



**HAL**  
open science

# Evolution et adaptation des Espeletiinae dans les Andes tropicales

Charles Pouchon

► **To cite this version:**

Charles Pouchon. Evolution et adaptation des Espeletiinae dans les Andes tropicales. Biodiversité et Ecologie. Université Grenoble Alpes, 2018. Français. NNT : 2018GREAV018 . tel-01989094

**HAL Id: tel-01989094**

**<https://theses.hal.science/tel-01989094v1>**

Submitted on 22 Jan 2019

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

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

## THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTE  
UNIVERSITE GRENOBLE ALPES**

Spécialité : **Biodiversité, Écologie, Environnement**

Arrêté ministériel : 25 mai 2016

Présentée par

**Charles POUCHON**

Thèse dirigée par **Sébastien Lavergne**  
et codirigée par **Jésus Mavarez**

préparée au sein du **Laboratoire d'Écologie Alpine (LECA)**  
dans l'École Doctorale **Chimie Science du Vivant**

# Évolution des Rosettes Géantes dans les Páramos

Thèse soutenue publiquement le **31 août 2018**,  
devant le jury composé de :

**Mr François POMPANON**

Professeur des Universités, UGA – Grenoble, Président

**Mr Jérôme CHAVE**

Directeur de Recherche, EDB – Toulouse, Rapporteur

**Mr Thomas HAEVERMANS**

Maitre de Conférence, ISYEB MNHN – Paris, Rapporteur

**Mr Nadir ALVAREZ**

Directeur de Recherche, MNHN – Genève, Examineur

**Mr Sébastien LAVERGNE**

Chargé de Recherche, LECA – Grenoble, Directeur de Thèse

**Mr Jésus MAVAREZ**

Chargé de Recherche, LECA – Grenoble, Co-directeur de Thèse



# Evolution of the Giant Rosettes in Tropical Alpine Ecosystems

ESPELETIINAE (ASTERACEAE; HELIANTHEAE)



by  
**CHARLES POUCHON**

Supervised by  
**SEBASTIEN LAVERGNE, SERGE AUBERT & JESUS MAVAREZ**



---

À Serge Aubert (1966–2015),

Un professeur, un voyageur, un naturaliste, et  
avant tout un ami, sans qui tout ce travail  
n'aurait jamais vu le jour.

Tu resteras toujours pour moi ce que  
Humboldt fut à la biogéographie ou encore  
Darwin à l'évolution.

Je ne t'en remercierais jamais assez.

# Content

<b>Content .....</b>	<b>4</b>
<b>Preface .....</b>	<b>7</b>
<b>INTRODUCTION – Evolution of Species, Tropical Alpine Ecosystems &amp; Espeletiinae.....</b>	<b>II</b>
<b>1   L'EVOLUTION DES ESPECES.....</b>	<b>12</b>
<b>2   LES RADIATIONS EVOLUTIVES .....</b>	<b>14</b>
<b>3   PARAMOS ET ROSETTES GÉANTES .....</b>	<b>19</b>
<b>4   RADIATION DES ESPELETIINAE: CONTEXTE PHYLOGENÉTIQUE ET PROBLÉMATIQUE .....</b>	<b>26</b>
<b>CHAPTER 1 – Phylogenomic Analysis of the Radiation of Espeletiinae in the Tropical Andes.....</b>	<b>32</b>
<b>1   INTRODUCTION.....</b>	<b>31</b>
The Exceptional Diversity of Espeletiinae (Asteraceae).....	33
Earlier Works of Phylogenetics and Evolution in Espeletiinae .....	35
<b>2   METHODS.....</b>	<b>36</b>
Sampling, DNA Extraction, and Sequencing.....	36
Assembly of Genomic Regions .....	37
Phylogenetic Analyses .....	38
Divergence Times and Diversification Rates of Espeletiinae .....	40
Reconstruction of Ancestral States .....	42
<b>3   RESULTS .....</b>	<b>42</b>
Phylogenetic Analysis of Plastomes.....	42
Phylogenomic Analysis of Nuclear Contigs .....	44
Divergence Times of Espeletiinae .....	46
<b>4   DISCUSSION.....</b>	<b>49</b>
Phylogeny and Gene Conflicts .....	50
A Phylogenomic Analysis of the Radiation of Espeletiinae .....	52
Morphological Evolution .....	54

A New Systematic Proposal for Espeletiinae ..... 56  
 5 | SUPPLEMENTARY INFORMATIONS .....57

**CHAPTER 2 –Specific, Morphological and Ecological Diversification of Espeletiinae.....66**  
 1 | INTRODUCTION..... 65  
 2 | METHODS..... 69  
     Morphological and Ecological Traits Acquisition ..... 69  
     Taxonomic Sampling, DNA Extraction and Shotgun Sequencing ..... 70  
     Alignments, Phylogenetic inference and Dating ..... 71  
     Morphological and Ecological correlation ..... 72  
     Diversification Analyses ..... 72  
 3 | RESULTS .....75  
     Phylogenetic inference ..... 75  
     Time-dependent diversification..... 76  
     Patterns of Ecological divergence..... 78  
     Patterns of phenotypic divergence ..... 78  
     Traits-dependent diversification ..... 79  
 4 | DISCUSSION ..... 80  
     Phylogenetic framework ..... 80  
     Ecological opportunity and paleotemperatures led to a burst of lineages diversification... 81  
     Phenotypic responses to changing ecological opportunities drove speciation..... 83  
 5 | CONCLUSION ..... 84  
 6 | SUPPLEMENTARY INFORMATIONS..... 85

**CHAPTER 3 – Genomic footprints of hybridisation in the evolution of Espeletiinae.....90**  
 1 | INTRODUCTION ..... 89  
 2 | METHODS..... 93  
     Sampling, Sequencing and Data Acquisition..... 93  
     Patterson’s D-statistic..... 95  
     Genetic clustering analyses ..... 96  
     Ecological divergence analyses..... 97  
 3 | RESULTS ..... 98  
     D-statistic inferences..... 98  
     Genetic structuration ..... 99  
     Ecological divergence ..... 101  
 4 | DISCUSSION .....103  
     Hybrid speciation hypothesis in *L. spectabilis*..... 103  
     Introgression between superpáramos species..... 107  
     Hybridisation and Evolution in Espeletiinae ..... 108  
 5 | SUPPLEMENTARY INFORMATIONS.....109

**DISCUSSION – Adaptive radiation of Espeletiinae in tropical alpine ecosystems..... II3**  
 I | RADIATION DES ESPELETIINAE DANS LES PARAMOS .....II3  
     Patrons Globaux ..... 113  
     Diversification au sein de la radiation vénézuélienne ..... 115

2	UNE RADIATION ADAPTATIVE ? .....	I20
3	CONCLUSION .....	I24
	Une diversification dictée par des facteurs biotiques et abiotiques.....	124
	Perspectives .....	125

<b>APPENDIX – Shotgun versus RAD-sequencing for phylogenetic issues in</b>		
<b>Espeletiinae.....I28</b>		
1	INTRODUCTION .....	I27
2	METHODS.....	I30
	Taxonomic Sampling.....	130
	ddRAD-sequencing Library preparation .....	130
	Post-sequencing processing.....	131
	Data Exploration and Phylogenetic Inference.....	132
3	RESULTS .....	I33
	Datasets characteristics .....	133
	Phylogenetic inferences and Loci Representation .....	134
4	DISCUSSION.....	I35
	Phylogenetic conflicting relationships .....	136
	Evolutionary reconstructions agreement .....	137
5	CONCLUSION .....	I38
6	SUPPLEMENTARY INFORMATIONS.....	I39
Literature cited .....		I4I



## Preface

“Les plantes semblent avoir été semées avec profusion sur la terre, comme les étoiles dans le ciel, pour inviter l’homme, par l’attrait du plaisir et de la curiosité, à l’étude de la nature: mais les astres sont placés loin de nous; il faut des connaissances préliminaires, des instruments, des machines, de bien longues échelles pour les atteindre et les rapprocher à notre portée. Les plantes y sont naturellement; elles naissent sous nos pieds et dans nos mains, pour ainsi dire; et si la petitesse de leurs parties essentielles les dérobe quelquefois à la simple vue, les instruments qui les y rendent sont d’un beaucoup plus facile usage que ceux de l’astronomie.”

Jean-Jacques Rousseau, 1778

Sur les pas de José Cuatrecasas (1903-1996), un grand botaniste espagnol du 20<sup>ème</sup> siècle, la Station alpine Joseph Fourier (SAJF), dirigée par Serge Aubert, et le Laboratoire d’Ecologie Alpine (LECA) de Grenoble, mené par Jésus Mavarez, ont lancé un projet ambitieux dans les écosystèmes tropicaux alpins des Andes du nord, les páramos, pour étudier les mécanismes d’évolution des plantes dans ces milieux caractérisés par les plus hauts taux de diversification connus au monde.

Les Espeletiinae, un groupe d’Astéracées endémique des páramos, forment le modèle d’étude principal de ce projet. En moins de 3 millions d’années, c’est-à-dire depuis la surrection des Andes du nord, un ancêtre unique a donné naissance à plus de 140 espèces de cette sous-tribue, morphologiquement très diverses et qui occupent aujourd’hui quasiment tous les habitats des páramos. On retrouve ainsi à l’intérieur de ce groupe aussi bien des arbres ramifiés que des plantes à rosettes acaules (à même le sol) ou encore à rosettes géantes caulescentes (avec un tronc) caractéristiques de cette sous-tribue. Cette dernière forme de vie apparaît être une adaptation clé dans les milieux alpins tropicaux du fait des nombreuses valeurs adaptatives qui lui sont associées (par exemple, la protection thermique des bourgeons au sein de la rosette et la rétention d’eau dans la moelle du tronc) et de sa présence au niveau d’autres écosystèmes alpins tropicaux comme chez les genres *Dendrosenecio* d’Afrique de

l'Est ou *Lobelia* des îles Hawaii. Les Espeletiinae constitueraient un exemple majeur de radiation adaptative dans le monde végétal. Cependant, liées à l'évolution très rapide de ces plantes, les relations phylogénétiques déjà étudiées à l'intérieur de ce groupe n'apparaissent aujourd'hui que très peu robustes en raison du manque de résolution des méthodes génétiques classiques utilisés jusqu'à présent. Cependant, la compréhension des mécanismes évolutifs à l'origine de cette diversification si rapide est aujourd'hui accessible grâce aux techniques de séquençage de nouvelles générations.

Le manuscrit suivant, résultant de la thèse que j'ai obtenue en octobre 2014 et dont le salaire est assuré par une bourse de l'école doctorale de l'université Grenoble Alpes, s'inscrit dans le cadre de ce projet en s'intéressant à l'évolution des Espeletiinae par l'utilisation de données génomiques issues de séquençage haut-débit. Mon intérêt pour ces plantes et ce projet a vu le jour dès la fin de ma troisième année de licence et au cours de ma première année de master. J'ai eu l'occasion de rencontrer Serge Aubert en tant que professeur puis en tant que responsable d'un premier stage effectué à la Station Alpine du Lautaret. Serge m'a très vite transmis sa passion pour la recherche et les découvertes scientifiques acquises au cours de ses nombreux voyages dans les Andes. À la fin de ma première année de master, il m'a ensuite proposé d'intégrer ce fabuleux projet avec un premier stage en co-encadrement avec Jésus Mavarez, puis un second au cours de mon master 2. Par la suite, c'est tout naturellement que j'ai rejoint ce projet de thèse. La disparition de Serge lors de la première année fut une épreuve très douloureuse pour toutes les personnes le connaissant en raison de son immense gentillesse et générosité. Serge était allé pour la première fois dans les páramos juste après avoir soutenu sa thèse. Afin de lui rendre hommage, ce projet, qui fut son dernier voyage scientifique, a continué avec Jésus Mavarez pour pouvoir aller au bout de son entreprise.

Je tenais tout naturellement à dédier ce travail à Serge, sans qui je n'en serais pas là actuellement, ainsi qu'à toute sa famille. Je remercie également Jésus Mavarez et Sébastien Lavergne d'avoir repris l'encadrement de ce travail, chose non aisée après un tel événement. Toutes les discussions et les prises de têtes que l'on a pu avoir n'ont rendu que meilleur l'aboutissement de ce manuscrit. Je remercie également tous mes amis et collègues du LECA pour les conseils, leurs aides et les bons moments passés avec vous : le bureau 313 tout d'abord avec Anthony, Clément, Céline, Fred, Monica et Tristan ; le bureau Darwin avec Marta, Stéphanie et Thibaut ; la plateforme AEEM avec Aurélie, Christian et Delphine pour les coups de mains lors de manipulations ; l'équipe XPAD que j'aurais beaucoup taquiné mais que j'aime finalement beaucoup (même si l'écotoxicologie n'est pas faite pour moi) ; Ludo et Roberto pour les debriefs les lendemains de match du FCG ; Maya et Julien pour ces cafés matinaux ; tous ceux du premier et plus particulièrement Arnaud et Stéphane ; toute l'équipe du foot et de rugby du LECA ; Florian Boucher pour les aides et discussions quant aux définitions mêmes des radiations évolutives ; Christophe Perrier (SAJF) pour tous les dessins réalisés dans ce manuscrit ; et bien sûr François qui m'a fortement aidé aussi bien pour l'enseignement que moralement dans les moments difficiles. Je tenais également à remercier les membres de mon comité de thèse pour tous leurs conseils et bienveillances quant à l'aboutissement de ce projet de thèse, à savoir Astrid Cruaud, Myriam Gaudeul, Frédéric Boyer et Sylvain Charlat. Je remercie également la société botanique de France et le Génoscope pour l'aide financière apportée pour la réalisation de séquençage haut-débit, ainsi

---

que le Muséum d'Histoire Naturelle de Grenoble pour la soutenance publique de ces travaux au sein de leurs locaux et tous les membres constituant du jury de thèse pour l'évaluation de ce travail, à savoir Laurence Desprès, Jérôme Chave, Thomas Haevermans, Nadir Alvarez et mes directeurs de thèse. Finalement je tenais à remercier ma famille, ma belle famille et surtout ma femme, Julie, pour m'avoir été d'un immense soutien tout au long de cette aventure.

# INTRODUCTION – Evolution of Species, Tropical Alpine Ecosystems & Espeletiinae

---

*“Nomina si nescis, perit cognitio rerum. Et nomina si perdas, certe distinctio rerum perditur.”* If you do not know the names of things, the knowledge of things themselves perishes; and, if you lose the names, the distinction of the things is certainty lost.

Edward Coke, 1628

Du début du 16<sup>ème</sup> jusqu’au 18<sup>ème</sup> siècle, aux travers de voyages et d’expéditions spectaculaires, de nombreux spécimens de végétaux viennent compléter les collections et les jardins européens. Face à ces découvertes croissantes, la classification des organismes selon des règles communes devient alors un réel besoin. Nommer *et classer ce que l’on observe fut ainsi l’œuvre initiée par le naturaliste suédois Carl Von Linné (1707-1778) à l’origine de la nomenclature binomiale des espèces dans son ouvrage Systema Naturae (Linné, 1758).* Depuis Aristote (384-322 av. J.C.), les organismes étaient classés selon un ordre de complexité d’après la scala naturae où chaque forme vivante, vu comme parfaite, avait un rang fixe précis au sein de cette échelle. Cette notion fixiste des espèces est retranscrite bien plus tard dans l’œuvre de Linné où les organismes qui se ressemblent, issus de la création selon un même plan, se voient regroupés dans un système hiérarchisé en sept catégories encore aujourd’hui utilisé (Règne – Embranchement – Classe – Ordre – Famille – Genre – Espèce). Nombreux de ses contemporains reconnaissent l’effort de Linné quant à cette hiérarchisation, notamment pour sa classification du règne végétal basée sur les ressemblances visibles des organes sexuels des fleurs. Néanmoins, empreinte de vision créationniste, la systématique linnéenne n’est pas acceptée par l’ensemble de ses successeurs.

La notion de transformisme apparaît ainsi au XVIII<sup>ème</sup> siècle notamment avec les travaux de Lamarck (1744-1829), fondant les prémices de l'évolutionnisme. En effet, en s'appuyant sur la paléontologie et les archives fossiles découvertes, Lamarck fonde une théorie en 1809, selon laquelle les espèces ne sont pas fixées mais peuvent se transformer en d'autres lignées en cours du temps (Lamarck, 1809). Deux principes viennent justifier cette théorie. Le premier repose sur le développement ou la perte des caractères au fil du temps en fonction de l'usage ou non de ces derniers par les organismes. Un célèbre exemple de ce fondement est la transformation du cou des girafes (Box 1). Le second principe de Lamarck fait appelle à ce que l'on connaît maintenant comme l'hérédité des caractères acquis, où selon laquelle toutes ces modifications de caractères peuvent se transmettre de descendant en descendant. D'après cette théorie, les organismes peuvent ainsi se transformer spontanément de la plus simple des choses à la plus compliquée. Face à cela, d'autres théories ont également été formulées au cours de cette période. En étudiant les fossiles géologiques, Cuvier (1769-1832) fait l'hypothèse à travers son mouvement de catastrophisme, que certaines espèces sont vouées à disparaître au cours du temps en raison de catastrophes ponctuelles, et que chaque période d'extinctions massives (catastrophes) est suivie d'une période de création de nouvelles espèces. Opposé à toutes idées de transformisme, cet anatomiste français ne manque pas de rappeler l'absence de formes fossiles intermédiaires soulignant cette transformation entre les différents embranchements du règne animal. Enfin, les idées du géologue Hutton (1727-1797) viennent se heurter à cette pensée de Cuvier en proposant que les changements géologiques observés sont le fruit d'un

## Glossaire

**Capacitance:** ratio volume de la moelle/aire des feuilles, traduisant la capacité de stockage de l'eau du tronc vis à vis du pouvoir transpirant des feuilles.

**Évolution convergente:** évolution indépendante de similarité fonctionnelle ou structurale chez des lignées génétiquement distantes, non basée sur une ascendance commune.

**Dérive génétique:** apparition de changements aléatoires dans la fréquence des gènes fréquente aux petites tailles de populations.

**Fréquences alléliques:** proportion d'un allèle sur le total des allèles d'un même locus au sein du pool de gène d'une population (voir GENES).

**Gènes:** séquences nucléotidique spécifique le long de chromosome de l'ADN codant pour une fonction particulière à un locus donné (voir LOCUS). Les formes alternatives sont appelées allèles.

**Locus:** Position d'un gène sur le chromosome.

**Monocarpique:** produisant un seul axe d'inflorescence ou période de floraison durant le cycle de vie.

**Mutation:** changement héritable soudain dans l'ADN pouvant altéré un gène.

**Niche:** rôle écologique d'une espèce dans une communauté.

**Paysage adaptatif:** Topographie représentant la fitness des organismes au travers de géotypes permettant d'occuper des pics adaptatifs particuliers de ce paysage (i.e. niches écologiques), séparés par des vallées qui leurs sont non-adaptatives.

**Phylogénie:** histoire évolutive des taxons, montrant les relations entre les organismes selon leurs différences et leurs similarités, représenté sous la forme d'un arbre phylogénétique .

**Polycarpique:** produisant plusieurs axes d'inflorescence ou périodes de floraison durant le cycle de vie.

**Spéciation:** mécanisme de formation d'une nouvelle espèce.

**Spéciation allopatrique:** spéciation dans laquelle la ségrégation et la différenciation des populations apparaît en raison d'un isolement géographique à l'origine d'un isolement reproductif complet.

**Spéciation écologique:** spéciation dans laquelle la ségrégation et la différenciation des populations apparaît en raison d'un isolement écologique chez des populations sympatriques.

**Spéciation parapatrique:** spéciation dans laquelle l'isolement reproductif entre populations n'est pas complet en raison d'aires de répartitions chevauchantes en zones hybrides .

**Sympatrie:** populations présentes dans une même aire géographique.

**Vicariante:** division de taxons sur la base de l'existence d'une barrière géographique naturelle.

processus graduel lent et continu en cours du temps et non pas de catastrophes. Cette idée sera finalement confortée par Lyell (1797-1875), ajoutant que les processus géologiques qui se sont produits dans le passé sont les mêmes actuellement à l'œuvre et que si de quelconques changements aussi lents soient-ils ont eu lieu, cela aurait pu causer des changements importants sur les organismes.

## 1 | L'ÉVOLUTION DES ESPÈCES

Au 19<sup>ème</sup> siècle, empreint de ces nombreuses théories, notamment celle de Lyell, et fort de son voyage au sein du *H.M.S Beagle*, Charles Darwin (1809-1882) apporte une réponse nouvelle quant à l'origine, la parenté, les différences entre espèces ou encore la répartition de la diversité du vivant. À travers son œuvre *De l'origine des espèces*, Darwin introduit la notion de transformation, qui sera ultérieurement changée en 'évolution' à partir de la 6<sup>ème</sup> édition, en soutenant que les espèces modernes sont apparentées à des espèces ancestrales différentes, ce qu'il appela '*descendance avec modification*' (Darwin, 1859). Selon lui, et contrairement au transformisme de Lamarck, cette évolution peut se représenter sous la forme d'un arbre généalogique où à chaque ramification, des lignées descendantes vont diverger à partir d'un ancêtre en des changements qui ne constituent pas forcément l'amélioration d'une forme précédente. Par une même ascendance, ces lignées vont alors partager un certain nombre de caractéristiques. Il décrit ainsi les espèces comme des individus interféconds d'une population, présentant des caractéristiques communes et capables de transformation. Afin de justifier ce processus évolutif, Darwin introduit la notion de sélection naturelle en parlant d'une évolution adaptative favorisant la survie du plus apte dans une lutte pour la survie. (Box 1).

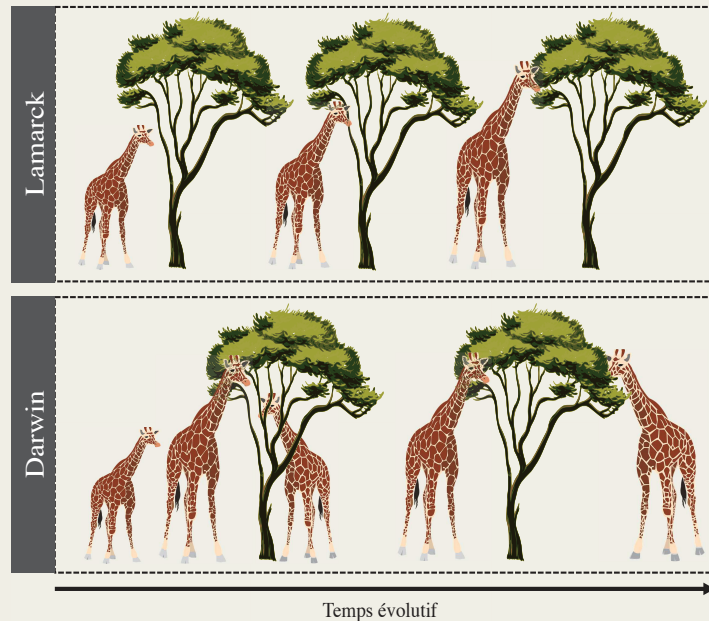
De plus, si cet environnement est changeant, des caractères jusqu'alors favorables peuvent devenir inappropriés. Si les individus d'une population se déplacent vers un nouvel environnement, la sélection naturelle va pouvoir favoriser de nouveaux caractères pouvant conduire à l'apparition d'une nouvelle espèce. Darwin souligne que ces changements vont à la fois varier au cours du temps, mais également dans l'espace, ce qui peut expliquer les différences dans la répartition des espèces modernes à la surface du globe. Cette notion de biogéographie fut particulièrement bien constatée lors de grandes expéditions au courant du 19<sup>ème</sup> siècle par de nombreux naturalistes. La description de la répartition des plantes en altitude, faite par Alexander von Humboldt (1769-1859) lors de son expédition en Amérique du Sud, en constitue un exemple majeur. La théorie darwinienne de descendance avec modification va permettre d'apporter une réponse quant à un phénomène jusqu'alors décrit mais sans aucune explication formelle. Des espèces apparentées ont tendance à vivre dans des environnements différents du globe tandis que la sélection naturelle va amener des espèces très différentes à occuper des environnements similaires. Darwin n'émettra toutefois que peu d'explications quant au processus d'apparition, de transmission, et de maintien de ces variations au sein des populations.

