughes & Eastwood, 2006; Drummond *et al.*, 2012; Nevado *et al.*, 2016), but almost nothing is known about the direction of the colonization that took place in the Andes. The derived placement of clades from the Eastern Cordillera with respect to clades from the Central Cordillera in the present phylogenetic tree suggest that colonization across the northern Andes could have been from Ecuador through the Colombian Central Cordillera and subsequently to the Eastern Cordillera, i.e. apparently from south to north. Although this pattern of diversification in the northern Andes remains to be properly tested, especially with inclusion of species from the Venezuelan Andes, such a pattern has not yet been convincingly documented for other high-elevation Andean plant groups with Northern Hemisphere origins.

The well-defined clades of species distributed across the Central Cordillera through Ecuador and the grouping of populations of *L. alopecuroides* from Ecuador and Colombia (Fig. 3) suggest that connections and gene flow may have been more frequent across this range. This pattern agrees with previous studies demonstrating species and phytogeographical affinities between the Colombian Central Cordillera and the Ecuadorian super-páramos (Sklenár & Balslev, 2007) and coincides with the phytogeographical provinces proposed by van der Hammen (1997), which were statistically corroborated by Londoño *et al.* (2014) for angiosperms as a whole across the Colombian Andes.

Thirdly, there is also evidence for dispersal limitation and geographical isolation at more local scales between individual páramos, as seen for *Lupinus* in the northern part of the Eastern Cordillera. Despite the geographical proximity and the inferred sister group relationship between the 'Pisba-Cocuy' and 'Ocetá' populations, the genetic separation of populations from these páramos (Fig. 3) is striking and is also marked by a suite of minor morphological differences (Table 1, Fig. 5C, D). The distributions of these two genetically distinct units correspond to different páramo complexes (as defined by Morales *et al.*, 2007): Cocuy, Pisba and Tota-Bijagual-Mamapacha (Ocetá). The Cocuy and Pisba complexes are sometimes considered to be connected (Morales *et al.*, 2007), forming a geographical continuum of páramos with no zones < c. 3300 m. This proximity and connectivity probably explain the genetic similarity observed in the cluster formed by accessions from Pisba and the single accession from Cocuy (NCO30) (Fig. 3B). In contrast, the Ocetá páramo is more isolated from the Pisba-Cocuy páramos by a valley at 2800 m elevation formed by the Sasa River along the Mongua-Chochal-San Ignacio municipalities in Boyacá. To date there are no known

studies that explain the intrinsic factors determining plant endemism in this particular páramo, but the results presented here suggest a degree of isolation between these areas sufficient to precipitate the marked genetic and minor morphological divergences observed here.

Although there is broad agreement about the recentness of the Andean *Lupinus* radiation in the Pleistocene from time-calibrated molecular phylogenetic analyses (1.19–3.50 Mya) (Hughes & Eastwood, 2006; Drummond *et al.*, 2012) and from fossil *Lupinus* pollen (2.25 Myr) (Hooghiemstra, 1983, calibrated with the age model of Torres *et al.*, 2013) post-dating the final uplift of the northern Andes (Gregory-Wodzicki, 2000; Hay *et al.*, 2002), estimating divergence times among species and clades in the Andean clade remains extremely challenging. This means that definitive interpretation of the detailed processes and the specific flux of dispersal and vicariant isolation involved remain speculative, especially given the dramatic impacts of Pleistocene climatic oscillations and consequent elevational shifts in vegetation zones (Simpson, 1975; Hooghiemstra & van der Hammen, 2004) on areas of available páramo habitat and their connectivity/fragmentation (Flantua *et al.*, 2014). For example, possible connections between the Central Cordillera and the southern portions of the Eastern Cordillera in Colombia (Simpson, 1975; Flantua *et al.*, 2014; Hooghiemstra & Van der Hammen, 2004) could have facilitated gene flow between these areas during colder periods of greater páramo connectivity followed by vicariant isolation and divergence during warmer periods. The low level of admixture observed between accessions of *L. alopecuroides* from Ecuador (PS10720 and PS10747) and *L. trianaanus* from Sumapaz (NCO_PSum1, NCO_PSum2, NCO_PSum3 and NCO41) (Fig. 3B) may be explained by the recent split of these two species and by possible Pleistocene connections between both cordilleras that facilitated limited gene flow.