Peu de temps après ces travaux, Gregor Mendel (1822-1884), un moine et botaniste austro-hongrois, formula une théorie sur l'hérédité avec la ségrégation et l'assortiment

### Box 1: Transformisme vs Sélection naturelle

La première théorie de l'évolution s'éloignant des idées fixistes fut établie par le naturaliste français Jean-Baptiste Lamarck (1744-1829), dans son livre *Philosophie zoologique* parue en 1809 (année coïncidente avec la naissance de Darwin). Selon lui, les organismes se transformeraient graduellement au cours du temps vers un état de complexité en raison des pressions fournies par l'environnement. Il explique également que la diversité des espèces serait liée à l'apparition spontanée d'espèces répétant ce processus de complexification à chaque fois, sans aucun lien de parenté. En prenant l'exemple des girafes, cette vision transformiste expliquerait que la taille du cou de ces animaux grandirait au fil du temps afin de pouvoir accéder aux ressources disponibles dans leur habitat.

En 1859, dans son ouvrage *De l'origine des espèces*, Charles Darwin (1809-1882) s'opposera aux idées de Lamarck en proposant une ascendance commune à tout les êtres vivants. Les variations observées chez les espèces seraient dictées par la sélection naturelle, où seuls les organismes les mieux adaptés survivent. Ces variations seraient ensuite transmises de générations en générations. Uniquement les girafes au long cou auraient survécu dans ce milieu, fixant ce caractère adaptatif au fil des générations.

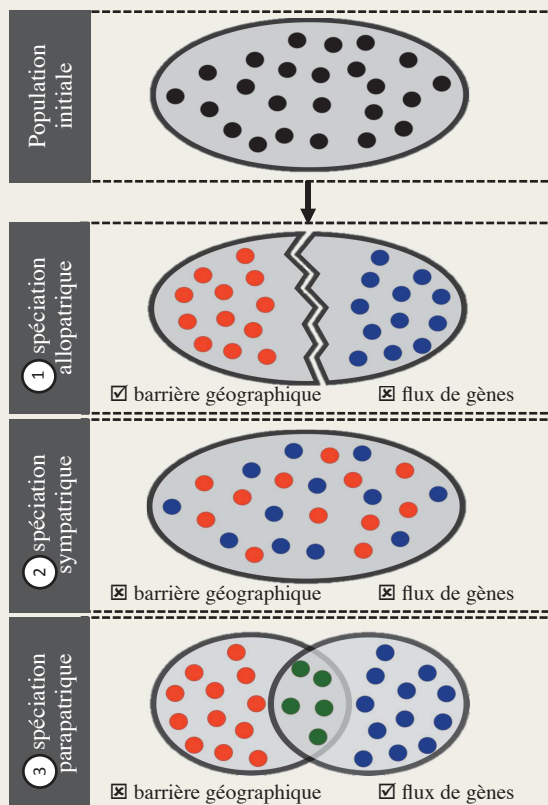


indépendant des caractères à partir de croisement réalisés sur des plantes. Selon lui, des unités héréditaires (référant à des caractères discrets, comme la couleur des fleurs ou surtout le tégument des graines) sont transmises par des lignées parentales et maintenues ensuite de génération en génération avec des effets de dominance ou non. Ce n'est finalement qu'au début du 20<sup>ème</sup> siècle que des chercheurs ont couplé ces unités mendéliennes à des positions au niveau des chromosomes, les GENES (voir Glossaire), faisant de l'ADN le support matériel de cette hérédité et de l'expression des caractères transmis. L'intégration des idées de Mendel et de Darwin, lié à ces découvertes génétiques, a permis de formuler une théorie de l'évolution dite synthétique reposant sur le travail de nombreux scientifiques tels que Wright (1889-1988), Fisher (1890-1962), Haldane (1892-1964), Dobzhansky (1900-1975), ou encore Mayr (1904-2005), et qui rejette définitivement l'hérédité des caractères acquis proposée par Lamarck. Selon ce principe, les populations vont évoluer en modifiant la fréquence des différentes versions de gènes (les allèles) présentes dans le pool génétique les constituant, pouvant même être à l'origine de nouvelles espèces. De plus, si la biologie évolutive, visant à comprendre ce processus dit de SPECIATION, repose sur les fondements même de la théorie darwinienne et des lois de l'hérédité, la sélection naturelle n'apparaît pas comme seule source de spéciation. Les FREQUENCES ALLELIQUES au sein des populations d'une espèce peuvent varier sous l'influence de nombreux autres processus tels que la MUTATION, le

## Box 2: Les modèles de spéciation

Le concept biologique de l'espèce fut décrit par Ernst Mayr en 1942 comme une population dont les individus sont en mesure de se reproduire les uns avec les autres et d'engendrer une descendance viable et féconde, au détriment d'individus d'autres populations. Cette notion repose sur le principe d'isolement reproducteur (processus empêchant les membres de deux espèces de produire des descendants viables et féconds), pouvant être à la fois prézygotique (barrières écologiques, temporels, éthologiques, mécaniques ou gamétiques) ou postzygotique (viabilité ou fécondité réduite des hybrides). Les modèles de spéciation vont dépendre du degré d'isolement géographique existant entre les individus. La subdivision géographique de deux populations va diminuer les flux géniques, et ainsi l'apport de diversité génétique. La dérive génétique (processus aléatoire dépendant de la taille des populations) va alors pouvoir rapidement fixer ou perdre des allèles au sein des populations séparées, en promouvant une différenciation génétique. La sélection naturelle peut aussi rapidement augmenter cette variation génétique en fixant des allèles codant pour des adaptations clés dans le milieu indépendamment des deux populations.

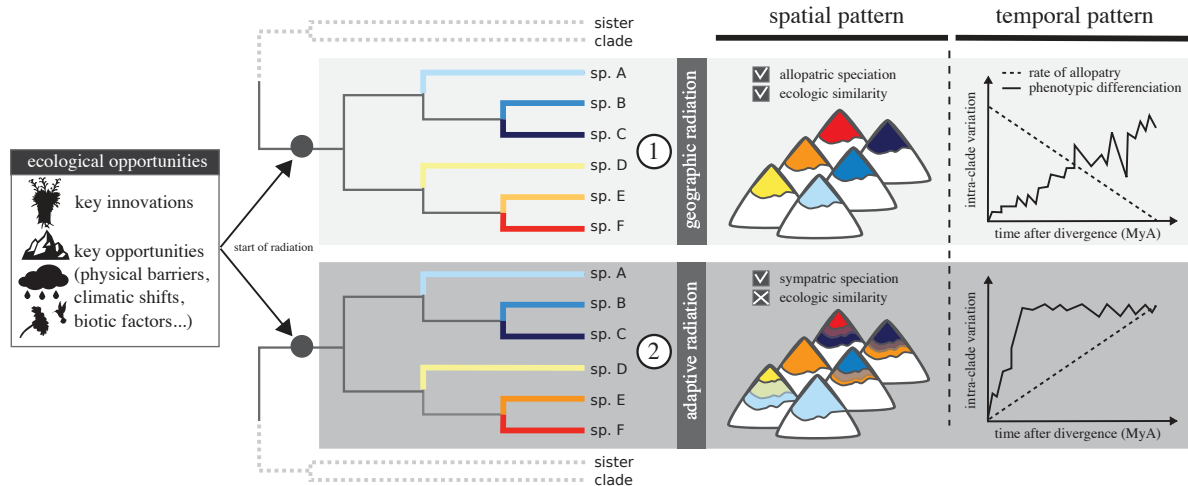
Selon le degré subdivision géographique, trois principaux modèles de spéciation ont été définis. (1) La spéciation allopatrique implique un isolement géographique entre populations où la dérive génétique et la sélection peuvent amener à la création de nouvelles espèces. (2) La spéciation sympatrique quant à elle implique une population panmictique, dont l'aire de répartition est chevauchante. Un isolement reproducteur est tout de même nécessaire pour qu'il y ait une spéciation. Celui-ci peut être de l'ordre d'une spécialisation écologique ou encore d'un changement de préférences sexuelles chez une des deux sous-populations. Enfin, (3) la spéciation parapatricque apparaît entre des populations géographiquement proches où le flux de gènes est possible au niveau de zones hybrides. Malgré l'apport de nouveaux gènes, ces deux populations adaptées à des milieux différents peuvent maintenir leur intégrité génétique sous l'effet de la sélection naturelle et pouvant ainsi fortement se différencier génétiquement en de nouvelles espèces.



phénomène aléatoire de **DÉRIVE GÉNÉTIQUE** ou encore les flux de gènes entre populations. Tout comme pour la sélection naturelle, la manière dont ces flux géniques sont répartis entre populations peut conduire à la formation de nouvelles espèces, en absence ou non d'un isolement reproducteur (Box 2).

## 2 | LES RADIATIONS EVOLUTIVES





**Figure 1.** Spatial and temporal predictions occurring in both (1) geographic and (2) adaptive radiations triggered by the evolution of key innovations and opportunities, pollination syndrome, physical barriers or climatic shifts.

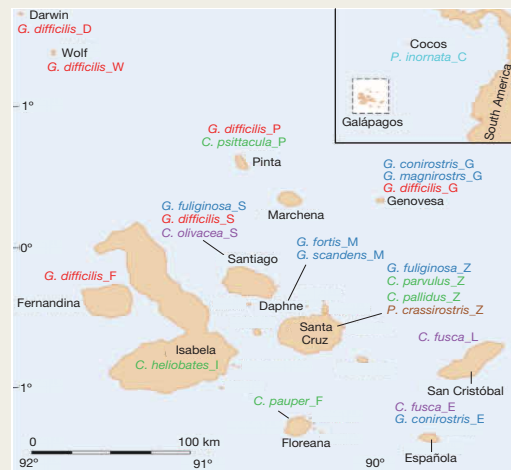
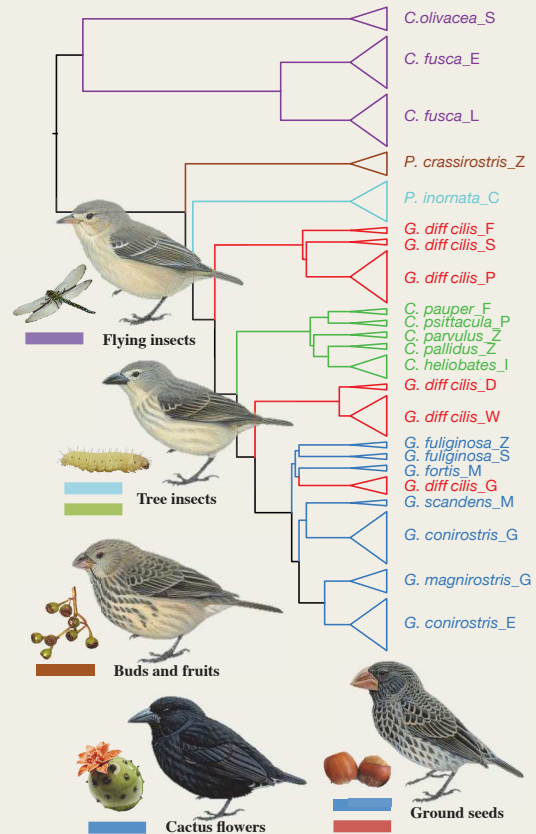
Comprendre l'origine des espèces nécessite de comprendre comment les variations génétiques se produisent. La microévolution témoigne de l'apparition de ces variations génétiques au sein d'une ou plusieurs populations issues d'une même espèce sur plusieurs générations sous l'influence des forces évolutives précédemment décrites. Toutefois, ces mêmes forces évolutives peuvent également se produire à une plus grande échelle, où les espèces elles-mêmes figurent cette fois-ci comme unité d'évolution, sous un processus dit de macroévolution (Reznick and Ricklefs, 2009). Les radiations évolutives, généralisant les processus de spéciation dans un contexte spatial et temporel plus important, en illustrent parfaitement ce phénomène (Gavrilets and Losos, 2009).

Les radiations évolutives en générale peuvent être définies comme une diversification d'espèces issues d'un même ancêtre commun sur une période de temps relativement courte par une augmentation des taux de spéciation et/ou une diminution des taux d'extinction (Schluter, 2000). Ces phénomènes de diversification se produisent en fonction d'opportunités écologiques pour les lignées au travers de facteurs biotiques ou abiotiques quand les ressources d'un habitat sont inutilisées par d'autres espèces. Les radiations peuvent arriver par exemple après la colonisation d'îles et de montagnes isolées ou de lacs, ou suites à des shifts climatiques permettant la migration d'espèces dans de nouvelles régions. Ces mécanismes extrinsèques traduisent une opportunité clé pour les taxons leur permettant de se diversifier (Donoghue and Sanderson, 2015) (Fig. 1). S'ils peuvent être de l'ordre abiotique, un ensemble de processus biotiques peuvent également intervenir de manière extrinsèque à la diversification des lignées comme le relâchement des pressions de compétition ou encore le changement de pollinisateur dans un habitat donné. D'une autre manière, ces radiations peuvent aussi survenir en fonction de facteurs intrinsèques comme l'évolution d'une innovation morphologique clé adaptée à de nouveaux milieux permettant aux lignées de coloniser un nouveau paysage adaptatif (Donoghue and Sanderson, 2015; Givnish, 2015)

## Box 3: Radiation des pinsons de Darwin

Les fortes variations morphologiques dans la forme des becs observées par Darwin sur plusieurs espèces de pinsons aux Galápagos, lors de son voyage à bord du *H.M.S Beagle* (1835), lui ont amené à développer ses idées et ses fondements mêmes sur la force de la sélection naturelle 14 ans après dans l'origine des espèces.

À partir de la fin du XX<sup>ème</sup> siècle, de nombreuses études ont fait le lien entre la diversité décrite par Darwin et les modèles de spéciation et d'évolution actuels, faisant de ces pinsons, communément appelés pinsons de Darwin, un exemple emblématique de radiations adaptative. L'apport de données moléculaire a permis par exemple de démontrer une ascendance commune de ces pinsons à partir d'une lignée ancestrale américaine unique. Cette population fondatrice aurait ainsi colonisé les îles Cocos et des Galápagos il y a 2-3 MyA où la surrection de ces dernières leur a fourni de nouvelles opportunités écologiques. Ces pinsons se seraient ensuite diversifiés de manière allopatrique en évoluant indépendamment sur chaque île, puis après des remises en contact secondaire, auraient évolué de façon sympatrique suite à un isolement écologique. Des observations moléculaires modernes ont démontré que ces espèces sympatriques se sont différenciées au niveau de la forme et de la taille du bec afin d'accéder à différents types de ressources en fixant génétiquement ces caractères. Les pinsons à bec court apparaissent ainsi adaptés à un régime de graines, ceux au bec long et arrondi à un régime frugivore et ceux au bec pointu à un régime insectivore. Le brassage génétique entraîné par l'hybridation aurait également joué un rôle majeur dans cette radiation en facilitant la réponse de la sélection naturelle par l'apport de nouvelles combinaisons génétique. L'exploration moderne de la biologie évolutive a permis de faire le lien entre les variations génétiques existantes et la sélection naturelle initialement décrite par Darwin. Toutefois, même après plus de 150 ans d'étude, l'évolution de ces pinsons dans les Galápagos continuera de délivrer encore quelques secrets sur les mécanismes macroévolutifs de spéciation, en raison d'une évolution divergente encore observée aujourd'hui à l'échelle microévolutive au sein de ces îles.



Références: Grant et Grant 2008, Lamichhane et al. 2015, Almén et al. 2016. Figures: adaptées de Lamichhane et al. 2015.

(Fig. 1). Finalement, la diversification des espèces ne s'avère pas seulement déterminée par un précurseur type à l'origine des lignées, qu'il soit biotique ou abiotique, ou encore extrinsèque ou intrinsèque. L'ensemble de ces facteurs peut agir de manière synergétique

dans la diversification des espèces (Bouchenak & Khelladi et al., 2015; Donoghue and Sanderson, 2015).

Différents types de radiations ont été définis en fonction des mécanismes de spéciation survenant après la colonisation des lignées dans ces nouveaux habitats suite à ces précurseurs. Les radiations adaptatives caractérisent une diversification des lignées en une variété de NICHES associée à une divergence adaptative pour des traits morphologiques ou physiologiques nécessaires à l'utilisation de ces habitats (Givnish, 2015; Glor, 2010; Losos and Mahler, 2010; Schluter, 2000). De nombreux exemples, emblématiques de radiations adaptatives, ont été soulignés chez les pinsons des Galápagos (Grant and Grant, 2008) (Box 3), les cichlides d'Afrique (Muschick et al., 2012; Seehausen, 2006), les marsupiaux australiens (Nilsson et al., 2010), les lézards des Caraïbes avec le genre *Anolis* (Losos, 2009; Wollenberg et al., 2013) ou encore chez l'alliance des *silverswords* (Carlquist et al., 2003) ou chez les *Lobelia* (Givnish et al., 2009) des îles d'Hawai. Le caractère 'explosif' de la diversification est souvent associé aux radiations adaptatives (Glor, 2010). Or, l'accélération du taux net de diversification des espèces vis à vis de lignées proches n'apparaît pas comme une condition *sine qua none* de ce type de radiation (Givnish, 2015; Rundell and Price, 2009). À titre d'exemple, les pinsons des Galápagos n'ont pas diversifié plus rapidement que leurs cousins d'Amérique continentale et des Caraïbes dont ils sont originaires (Burns et al., 2014). Ce processus repose fondamentalement sur la sélection naturelle divergente entre les lignées pour différents types de ressources disponibles dans un environnement donné. Dans ce cas de figure, les lignées se retrouvant en SYMPATRIE et donc en compétition dans le milieu vont se spécialiser écologiquement afin de partager les niches disponibles, créant une SPECIATION ECOLOGIQUE (Schluter, 2000). Sous un modèle de disparité adaptative, la théorie de la radiation adaptative prédit que l'espace écologique disponible va progressivement diminuer au cours du temps jusqu'à une occupation totale des opportunités écologiques du milieu par les espèces (Losos and Mahler, 2010). Au même titre, la différenciation morphologique entre lignées va rapidement augmenter sous effet de la sélection naturelle divergente pour ensuite se stabiliser (Fig. 1). Le taux d'évolution morphologique et des lignées au cours du temps s'avèrent ainsi inversement corrélé à la disponibilité des niches écologiques dans le milieu.

La diversification des lignées durant une radiation adaptative va dépendre des opportunités écologiques fournies par le milieu et de la variation génétique des espèces. Seehausen (2004) souligne néanmoins que cette diversification est également dépendante dans la capacité des espèces à s'hybrider. En effet, l'hybridation, qui peut être un frein à l'évolution des espèces, est aujourd'hui reconnue comme étant une source importante de nouveauté génétique, phénotypique et adaptative chez les espèces par le biais de (1) l'introgression adaptative chez les parents et de (2) la recombinaison génétique chez les hybrides. Le phénomène (1) d'introgression adaptative correspond au transfert d'allèles adaptatifs entre les espèces qui, codant pour des traits avantageux, peuvent interagir positivement dans le background génétique de ces dernières et accélérer leur adaptation (Abbott et al., 2013; Barton, 2001; Hedrick, 2013; Mallet, 2005). À l'opposé, la recombinaison génétique du patrimoine parentale dans les génomes des hybrides peut également conduire à l'apparition de nouveautés phénotypiques pouvant être fortement adaptatives chez ces populations hybrides (Elgvin et al., 2017; Holzman and Hulsey, 2017;

Keller et al., 2013; Mavárez et al., 2006). Durant la colonisation d'un nouveau PAYSAGE ADAPTATIF, une hybridation intensive peut alors être à l'origine d'une radiation adaptative par la capacité des espèces à s'adapter à de nouvelles niches écologiques encore inutilisées (Seehausen, 2004). D'une autre manière, des évènements occasionnels et localisés d'hybridation peuvent aussi maintenir et accélérer la diversification adaptative des espèces au cours de radiation par le biais de ces nouveautés adaptatives leur permettant de s'établir dans des niches écologiques encore vacantes (Seehausen, 2004). Ce phénomène, précurseur et/ou acteur aux radiations adaptatives, a ainsi été identifiée dans de nombreuses radiations, comme chez les pinsons des Galápagos (Lamichhaney et al., 2015), les papillons du genre *Heliconius* (Pardo-Diaz et al., 2012) ou encore chez les cichlides africains (Meier et al., 2017) à titre d'exemple.

D'autres facteurs peuvent également accélérer ou diminuer le taux spéciation/extinction sans une divergence adaptative entre les lignées comme une dispersion limitée des espèces, la venue de barrières naturelles aux flux de gènes, l'hybridation ou encore la sélection sexuelle (Donoghue and Sanderson, 2015; Rundell and Price, 2009). La similarité écologique des espèces peut être à l'origine de radiations non-adaptative comme la radiation géographique (Rundell and Price, 2009), où la SPECIATION ALLOPATRIQUE ou PARAPATRIQUE, peut excéder la divergence adaptative (Fig. 1 ; Box 2). Les radiations non-adaptatives ont été définies par Gittenberger (1991) comme la diversification d'espèces à partir d'un ancêtre commun, non accompagnée par une différenciation de niche écologique. Cette similarité écologique des espèces résulte d'une exclusion compétitive mutuelle qui serait la principale cause d'un isolement géographique (Rundell and Price, 2009). Selon ce principe, le taux d'allopatrie des espèces apparaît très élevée entre ces dernières après leur divergence. Néanmoins ce taux tend à diminuer au cours du temps par un retour à la sympatrie entre ces espèces suite à de la migration (Fig. 1). Les espèces allopatriques, pouvant également diverger écologiquement sous l'effet d'adaptations locales (Rundell and Price, 2009), peuvent par cette remise en contact secondaire augmenter la divergence écologique dans ces zones de sympatrie à partir d'un renforcement de l'isolement reproducteur ou de déplacement de caractères entre espèces et activement participer aux radiations adaptatives (Aguilée et al., 2012; Gavrilets and Losos, 2009; Losos and Mahler, 2010; Losos and Ricklefs, 2009).

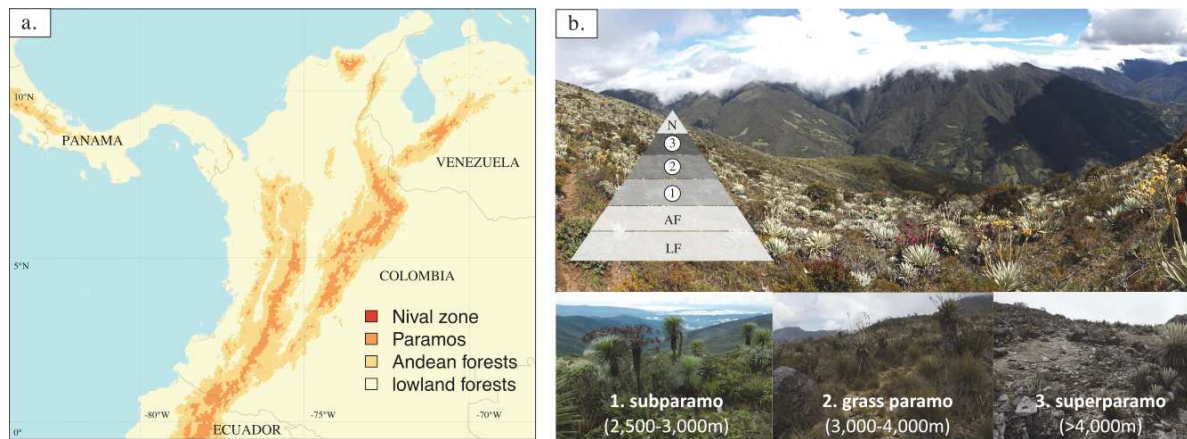
Si les précurseurs peuvent agir de manières synchrones à la diversification des lignées, ces mécanismes de spéciation allopatrique, sympatrique et de divergence adaptative apparaissent liés lors de radiations adaptatives, et ce à différentes échelles temporelles et spatiales (Bouchenak-Khelladi et al., 2015; Donoghue and Sanderson, 2015). De plus, certaines radiations contiennent des composantes à la fois adaptatives ou non, comme il a été démontré chez la radiation des *Tetragnatha* dans les îles d'Hawaii par la présence d'espèces sympatriques très différenciées écologiquement et d'espèces allopatriques ou parapatrique écologiquement similaires (Rundell and Price, 2009). Les radiations évolutives apparaissent ainsi comme un processus complexe dépendant à la fois du contexte géographique, temporel, environnemental, et des nombreuses opportunités écologiques disponibles. Nombreux de ces phénomènes ont été démontré chez une grande variété d'espèces et d'habitats, où les écosystèmes de hautes montagnes, subissant les mêmes pressions que les habitats insulaires,

offrent des conditions divergentes fortes pour la radiation des espèces, notamment chez les plantes dans les páramos (Hughes and Atchison, 2015; Madriñán et al., 2013). Toutefois, un certain nombre de ces radiations adaptatives et des processus d'innovation morphologique décrits reposent sur des évidences micro-évolutives qui ne peuvent pas être étudiées à l'échelle macro-évolutive. De nombreuses questions demeurent ainsi quant à notre compréhension des mécanismes de spéciation et d'interaction entre la micro- et la macroévolution (Bell, 2016). Et si ces radiations s'avèrent bien décrites chez les animaux comme chez les pinsons de Darwin (Almén et al., 2016; Chaves et al., 2016), ou encore les cichlides africains (Brawand et al., 2014; Ivory et al., 2016; Takahashi and Koblmüller, 2011), on pourrait alors se poser la question suivante : a-t-on jamais démontré des radiations adaptatives chez les plantes au niveau de patrons macro-évolutifs ?

### 3 | PARAMOS ET ROSETTES GÉANTES

Les montagnes des Andes du Nord abritent un écosystème tropical alpin connu sous le nom de páramo, présent de la limite supérieure de la forêt (à environ 3,000-3,500 m d'altitude) jusqu'aux neiges permanentes (à environ 5,000 m). Cet écosystème, occupant environ 35,000 km<sup>2</sup>, s'étend entre les latitudes 11°N et 8°S au niveau des Andes du Venezuela, de la Colombie et de l'Équateur, avec quelques zones plus extrêmes au Nord du Pérou et au Panama/Costa Rica (Fig. 2a). Les écosystèmes tropicaux alpins, présents également en Amérique du Sud, à Hawaii, en Nouvelle-Guinée ou encore en Afrique de l'Est, sont caractérisés par de forts gradients altitudinaux où la fréquence et l'intensité des conditions climatiques changent rapidement sur quelques centaines de mètres, tels que de fortes variations de températures ou de forts taux de radiations solaires (Luteyn, 1999). À la différence des biomes alpins tempérés, ces écosystèmes alpins tropicaux ne présentent pas de saisonnalités thermiques au cours de l'année mais de fortes oscillations de température entre la journée et la nuit. Climatiquement ces écosystèmes sont caractérisés par de fortes précipitations annuelles (~1,000-3,000mm) et de faibles températures annuelles moyennes (environ 8-12°C) où les variations journalières peuvent être considérées comme l'été pendant la journée et comme l'hiver pendant la nuit (Hedberg, 1964). Ces conditions *a priori* hostiles ont rendues nécessaires l'évolution d'adaptations particulières pour les organismes ayant colonisé ces habitats, modelant ainsi un environnement et une végétation type au sein des ces biomes.

Selon les communautés végétales se retrouvant le long du gradient altitudinal et les conditions thermiques présentes (principalement l'humidité et la température), trois zones de végétations ont ainsi été définies au niveau des páramos (Cuatrecasas, 2013; Luteyn, 1999) (Fig. 2b). S'étendant de 2,500 à 3,000m, au dessus de la limite supérieur de la forêt Andine (~2500m), la végétation est dominée par des graminées et des arbustes aux feuilles coriaces et sclérophylles. Ce milieu, appelé (1) subpáramo, assez dense, représente une zone de transition entre la forêt humide et le páramo propre plutôt sec. Avec l'augmentation de l'altitude, cette végétation est remplacée graduellement par de plus petites herbes pérennes, de plus petits arbustes, des graminées, des mousses, des lichens, des plantes en coussin hautement adaptée



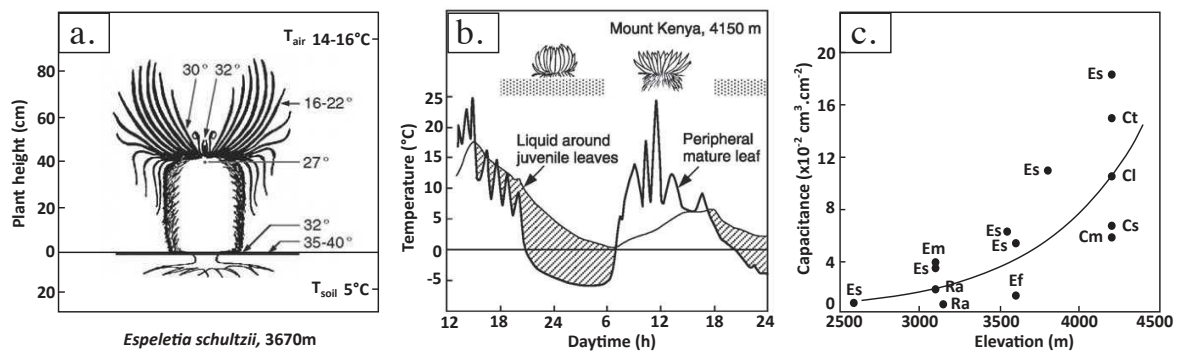
**Figure 2.** Distribution (a) of páramos habitat into northern tropical Andes and (b) altitudinal vegetation zones of páramos ranging from shrubby to open grassland consisting on (1) subpáramos, (2) páramo proper and (3) superpáramo. Photos by Serge Aubert.

au milieu alpins (Aubert et al., 2014), de petites rosettes mais également par une forme emblématique et dominante, la rosette géante caulescente (avec un tronc), connue également sous le terme de ‘rosette pachycaule’ ou de ‘caulirosula’. La zone du (2) páramo forme une prairie alpine, s’étendant entre 3,000 et 4,000m d’altitude, plus sèche et plus ouverte que celle du subpáramo, mais conservant un sol plutôt profond, riche en matière organique et constamment humide (Luteyn, 1999). Au delà de 4,000m d’altitude jusqu’aux neiges permanentes, le sol est beaucoup plus fin, avec un fort pourcentage de sable et de roches, et très peu de matière organique, entraînant par conséquent une faible rétention d’eau. Cette zone apparaît fortement impactée par les fluctuations thermiques journalières où le sol gèle la nuit, et où les températures sont plus faibles avec des taux de radiations aux UV plus forts. Ces conditions impactent profondément la croissance racinaire, la germination et le type de végétation. La végétation de ce milieu, appelé (3) superpáramo, est ainsi plus clairsemée avec peu d’espèces. Celle-ci est représentée par une flore cryophyle dont la pubescence des feuilles est plus importante, avec des plantes plutôt petites, à l’exception de certaines rosettes géantes caulescentes, constituant un réel paradoxe vis à vis d’une taille plus petite attendue à haute altitude (Körner, 2003; Smith, 1981).

Contrairement aux habitats alpins tempérés où les communautés végétales semblent adaptés principalement au couvert neigeux par l’évolution de formes de vie tamponnant les variations thermiques saisonnières, la forme en rosette caulescente apparaît être une adaptation clé au cycle thermique diurnal propre aux écosystèmes tropicaux alpins en raison des valeurs adaptatives qui lui sont associées (Rada et al., 1985). Cette forme est représentée par une rosette aux feuilles plutôt larges supportée par un tronc non ramifié ou très peu (Fig. 3). Les feuilles compactes, arrangées en spirale, forment une couronne dense autour des méristèmes apicaux. Ces feuilles sont également dotées d’une gaine résistante qui va permettre de recouvrir le tronc au cours de la croissance de la plante de feuilles mortes ou uniquement des gaines restantes. Des nombreuses études, portant à une échelle micro-évolutive, ont démontré que la couronne de feuilles apporte une protection thermique des bourgeons de croissance et des jeunes feuilles au sein de la rosette (Beck et al., 1982;

Monasterio and Sarmiento, 1991; Smith, 1981) (Fig. 3a). Cette couronne va également se rétracter la nuit et s'ouvrir le jour, où les feuilles pubescentes dotées d'une capacité de surfusion permanente vont permettre de tolérer les fluctuations journalières et le gel (Beck et al., 1982; Smith, 1974) (Fig. 3b). Si elle n'a que très peu d'effet sur l'absorption des radiations solaires contrairement aux attendus dans les milieux alpins tempérés, la pubescence foliaire a un effet fort sur la régulation thermique au sein des feuilles au cours des fluctuations thermiques journalières importantes dans ces milieux tropicaux alpins (Meinzer and Goldstein, 1985; Meinzer et al., 1985). Les rosettes géantes ont aussi une CAPACITANCE d'eau élevée afin de lutter contre la sécheresse (notamment au niveau du superpáramo). Il a été démontré que cette capacitance augmente de façon exponentielle avec l'altitude en relation avec la surface de transpiration foliaire et le volume d'eau compris dans la moelle du tronc, facilitant les échanges de gaz quand la température est encore faible ou durant les période sèches (Goldstein et al., 1984) (Fig. 3c). Enfin, les feuilles mortes présentes sur le tronc vont permettre une protection thermique de l'eau stockée dans la moelle et jouer un rôle important dans le cycle des nutriments (Monasterio and Sarmiento, 1991). Finalement, ce morphotype représente également une ÉVOLUTION CONVERGENTE au niveau de ces écosystèmes alpins tropicaux, renforçant la trajectoire adaptative d'un tel phénotype à ces milieux à une échelle macro-évolutive. Les rosettes géantes caulescentes ont évolué de manières indépendantes chez les genres *Dendrosenecio* (Asteraceae) et *Lobelia* (Campanulaceae) d'Afrique de l'Est, *Lobelia* et *Argyroxiphium* (Asteraceae) des îles Hawaï, ou encore *Puya* (Bromeliaceae) et *Espeletia* (Asteraceae : Heliantheae) dans les páramos (Fig. 4).

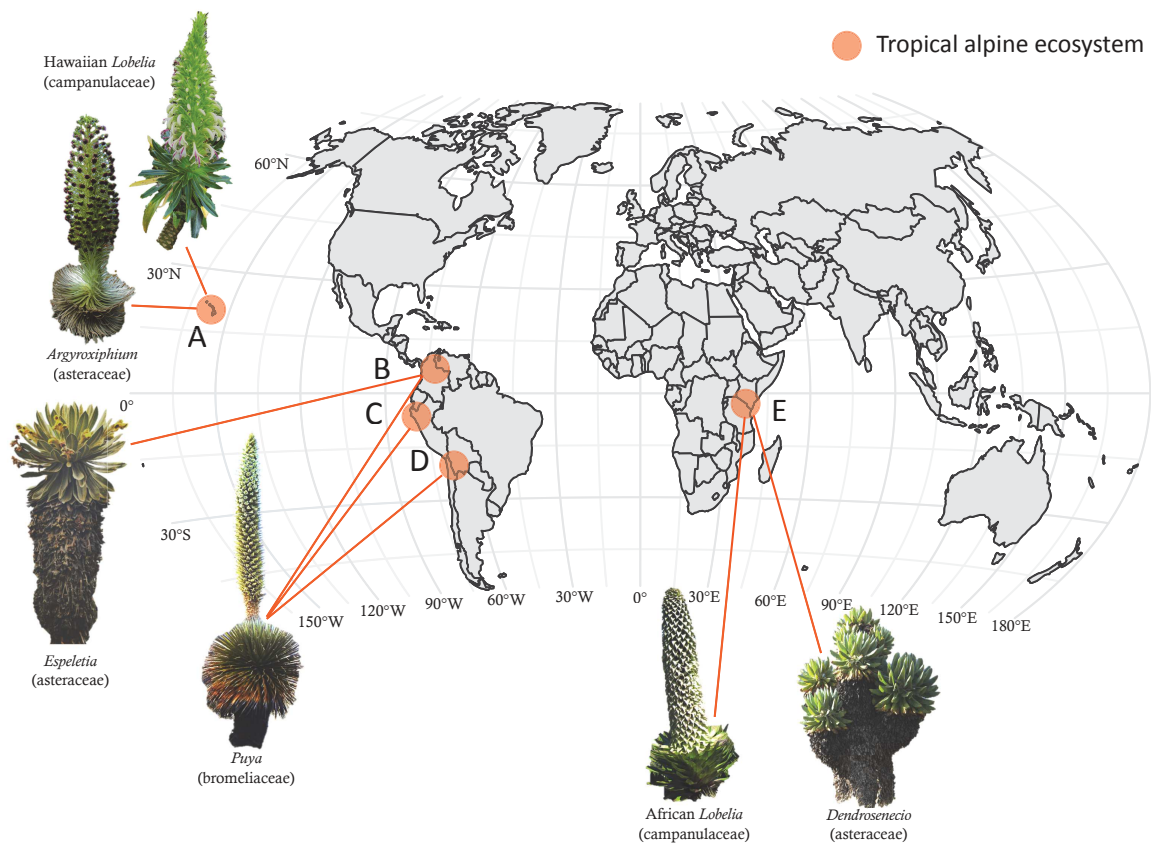
Malgré une superficie assez restreinte et des conditions de vie difficiles, les páramos totalisent près de 4,700 espèces, représentant environ 10 à 20% de la richesse de la flore andine (Rangel-Ch., 2015), avec des taux d'endémisme incroyablement élevés. Selon Madriñan et al. (2013), les páramos abritent également les radiations de plantes les plus rapides au monde, générant le plus d'espèces de tous les biomes alpins avec les habitats himalayens. Deux principaux mécanismes, géologique et climatique, pouvant expliquer une telle diversité ont été débattus et opposés dans la littérature (Rull, 2011a, 2011b). En premier, (1) l'orogénèse des Andes est considéré comme un événement important dans la diversification et l'évolution des organismes dans les régions néo-tropicales (Antonelli and Sanmartín, 2011a; Antonelli et al., 2009; Hoorn et al., 2013). Au niveau des Andes du Nord, la surrection des Andes a débuté il y a 25 million d'années (MyA) et s'est intensifiée durant le Miocène entre 10-5 MyA où un proto-páramo a commencé à se développer avec une végétation plutôt arbustive émergeant au dessus de la forêt andine à des altitudes plus basses que celles retrouvées aujourd'hui (~2,000-3,000m) (Hooghiemstra and Van der Hammen, 2004a). À ce moment là, les Andes du Nord éteignait 40% de leur élévation actuelle (Garzzone et al., 2008; Gregory-Wodzicki, 2000; Hoorn et al., 2010). La région des páramos a finalement atteint son élévation actuelle au cours des 5.0-2.5 MyA où une végétation type paraméenne est retrouvée au niveau de fossiles de pollen (Gregory-Wodzicki, 2000; Hooghiemstra and Van der Hammen, 2004a). En se soulevant, les Andes ont apportées de nouvelles opportunités écologiques en augmentant l'hétérogénéité des habitats pouvant favoriser les radiations évolutives dans les montagnes nouvellement créées, et permettre un isolement géographique favorisant la spéciation allopatrique (Antonelli and Sanmartín,



**Figure 3.** Adaptation of caulescent rosettes to the tropical alpine climates. (a.) Such form provides thermal protection to the apical buds due to the crown of green leaves and prevents freezing of water stored on the large pith due to the coverage of marscescent (dead) leaves surrounding the stem. (b.) Buds are also protected by the closure of the rosette leaves to cope with negative temperatures occurring at night. (c.) Other example concerns the water capacitance, i.e. the relationship between the volume of water kept in the pith and leaf surface, which increases according to elevation. Gas exchanges are facilitated during early morning hours when low soil temperatures result in high root resistance for water uptake, and during drought periods. Graphics were modified from Körner (2003) for (a.) and (b.), and from Goldstein et al. (1984) for (c.).

2011a). De plus, cette surrection a également conduit à la création d'un corridor biotique entre l'Amérique du Nord et du Sud, à la formation d'une barrière VICARIANTE entre l'Est et l'Ouest du continent et aussi à de nombreux changements de conditions environnementales en dehors des Andes comme en Amazonie, pouvant être moteur de spéciation (Antonelli and Sanmartín, 2011a). D'un autre point de vue, (2) la dynamique des glaciations s'étant produit pendant le Pléistocène aurait également fourni de nombreuses opportunités écologiques à la diversification des plantes dans les páramos (Hooghiemstra et al., 2006; Londoño et al., 2014; Vuilleumier and Monasterio, 1986). D'après les données fossiles de pollen, la végétation des páramos aurait subis d'importants changements dans la composition, la distribution et la structuration des communautés végétales en raison des fluctuations thermiques déterminées par cette dynamique climatique (Hooghiemstra and Van der Hammen, 2004a; Torres et al., 2013). Durant les périodes glaciaires, les páramos étaient plus larges et interconnectés entre eux. À l'inverse, les périodes interglaciaires ont favorisé l'isolement des páramos, comme un archipel d'îles continentales (Hooghiemstra et al., 2006). Ces mouvements de l'habitat ont ainsi entraîné des expansions et des contractions des communautés végétales entraînant une spéciation allopatrique durant l'isolement, et des remises en contact secondaires d'espèces durant l'expansion favorisant un isolement écologique des espèces (Antonelli and Sanmartín, 2011a; Madriñán et al., 2013; Rull, 2011a, 2011b). La majorité des évènements de spéciations dans les páramos, soulignés par Madriñán et al. (2013) s'est produit durant le Pléistocène. Ces auteurs suggèrent ainsi que ces oscillations climatiques, produisant des expansions/rétractions des populations, peuvent être une force motrice clé à l'origine des taux de diversification exceptionnels de la flore des páramos. En reconstruisant des PHYLOGENIES moléculaires à partir de données génétiques disponibles sur des espèces endémiques des páramos et des estimations des âges établies par Madriñán et al. (2013), on peut également apercevoir que si

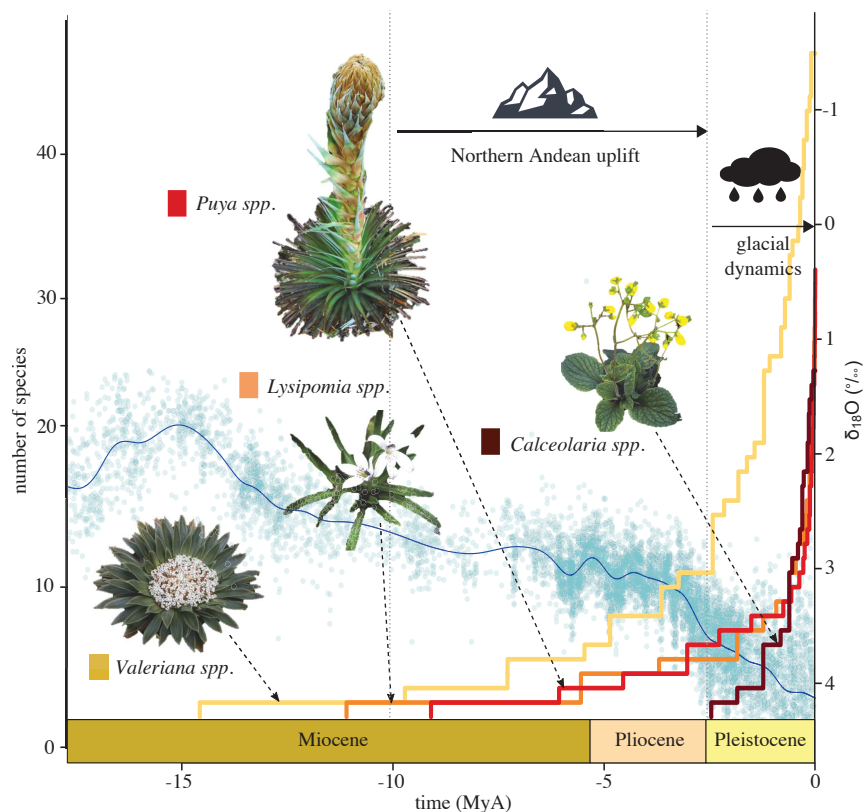




**Figure 4.** Evolutionary convergence of caulescent rosettes across tropical alpine ecosystems including: (A) Hawaiian mountains, (B) Andean páramos, (C) Peruvian jalca, (D) Central Andean puna grassland and (E) Afroapine habitats. Such repeated evolution translates adaptive values of this life-form into these habitats.

certaines radiations ont commencées suite à la formation géologique des páramos, la dynamique intense de diversification semble bien s'être produite pendant les fluctuations thermiques des activités glaciaires lors du Pléistocène (Fig. 5).

En raison de son histoire géologique très récente, de sa modélisation topographique au cours des activités glaciaires et de ses conditions de vie particulières nécessitant des adaptations clés, les páramos fournissent un cadre remarquable d'étude évolutive de l'origine de la biodiversité néo-tropicale et des modes de diversifications au sein des habitats de hautes montagnes. Néanmoins, l'importance relative des facteurs à l'origine d'une telle diversification chez ces taxons reste encore floue par le manque de phylogénies moléculaires robustes en raison d'une diversification très rapide et d'un temps insuffisant pour l'accumulation de variations génétiques (Glor, 2010; Rokas and Carroll, 2006; Wiens et al., 2008). La contribution des phylogénies moléculaires apparaît aujourd'hui comme un outil remarquable pour tester ces principes de diversification et mettre en évidence l'implication ou non des radiations adaptatives dans la diversification des (Agrawal et al., 2009; Glor, 2010) (Box 4). En effet, les phylogénies peuvent vérifier le critère d'une multiplication d'espèces et d'une descendance commune où la théorie de la radiation prédit une MONOPHYLIE chez les taxons étudiés (Schluter, 2000). De plus, au travers d'analyses comparatives, couplant les reconstructions phylogénétiques à des modèles statistiques d'évolution de traits



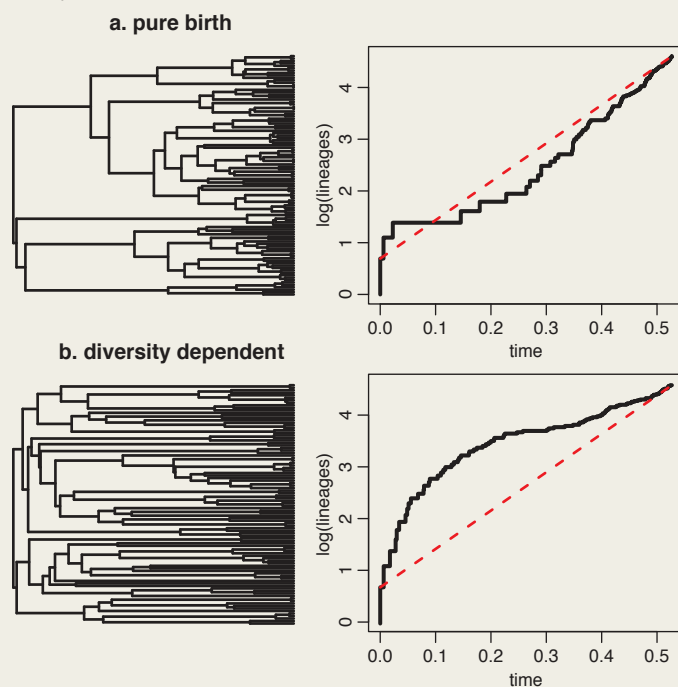
**Figure 5.** Lineages Through Time (LTT) plot estimated for páramo plant radiations. For each group, genomic data were collected from GenBank only for species distributed into páramos thanks to the Páramo Plants Online web database (Madriñan et al. 2015). Phylogenies were then inferred from the concatenation of available regions using RAxML-HPC (Stamatakis 2006) with 500 bootstrap replicates. Chronograms were inferred using age estimates from Madriñan et al. (2013) and the R function *chronopl* of ape package. Blue points and line shows high-resolution deep-sea oxygen records ( $\delta_{18}\text{O}$ ) estimating paleotemperatures of sea surface (Zachos et al. 2001). According to the LTT plot, most of the speciation events occurred during the Pleistocene when the climate was subject to important glacial dynamics.