There are few observational or quantitative data on the flowering and pollination biology, breeding systems and seed dispersal potential in Andean *Lupinus*. However, the limited data on genetic variation within and among species and populations suggest high genetic uniformity within populations that form well-defined clades and clusters with little admixture between them (Fig. 3B). A recent study by Vásquez *et al.* (2016), using microsatellite markers to examine patterns of genetic variation across ten populations of *L. alopecuroides*, also found low levels of genetic diversity within populations and high genetic differentiation between populations [albeit that our results suggest that one of the populations sampled in their study (COCUY) does not in fact belong in *L. alopecuroides*; instead it should be assigned to

the putative new ‘Ocetá’-‘Pisba-Cocuy’ species identified here]. Vázquez *et al.* (2016) suggested three main demographic factors (founder effects, lack of gene flow and/or autogamy) as possible explanations for these patterns. Inbreeding has been suggested as one of the factors causing genetic uniformity within populations of another high elevation Andean plant, the perennial rosette *Puya raimondii* Harms. (Bromeliaceae) from the Peruvian Puna (Sgorbati *et al.*, 2004) and may play a role in *Lupinus*, given the small population sizes of some of the high elevation FIR species.

ADAPTIVE CONVERGENT EVOLUTION OF GROWTH FORMS

Alongside this marked geographic structure across the Andes at multiple scales, there is also evidence for multiple independent occurrences of the rosette growth form (FIR) scattered across the different Andean subclades (Fig. 4), in the northern Andes (*L. alopecuroides*-*L. trianaus* and the ‘Pisba-Cocuy’-‘Ocetá’ clades) and further south (*L. weberbaueri*, *L. chrysanthus*, *L. pinguis*, *L. nubigenus*, *L. puroserriceus*, *L. huaronensis* and *L. tominensis*) (Fig. 1). This suggests that the rosette growth form evolved independently multiple times in different parts of the Andes within the overall Andean *Lupinus* radiation and that the morphological similarity among species/morphotypes is the result of convergent evolution of growth forms adapted to the unusual and extreme daily freeze-thaw conditions found at high elevations across the tropical Andes.

Each of these independent origins of FIR species is associated with a group of species spanning a wide spectrum of growth forms most of which occur at lower elevations in these regionally disjunct groups in the Eastern Cordillera, the Central Cordillera and the south-central Andean clade (Supporting Information, Fig. S1). This pattern of convergently evolved FIR growth forms that occupy similar high-elevation habitats and are independently derived in separate subclades of lower elevation shrubs and treelets raises the intriguing possibility that Andean *Lupinus* comprises a series of replicated sub-radiations (Losos, 2010) similar to those seen in other widely studied adaptive radiations, such as the *Anolis* lizard radiations on different Caribbean islands (Mahler *et al.*, 2013) or the pelagic-benthic diversification of species associated with fish radiations in lakes (Schluter, 2000).

For Andean *Lupinus*, these ideas remain to be more fully explored both in terms of more complete sampling of the c. 85 species and more rigorous testing for convergence as a result of diversification on similar adaptive landscapes. Even in the northern Andes there are still some important taxon sampling gaps and most notably the Colombian FIR species *L. carrikeri* endemic to the Sierra Nevada de Santa Marta (Fig. 1D) and the small

set of *Lupinus* spp. from the Andes of Venezuela, none of which have been included in phylogenetic studies to date.

SPECIES DELIMITATION: A TAXONOMIC ACCOUNT OF THE FIR SPECIES FROM THE NORTHERN ANDES

The phylogenetic, multi-species coalescent (Supplementary Information Fig. S2) and genetic structure analyses (Fig. 3) all identified four distinct genetic clusters that also show variable degrees of morphological distinction (Fig. 5, Table 1). Of course, in addition to robustly supported genetic divergence (reciprocal monophyly and genetic structure), morphology is fundamental in systematic biology for the discrimination of taxa based on well-defined diagnostic characters. The integration of morphological and genetic data validates the existence of well-defined OTUs and provides explicit hypotheses for testing putative species boundaries. This sort of integrative taxonomic framework drawing on morphology, genetics, geography and ecology provides a powerful approach for species delimitation especially for alliances of closely related species. For the FIR lupin alliance from the northern Andes, four morphologically distinct OTUs (Table 1, Fig. 5) were validated by the phylogenetic and genetic structure analysis (Fig. 3) and also show allopatric distributions and distinct ecological characteristics. These results confirm the status of two species that have been long confused (*L. alopecuroides* and *L. trianaus*) and reveal an additional undescribed species.