(morphologiques ou écologiques), des corrélations phénotype-environnement, ou bien des caractères les plus divergents entre espèces sœurs, peuvent être établies pour diagnostiquer l'aspect adaptatif ou non des radiations. Enfin, le côté temporel de ces diversifications peut également être testé à partir des phylogénies afin d'établir des taux de diversification des traits ou des lignées au cours du temps (Fig. 5). Au niveau des páramos, bien que des exemples spectaculaires de ces radiations ont récemment été démontrés chez les *Bartsia* (Uribe-Convers and Tank, 2015), Espeletiinae (Diazgranados and Barber, 2017), *Hypericum* (Nürk et al., 2013), *Jamesonia* + *Eriosorus* (Sánchez-Baracaldo and Thomas, 2014), *Lupinus* (Drummond et al., 2012b) ou *Valeriana* (Bell et al., 2015), peu d'études présentent des phylogénies robustes. Toutefois, l'avènement des nouvelles technologies de séquençage (les NGS) offre maintenant de nouvelles possibilités en phylogénomique permettant de contourner ce manque de données et d'apporter de nouvelles perspectives quant à la diversification de ces plantes dans les páramos, comme il a été démontré chez les *Lupinus* (Nevado et al., 2016), les Lobelioideae (Lagomarsino et al., 2016) ou *Diplostephium* (Vargas et al., 2017).

## Box 4: Phylogénie et diversification

Les taux de diversification des lignées sont mesurés comme la combinaison des taux d'extinction et de spéciation au cours du temps. Pour la majorité des groupes vivant sur la planète, il est difficile d'accéder à des registres fossiles pour pouvoir estimer cette mesure de diversification au travers de mesures d'extinctions chez les taxons. Dans ce contexte, les phylogénies, traduisant l'histoire évolutive de la spéciation et de l'extinction des espèces et permettant d'accéder à des événements passés de diversification, constituent une approche alternative très répandue pour étudier la diversification des espèces. L'étude phylogénétique de la diversification se concentre sur deux aspects principaux issues des arbres phylogénétiques: (1) la date des événements de divergence où les distance phylogénétique sont traduites en unité de temps de divergence, et (2) la topologie des espèces sur l'arbre phylogénétique généré. Le principe de cette approche est basé sur la comparaison des phylogénies empiriques obtenues à des phylogénies théoriques reconstruites sous divers modèles de diversification sous des approches statistiques. Le modèle le plus simple est un modèle *birth-death* où les taux de spéciation et d'extinction sont fixés et homogènes entre les espèces. Sous ce modèle la diversification des clades est attendue à croître de manière exponentielle au cours du temps. Le modèle *pure-birth* implique un taux de spéciation unique et constant au cours du temps sans aucune extinction. Le nombre de lignée au cours du temps, obtenus à partir des phylogénies aux travers de LTT plot, augmente également de manière exponentielle au cours du temps (ou linéaire à l'échelle log) (Fig. a.), au même titre que le taux de diversification sous un modèle *birth-death*. La diversification des lignées n'apparaît cependant pas uniquement constante au cours du temps, et peut résulter de nombreux facteurs biotiques ou abiotiques, comme en peuvent témoigner les radiations évolutives. Pour tenir compte de ces processus, certains modèles de diversification, associés à des données de traits et des données paléo-environnementale, vont considérer une dépendance du taux de diversification en fonction du temps ou de l'environnement. Au même titre, d'autres modèles vont inclure une dépendance de la diversité spécifique des clades dans la variation des taux de diversification au cours du temps, prédisant une diversification exponentielle des espèces au début de la divergence suivie d'un ralentissement de la diversification jusqu'à une saturation (Fig. b.). Ce processus constitue un attendu théorique au cours de radiations adaptatives, au niveaux desquelles les espèces vont fortement diversifier jusqu'à une saturation des niches écologiques réduisant les opportunités écologiques disponible dans le milieu. Finalement, d'autres modèles de

diversification ont également été établie afin de tolérer un équilibre entre les taux de migrations et d'extinction ou de tenir compte d'événements de spéciation géographique et écologique, ou de dépendance de traits morphologiques. L'étude de tels modèles de diversification phylogénétiques permet ainsi de tester différentes hypothèses quant au temps et au mode de diversification des espèces: à savoir par exemple comment cette diversification varie au cours du temps, si elle peut être diversité-dépendante, homogène entre les clades ou encore adaptative entre les lignées.



Références: Glor 2010, Morlon 2014. Figure: LTT plots obtenues par simulations de phylogénies sous modèles *pure-birth* et de diversité dépendance à partir du package R *phytools* (Revell 2012).

## 4 | RADIATION DES ESPELETIINAE: CONTEXTE PHYLOGÉNÉTIQUE ET PROBLÉMATIQUE

Les Espeletiinae Cuatrec. (Asteraceae: Heliantheae : Millerieae Lindl.), connues également sous le nom du complexe *Espeletia* ou encore localement comme *frailejones*, constitue un groupe particulièrement emblématique et endémique des écosystèmes des páramos. Ce groupe a été décrit et collecté pour la première fois par Mutis en 1792, et fut officiellement publié par Humboldt et Bonpland en 1809 dans leur ouvrage *Plantae Aequinoctiales* suite à leur expédition en Amérique du Sud, sous le genre *Espeletia* H. & B. (Humboldt and Bonpland, 1809). Néanmoins, la majeure partie de la description morphologique, écologique, systématique ou encore évolutive de ces plantes fut apportée presque deux siècles plus tard par Cuatrecasas (1976, 1986, 2013). Ce groupe, formé actuellement de 144 espèces (Cuatrecasas, 2013; Diazgranados, 2012), est divisé en huit genres selon la forme de vie des espèces et la structure de leur inflorescence, présentant une incroyable combinaison de traits morphologiques en probablement moins de 4 MyA (Cuatrecasas, 1986, 2013) (Fig. 6). Les Espeletiinae sont majoritairement représentés par une forme en rosette caulescente chez les genres *Espeletia* (72 spp.), *Espeletiopsis* Cuatrec. (24 spp.), *Coespeletia* Cuatrec. (7 spp.), *Ruilopezia* Cuatrec. (24 spp.) et *Paramiflos* Cuatrec. (1 sp.), impliquant de nombreuses variations morphologiques au niveau végétatif dans la taille du tronc (<10 cm à 15 m), ou la forme et la surface des feuilles (~5 cm<sup>2</sup> à 500 cm<sup>2</sup>) par exemple, et au niveaux reproductif dans la position de l'inflorescence (terminal chez *Ruilopezia* vs. latéral et opposé chez *Espeletia* ou latéral et alterné chez *Coespeletia/Espeletiopsis/Paramiflos*), ou dans le nombre de capitules (de 1 à > 1,000) (Cuatrecasas 2013). On retrouve également chez ces plantes, une forme arborescente (feuilles courbées avec une gaine tubulaire, tronc avec un faible volume de moelle) caractéristique des genres *Carramboia* Cuatrec. (4 spp.), *Libanothamnus* Ernst (11 spp.) et *Tamania* Cuatrec. (1 sp.) avec des variations dans l'intensité de ramification du tronc principale, et dans la structure de l'inflorescence (alternée et terminale chez *Libanothamnus/Tamania* vs. opposée et latérale chez *Carramboia*) (Cuatrecasas, 2013). L'ensemble de cette diversité de traits permettrait à ces plantes de s'adapter aux conditions particulièrement extrême de ces habitats tropicaux alpins d'autant plus que les Espeletiinae sont très répandus au niveau des páramos en Colombie (~86 spp.) et au Venezuela (~57 spp.), et présentent une distribution écologique très variée. On retrouve ainsi ces plantes aussi bien dans la forêt andine à de plus faible altitudes (~2,000 m) qu'à la limite des glaciers (~4,800 m), dans des milieux secs, très humides ou rocailleux, ou au niveau de pentes et de crêtes. Plus de 120 espèces apparaissent endémique de l'étage du páramo s'étendant entre le subpáramo arbustif, les prairies du páramo et le superpáramos.

Tandis que de nombreux taxons des páramos semblent originaires d'Amérique du Nord, suite à la surrection de l'Isthme du Panama à la fin du Pliocène (3-5 MyA) (Antonelli and Sanmartín, 2011a), comme les genres *Lupinus*, *Valeriana*, *Calceolaria*, *Gentianella* ou *Draba*, la radiation des Espeletiinae trouverait une origine néo-tropicale au même titre que *Aragoa*, *Lysipomia*, *Diplostephium*, ou *Jamesonia* + *Eriosorus* (Sklenář et al., 2011). La première hypothèse évolutive sur ce complexe fut émise par Smith et Koch (1935), où seulement 30/144 espèces étaient référencées. D'après les auteurs, ces plantes auraient évolué

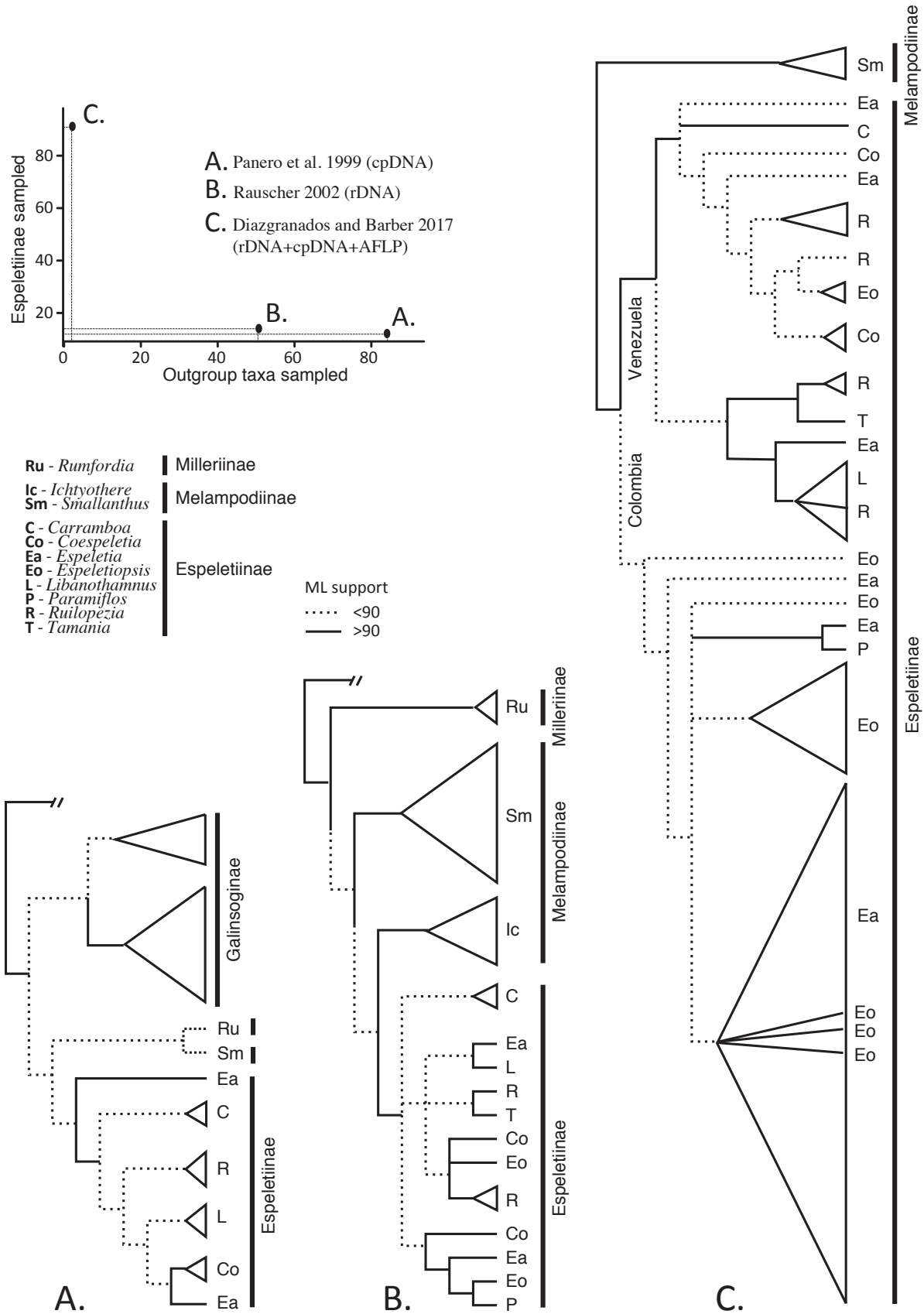


**Figure 6.** Classification of Espeletiinae into eight genera according to their vegetative and reproductive traits (Cuatrecasas 2013). Caulescent rosettes are characterized by lateral inflorescence with opposite structure in (1) *Espeletia* and (3) *Coespeletia* or with alternate structure in (2) *Espeletopsis* and (5) *Paramiflos*, and by terminal inflorescence in (4) *Ruilopezia* species. Tree-form is presented in (6) *Carramboa* species including opposite and lateral inflorescence structure, and in (7) *Tamania* and (8) *Libanothamnus* species composed of terminal and alternate structure. Photos (1-4, 6-8) by Serge Aubert and (5) by the Institut Humboldt.

à partir d'un ancêtre commun au genre néo-tropical *Smallanthus*, en une forme ramifiée avec une inflorescence multiple (correspondant actuellement au genre *Libanothamnus*) jusqu'à une simplification de cette forme en rosette. Cette transition ce serait faite de formes intermédiaires MONOCARPIQUE présentant une texture de branche sub-ligneuse (retrouvées chez quelques espèces de *Ruilopezia*) commune aux arbres à des formes monocarpiques ou POLYCARPIQUES incluant des variations dans la taille des capitules, de l'inflorescence et de la structure de la tige ou du tronc (Smith and Koch, 1935). Cette relation étroite entre *Espeletia/Smallanthus* fut ensuite soulignée par Stuessy (1973) et Robinson (1981) plaçant les Espeletiinae proche des Melampodiinae (incluant *Smallanthus*, *Acanthospermum*, *Ichthyothere*, *Lecocarpus* et *Melampodium*) et des Milleriinae (incluant *Axiniphyllum*, *Rumfordia*, *Sigesbeckia*, *Milleria* et *Guizotia*). Peu de temps après, Cuatrecasas émit une seconde hypothèse quant à l'origine et l'évolution des Espeletiinae dans les páramos en confortant l'idée d'une évolution de ces plantes à partir d'un ancêtre arbustif présents dans les forêts néo-tropicale (Cuatrecasas, 1986). Contrairement aux idées de Smith et Koch (1935), selon

lui, l'évolution de la rosette se serait produite plusieurs fois à partir de lignées arbustives et non comme un évènement unique. La radiation aurait été alors initiée au Venezuela durant la fin du Pliocène et début du Pléistocène (2-4 MyA) en raison d'une plus grande diversité morphologique (7/8 genres morphologiques sont présent au Venezuela et 4/8 en Colombie). La migration de ces plantes se serait faite en premier verticalement en colonisant les páramos à de plus hautes altitudes à partir des forêts andines et horizontalement par la suite par des transferts de páramos en páramos et du Venezuela à la Colombie pendant les périodes glaciaires du Pléistocène favorisant la connections entre les páramos. Le manque de *pappus* impliquant une dispersion gravitationnelle (Cuatrecasas, 1986, 2013) et la présence de pollinisateurs dispersant majoritairement à de courtes distances chez ces plantes (principalement des abeilles) (Berry and Calvo, 1989; Fagua and Gonzalez, 2007; Rundel et al., 1994) suggèrent néanmoins une faible capacité de dispersion entre les différents páramos et les cordillères, laissant des doutes sur l'histoire évolutive de ce complexe.

Les premières phylogénies moléculaires réalisées visant à tester ces hypothèses évolutives et conforter la systématique de Cuatrecasas (1986, 2013), viennent confirmer l'origine néo-tropicale de ces plantes (Panero et al., 1999; Rauscher, 2002). Ces études démontrent une relation phylogénétique robuste des Espeletiinae avec les genres *Smallanthus*, *Rumfordia* et *Ichthyothere* de la sous-tribu des Melampodiinae, tout en soulignant la monophylie du complexe, jusqu'alors non vérifiée. Néanmoins, les relations phylogénétiques à l'intérieur du complexe n'apparaissent pas résolues en raison d'un faible échantillonnage taxonomique réalisé et du faible signal phylogénétique fournit par ces marqueurs due à l'évolution très récente de ces plantes (Panero et al., 1999; Rauscher, 2002) (Fig. 7). Plus récemment, Diazgranados et Barber (2017) ont augmenté considérablement cet échantillonnage dans de nouvelles phylogénies (Fig. 7). Néanmoins, si l'hypothèse de Cuatrecasas (1986) d'un centre présumé de diversification au Venezuela ne semble pas être supportée par l'observation de deux radiations indépendantes en Colombie et Venezuela à l'origine du groupe, le signal phylogénétique reste faible pour élucider les questions taxonomiques dans le groupe (Diazgranados and Barber, 2017) (Fig. 7). La compréhension de des relations évolutives au sein des Espeletiinae s'avère ainsi particulièrement difficile par le manque de résolution phylogénétique laissant de nombreuses questions encore peu renseignées quant à la systématique établie par Cuatrecasas, aux relations phylogénétiques ou encore à l'évolution morphologique, spécifique et écologique dans la radiation des Espeletiinae. De plus, de nombreux événements d'hybridation ont été référencés chez ces plantes (Diazgranados, 2012; Morillo and Briceño, 2007; Rauscher, 2000). Si ce processus, fréquemment répandu lors des radiations évolutives récentes peut être moteur de spéciation (Abbott et al., 2013; Mallet, 2007; Seehausen, 2004), il peut également fortement biaiser les méthodes de reconstructions phylogénétiques conventionnelles (Degnan and Rosenberg, 2009; Pease et al., 2016), rendant encore plus difficile l'obtention d'un signal phylogénétique robuste chez ces plantes. L'arrivée des nouvelles générations de séquençages (les NGS) a toutefois permis de résoudre des radiations récentes, à partir du développement de nouvelles techniques de génomique et de nouvelles méthodes d'inférence phylogénétiques (Fernández-Mazuecos et al., 2018; Léveillé-Bourret et al., 2018). Ces nouvelles stratégies permettent d'accéder à un grand nombre de fragments moléculaires chez des taxons non modèles utiles pour les phylogénies par des méthodes de réduction de génome comme le *RAD-sequencing*



**Figure 7.** Phylogenetic reconstructions of Espeletiinae made by (A) Panero et al. (1999), (B) Rauscher (2002) and (C) Diazgranados and Barber (2017). Dashed branches on phylogenetic trees highlight low support (ML < 90). Upper panel shows ratio of outgroup/ingroup species sampled into phylogenies.

(Eaton and Ree, 2013; Wagner et al., 2013), de capture de séquences cibles (Bragg et al., 2016; Lemmon et al., 2012) ou des méthodes de *genome skimming* (Malé et al., 2014; Straub et al., 2012). L'accès à cette multitude de fragments a également rendu possible le développement de techniques nouvelles de reconstruction phylogénétique prenant en compte des biais possibles liés à l'évolution récente des taxons comme l'hybridation (Kubatko, 2009; Mirarab et al., 2014; Solís-Lemus et al., 2017). L'utilisation de ces techniques de *genome skimming* et de *RAD-sequencing* en association à ces méthodes phylogénétique a permis, par exemple, de résoudre très récemment la radiation du genre *Diplostephium* dans les páramos jusque là non élucidée (Vargas et al., 2017).

Dans ce contexte, ce travail a pour objectif de répondre à ces questions portant sur l'évolution des Espeletiinae en se focalisant sur l'utilisation de données génomiques issues de NGS afin de (1) reconstruire l'histoire phylogénétique et clarifier la systématique du complexe, (2) d'étudier les mécanismes de diversification écologique et spécifique impliqués dans la radiation évolutive de ces plantes, et enfin (3) d'évaluer l'importance de l'hybridation dans leur évolution. Chacun de ces trois axes représente les différents chapitres de ce travail de thèse. Le chapitre 1, constituant le point de départ des autres parties, se focalise sur l'utilisation de *genome skimming* avec le développement analytique d'un pseudo génome nucléaire de référence afin de résoudre la phylogénie des différents genres du complexe et d'apporter des premières réponses quant à la taxonomie et à l'évolution de ces plantes dans les páramos. L'influence de l'hybridation sur la phylogénie des genres, à une large échelle, est également évaluée afin de tester la robustesse des arbres générés et de détecter des signaux d'introgressions à la base de la radiation, pouvant être moteur à la diversification. Ce chapitre fait l'œuvre d'une publication dans *Systematic Biology* (Pouchon et al., 2018). Les deux autres chapitres se repose sur des manuscrits d'article. Le second chapitre, en lien direct avec le premier, a pour but de déterminer les différentes composantes explicatives de la radiation de ces plantes dans ces écosystèmes, à savoir l'influence du climat, de l'orogénèse des Andes ou encore de l'évolution de traits adaptatifs sur les patrons de spéciations, en répondant aux différentes questions soulevées dans ce premier chapitre. Pour cela, des données de type *shotgun* ont également été générées et utilisées en améliorant l'effort d'échantillonnage taxonomique pour la phylogénie. Le dernier chapitre se focalise quant à lui sur l'importance de l'hybridation dans l'évolution de ces plantes, à une échelle plus fine cette fois-ci sur des couples d'espèces, en ciblant et en testant d'un point de vue génétique des événements identifiés dans le premier chapitre. Ce chapitre 3 combine l'utilisation de données *shotgun* avec cette fois-ci l'utilisation de données issues de *RAD-sequencing* à cette échelle. Une annexe propose également de comparer l'utilisation de ces deux techniques dans l'inférence phylogénétique de cette radiation. Finalement, l'ensemble des principaux résultats obtenus est résumé en une conclusion générale à la fin de ce manuscrit où des perspectives sont également énoncées. Bien que le premier chapitre amène à revoir la taxonomie du groupe, pour une question de visibilité et de clarté, la nomenclature de Cuatrecasas (2013) sera conservée tout au long de ce manuscrit.



---

# CHAPTER 1 – Phylogenomic Analysis of the Radiation of Espeletiinae in the Tropical Andes

---

*“No zone of alpine vegetation in the temperate or cold parts of the globe can well be compared with that of the Páramos in the tropical Andes.”*

Alexander von Humboldt, 1850

## 1 | INTRODUCTION

There has been a persistent debate about the causes of the outstanding species richness in northern South America, and particularly in the Andean-Amazonian region (reviewed in Antonelli and Sanmartín, 2011; Hoorn et al., 2010, 2011; Luebert and Weigend, 2014; Madriñán et al., 2013; Rull, 2011a, 2011b). Two broad types of explanations have been traditionally proposed. On one hand, the Miocene/Pliocene Andean orogeny is generally considered to be an important driver of species diversification in the Neotropics, as it provided new mountain habitats and deep valleys that increased habitat heterogeneity and acted as both intra and inter Andean vicariant barriers, e.g. *Hedyosmum* (Antonelli and Sanmartín, 2011b), *Macrocarpaea* (Struwe et al., 2009) and Rubiaceae (Antonelli et al., 2009). On the other hand, the Pleistocene climatic (glacial) cycling undoubtedly impacted the adaptation, diversity and distribution of organisms, particularly at high elevation habitats in the Northern Andes (a.k.a. paramo) (Hooghiemstra et al., 2006; Vuilleumier and Monasterio, 1986). Thus, during the last 2.4 Myr, glacial periods during which paramos were relatively larger and inter-connected alternated with interglacial periods during which paramos tended to be relatively smaller and isolated “sky islands” (Vuilleumier and Monasterio, 1986). These

repeated cycles of ecological connectivity and spatial isolation could have acted as a “species pump,” increasing significantly the speciation rates in high-elevation Andean taxa as consequence of the joint action of allopatry, diversifying selection and adaptation, e.g., *Halenia* (von Hagen and Kadereit, 2003), *Hypericum* (Nürk et al., 2013) and *Lupinus* (Drummond et al., 2012b). More recently, biotic factors such as plant-pollinator interactions have also been considered as important drivers in the diversification of Andean taxa (Abrahamczyk et al., 2014; Lagomarsino et al., 2016; Schmidt-Lebuhn et al., 2007).

A proper assessment of the relative importance of these factors on the evolution and diversification of Andean taxa requires integrative analyses that combine morphologic, geographic, and ecological data into a time-calibrated species-level phylogenetic framework. Several examples of spectacular high-elevation Andean diversification have been documented recently, e.g. *Bartsia* (Uribe-Convers and Tank, 2015), Espeletiinae (Diazgranados and Barber, 2017; Rauscher, 2002), *Hypericum* (Nürk et al., 2013), *Jamesonia-Eriosorus* (Sánchez-Baracaldo and Thomas, 2014), *Lupinus* (Drummond et al., 2012) and *Valeriana* (Bell et al., 2015), but only a few have provided fully-resolved phylogenetic trees, e.g. *Lupinus* (Nevado et al., 2016), *Diplostephium* (Vargas et al., 2017). In fact, the lack of resolution has been so far a prominent feature of most phylogenetic studies of species-rich high-elevation Andean plant taxa, surely due to both the rapid diversification yielding low phylogenetic signal in molecular data and the low resolution of the molecular markers used (Maddison and Knowles, 2006). Some of these problems can now be overcome thanks to the advent of next-generation sequencing (NGS), which offers unprecedented possibilities in phylogenomics by increasing greatly both taxonomic coverage and the amount of sequence data, while remaining at a relatively low cost.

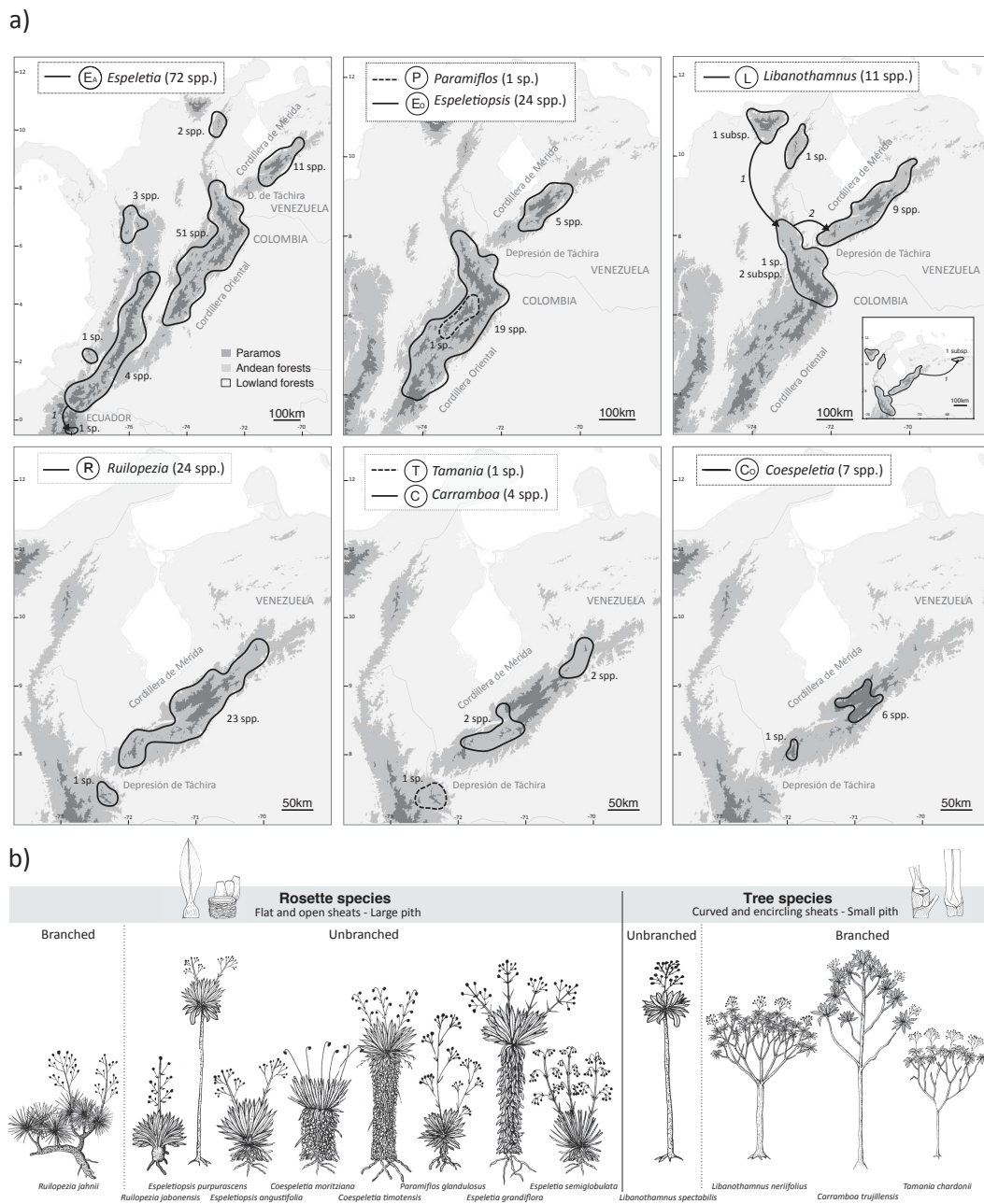
The increase in sequence data has also reinforced the view that phylogenetic trees built from different genes sometimes show conflicting branching patterns due to incomplete lineage sorting (ILS), gene flow or gene duplication/loss (Degnan and Rosenberg, 2009; Maddison and Knowles, 2006). Some of these phenomena are presumably important among plant groups that have experienced rapid and recent diversification. For instance, ILS is expected to be common because the short phylogenetic branches exhibited by rapidly diverging taxa provide insufficient time for lineages to coalesce within species between speciation events. Recurrent gene flow is also expected to be important after the secondary contact of interfertile taxa that diverged through ecological or geographical isolation, and that frequently shift their distribution ranges due to climate oscillations. Fortunately, several methods based on the multispecies coalescent have been developed recently to reconstruct the evolutionary history of species in the presence of gene-tree discordance, e.g. ILS (ASTRAL, Mirarab et al., 2014) and hybridization (SNaQ, Solís-Lemus et al., 2016). Other sources of phylogenetic discordance, such as gene duplication/loss, are often treated upstream during the selection of studied loci, in order to maximize the proportion of orthologous markers in the sequence dataset (Dalquen et al., 2013; Gabaldón and Koonin, 2013). Despite these methodological advances, there is still a persistent debate about the best strategy for species tree inference (Edwards et al., 2016; Springer and Gatesy, 2016), with some studies showing that coalescent-based multispecies methods that use the ensemble of markers perform better than the concatenation of those markers into a single sequence (Kubatko and Degnan, 2007; Roch and Steel, 2015), whereas other studies have showed the opposite under some

circumstances, e.g. low phylogenetic signal (Bayzid and Warnow, 2013; Mirarab et al., 2014). In consequence, most studies use both approaches and perform statistical tests to infer the extent of conflict between them and among genes, e.g. Nevado et al. (2016).

### The Exceptional Diversity of Espeletiinae (Asteraceae)

The subtribe Espeletiinae comprises about 144 species characterized by a unique combination of: diploid karyotype ( $N = 19$ ), spiral phyllotaxis, xeromorphic structure and organization of leaves and inflorescences, fertile ray flowers, functionally male disc flowers and, with a single exception, lack of specialized seed dispersal mechanisms (Cuatrecasas, 1986, 2013). These plants are distributed in the high Andes of Venezuela, Colombia and Northern Ecuador, with an important diversity in the Colombian Cordillera Oriental (~76 spp.) and the Venezuelan Cordillera de Mérida (~59 spp.) (Fig. 1a). Other areas within the distribution of Espeletiinae harbor considerably lower species diversity, a pattern that has been attributed to their isolation from the presumed center of origin in the Venezuelan Cordillera de Mérida and the lack of long-distance dispersal mechanisms in these plants (Cuatrecasas, 2013). Espeletiinae originated from a common ancestor likely around 3–5 Ma (Madrinán et al., 2013; Rauscher, 2002) and are thus considered as one of the fastest examples of continental plant diversification. In this relatively short time they have accumulated a remarkable ecological variation in regards with 1) elevation ranges: from upper montane cloud forests at about 2500 m to the edge of glaciers at 4800 m, 2) moisture requirements: from wet paramo bogs to xeric periglacial talus slopes and rocky outcrops, 3) light conditions: from montane forest understory to open grasslands of paramo, and 4) reproductive strategies: with polycarpic species that reproduce repeatedly across adult life and monocarpic species that reproduce only once before death. In addition, Espeletiinae also exhibit a large morphological disparity: caulescent (stemmed) rosettes with dense foliar pubescence and thick layers of dead leaves covering the stem, caulescent rosettes with leafless stems, stem-less and dwarf rosettes, large and profusely branched trees, dichotomous trees, and shrubs (Cuatrecasas, 1986, 2013) (Fig. 1b).

Throughout this article, the rosette growth-form is defined as a compact crown of green leaves arranged spirally around the apical meristem, each leaf possessing a flat sheath tightly appressed under and over the sheaths of nearby leaves. The stem consists of a large central pith surrounded by secondary vascular tissues, entirely covered in its apical part by the sheaths of the green leaves. Most rosette species have a unique, unbranched stem, but a few can be slightly branched nearby ground level, each branch producing a terminal crown of rosette leaves. In contrast, tree/shrub growth-forms have a distinct woody trunk with a small or absent pith in comparison with the surrounding wood. They are profusely branched at a considerable distance from the ground and show a less compact leaf arrangement, although with some crowding at the end of the branches. The leaf sheaths of these forms are either curved with free margins and embracing part of the branch, or tubular and encircling the branch entirely. All Espeletiinae fit into these two broad types, with only two known exceptions: *Libanothamnus spectabilis* is a tall caulescent rosette with curved and embracing sheaths, while *Tamania chardonii* is a treelet with relatively flat leaf sheaths (Fig. 1b).



**Figure 1.** a) Species diversity and approximate distributions of the eight genera of Espeletiinae. Arrows provide numbers of shared species between main geographic areas. b) Schematic representations of main growth-forms.

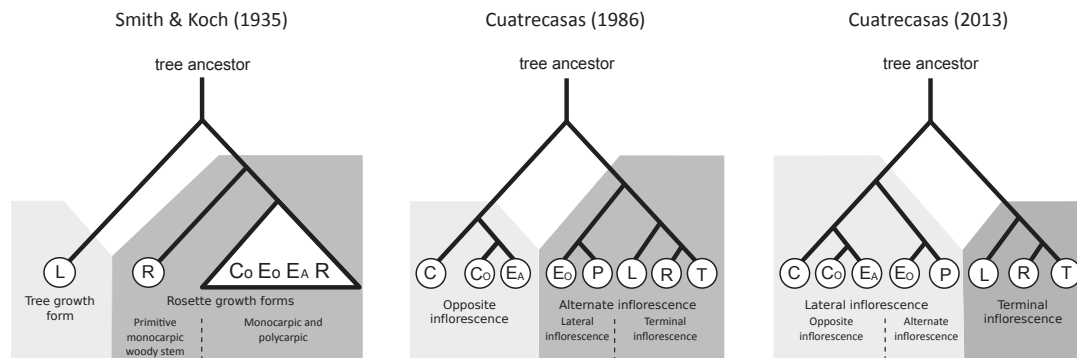
The latest systematic treatment of Espeletiinae divides the subtribe in eight genera regrouped in two large clades split according to the position of the inflorescence (Cuatrecasas, 2013; Diazgranados, 2012): 1) with terminal inflorescences in the trees *Libanothamnus* Ernst (11 spp) and *Tamania* Cuatrec. (1 sp), and the rosettes *Ruilopezia* Cuatrec. (24 spp) and 2)

with lateral inflorescences in the trees *Carramboa* Cuatrec. (4 spp) and the rosettes *Coespeletia* Cuatrec. (7 spp), *Espeletia* Mutis ex Bonpl. (72 spp), *Espeletiopsis* Cuatrec. (24 spp) and *Paramiflos* Cuatrec. (1 sp). This system contrasts with a previous one by the same author (Cuatrecasas, 1986), in which the same genera were grouped according to the structure of the inflorescence: 1) with alternate (monochasial) inflorescences in *Libanothamnus*, *Tamania*, *Espeletiopsis*, *Paramiflos* and *Ruilopezia* and 2) with opposite (dichasial) inflorescences in *Carramboa*, *Espeletia* and *Coespeletia* (the alternate inflorescences in *Coespeletia* would be a secondary condition). In this context, molecular phylogenetic methods can be used to assess the relative importance of the position and the structure of the inflorescence on the evolution and the systematic treatments of the group.

### Earlier Works of Phylogenetics and Evolution in Espeletiinae

Despite the long-standing interest generated by their exceptional morphological and ecological diversity, the diversification dynamics of Espeletiinae and its potential drivers have remained largely unknown. The earlier works of Panero et al. (1999) on cpDNA and Rauscher (2002) on ITS1-5.8s-ITS2 dealt more with the closest relatives and positioning of the subtribe Espeletiinae within the tribe Heliantheae than with its internal relationships and recent evolution. Indeed, the evolutionary relationships within Espeletiinae have been particularly difficult to disentangle due to lack of phylogenetic resolution (Diazgranados and Barber, 2017; Rauscher, 2002). Nonetheless, a certain number of hypotheses regarding patterns of molecular, morphologic and geographic evolution in Espeletiinae have been formulated. For instance, first Smith and Koch (1935), and later Cuatrecasas (1986) considered that the group evolved from a branched woody tree or shrub ancestor, similar to the tree genus *Libanothamnus* in the Cuatrecasas' classification, an idea that reflects the supposed ancestral habit of the tribe Heliantheae (Cuatrecasas, 2013; Smith and Koch, 1935). However, as noted by Rauscher (2002), most of the closest relatives of Espeletiinae are herbs, therefore, if the most recent common ancestor of the subtribe was a woody tree, it must have been the result of a recent evolution from herbaceous ancestors near the origin of the subtribe.

It has also been hypothesized that the radiation of Espeletiinae must have started in the Venezuelan Cordillera de Mérida given the greater morphological diversity occurring there (Cuatrecasas, 1986, 2013; Smith and Koch, 1935) or a supposed older age of species in this range (Diazgranados and Barber, 2017). According to this view, Espeletiinae lineages that are non-endemic to the Cordillera de Mérida, mainly the trees *Libanothamnus* and the rosettes *Espeletia* and *Espeletiopsis* (Fig. 1a), would have later crossed the ~30 km wide Táchira depression and spread to the Colombian Cordillera Oriental, presumably during Pleistocene glacial periods that favored connectivity among paramos. However, such view has no support in previous phylogenetic (Diazgranados and Barber, 2017; Rauscher, 2002) or paleobotanical reconstructions of the distributions of paramos during the Pleistocene (Vuilleumier and Monasterio, 1986). Finally, some conflicting views have as well emerged in regard to the evolution of the caulescent rosette growth-form within Espeletiinae. Smith and Koch (1935) suggested that it evolved from the tree ancestor only once, followed by repeated evolutions of monocarpic and polycarpic rosette growth-forms. On the other hand, Cuatrecasas (1986, 2013) suggested that the caulescent rosette evolved multiple times from the tree ancestor



**Figure 2.** Hypotheses of morphological evolution in Espeletiinae (codes for genera as in Fig. 1).

lineage and without repeated evolution of monocarpy/polycarpy (Fig. 2). Unfortunately, the lack of phylogenetic resolution in previous analyses has prevented a conclusive test of these ideas, e.g. Rauscher (2000); Diazgranados and Barber (2017).

In this work, we produce the first fully resolved phylogenetic tree of all the morphological groups of Espeletiinae, based on the analysis of genome shotgun sequence data. The study provides sequences for both the plastome and a large number of nuclear contigs obtained with an original *de novo* assembly procedure that does not require a reference genome. We use this data to run phylogenomic analyses to test previous ideas on the diversification of this emblematic group of Andean plants, particularly with regard to the following questions: 1) When did the radiation of Espeletiinae start? Was it associated with Andean Pliocene tectonics, Pleistocene climatic cycles or both? 2) What was the ancestral morphological growth-form of Espeletiinae? 3) Is there evidence for long-distance dispersal events between the Colombian Cordillera Oriental and the Venezuelan Cordillera de Mérida, particularly among rosette lineages? 4) Do the different growth-forms in the subtribe represent unique evolutionary events or instead multiple/repeated evolutions of morphologically similar growth-forms? Did these changes occur in a predictable way, e.g. from trees to rosettes and from lower to higher elevations?

## 2 | METHODS

### Sampling, DNA Extraction, and Sequencing

A total of 41 samples were taken from an equal number of Espeletiinae taxa in the Andes of Colombia and Venezuela, representing the eight previously recognized genera and all the known morphological diversity of the subtribe in terms of growth-form (trees and rosettes), inflorescence structure (opposite and alternate), inflorescence position (terminal and lateral) and reproductive strategy (monocarpic and polycarpic) (Supplementary Table S1 at

<http://dx.doi.org/10.5061/dryad.m1t32>). Species were identified following the keys provided in Cuatrecasas (1996, 2013). Following Rauscher (2002), *Smallanthus* aff. *pyramidalis* and *Rumfordia penninervis* in the closely related subtribe Milleriinae were also sampled and considered as outgroups for phylogenetic reconstructions. To assess the possible impact of using a single sample per taxon on the phylogenetic reconstruction of such a recent evolutionary group, two samples were used in species exhibiting considerable intra-specific variation, such as *Carramboa badilloi* (*C. badilloi* var. *badilloi* and *C. badilloi* var. *pittieri*) and *Espeletia schultzii* (*E. schultzii* var. *schultzii* and *E. schultzii* var. *mucurubana*).

Leaf samples were collected, shaved with a razor blade to eliminate as much foliar pubescence as possible and placed in airtight plastic bags with silica-gel. To increase yield, DNA extractions were done with a modified protocol of the DNeasy Plant Mini Kit (QIAGEN), using 40% more than recommended grinded dried leaf tissue (and reagents) and overnight incubations at 65°C for the cell lysis step. Genomic shotgun libraries were prepared and sequenced according to the manufacturer's instructions in an Illumina HiSeq 2000 at FASTERIS (Geneva, Switzerland).

### Assembly of Genomic Regions

Paired-end 2 × 100 bp Illumina reads were filtered out using a minimum Phred score quality threshold value of 20 (i.e., 99% correct base calling) with FASTX ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). The remaining reads were used for a de novo assembly of the plastid genomes and nuclear contigs (referred to as Total Number of Reads, TNR, in Supplementary Table S1).

*Plastomes*— Assembly of plastid genomes was performed using the Organelle Assembler described in <http://metabarcoding.org/asm>. The sequence seeds necessary for the *buildgraph* command consisted of chloroplast proteins from *Arabidopsis thaliana* provided with the assembler. In order to optimize the assembly procedure, the *cutlow* coverage value was readjusted for each species after inspection of its assembly graph. Plastomes that could not be assembled with this approach were obtained by aligning the genomic reads of concerned species onto a reference composed by the *de novo* assembled plastomes using the BWA v0.7.5 (Li and Durbin, 2009). We considered only reads with maximal MAPQ scores of 60 for variant calling using *SAMtools* v0.1.19 and *BCFtools* v0.1.19 (Li et al., 2009). To ensure calling quality, SNPs were filtered using the *vcfutils* script (provided by *SAMtools*) with a root-mean-square of the SNP mapping quality  $\geq 15$  and with a minimum coverage of 10 per sample.

*Nuclear contigs*— Given the unavailability of reference genomes in the Heliantheae at the time of this study (July 2016), we used a *de novo* assembly for nuclear contigs of Espeletiinae according to the pipeline described schematically in Supplementary Fig. S1. In short, to isolate nuclear sequences, BWA was used to remove all the reads of each species that aligned to the mitochondrial genome of the sunflower (*Helianthus annuus*, GenBank accession number KF815390) and the assembled plastomes of Espeletiinae. Reads that aligned to an available database of the repeated elements of the sunflower genome (<http://www.agr.unipi.it/ricerca/plant-genetics-and-genomics-lab/sequence-repository.html>) were also removed, in order to avoid problems during SNP base calling steps (step 1,

Supplementary Fig. S1). All remaining unmapped reads were considered as potential nuclear reads (referred to as PNR in Supplementary Table S1) and assembled into contigs using *Velvet* v.1.2.10 (Zerbino and Birney, 2008) (step 2, Supplementary Fig. S1). To optimize the assembly, the hash length (or k-mer length) was automatically selected using *VelvetOptimiser* v.2.2.5 (<https://github.com/tseemann/VelvetOptimiser>), according to the N50 value. Coverage *cutoff* and expected coverage were automatically selected by *Velvet* but only contigs >200 bp were considered (Contigs, Supplementary Table S1). Contigs that still matched mitochondrial (*Helianthus*) and plastid (Espeletiinae) sequences were removed using a Python script. Only the remaining sets of nuclear contigs were retained for further analyses (referred to as Final Number of Contigs, FNC, in Supplementary Table S1).

In order to find shared nuclear regions, nuclear contigs from all species were pooled and assembled into composite multispecies contigs (hereafter metacontigs) using *VelvetOptimiser* (step 3, Supplementary Fig. S1). Only metacontigs assembled from at least three different species were retained. Next, the nuclear reads of all species were pooled and aligned to this set of metacontigs, from which coverage and position were extracted using *BEDtools* v2.17.0 (Quinlan and Hall, 2010). In order to identify and eliminate repeated elements that could still be present in the data, metacontigs with at least one region within the 10% most covered FNC were arbitrarily removed from the list using *SAMtools* and a custom Python script (step 4, Supplementary Fig. S1). The nuclear reads of each species were aligned to the remaining metacontigs using *BWA*, retaining only those that mapped to a single metacontig with a maximal MAPQ score of 60 for the variant calling using *SAMtools* and *BCFtools* (step 5, Supplementary Fig. S1). SNPs were further filtered with *vcfutils* using a root-mean-square  $\geq 15$  and a minimum SNP coverage of 10 per marker and per species. Finally, to guarantee the quality of alignments, only metacontigs with less than 5% of ambiguities were kept for further phylogenetic analyses (step 6, Supplementary Fig. S1). The final number of composite nuclear metacontigs in the reference list was 9880.