The evidence presented here strongly supports the recognition of *L. alopecuroides*, restricted to the Central Cordillera and northern Ecuador and *L. trianaus*, restricted to the southern portion of the Eastern Cordillera and specifically the páramos of Chingaza and Sumapaz and including the type locality at Páramo Cruz Verde close to Bogotá, as distinct species (Fig. 5A and B). In addition, *L. nubigenus* (Fig. 1K) formerly treated as a dwarf variety of *L. alopecuroides* (Weddell, 1861), is here shown to be a member of the ‘Central Andes’ clade and not closely related to *L. alopecuroides*, confirming that this species is indeed distinct in line with the distinctive inflorescence types of these two species [*L. nubigenus* with a much shorter inflorescence held below the leaves (Fig. 1K) and *L. alopecuroides* with a distinctive longer, ‘foxtail’ inflorescence exerted well above the leaves (Figs 1B and 5A)]. The combined evidence also supports the recognition of the genetically distinct populations from the northern portion of the Eastern Cordillera, ‘Pisba-Cocuy’ and ‘Ocetá’, as a single species (Fig. 5C, D). Although genetic analyses (Fig. 3) including the multi-species coalescent (Supplementary Information Fig. S2) show that

these populations are genetically distinct, failure to distinguish between structure as a result of population level processes and that because of species divergence can lead to over-splitting (Sukumaran & Knowles, 2017). The morphological differences between these two entities are minor and the fact that they come from geographically closely adjacent páramos supports their recognition as a single, as yet undescribed species (Contreras-Ortiz *et al.*, in prep).

CONCLUSIONS

Previous studies based on 9kb of DNA sequence data were unable to infer species relationships in the Andean *Lupinus* radiation. RADseq data, recently used to infer the origin of domestication of the Andean crop species *L. mutabilis* (Atchison *et al.*, 2016), here provide robustly supported resolution of species relationships despite the recent divergence of this radiation, providing new insights into the evolution of growth forms, species relationships, biogeography and species limits of the northern Andean FIR *Lupinus* spp. The convergent evolution of the FIR growth form and the strong geographical structure at different scales suggest joint contributions of adaptive (ecological) and non-adaptive (geographical) components in the diversification of the Andean lupins.

Although more rigorous testing for convergence as a result of diversification on similar adaptive landscapes would be desirable, we provide first insights into the evolutionary trajectories in the Andean *Lupinus* radiation that raises the intriguing possibility that Andean *Lupinus* comprises a series of geographically restricted and replicated sub-radiations. Geographical patterns suggest dispersal limitation as a driver of diversification alongside vicariant processes potentially linked to Pleistocene glaciations that, with topographical features in the Andes such as inter-mountain valleys and low passes (e.g. the Huancabamba Depression and Río Magdalena valley), may have played an important role in generating the contemporary genetic and biogeographic patterns revealed here for the northern Andes. The combined evidence also supports the recognition of *L. nubigenus*, *L. alopecuroides* and *L. trianaus* as distinct species and a putative new species encompassing the 'Pisba-Cocuy' and 'Ocetá' populations.

ACKNOWLEDGEMENTS

This work was supported by funds from the Swiss National Science Foundation (Grant 31003A_135522 to CEH), a Seed Grant from the Facultad de Ciencias, Universidad de los Andes and the van der Hammen

Program of Support of the Jardín Botánico José Celestino Mutis of Bogotá. We thank Mario Contreras, Luis H. Contreras, Ana Ortiz, Alejandro Contreras, Daniel Rojas, Nicole Yepes, Nicolás Rodríguez, Nicolai Nürk, Mauricio Andres Rojas and Gabriel Ortiz for assistance on field trips and with data collection, Daniel Cadena, Andrés J. Cortés, Mauricio Diazgranados, Adolfo Jara and Fausto Rodríguez for discussion of ideas and Jorge López for help with the map in Figure 2, the national authorities in Bolivia, Colombia, Ecuador and Peru for permission to collect plant material, the Botanical Garden and Department of Systematic and Evolutionary Botany at the University of Zurich for workspace and facilities provided during NCO's visit, the ScienceCloud service of S3IT, University of Zurich for computational support, Paul Etter and Eric Johnson of SNPSaurus, Oregon, for guidance and help with RAD sequencing and Petr Sklenář for collecting and contributing plant material from Ecuador used in this study.