## Phylogenetic Analyses

*Plastomes*—Plastid sequences were aligned using *MAFFT* v7.222 (Katoh 2002). Ambiguous sites from the alignment matrices were trimmed with *trimAl* using the heuristic algorithm *automated1* (Capella-Gutiérrez et al., 2009), which uses the distributions of gaps and similarities to determine the thresholds for removing the poorly aligned sites of an alignment. Phylogenies were next reconstructed using Bayesian Inference (BI) and Maximum Likelihood (ML) approaches. BI analyses were conducted using a standard non-clock approach with *ExaBayes* v.1.4.1 (Aberer et al., 2014). We used the default values for the temperatures of chain heating and the number of Metropolis-coupled Markov chain Monte Carlo chains for the replicate runs with 500,000 generations. Convergence and mixing were assessed using summary statistics from the *postProcParam* command. We ensured a minimum value of 200 for the effective sample sizes of each parameter, and the similarity within and between chain-variance with a potential scale reduction factor close to 1 and below 1.1. 50% majority-rule consensus phylograms and posterior probabilities were obtained using the *consense* command with a burnin fraction of 25%. ML analyses were performed using *RAxML-HPC* (Stamatakis, 2006). The best-fitting model of evolution, GTR+G+I, was determined with the Akaike



Information Criterion (AIC) as implemented in the R package *Phangorn* (Schliep, 2011) (Supplementary Table S2). 1000 ML bootstrap replicates were performed.

*Nuclear metacontigs*— Phylogenies based on the nuclear dataset were built with two approaches: using concatenated alignments of metacontigs and coalescent- based multispecies tree methods. In order to use the same data in the two phylogenetic approaches, only nuclear metacontigs shared by at least four taxa were used (i.e., quartets), which is a requisite of the coalescent-based approach. This dataset, consisting of all metacontigs shared by one outgroup and at least three other taxa, contained 1877 metacontigs and 933,394 nucleotide positions.

*Concatenated alignment approach.* The alignments of metacontigs were concatenated into a single sequence, trimmed with *trimAl* and used for BI and ML phylogenetic analyses performed with the same software described above. The ML analysis used a GTRCAT substitution model and 1000 bootstrap replicates.

*Coalescent-based multispecies tree approach I: ASTRAL.* In order to assess for the sensitivity of the concatenated approach to conflicts among gene trees due to ILS, phylogenetic reconstructions were also based on a coalescent-based multispecies tree approach as implemented in *ASTRAL II* (Mirarab and Warnow, 2015). This software has been shown to enable highly accurate species tree estimation in the presence of ILS. 1877 gene trees were first estimated using *RAxML* under the GTRCAT model with 200 bootstrap replicates performed on the alignment of each metacontig. To assess for gene conflicts among metacontigs, we used the multilocus bootstrapping procedure from *ASTRAL II* with 1000 replicates combining each *RAxML* best tree and its bootstrap replicates. A majority-rule consensus tree was generated with the R package *ape* (Popescu et al., 2012).

*Coalescent-based multispecies tree approach II: SNaQ.* We used another approach to discriminate between ILS and gene flow as possible sources of phylogenetic discordance, implemented in the software *SNaQ* (Solís-Lemus et al., 2017). *SNaQ* performs maximum pseudolikelihood estimation of phylogenetic networks from multilocus data using the coalescent-based multispecies model and quartet- based concordance analyses. This software allows for the simultaneous estimation of both a major phylogenetic topology depicting the main evolutionary history (“major edge”) and past events of hybridization (“minor edges”), while integrating optimization of branch lengths, inheritance probabilities and a heuristic search in the space of phylogenetic networks based on a starting topology and on the observed quartet concordance factors (CF). A relatively fast procedure to run *SNaQ* consists in using the *ASTRAL* tree as the starting topology and a table of the CF values obtained from the best trees inferred with *RAxML* for each metacontig. Unfortunately, modelling of the multispecies coalescent with gene flow has a heavy computational cost so that the application of *SNaQ* (and other equivalent software) is severely hampered with more than ~25 taxa (Hejase and Liu, 2016). To deal with this limitation, we used the fast procedure for *SNaQ* on 29 datasets with <25 taxa each (Supplementary Fig. S2). The first dataset, consisting of at least two members of each major clade identified in *ASTRAL* plus an outgroup, was used to identify the main topology of the phylogenetic tree connecting all major clades. The remaining 28 datasets, each consisting of all members of two major clades identified in *ASTRAL* (e.g., I+II, I+III, ... VII+VIII), were used to identify within clade topologies and to detect all possible hybridization events within and between major clades. A supertree was built with *Phangorn*

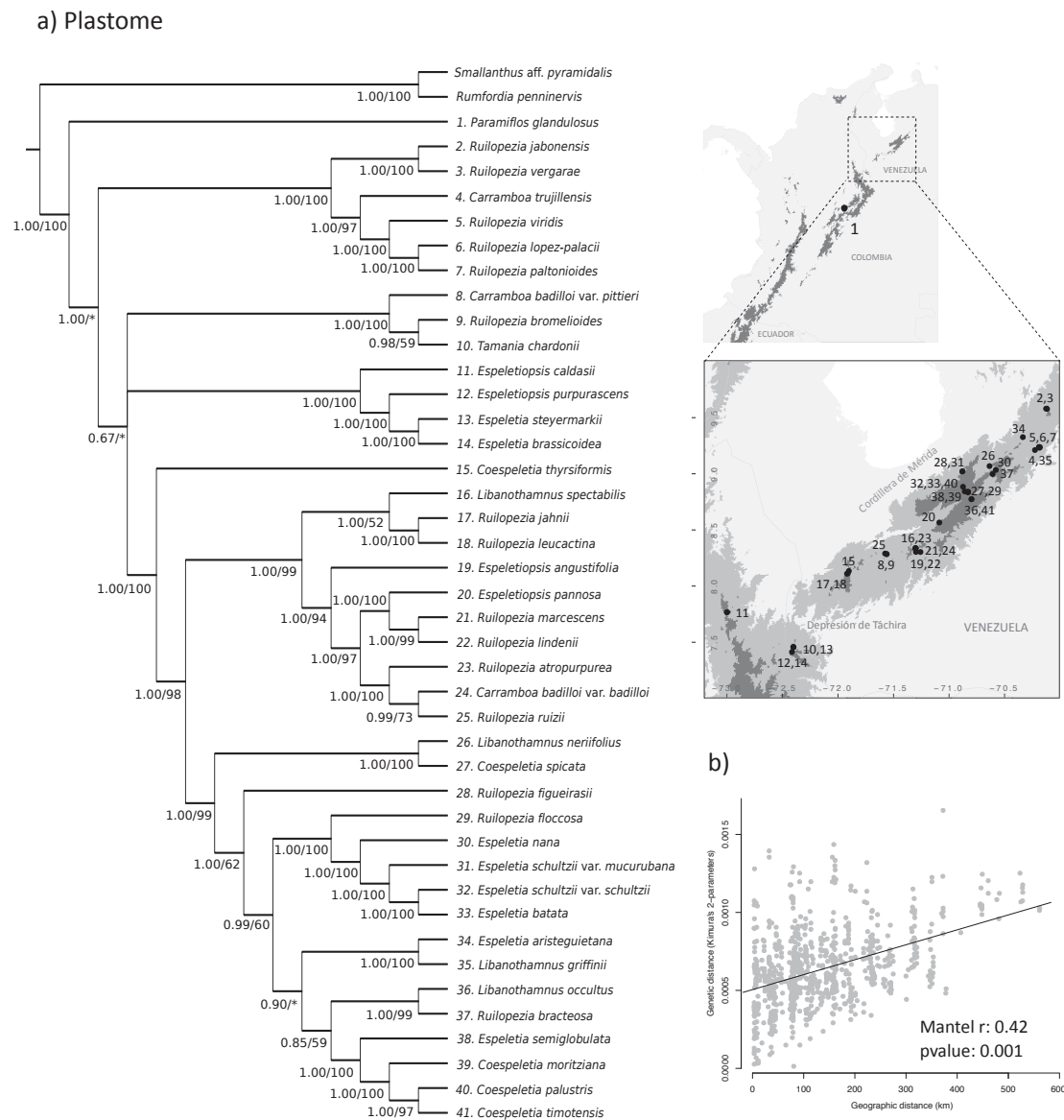
combining all 29 *SNaQ* trees. *SNaQ* was run on all 1877 *RAxML* best trees, initiated in every case with a starting topology based on the *ASTRAL* species tree pruned to the taxa list of the specific dataset. Five to ten runs were used for each value of the maximum expected number of hybridization events (parameter *h*, increased sequentially from 0 to 5). The optimal *h* was chosen using a heuristic approach by plotting pseudo-likelihoods against *h* values, following *SNaQ*'s authors recommendation (Cécile Ané, comm. pers).

*Phylogenetic tests*— To assess the significance of discordances between phylogenetic trees built with the concatenated and the coalescent-based multispecies approaches, we used *RAxML* to calculate site-by-site likelihoods for both a bootstrap set of 100 unconstrained phylogenetic trees and bootstrap sets of 100 trees generated for each alternative constrained topologies. This was used to perform Shimodaira and Hasegawa (1999) test as implemented in *CONSEL* (stat.sys.i.kyoto-u.ac.jp/prog/consel). These tests were also used to compare the phylogenetic hypotheses depicted in Fig. 2 and the monophyly of Cuatrecasas' genera of Espeletiinae (Cuatrecasas, 1986, 2013). For the Smith and Koch (1935) hypothesis, a pruned tree was produced by retaining only the 14 species known at that time that matched with our species sampling. In addition, the correlation between genetic (plastid or nuclear) and geographic distances was assessed using the R package *Ecodist* (Goslee and Urban, 2007).

## Divergence Times and Diversification Rates of Espeletiinae

Divergence times of Espeletiinae were estimated following a two-step procedure involving: 1) estimation of age of the crown node using plastid genomes of additional Asteraceae with ages known from literature and 2) calibration of the divergence time analysis within Espeletiinae using the nuclear data and the crown node age estimated in (1).

(i) The age of the crown node of the subtribe was estimated with a phylogenetic analysis of plastid sequences of Espeletiinae and some members of the Asteraceae tribes Millerieae, Heliantheae, Anthemidae and Cichorieae (Supplementary Table S3), using known ages extracted from fossil data (pollen) and/or previous molecular divergence analyses. Three calibration dates were employed (Tremetsberger et al., 2013): 22.0–28.4 Ma interval for the age of Cichorieae based on fossil pollen, 10.6–17.8 Ma interval for the divergence *Sonchus/Crepis* and 6.9–16.5 Ma interval for the divergence *Scorzonera/Tragopogon*. Plastomes were sequenced, assembled and analyzed as described above for the plastid BI phylogeny of Espeletiinae. In order to estimate a time-calibrated phylogeny, we used the relaxed-clock penalized likelihood (PL) method implemented in *r8s* v1.7.1 (Sanderson, 2003). As implemented in Buerki et al. (2010), the uncertainty in the estimation of divergence times was assessed by generating 1000 trees during the Bayesian Markov chain Monte Carlo process, discarding the first 100 (10% burnin) and retaining the remaining 900 trees. The cross-validation method was used on each tree replicate to determine the optimal level of rate-smoothing for the PL analyses by fitting smoothing parameters with 100 estimates (starting with a log<sub>10</sub> value of 1 and increments of 0.1). Dating analyses were then performed on each tree according to the optimal smoothing value. Mean values and 95% confidence intervals for each node age were estimated from the chronograms of the 900 time-calibrated trees using *TreeAnnotator* (Drummond et al., 2012a).



**Figure 3.** a) Phylogeny of Espeletiinae inferred from analyses of whole plastomes. Phylogenetic supports are represented for each node as Bayesian inference (BI/-) and maximum likelihood (-/ML) values (\*indicate support <50). Maps show sampling localities in the Colombian and Venezuelan Andes. b) Mantel tests of the correlation between geographic and genetic (plastid) distances in Espeletiinae.

(ii) A time-calibrated analysis of divergences within the subtribe was produced in *r8s* using the concatenated nuclear BI phylogenetic tree, and minimum and maximum age estimations for the crown node of Espeletiinae obtained from the previous analysis. Uncertainty in the estimation of node ages was assessed as above but with 200 trees randomly selected from the Bayesian Markov chain Monte Carlo stationary distribution. Cross-validation was conducted for 100 estimates starting with a log<sub>10</sub> value of -5 and increments of 0.1. Mean values and 95% confidence intervals of the divergence dates were estimated from the 200 time-calibrated trees in *TreeAnnotator*.

Diversification rates of Espeletiinae were analysed by testing the fit to models with density-dependence birth-death processes using the R package *DDD* (Etienne et al., 2012), and calculating rates of speciation, extinction and carrying capacities. Lineages-through-time (LTT) plots were built with the R package *ape* using the 200 nuclear chronograms.

### Reconstruction of Ancestral States

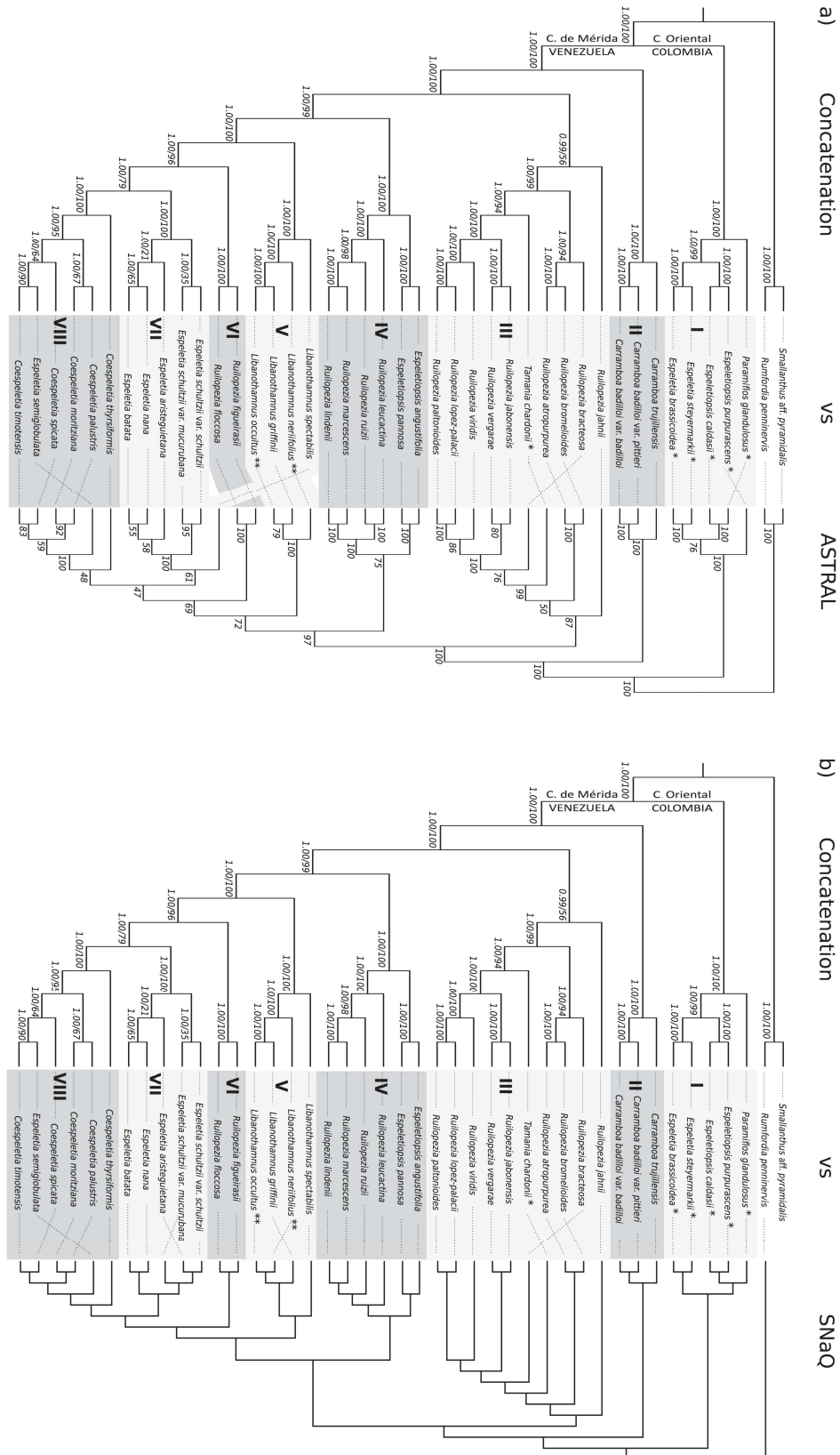
Ancestral state reconstruction was performed on the consensus of the nuclear chronograms for four morphological traits considered as binary characters: growth-form (GF, rosette/non-rosette), branching of the stem (STEM, branched/unbranched), inflorescence position on stems (IPOS, lateral/terminal) and inflorescence structure (ISTR, alternate/opposite). Information about these characters was collected from Cuatrecasas (2013). Inferences about trait evolution were done using Markov models (Mk) in the R package *Diversitree* (FitzJohn, 2012). Two possible discrete Markov models were fitted and compared using AIC values: a full model with different transition rates (ARD) and a constrained model with equal transition rates (ER). For the growth-form trait, the tree root was constrained to have the non-rosette character state owing to the absence of rosettes among close relatives of Espeletiinae within Heliantheae. No other constraints were imposed in the analyses; however, due to limited sampling of outgroup taxa, only Espeletiinae were considered for inferences about the evolution of the two inflorescences traits. The best model was used to run simulations of stochastic character mapping in order to infer most likely scenarios of trait evolution (Supplementary Table S4).

## 3 | RESULTS

Illumina sequencing returned on average  $12.8 \times 10^6$  (sd. =  $6.6 \times 10^6$ ) quality-filtered reads per species, of which 0.77% (91,055 reads, sd. = 71,839) served to assemble plastid genomes and 83.9% ( $10.7 \times 10^6$  reads, sd. =  $55.5 \times 10^6$ ) were used to build nuclear contigs (Supplementary Table S1). The nuclear data allowed the assembly of on average 7883 (sd. = 3048) contigs per species, with an average N50 of 389 bp (sd. = 32).

### Phylogenetic Analysis of Plastomes

We were able to assemble *de novo* 25 complete plastomes from 23 Espeletiinae species and the two outgroups. Plastomes of the remaining 18 species were obtained by mapping their reads on the consensus sequence of the *de novo* assemblies. The alignment contained 151,388 nucleotide sites (mean coverage of 60.1), of which 2.46% were variable and 0.60% were parsimony informative (Supplementary Table S5). The BI phylogenetic tree showed well-supported clades (mean BPP = 0.98, sd. = 0.06, Fig. 3) and only one basal polytomy. The ML phylogenetic tree was topologically similar (SH-test,  $P=0.477$ ), although with slightly lower bootstraps values (mean PP = 89.12, sd. = 19.9, Fig. 3). In spite of this high phylogenetic signal, plastid-based analyses do not support the monophyly of morphological growth-forms, since rosettes and branched species are present in all major clades. The same morpho-



**Figure 4.** Comparisons of nuclear phylogenies of Espeletiinae. a) Concatenation vs. ASTRAL. b) Concatenation vs. SNaQ. Node supports are given for concatenation as Bayesian inference (BI/-) and maximum likelihood (-/ML), and for ASTRAL as bootstrap probabilities. Species distributions: Venezuelan Andes (no marks), Colombian Andes (\*), both countries (\*\*).

phylogenetic discordances were also observed for both the structure and the position of the inflorescences. Accordingly, analyses of plastomes do not support the monophyly of any of the genera proposed by Cuatrecasas (1986, 2013). Finally, there is a strong spatial influence in the clustering of plastid diversity in Espeletiinae, as shown by the highly significant correlation between geographic and genetic (plastid) distances (Mantel test,  $P = 0.001$ , Fig3b)

### Phylogenomic Analysis of Nuclear Contigs

*Concatenated alignment approach*—The BI phylogenetic reconstruction appeared fully resolved with only one node with a BPP <1.0 (Fig. 4). The ML reconstruction provided an identical topology (SH-test,  $P = 0.928$ , Table 1) although with slightly lower values for node support (PP = 91.3, sd. = 18.7). These analyses indicated an early split of Espeletiinae between two clades composed of lineages from the Colombian Cordillera Oriental and the Venezuelan Cordillera de Mérida, respectively. The Colombian lineage (clade I) is composed of three groups of rosette growth-forms in the genera *Espeletia*, *Espeletiopsis* and the monotypic *Paramiflos glandulosus*. The remaining lineages (clades II to VIII) also formed a strongly supported clade with an early divergence of the lineage composed of tree growth-forms in the genus *Carramboa* (clade II). The tree growth-form species currently grouped in the genus *Libanothamnus* formed a monophyletic group (clade V), not related to the *Carramboa* lineages but rather nested within a larger clade composed exclusively by rosette growth-forms. The third tree growth-form lineage, the monotypic genus *Tamania* (*T. chardonii*), was not closely related to any other tree lineage but rather nested within a clade of rosette growth-forms (clade III). These results suggest therefore that taxa exhibiting the tree growth-forms in Espeletiinae are polyphyletic. Phylogenetic analyses indicated that the monocarpic rosettes in the genus *Ruilopezia* are polyphyletic as well (clades III, IV and VI). Accordingly, the SH-test showed significant support for the unconstrained BI tree compared to the constrained topology in which *Ruilopezia* is forced to be monophyletic ( $P < 0.001$ , Table 1). Furthermore, with only one exception (*Espeletia semiglobulata*, see below) the rosette growth-forms in the Venezuelan members of the genera *Espeletia* (clade VII) and *Espeletiopsis* (within clade IV) represented two monophyletic groups respectively, but polyphyletic at the scale of the subtribe as they appeared unrelated to Colombian taxa grouped under the same generic names. The SH-tests showed significant support for the unconstrained BI trees compared to the constrained topologies in which *Espeletia* and *Espeletiopsis* are forced to be monophyletic ( $P < 0.001$  in both cases, Table 1). Finally, all species in the high-elevation genus *Coespeletia* appeared grouped within a clade that also included *E. semiglobulata* (clade VIII). The SH-test showed significant support for the unconstrained tree compared to the constrained topology in which *Coespeletia* is forced to be monophyletic by placing *E. semiglobulata* within the Venezuelan *Espeletia* (clade VII) ( $p = 0.037$ , Table 1). Finally, conspecific taxa appeared related to each other within species *C. badilloi* and *E. schultzii*, supporting the strength of our approach at recovering phylogenetic relationships even between closest relatives.