REFERENCES

- Aïnouche AK, Bayer RJ. 1999. Phylogenetic relationships in *Lupinus* (Fabaceae: Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. *American Journal of Botany* **86**: 590–607.
- Aïnouche A, Bayer RJ, Misset MT. 2004. Molecular phylogeny, diversification and character evolution in *Lupinus* (Fabaceae) with special attention to Mediterranean and African lupines. *Plant Systematics & Evolution* **246**: 211–222.
- Ackerley D. 2009. Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proceedings National Academy of Sciences* **106**: 19699–19706.
- Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews. Genetics* **17**: 81–92.
- Atchison GW, Nevado B, Eastwood RJ, Contreras-Ortiz N, Reynel C, Madriñán S, Filatov DA, Hughes CE. 2016. Lost crops of the Incas: origins of domestication of the Andean pulse crop tarwi, *Lupinus mutabilis*. *American Journal of Botany* **103**: 1592–1606.
- Ayers TJ. 1999. Biogeography of *Lysipomia* (Campanulaceae), a high elevation endemic: an illustration of species richness at the Huancabamba Depression, Peru. *Arnaldoa* **6**: 13–28.
- Bell CD, Donoghue MJ. 2005. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms Diversity and Evolution* **5**: 147–159.
- Berry PE. 1982. The systematics and evolution of *Fuchsia* sect. *Fuchsia* (Onagraceae). *Annals of The Missouri Botanical Garden* **69**: 1–198.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* **30**: 2114–2120.

- Bonaccorso E. 2009.** Historical biogeography and speciation in the neotropical highlands: molecular phylogenetics of the jay genus *Cyanolyca*. *Molecular Phylogenetics and Evolution* **50**: 618–632.
- Cadena CD, Klicka J, Ricklefs RE. 2007.** Evolutionary differentiation in the Neotropical montane region: molecular phylogenetics and phylogeography of Buarremon brush-finches (Aves, Emberizidae). *Molecular Phylogenetics and Evolution* **44**: 993–1016.
- Carlquist S. 1974.** *Island biology*. New York & London: Columbia University Press.
- Cavender-Bares J, González-Rodríguez A, Eaton DA, Hipp AA, Beulke A, Manos PS. 2015.** Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. *Molecular Ecology* **24**: 3668–3687.
- Chacón J, Madriñán S, Chase MW, Bruhl JJ. 2006.** Molecular phylogenetics of *Oreobolus* (Cyperaceae) and the origin and diversification of the American species. *Taxon* **55**: 359–366.
- Chaves JA, Smith TB. 2011.** Evolutionary patterns of diversification in the Andean hummingbird genus *Adelomyia*. *Molecular Phylogenetics and Evolution* **60**: 207–218.
- Chen L-Y, Wang Q-F, Renner SS. 2016.** East Asian Lobelioideae and ancient divergence of a giant rosette *Lobelia* in Himalayan Bhutan. *Taxon* **65**: 293–304.
- Conterato IF, Schifino-Wittmann MT. 2006.** New chromosome numbers, meiotic behaviour and pollen fertility in American taxa of *Lupinus* (Leguminosae): contributions to taxonomic and evolutionary studies. *Botanical Journal of the Linnean Society* **150**: 229–240.
- Cosacov A, Sersic AN, Sosa V, De-Nova JA, Nylinder S, Cocucci AA. 2009.** New insights into the phylogenetic relationships, character evolution, and phytogeographic patterns of *Calceolaria* (Calceolariaceae). *American Journal of Botany* **96**: 2240–2255.
- Davey JL, Blaxter MW. 2010.** RADseq: next-generation population genetics. *Briefings in Functional Genomics* **9**: 416–423.
- Davey JW, Cezard T, Fuentes-Utrilla P, Eland C, Gharbi K, Blaxter ML. 2013.** Special features of RAD sequencing data: implications for genotyping. *Molecular Ecology* **22**: 3151–3164.
- Diazgranados M, Barber JC. 2017.** Geography shapes the phylogeny of frailejones (Espeletiinae Cuatrec., Asteraceae): a remarkable example of recent rapid radiation in sky islands. *PeerJ* **5**: e2968.
- Doyle JJ, Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.
- Drummond CS, Eastwood RJ, Miotto ST, Hughes CE. 2012.** Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): testing for key innovation with incomplete taxon sampling. *Systematic Biology* **61**: 443–460.
- Earl DA, vonHoldt BM. 2012.** STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.
- Eaton DA, Ree RH. 2013.** Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* **62**: 689–706.
- Eaton DA. 2014.** PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics (Oxford, England)* **30**: 1844–1849.
- Eaton DA, Spriggs EL, Park B, Donoghue MJ. 2016.** Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology* 1–14.
- Emerson KJ, Conn JE, Bergo ES, Randel MA, Sallum MA. 2015.** Brazilian *Anopheles darlingi* root (Diptera: Culicidae) clusters by major biogeographical region. *PLoS One* **10**: e0130773.
- Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.
- Flantua S, Hooghiemstra H, Van Boxel JH, Cabrera M, González-Carranza Z, González-Arango C. 2014.** Connectivity dynamics since the Last Glacial Maximum in the northern Andes: a pollen-driven framework to assess potential migration. In: Stevens WD, Montiel OM, Raven P, eds. *Paleobotany and biogeography: a festschrift for Alan Graham in his 80th year*. St. Louis: Botanical Garden Press, 98–123.
- Givnish TJ. 2010.** Giant lobelias exemplify convergent evolution. *BMC Biology* **8**: 1–4.
- Givnish TJ. 2015.** Adaptive radiation versus ‘radiation’ and ‘explosive diversification’: why conceptual distinctions are fundamental to understanding evolution. *The New Phytologist* **207**: 297–303.
- Givnish TJ, Millam KC, Mast AR, Paterson TB, Theim TJ, Hipp AL, Hess JM, Smith JF, Wood KR, Sytsma KJ. 2009.** Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society of London B: Biological Sciences* **276**: 407–416.
- Gohli J, Leder EH, Garcia-Del-Rey E, Johannessen LE, Johnsen A, Laskemoen T, Popp M, Liffield JT. 2015.** The evolutionary history of Afrocanarian blue tits inferred from genomewide SNPs. *Molecular Ecology* **24**: 180–191.
- Gregory-Wodzicki KM. 2000.** Uplift history of the central and northern Andes: a review. *GSA Bulletin* **112**: 1091–1105.
- Gutiérrez-Pinto N, Cuervo AM, Miranda J, Pérez-Emán JL, Brumfield RT, Cadena CD. 2012.** Non-monophyly and deep genetic differentiation across low-elevation barriers in a Neotropical montane bird (*Basileuterus tristriatus*; Aves: Parulidae). *Molecular Phylogenetics and Evolution* **64**: 156–165.
- Hay WW, Soeding E, DeConto RM, Wold CN. 2002.** The Late Cenozoic uplift – climate change paradox. *International Journal of Earth Sciences* **91**: 746–774.
- Hedberg I, Hedberg O. 1979.** Tropical-alpine life forms of vascular plants. *Oikos* **33**: 297–307.
- Hickey L. 1973.** Classification of the architecture of dicotyledonous leaves. *American Journal of Botany* **60**: 17–33.