In agreement with the results above, phylogenetic reconstructions allowed the rejection of all the scenarios previously considered for the evolution of Espeletiinae (Fig. 2). Thus, SH-tests performed rejected Smith and Koch's (1935) hypothesis proposing the

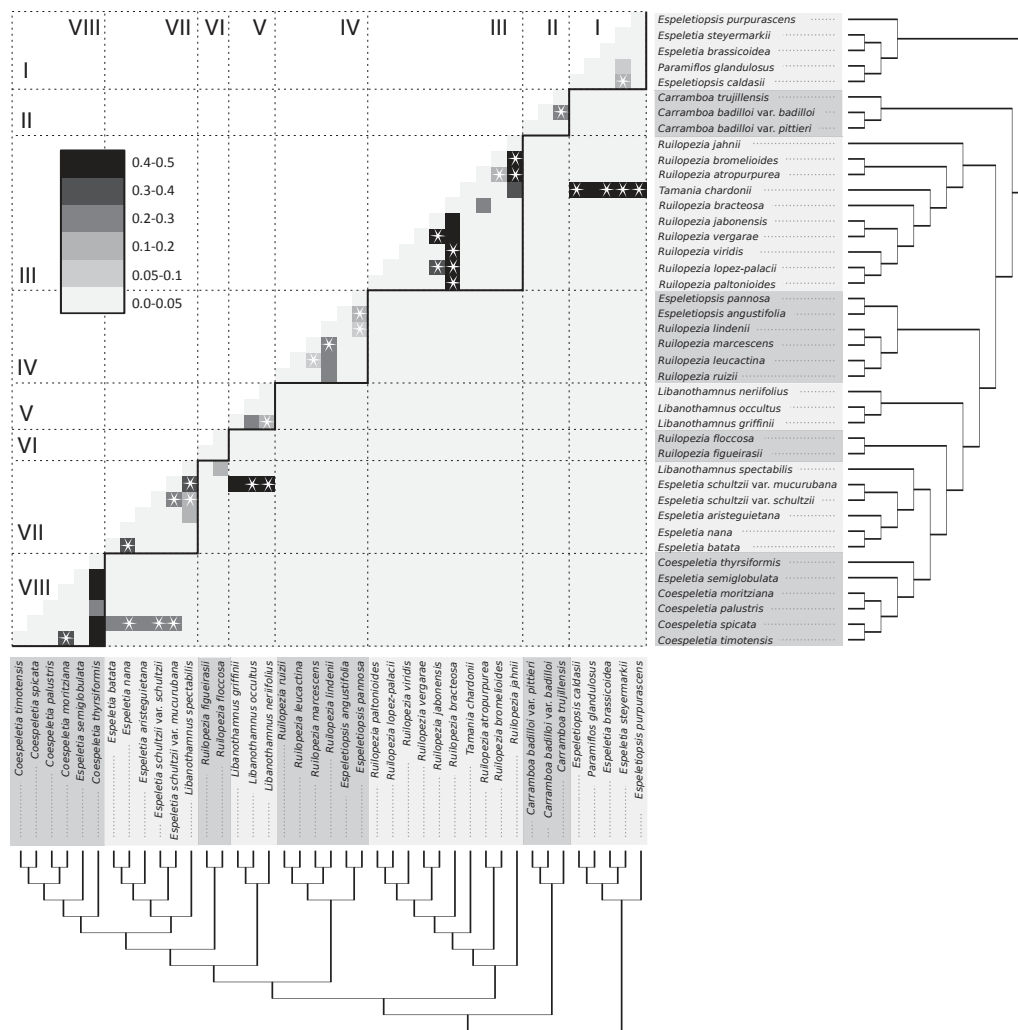
**Table 1.** Phylogenetic comparisons among alternative hypotheses of evolution in Espeletiinae.

Phylogenetic topology	rank	$\Delta_{\log}L$	SH-pvalue
BI concatenated nuclear	1	best	0.990
ML concatenated nuclear	2	21.6	0.928
SNaQ species tree	3	678.2	0.237
ASTRAL species tree	4	1145.7	0.065
Monophyly of <i>Coespeletia</i>	5	1266.9	0.037
Monophyly of <i>Ruilopezia</i>	6	3749.1	<0.001
Monophyly of <i>Espeletiopsis</i>	7	23125.4	<0.001
Monophyly of <i>Espeletia</i>	8	24102.6	<0.001
Cuatrecasas 2013	9	43111.1	<0.001
Cuatrecasas 1986	10	43655.2	<0.001
ML plastid	11	57295.8	<0.001
BI pruned nuclear	1	best	1.000
Smith and Koch 1935	2	7335.4	<0.001

monophyly of the rosette growth-form ( $P < 0.001$ , Table 1). Cuatrecasas' hypotheses about the reciprocal monophyly of alternate/opposite inflorescence structures (Cuatrecasas 1986) and terminal/lateral inflorescence positions (Cuatrecasas 2013) were also highly unlikely given our phylogenomic reconstructions ( $P < 0.001$  in both cases, Table 1). In addition, unlike plastid data, Mantel tests on nuclear data show no significant correlation between genetic and geographic distances ( $P = 0.393$ ).

*Coalescent-based multispecies tree approaches*—The tree topology obtained with *ASTRAL* (Fig. 4a) was not significantly different to concatenation according to the SH-test ( $P = 0.065$ , Table 1). It had, however, lower node support values (mean PP = 86.1, sd. = 17.8), and it exhibited five topological differences. Four of these differences involved within-clade changes without effect on the composition of clades: the position of terminal lineages *P. glandulosus* and *E. purpurascens* (clade I), *R. bracteosa* (clade III) and *E. semiglobulata* (clade VIII). The fifth difference involved an inter-clade change in regard with the position of the unbranched tree *Libanothamnus spectabilis*, which moved from a tree growth-form clade (clade V) to the base of a rosette growth-form clade (clade VII).

The phylogenetic tree recovered with *SNaQ* (Fig. 4b) did not differ significantly from the concatenated tree according to the SH-test ( $P = 0.237$ , Table 1), although it exhibited five differences, all concerning within-clade changes: the position of terminal lineages *P. glandulosus* (clade I), *R. bracteosa* (clade III), *L. griffinii* (clade V), *E. aristeguietana* (clade VII) and *E. semiglobulata* (clade VIII). It is worth noting that *SNaQ* analyses did not place *L. spectabilis* within the Venezuelan *Espeletia* (clade VII), but at the base of the *Libanothamnus* clade (clade V). However, *SNaQ* detected significant evidence of gene flow between this species and some rosettes in clade VII, suggesting that hybridization may indeed have played a role during the speciation of *L. spectabilis*. In fact, *SNaQ* allowed detection of an important number of inter-specific gene flow events, but most represented hybridizations between very closely related species within major clades and living in sympatry (Fig. 5). *SNaQ* detected as



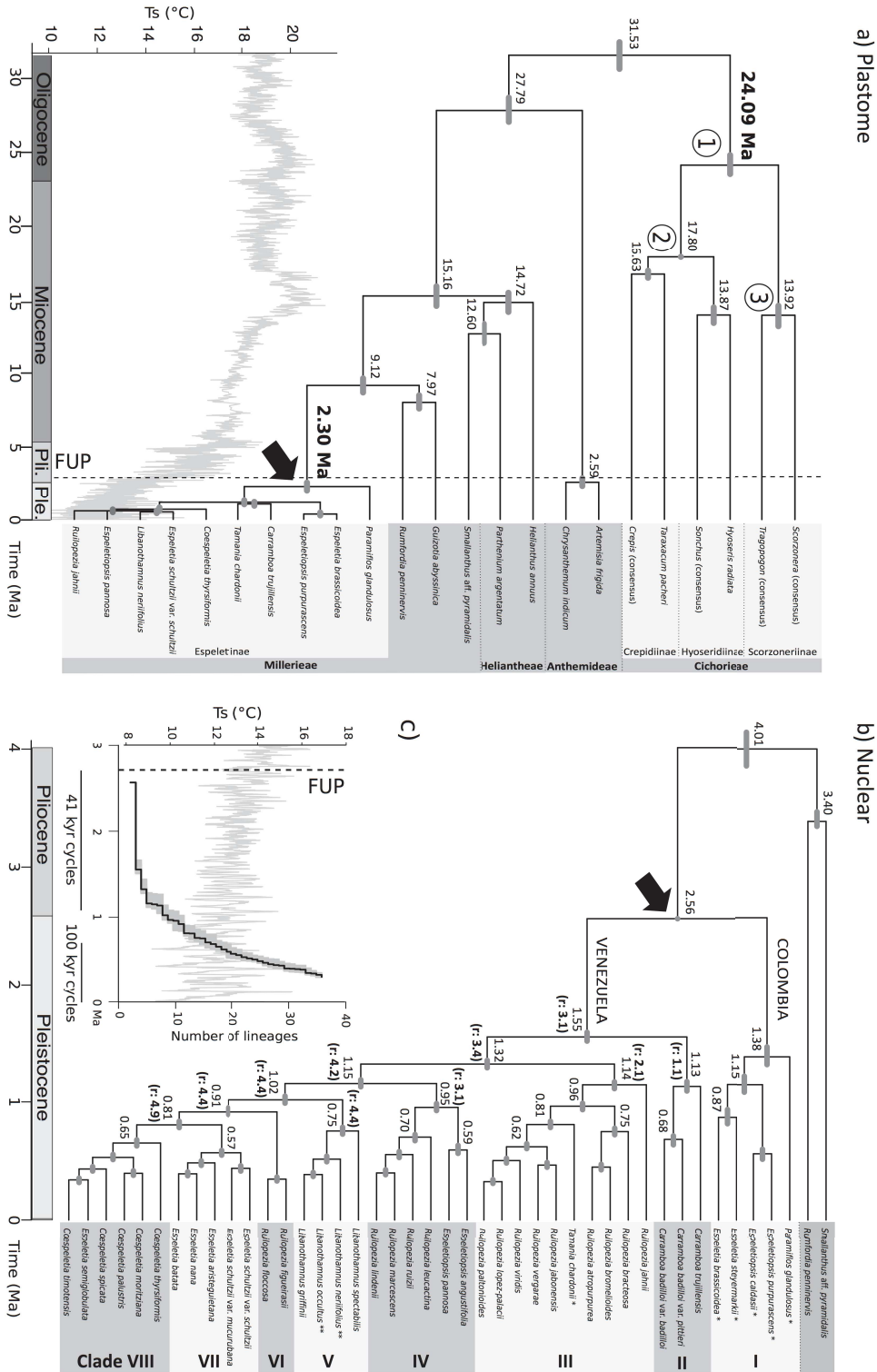
**Figure 5.** Gene flow events among Espeletiinae detected by *SNaQ*. Grey tones show the percentage of nuclear contigs with significant evidence of gene flow between the concerned taxa. Asterisks indicate sympatric/parapatric taxa. The black line separates intra and inter-clade gene flow events. Main clades defined according to the *ASTRAL* gene tree.

well some inter-clade hybridization events: 1) between *T. chardonii* and Colombian rosettes in clade I which occur in sympatry in the eastern part of the Colombian Cordillera Oriental, 2) between *E. schultzii* var. *mucurubana* and sympatric *Libanothamnus spp* (clade V), and 3) between *C. spicata* and mostly sympatric *Espeletia spp* (clade VII).

### Divergence Times of Espeletiinae

The age recovered for the crown node of Espeletiinae in plastid dataset was 2.3 Ma (2.0–2.6), suggesting that the divergence of the group started quite quickly after paramos reached their





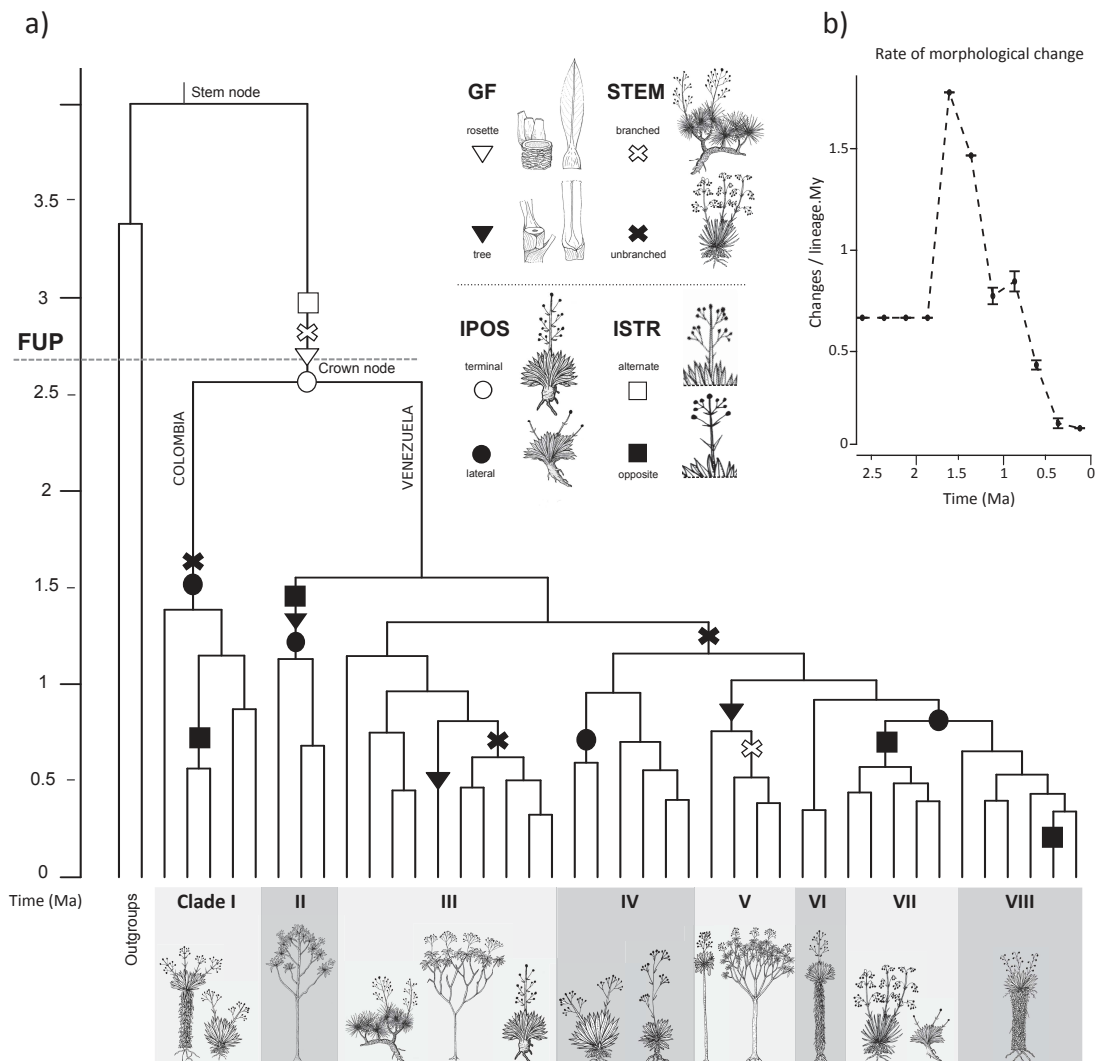
**Figure 6.** Divergence times of Espeletiinae based on (a) whole plastomes and (b) BI analysis of concatenated nuclear contigs. Age estimations for the crown node of Espeletiinae obtained in (a) were used to calibrate the analysis in (b) (black arrows). Bars on nodes represent 95% confidence intervals. Numbers in parentheses are diversification rates estimated with DDD (see text). c) Lineage through time (LTT) plot of the diversification of Venezuelan Espeletiinae for 200 time-calibrated trees (grey) and the median time-calibrated tree (black line). In (a) and (c), the dashed vertical line represents the age for the final uplift of the paramo ecosystem (FUP ; Gregory-Wodzicki, 2000) and the light grey lines show sea surface temperatures (Zachos et al., 2001). Calibration nodes are highlighted with circles.

modern elevation (~2.7 Ma, Gregory-Wodzicki, 2000) (Fig. 6a). Divergence time estimations of Espeletiinae based on the nuclear dataset provided narrow age estimates for all nodes, a few tens of thousands of years on average (mean 63.0 kyr, sd. 46.3) (Fig. 6b). The age of the crown node of Espeletiinae was estimated with this analysis at 2.6 Ma, which is slightly older than the plastid-based age estimation and even closer to the proposed time for the final uplift of paramos. In the Venezuelan radiation, the most thoroughly sampled in this study, no further diversification occurred for about a million years following the initial split of the subtribe, after which the number of lineages accumulated very rapidly and concomitantly with the increase in the amplitude of climate cycles of the second-half of the Pleistocene (6c).

The analysis of diversification in the Venezuelan Espeletiinae (clades II–VIII) showed that the best model corresponds to a pure-birth with an average speciation rate = 3.09 spp/Myr (Supplementary Table S6). However, there is a high degree of rate variation among clades in this radiation, with a minimum rate in trees in clade II (1.12), intermediate rates in monocarpic caulescent rosettes in clades III (2.06) and IV (3.09), and trees in clade V (4.37), and a maximum rate observed in polycarpic caulescent rosettes in the highest elevations of the Cordillera de Mérida, clade VII– VIII (4.86) (Fig. 6b, Supplementary Table S6).

#### *Morphological Evolution of Espeletiinae*

According to AIC values, the ER model was used to run evolutionary simulations for traits GF, STEM, and ISTR, whereas the ARD model was used for the IPOS trait (Supplementary Table S4). The reconstructions of ancestral states, displayed in Fig. 7a and Supplementary Fig. S3, indicated that the most recent common ancestor of Espeletiinae likely had a woody rosette-like growth-form with a branched stem and a terminal inflorescence with alternate structure. This analysis showed that the tree growth-form evolved from rosette ancestors three times: in the ancestors of clades II, V and within clade III (GF trait, triangles). The opposite trend, from trees to rosettes, does not seem to have occurred within Espeletiinae. The branching in the ancestral rosette-form seems to have been lost once in the Colombian Andes (clade I) and twice in the Venezuelan Andes; in the common ancestor of clades IV–VIII and within clade III (STEM trait, crosses). The reverse trend did occur once, i.e. the re-evolution of branched forms from unbranched ancestors in clade V. According to these results, with the exception of *L. spectabilis*, the members of clade V are the consequence of two major morphological changes: from unbranched rosettes to branched trees. The change from terminal to lateral inflorescences has occurred four times in the group: in Colombia, and in the Venezuelan clades II, VII + VIII, and within clade IV (IPOS trait, circles). The opposite trend, from lateral to terminal inflorescences, has not occurred. Altogether with the analysis of trait STEM, these results indicate that the monocarpic reproductive strategy of clades IV, VI, some members of clade III and *L. spectabilis* (clade V) is the result of the loss of branching in lineages that already possessed a terminal inflorescence and not the other way around. Finally, the transition from alternate to opposite inflorescences occurred at least four times, once in Colombia (a number that might increase upon analysis of more Colombian taxa), and in Venezuelan clades II, VII and *Espeletia semiglobulata* (clade VIII) (ISTR trait, squares). Taken together, these results indicate complex, repetitive, and rarely bi-directional trends in the morphological evolution of Espeletiinae. This analysis also showed that after a period of relative stasis following the start of the Espeletiinae radiation, the rate of morphological



**Figure 7.** a) Reconstruction of ancestral states of four morphological traits in Espeletiinae. The horizontal dashed line represents the age for the final uplift of the paramo ecosystem (FUP, Gregory-Wodzicki 2000. GF = growth-form; STEM = branching type; IPOS = inflorescence position; ISTR = inflorescence structure. b) Temporal changes in the rate of morphological evolution in Espeletiinae.

change per lineage and per million years increased considerably between 1.75 and 1.25 Ma, after which it decreased rather steadily towards the minimum rates observed in the last 0.5 Myr (Fig. 7b).

## 4 | DISCUSSION

We developed here a new approach of phylogenomic reconstruction that uses informative data identified from alignments of shotgun sequences without the need for a reference

genome in the group of interest and that requires neither annotation nor synteny for the identification of conserved regions across taxa. Although with some important differences, it is conceptually close to the whole-genome resequencing of multiple individuals from a population to a low depth (lcWGR) described in Fuentes-Pardo and Ruzzante (2017). It is also close to the one of Schwartz et al. (2015), who used similar methods and available NGS data to resolve the phylogenetic relationships within apes and mammals. In practical terms, our approach is analogous to Genotyping by Sequencing (GBS) and Restriction-site Associated DNA sequencing (RAD-seq) in that it provides a large number of short and anonymous nuclear markers (Cronn et al., 2012). Compared to RAD-seq or GBS approaches, ours has the advantages of requiring minimum sample preparation before sequencing, producing full organelle sequences, yielding data more evenly scattered across the nuclear genome and, possibly, generating more data. On the other hand, it requires more sequencing effort in order to keep reliable read coverage levels, particularly for nuclear markers, as well as more intensive bioinformatics methods during the post-sequencing locus selection steps. From a phylogenetic perspective, our approach also shares the rationale of studies based on GBS and RAD-seq, which is that even if the phylogenetic signal of individual markers may be small, the phylogenetic signal of the sum can indeed be very large, provided that markers are conserved across taxa, i.e. orthologous (Eaton and Ree, 2013; Rubin et al., 2012).

The issues of orthology and paralogy are central to molecular phylogenetics and deserve further comment. Whenever possible, orthology and paralogy should be assessed using a reference genome (Koonin, 2005), but several theoretical and empirical studies have shown that, in the absence of this resource, the clustering by sequence similarity provides a simple and generally effective way to identify orthologous markers (Dalquen et al., 2013; Gabaldón and Koonin, 2013; Rubin et al., 2012; but see Smith and Pease, 2017). However, orthology assessment by similarity clustering is clearly not a perfect method given that some sequences may be similar but not orthologous (e.g., gene duplication), or orthologous but not similar (e.g., sequence divergence). Nevertheless, the importance of these processes is expected to increase with divergence time. Hence, as long as the study deals with groups of recently diverged taxa such as Espeletiinae, clustering by similarity should perform particularly well (Rubin et al., 2012). Yet, in order to overcome biases in similarity scores and to maximize the number of presumably orthologous markers in our dataset we excluded all species' sequences matching more than one contig, species' contigs matching more than one metacontig and metacontigs with more than 5% ambiguous sites. Further work will however be needed in order to assess thoroughly the issues of orthology and paralogy, as well as the overall benefits/shortfalls of this approach and its performances in regard with different phylogenetic scales, genome sizes and complexities.

## Phylogeny and Gene Conflicts

*Nuclear concordance*— In spite of the conceptual differences underlying concatenation and coalescent-based approaches, with the only exception of the position of *L. spectabilis* in the *ASTRAL* tree, no evidence for important conflicts among nuclear genes due to ILS or inter-specific gene flow could be detected in relation with the phylogenetic relationships among major clades of Espeletiinae. In the case of ILS, it should be noted that the incidence of this

phenomenon decreases with phylogenetic branch lengths measured in coalescent time units (i.e.,  $T/Ne$ , where  $T$  is the number of generations and  $Ne$  is the effective population size). Thus, the small number of generations between speciation events expected in Espeletiinae due to their young age does not necessarily imply a short branch in coalescent time units provided that their effective population sizes are small (Degnan and Rosenberg, 2009). There are no available estimations of effective population sizes in Espeletiinae, but given that an important number of species are only known from very small areas or even single paramo summits (Cuatrecasas, 2013), it is possible that actual values of this parameter be indeed quite small in these plants. In regard with gene flow, our study did find significant evidence of this process among the taxa examined, but most cases represent recent events of nuclear introgression between closely related and sympatric species within terminal clades, which is not expected to generate much phylogenetic conflicts for the relationships between clades or among the older nodes of the Espeletiinae tree (Fig. 5). Similar results have been recently found in phylogenomic analyses of diversification in the genera *Heuchera* (Folk et al., 2017) and *Diplostephium* (Vargas et al., 2017), in which low levels of phylogenetic signal conflict were observed among nuclear markers despite frequent inter-specific hybridization. More generally, the ensemble of these results agree with a growing body of empirical and theoretical studies which found that concatenation and coalescent-based methods can recover the true tree topology under low to moderate levels of ILS and/or gene flow (Bayzid and Warnow, 2013; Mirarab et al., 2014; Solís-Lemus et al., 2016; Stephens et al., 2015).