- Hooghiemstra H. 1983.** *Vegetational and climatic history of the Bogotá high plain, Colombia*. PhD Thesis, University of Amsterdam, the Netherlands.
- Hooghiemstra H, van der Hammen T. 2004.** Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **359**: 173–180.
- Hughes CE, Atchison GW. 2015.** The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *The New Phytologist* **207**: 275–282.
- Hughes CE, Nyffeler R, Linder HP. 2015.** Evolutionary plant radiations: where, when, why and how? *The New Phytologist* **207**: 249–253.
- Hughes CE, Eastwood R. 2006.** Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 10334–10339.
- Hughes CE. 2016.** The tropical Andean plant diversity powerhouse. *The New Phytologist* **210**: 1152–1154.
- Jabaily RS, Sytsma KJ. 2013.** Historical biogeography and life-history evolution of Andean *Puya* (Bromeliaceae). *Botanical Journal of the Linnean Society* **171**: 201–224.
- Käss E, Wink M. 1997.** Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the *rbcL* gene and ITS 1 + 2 regions of rDNA. *Plant Systematics and Evolution* **208**: 139–167.
- Knight R, Maxwell P, Birmingham A, Carnes J, Caporaso JG, Easton BC, Eaton M, Hamady M, Lindsay H, Liu Z, Lozupone C, McDonald D, Robeson M, Sammut R, Smit S, Wakefield MJ, Widmann J, Wikman S, Wilson S, Ying H, Huttley GA. 2007.** PyCogent: a toolkit for making sense from sequence. *Genome Biology* **8**: R171.
- Kolář F, Dušková E, Sklenář P. 2016.** Niche shifts and range expansions along cordilleras drove diversification in a high-elevation endemic plant genus in the tropical Andes. *Molecular Ecology* **25**: 4593–4610.
- Lagomarsino LP, Condamine FL, Antonelli A, Mulch A, Davis CC. 2016.** The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *The New Phytologist* **210**: 1430–1442.
- Londoño C, Cleef AM, Madriñán S. 2014.** Angiosperm flora and biogeography of the páramo region of Colombia, northern Andes. *Flora - Morphology, Distribution, Functional Ecology of Plants* **209**: 81–87.
- Losos JB, Malher L. 2010.** Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. In: Bell MA, Futuyma DJ, Eanes WF, Levinton JS, eds. *Evolution since Darwin: the first 150 years*. Sunderland: Sinauer Associates, 381–420.
- Losos JB. 2010.** Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist* **175**: 623–639.
- Losos JB, Ricklefs RE. 2009.** Adaptation and diversification on islands. *Nature* **457**: 830–836.
- Luteyn JL. 1999.** *Páramos: a checklist of plant diversity, geographical distribution, and botanical literature*. New York: The New York Botanical Garden Press.
- Maddison WP, Maddison DR. 2017.** *Mesquite: a modular system for evolutionary analysis*. Version 3.2. Available at: <http://mesquiteproject.org> (accessed 10 May 2017).
- Madriñán S, Cortés AJ, Richardson JE. 2013.** Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Frontiers in Genetics* **4**: 192.
- Mahler DL, Ingram T, Revell LJ, Losos JB. 2013.** Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science (New York, N.Y.)* **341**: 292–295.
- Morales M, Otero J, Van der Hammen T, Torres A, Cadena C, Pedraza C, Rodríguez N, Franco C, Betancourth JC, Olaya E, Posada E, Cárdenas L. 2007.** *Atlas de páramos de Colombia*. Bogotá: Instituto de Investigación de Recursos Biológicos Alexander von Humboldt.
- Nevado B, Atchison GW, Hughes CE, Filatov DA. 2016.** Widespread adaptive evolution during repeated evolutionary radiations in New World lupins. *Nature Communications* **7**: 12384.
- Nürk NM, Scheriau C, Madriñán S. 2013.** Explosive radiation in high Andean *Hypericum*-rates of diversification among New World lineages. *Frontiers in Genetics* **4**: 175.
- Pante E, Abdelkrim J, Viricel A, Gey D, France SC, Boisselier MC, Samadi S. 2015.** Use of RAD sequencing for delimiting species. *Heredity* **114**: 450–459.
- Parker TA, Schulenberg TS, Graves GR, Braun MJ. 1985.** The avifauna of the Huancabamba Region, northern Peru. *Ornithological Monographs* **36**: 169–197.
- Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Rosser N, Phillimore AB, Huertas B, Willmott KR, Mallet J. 2012.** Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America. *Biological Journal of the Linnean Society* **105**: 479–497.
- Russello MA, Waterhouse MD, Etter PD, Johnson EA. 2015.** From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ* **3**: e1106.
- Sanderson MJ. 1998.** Reappraising adaptive radiation. *American Journal of Botany* **85**: 1650–1655.
- Särkinen T, Pennington RT, Lavin M, Simon MF, Hughes CE. 2012.** Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. *Journal of Biogeography* **39**: 884–900.
- Scherson RA, Vidal R, Sanderson MJ. 2008.** Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. *American Journal of Botany* **95**: 1030–1039.
- Schluter D. 2000.** *The ecology of adaptive radiation*. New York: Oxford University Press.
- Seehausen O. 2006.** African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B* **273**: 1987–1998.

- Seehausen O.** 2015. Process and pattern in cichlid radiations – inferences for understanding unusually high rates of evolutionary diversification. *The New Phytologist* **207**: 304–312.
- Sgorbati S, Labra M, Grugni E, Barcaccia G, Galasso G, Boni U, Mucciarelli M, Citterio S, Benavides Iramátegui A, Venero Gonzales L, Scannerini S.** 2004. A survey of genetic diversity and reproductive biology of *Puya raimondii* (Bromeliaceae), the endangered queen of the Andes. *Plant Biology (Stuttgart, Germany)* **6**: 222–230.
- Simpson BB.** 1974. Glacial migrations of plants: island biogeographical evidence. *Science (New York, N.Y.)* **185**: 698–700.
- Simpson BB.** 1975. Pleistocene changes in the flora of the high tropical Andes. *Paleobiology* **1**: 273–294.
- Sklenář P, Balslev H.** 2007. Geographic flora elements in the Ecuadorian superpáramo. *Flora* **202**: 50–61.
- Smith CP.** 1930–1953. *Species lupinorum*. Saratago: Privately published.
- Stamatakis A.** 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* **30**: 1312–1313.
- Sukumaran J, Knowles LL.** 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences* **114**: 1607–1612.
- Torres V, Hooghiemstra H, Lourens L, Tzedakis PC.** 2013. Astronomical tuning of long pollen records reveals the dynamic history of montane biomes and lake levels in the tropical high Andes during the Quaternary. *Quaternary Science Reviews* **63**: 59–72.
- Uribe-Convers S, Settles ML, Tank DC.** 2016. A phylogenomic approach based on PCR target enrichment and high throughput sequencing: resolving the diversity within the South American species of *Bartsia* L. (Orobanchaceae). *PLoS One* **11**: e0148203.
- Uribe-Convers S, Tank DC.** 2015. Shifts in diversification rates linked to biogeographic movement into new areas: an example of a recent radiation in the Andes. *American Journal of Botany* **102**: 1854–1869.
- Vargas OM, Ortiz EM, Simpson BB.** 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *The New Phytologist* **214**: 1736–1750.
- van der Hammen T.** 1997. Ecosistema terrestre: páramo. In: Chaves M, Arango N, eds. *Informe nacional sobre el estado de la biodiversidad*. Bogotá: Investigación de Recursos Biológicos Alexander von Humboldt, 10–37.
- Vásquez DLA, Balslev H, Hansen MM, Sklenář P, Romoleroux K.** 2016. Low genetic variation and high differentiation across sky island populations of *Lupinus alopecuroides* (Fabaceae) in the northern Andes. *Alpine Botany* **126**: 135–142.
- von Hagen KB, Kadereit JW.** 2001. The phylogeny of *Gentianella* (Gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. *Organisms Diversity & Evolution* **1**: 61–79.
- Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O.** 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology* **22**: 787–798.
- Weddell HA.** 1861. *Lupinus alopecuroides* var. *nubigenus* (Kunth) Wedd. In: Weddell HA. *Chloris andina*. Paris: P. Bertrand, **2**: 250.
- Weigend M.** 2002. Observations on the biogeography of the Amotape-Huancabamba Zone in northern Peru. *The Botanical Review* **68**: 38–54.
- Weigend M.** 2004. Additional observations on the biogeography of the Amotape-Huancabamba zone in northern Peru: defining the south-eastern limits. *Revista Peruana de Biología* **11**: 127–134.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Maximum likelihood phylogenetic tree of Andean *Lupinus* generated using RAxML based on the 124acc data matrix that comprises 2 151 174 SNPs. The tree is rooted with *Lupinus polyphyllus* RS13568. FIR species/accessions are in bold. Bootstrap support values (100 replicates) are shown at the nodes.

Figure S2. Majority-rule consensus species tree with bootstrap support (100 replicates) constructed by joining the quartets inferred from the full SNP alignment (291 172 SNPs) of 32 samples (all 35 960 possible quartets) from the northern Andes clade using *tetrad*. Tree rooted for visualization with *Lupinus polyphyllus* RS13568.

Table S1. Voucher data and NCBI SRA numbers for accessions of FIR and non-FIR species and morphotypes included in phylogenetic and genetic structure analyses. FIR taxa included in the 30acc (*) and 31acc (**) matrix for genetic structure analysis.