Our finding of a low impact of gene flow on the phylogenetic inference of main clades of Espeletiinae should not be interpreted as implying a diminished importance of hybridization in the evolution of this group. Hybridization can clearly be a strong evolutionary force with profound impacts in the processes of speciation (Mallet, 2007; Mavárez and Linares, 2008) and adaptive radiation (Schwarzer et al., 2012; Seehausen, 2004). In Espeletiinae, inter-specific hybridization may be in part responsible for the important morphological variation observed within species known to hybridize frequently, e.g. *E. schultzii*, *E. schultzii* var. *mucurubana*, *C. spicata* (Berry et al., 1988; Cuatrecasas, 2013). Moreover, some species exhibiting mixtures of traits from different morphologic groups such as *L. spectabilis* (a rosette with closed sheaths), but also *T. chardonii* (a tree with open sheaths) and *C. thyriformis* (a *Coespeletia*-like rosette with *Espeletia*-like inflorescences), also appear to be genetically admixed in our analyses (Fig. 5). Further work will however be necessary to clarify whether this is the result of hybrid speciation or post-speciation events of introgressive hybridization.

*Cyto-nuclear discordance*— In contrast with the phylogenetic signal homogeneity among nuclear markers discussed above, this study showed that plastid and nuclear phylogenies are highly discordant in Espeletiinae (SH-test,  $P < 0.001$ ). In fact, as can be seen in the tanglegram in Supplementary Fig. S5 available on Dryad, the backbones of the two phylogenies and even a large fraction of their terminal nodes are very different. Yet, strong cyto-nuclear discordances as the observed here are relatively common among plants (Bruun-Lund et al., 2017; Liu et al., 2017; Schmickl et al., 2016), including some paramo taxa (e.g. *Diplostephium*, Vargas et al., 2017), and they are also usually attributed to ILS and reticulate evolution processes such as interspecific hybridization (Rieseberg and Soltis, 1991). Under

the light of the *ASTRAL/SNaQ* results discussed above, and given that the effective population size of plastids is 1/4 of the nuclear genome due to its haploid nature and maternal inheritance, ILS does not seem to be a likely explanation for the cyto-nuclear discordance observed in Espeletiinae. Furthermore, ILS is a random process that should not necessarily lead to the geographic footprint in the phylogenetic clustering of plastid markers that we demonstrated here. Indeed, the joint results of the phylogenetic analyses, *SNaQ* and isolation- by-distance indicate that the cyto-nuclear discordance in Espeletiinae is probably caused by differences between plastid and nuclear markers in regard with the strength of barriers to inter-specific gene flow. Thus, plastids seem to be exchanged between geographically close taxa regardless of their morphologic or nuclear phylogenetic distance, while nuclear gene flow appears mostly restricted to sympatric and closely related species within major phylogenetic groups (Figs. 3 and 5). Such heterogeneity in the permeability of barriers is not unexpected given the important differences in size, composition and inheritance of plastid and nuclear genomes (Rieseberg and Soltis, 1991; Tsitrone et al., 2003). For instance, relatively common biological phenomena such as sex-biased hybridization, unidirectional backcrossing and cytoplasmic male sterility can allow the complete transfer of uniparentally-inherited markers across species with little or no traces of concomitant nuclear exchange (Folk et al., 2017; Tsitrone et al., 2003). Since all these phenomena depend on inter-specific hybridization, they generate patterns of cyto-nuclear discordance that are usually structured geographically, e.g. Acosta and Premoli (2010). Although no analysis has been specifically designed to study these processes in Espeletiinae, the data in Berry et al. (1988) indicates that some inter-specific crosses can be completely sterile in one direction and partially fertile in the other, and we have observed exclusively unidirectional backcrossing in a *E. schultzei/C. moritziana* hybrid zone (Mavárez et al. unpublished data).

## A Phylogenomic Analysis of the Radiation of Espeletiinae

The dataset of nuclear contigs generated in this study, consisting of 1877 metacontigs with >31,000 informative sites, provided enough polymorphic genomic data to recover for the first time a fully-resolved phylogeny of major morphological groups of Espeletiinae (Fig. 4). Some of these relationships agree with previous views of the evolution of the group, while others appear to be rather in conflict. For example, in agreement with the recent work of Diazgranados and Barber (2017), Espeletiinae experienced two geographically delimited and mostly disconnected radiations in Venezuela and Colombia, respectively, instead of an initial radiation in Venezuela followed by multiple trans-cordilleran dispersal events towards Colombia, as proposed by Smith and Koch (1935) and Cuatrecasas (1986, 2013). More specifically, our study shows that dispersal across the Táchira depression has indeed occurred quite rarely in Espeletiinae, always from Venezuela to Colombia, each time involving lineages with the tree growth- forms and occurring at relatively lower elevations. These events may have concerned *L. neriifolius* (min. 1800 m), *L. occultus* (min. 2700 m) and *T. chardonii* (min. 2200 m), which can sometimes thrive in the warmer and wetter conditions of the upper Andean forest habitats (Supplementary Fig. S4 available on Dryad) (Cuatrecasas, 2013). However, the actual number of dispersal events could rise once all the phylogenetic relationships within Venezuelan clade V are fully resolved, because some of its members are

distributed in the Colombian Cordillera Oriental (*L. tamanus*), in its northward extension the Sierra de Perija (*L. divisoriensis*) or in the Sierra Nevada de Santa Marta (*L. occultus glossophyllus*).

Our study also confirms a result implicit in Diazgranados and Barber (2017), which is that there have been no trans-cordilleran dispersal events in the Cuatrecasas' genera *Espeletia* or *Espeletiopsis*. These taxa represent independent evolutions in each country of the polycarpic caulescent rosette with opposite or alternate inflorescences, respectively. This result explains one of the most puzzling mysteries unsolved by previous scenarios of evolution in Espeletiinae (Cuatrecasas, 1986, 2013): how did the dispersal-limited paramo-endemic *Espeletia* and *Espeletiopsis* manage to cross the >30 km of land below 2000 m in the Táchira depression between the southern border of the Venezuelan Cordillera de Mérida and the eastern border of the Colombian Cordillera Oriental which, even during glacial maxima, never sheltered a paramo habitat? The answer appears now simple: they never did.

The two radiations of Espeletiinae in Colombia and Venezuela differ notably in terms of morphological diversity—much larger in Venezuela, a pattern that Diazgranados and Barber (2017) suggested to result from a supposed older age of the Venezuelan Espeletiinae lineages. Although our species sampling remains partially incomplete, our results suggest otherwise, as the onsets of the radiation in each country appear quite simultaneous (Fig. 6b), with the possibility of the Colombian radiation becoming even older following the analysis of more Espeletiinae taxa from this country. This agrees with thermochronological studies of the exhumation of the Northern Andes that show similar ages and uplift dynamics for the Venezuelan Cordillera de Mérida and the Colombian Cordillera Oriental (Bermúdez et al., 2010; Hoorn et al., 2010).

*Rapid Radiation During the Pleistocene*— We estimated the age for the crown node of Espeletiinae at 2.3 (2.0–2.6) Ma (Fig. 6a), which is both much more recent and precise than the previous estimate of 4.0 (2.4–5.9) Ma provided by Madriñan et al. (2013). According to this result, the diversification dynamics of Espeletiinae appears to have been driven most likely by Pleistocene climatic oscillations than by Neogene tectonics. Further support for this idea comes from our finding that, in the better-sampled Venezuelan radiation, no diversification occurred for about the first million years of its evolution, during a period in which glacial dynamics were dominated by the low-amplitude cycles of the Early Pleistocene. Later, the rate of diversification increased dramatically in this radiation, when climate variability changed to the large cycles characteristic of the Late Pleistocene (Hönisch et al., 2009). On the other hand, even though the Andes had reached current elevations by about 2.7 Ma, cold-adapted high-elevation paramo lineages (clades VII and VIII) appeared in the Venezuelan Cordillera de Mérida only by 0.65 Ma, when glacial cycles reached the lowest temperatures. Moreover, although still preliminary, the estimated age for the crown node of the Colombian Espeletiinae (1.38 Ma), composed entirely of caulescent rosettes, seems quite synchronous with that of the most recent common ancestor of all Venezuelan caulescent rosettes (1.32 Ma, clades III-VIII, Fig. 6b). This suggests that similar time-coordinated processes could have been at play at the onset of the radiations in the two cordilleras. Further studies specifically designed to understand the drivers of diversification in species-rich paramo taxa such as *Lupinus* (Nevado et al., 2016), *Hypericum* (Nürk et al., 2013),

*Diplostegium* (Vargas et al., 2017) or the Colombian Espeletiinae will however be necessary before a link with Pleistocene climatic instability or Miocene/Pliocene tectonics can be firmly established.

Models of diversification rates on phylogenetic trees show that with an average rate  $r = 3.09$  spp/Myr, the Venezuelan radiation of Espeletiinae ranks among the fastest evolving plant groups on Earth, just behind *Dianthus* ( $r = 2.20\text{--}7.60$ , Valente et al., 2010) and Andean *Lupinus* ( $r = 1.56\text{--}5.21$ , Hughes and Atchison, 2015), and perhaps other Andean taxa that deserve further studies (Lagomarsino et al., 2016; Madriñán et al., 2013). However, there is ample rate heterogeneity among clades in the Venezuelan radiation, suggesting that an average rate does not necessarily reflect changes in the diversification dynamics neither through space nor across lineages. For example, the diversification rate in the unbranched polycarpic caulescent rosettes in the highest elevations of the Cordillera de Mérida (clades VII–VIII) is faster than of trees in clades II or V, and than of monocarpic caulescent rosettes in clades III and IV. This suggests an interesting link between growth-form, elevation and diversification rate in Espeletiinae, a pattern also observed in the Andean *Lupinus* and that has been attributed to a possible ecological release resulting from reduced inter-specific competition in recently formed habitats in the high Andes (Hughes and Atchison 2015; Nevado et al., 2016).

## Morphological Evolution

*The ancestor of Espeletiinae*—Both Smith and Koch (1935) and Cuatrecasas (1986) suggested that the ancestor of Espeletiinae was an evergreen perennial shrub/tree exhibiting the xeromorphic foliage with spiral phyllotaxy characteristic of the subtribe, likely similar to the trees *Libanothamnus*. Our study suggests, however, an alternative scenario, with a branched rosette with imbricated open/flat leaf sheaths as the crown ancestor of the subtribe (Fig. 7a). This result is coherent with our finding of a Pleistocene onset for the radiation of Espeletiinae, because an older age would imply that the evolution of this rosette ancestor pre-dated the paramo habitat, in contradiction with the known distributions and ecological restrictions of this growth-form in the subtribe and other Northern Andes plants (Cuatrecasas, 2013). The synchronicity between the evolution of Espeletiinae and the rising of paramo conditions, together with the fact that the diversification of the subtribe started from a rosette, provide additional support to the widely accepted notion that this growth-form, and particularly the caulescent type, represents a key morphological trait or an ensemble of traits that triggered the adaptive radiation of Espeletiinae (Monasterio and Sarmiento, 1991). This is also likely the case of other caulescent rosette taxa across the World, e.g. *Dendrosenecio* and *Lobelia* in the East African highlands, *Argyroxiphium*, *Wilkesia* and the lobeliads in the Hawaiian mountains (Givnish, 2010; Givnish et al., 2009).

*Trees*—Our study revealed that there have been three independent evolutions of the tree growth-form in the subtribe; two from branched rosette ancestors in clade II (*Carramboa*) and within clade III (*Tamania*), and one from unbranched rosettes in clade V (*Libanothamnus*) (Fig. 7a). In opposition to the dominant ideas about the evolution of this group, the reverse trend from trees to rosettes does not seem to have occurred. Furthermore, given that rosettes



are distributed at significantly higher elevations and with a marked preference for the colder/drier conditions of the paramo habitat, the evolution of the tree growth-form in the subtribe appears to be associated with a downward colonization of the warmer/wetter conditions in the forest-paramo ecotone and upper Andean forest, in which tree forms are dominant (Supplementary Fig. S4 available on Dryad). This trend contradicts another common sentiment regarding the evolution of Espeletiinae and other paramo plants, which claims that habitat colonizations have occurred from lowlands to highlands (Cuatrecasas, 1986). A somewhat similar trend has been proposed recently for the evolution of the Neotropical genus *Diplostephium*, in which lower- elevation tree species appear to have evolved from higher elevation shrubby ancestors (Vargas and Madriñán, 2012).

*Rosettes*— This study also highlighted three independent evolutions of the unbranched caulescent rosettes in Espeletiinae; one in the Colombian radiation (clade I) and two in the Venezuelan; within clade III and in the ancestor of clades IV–VIII (Fig. 7a). This conclusion is unlikely to change upon analysis of more Colombian rosette taxa, because all species in this country belong to the unbranched caulescent type and appear to be monophyletic (Diazgranados and Barber, 2017; this study). The numerous parallelisms between the giant caulescent rosettes in Espeletiinae of the high Andes of Colombia and Venezuela in regard to, for instance, plant size and form, foliar pubescence and insulation of the stem with dead leaves, provide strong support to the idea of a convergent evolution of this growth-form in response to the cold, dry, and irradiated conditions that prevail at high elevations in both countries. These traits confer protection against freezing of the apical meristem, reduced water loss through transpiration, increased volume of water stored in the large pith of an insulated stem, etc. (Monasterio and Sarmiento, 1991; Smith, 1981). A similar idea has been recently put forward to explain correlations between geographic origin and metabolic fingerprints in Colombian and Venezuelan *Espeletia* caulescent rosettes by Padilla-González et al. (2017). Analogous arguments were considered by Loeuille et al. (2015) to explain the independent evolution of three lineages of caulescent rosettes in the subtribe Lychnophorinae (Asteraceae) endemic to the herbaceous/shrubby vegetation known as *campos rupestres* in the Brazilian Cerrado, which the authors describe as a stressful habitat with paramo-like conditions such as intense solar radiation, important fluctuations in daily temperature and seasonal water scarcity, in addition to frequent fires.

We note however that some of the tallest caulescent rosettes in Espeletiinae, such as *Ruilopezia marcescens* (clade IV), *R. paltonioides* (clade III) and *Espeletiopsis purpurascens* (clade I) (Fig. 1b), do not live in the proper paramo but at lower elevations in the forest-paramo ecotone. The gigantism in these species with naked stems, reduced foliar pubescence and comparatively smaller piths cannot be considered as a protection against freezing and/or droughts, but perhaps as the response of the shade-intolerant Espeletiinae (Baruch and Smith, 1979) selection favoring rapid vertical growth under the conditions of strong competition for light typical of forest habitats.

*Monocarpy*— In theory, monocarpy in Espeletiinae could be achieved through two paths: by a change from lateral to terminal inflorescences in unbranched species or by the loss of branching in species with terminal inflorescences. Our study only found examples of the

latter, which produced at least four independent evolutions of the monocarpny in clades III, IV, V (*L. spectabilis*) and VI (Fig. 7a). A detailed analysis of the causes of changes of reproductive strategy in Espeletiinae lies beyond the scope of this study, but the pattern of repeated evolution suggests that some form of selection could have been at play. Indeed, monocarpic species in the subtribe exhibit several traits expected to be favored by natural selection as part of a “monocarpic syndrome”, such as large inflorescences (Cuatrecasas, 2013), big seeds (Smith, 1981) and synchronized reproduction (Mavarez pers. obs). On the other hand, monocarpny in Espeletiinae is common among unbranched rosettes but rare or inexistent among branched rosettes and trees. This association between branching pattern and reproductive strategy holds for all known long-lived monocarpic rosettes, which has been suggested to represent a predisposition for monocarpny in plants with this growth-form (Jabaily and Sytsma, 2013; Young and Augspurger, 1991). Further studies will therefore be necessary in order to depict whether monocarpny in Espeletiinae evolves as a direct consequence of selection or as a by-product of other processes, e.g. loss of branching.

### A New Systematic Proposal for Espeletiinae

Our study shows that, as currently circumscribed, the three largest genera in the Cuatrecasas' classification system of Espeletiinae, *Espeletia*, *Ruilopezia* and *Espeletiopsis*, representing about 84% of the subtribe, are clearly polyphyletic. Furthermore, the genus *Coespeletia* is paraphyletic in regard to *E. semiglobulata*, while the two monotypic genera *Tamania* and *Paramiflos* are nested within one group of *Ruilopezia* and the Colombian *Espeletiopsis*, respectively. Only the genera *Carramboia* and *Libanothamnus* have been resolved as monophyletic. As a consequence, the classification system with eight-genera of Espeletiinae (Cuatrecasas, 1976, 1986, 2013) appears largely artificial and must be modified. One such modification, following a “splitter approach”, could be: 1) preserve the monophyletic genera *Carramboia* and *Libanothamnus*; 2) modify the definition of *Coespeletia* so that it can accept *Espeletia semiglobulata*; 3) create a new genus for the Venezuelan *Espeletiopsis*; 4) create a new genus for the Venezuelan *Espeletia*; 5) decide whether *Tamania* is preserved or merged into *Ruilopezia*; and 6) split *Ruilopezia* in three or four genera depending on the decision taken in (5). Additional changes might be necessary for the Colombian radiation, particularly in regard to the status of *Paramiflos*. The problem with this modification is that it would require a considerable amount of time and effort in order to identify the synapomorphies that define the new taxa, as well as the creation of many new name combinations.

We propose a more workable and stable solution, adopting the view that only the genus *Espeletia* Mutis ex Bonpl. should be recognized in the subtribe, with the other seven genera considered as (heterotypic) synonyms. Given that the majority of the species were originally described under the genus *Espeletia* (~130 spp.), the change proposed here implies the resurrection of species names discarded by Cuatrecasas (1976, 2013) and about a dozen new combinations. Moreover, several influential workers on the ecology and physiology of the group did not accept Cuatrecasas' nomenclature and a large number of important studies have been published using only *Espeletia* (Berry and Calvo, 1989; Goldstein et al., 1984, 1985; Monasterio and Sarmiento, 1991; Smith, 1981). Following Rauscher (2002), to

acknowledge for the large specific and morphological diversity in the genus, the ensemble of the species can also be collectively referred to as the “*Espeletia* complex”.

## 5 | SUPPLEMENTARY INFORMATIONS

**Supp. Table 1.** Sampling data and results of the *de novo* assembly. TNR: total number of DNA reads. CR: plastid reads. PNR: potential nuclear reads. Contigs: potential nuclear contigs. N50: average N50 contig length. Cov: average contig coverage. FNC: final number of nuclear contigs after removal of remaining plastid and mitochondrial contigs.

Sampled taxa	Longitude	Latitude	Locality	TNR	CR	PNR	Contigs	N50	Cov	FNC
<i>Carramboa badilloi</i> var. <i>badilloi</i>	-71.3033333	8.3358611	San José (VEN)	4984893	40102	4184916	4473	368	8.860	4459
<i>Carramboa badilloi</i> var. <i>pittieri</i>	-71.5636944	8.2833888	El Molino (VEN)	16054991	210883	13287364	7563	370	13.00	7546
<i>Carramboa trujillensis</i>	-70.2278634	9.2099248	Guaramacal (VEN)	8127220	77174	6854071	4737	402	13.89	4720
<i>Coespeletia moritziana</i>	-70.8308036	8.8361219	Piedras Blancas (VEN)	8136454	140607	6752907	7291	359	9.036	7275
<i>Coespeletia palustris</i>	-70.7996860	8.7724805	Piedras Blancas (VEN)	13521468	299939	11113903	9538	374	12.92	9522
<i>Coespeletia spicata</i>	-70.8308036	8.8361219	Piedras Blancas (VEN)	35903204	228916	30379015	15902	378	15.35	15884
<i>Coespeletia thysiformis</i>	-71.9038361	8.1331580	Batallón (VEN)	8117807	53562	6763002	6311	400	12.10	6292
<i>Coespeletia timotensis</i>	-70.8308036	8.8361219	Piedras Blancas (VEN)	22914069	321950	18677113	9302	399	15.88	9286
<i>Espeletia aristeguietana</i>	-70.3367777	9.3248055	La Cristalina (VEN)	10939974	69907	9227171	6995	395	12.90	6978
<i>Espeletia batata</i>	-70.8757138	8.8829417	Piedras Blancas (VEN)	7925776	116657	6614794	4173	395	10.81	4161
<i>Espeletia brassicoidea</i>	-72.4164166	7.4121388	Tamá (VEN-COL)	7201499	83340	5934785	8600	344	9.000	8582
<i>Espeletia nana</i>	-70.5801666	9.0325113	Tuñame (VEN)	8224828	98012	6993581	7318	356	8.969	7300
<i>Espeletia schultzei</i> var. <i>mucurubana</i>	-70.8812610	9.0200056	Guaramacal (VEN)	10477041	79055	8862128	8696	349	8.644	8680
<i>Espeletia schultzei</i> var. <i>schultzei</i>	-70.8757138	8.8829417	Piedras Blancas (VEN)	11155891	80609	9344535	6579	384	11.27	6565
<i>Espeletia semiglobulata</i>	-70.8308036	8.8361219	Piedras Blancas (VEN)	6202382	102116	5140647	5847	376	8.754	5831
<i>Espeletia steyermarkii</i>	-72.4026666	7.4543888	Tamá (VEN-COL)	11429995	94201	9545357	7772	386	12.04	7753
<i>Espeletiopsis angustifolia</i>	-71.3033333	8.3358611	San José (VEN)	16507894	73368	13758950	9306	393	13.04	9289
<i>Espeletiopsis caldasii</i>	-73.0017321	7.7655552	Santurbán (COL)	15045803	34175	11871434	10947	429	13.20	10931
<i>Espeletiopsis pannosa</i>	-71.0893911	8.5647806	Los Nevados (VEN)	14077121	25557	12165567	6851	436	15.30	6838
<i>Espeletiopsis purpurascens</i>	-72.4051388	7.4559166	Tamá (VEN-COL)	4237288	19349	3557917	5898	347	8.596	5884
<i>Libanothamnus griffinii</i>	-70.1926775	9.2369722	Guaramacal (VEN)	4231602	38070	3551831	4383	367	8.040	4369
<i>Libanothamnus neriiifolius</i>	-70.6371725	9.0670111	Tuñame (VEN)	10670853	36801	8990353	8239	345	9.480	8223
<i>Libanothamnus occultus</i>	-70.7996860	8.7724805	Piedras Blancas (VEN)	5535336	15085	4674931	6409	340	8.429	6393
<i>Libanothamnus spectabilis</i>	-71.3033333	8.3358611	San José (VEN)	7060812	86032	6000035	7204	343	8.420	7186
<i>Paramiflojos glandulosus</i>	-73.3771527	5.7183028	Arcabuco (COL)	18501072	244925	14858624	10115	410	12.03	10101
<i>Ruilopezia atropurpurea</i>	-71.3039212	8.3351211	San José (VEN)	17430030	26616	14955860	7728	394	13.66	7710
<i>Ruilopezia bracteosa</i>	-70.6093611	8.9961666	Tuñame (VEN)	8684341	53237	7457490	5428	406	12.57	5412
<i>Ruilopezia bromelioides</i>	-71.5636944	8.2833888	El Molino (VEN)	18849924	93938	15825965	10529	370	13.42	10512
<i>Ruilopezia figueirasii</i>	-70.8812611	9.0200056	Piedras Blancas (VEN)	20081214	111378	17058280	7794	455	22.92	7777
<i>Ruilopezia floccosa</i>	-70.8600690	8.8417778	Piedras Blancas (VEN)	6297993	118824	5269967	4581	450	12.04	4565
<i>Ruilopezia jabonensis</i>	-70.1191666	9.5765333	Guaramacal (VEN)	12855749	64823	11011693	6717	444	13.57	6699
<i>Ruilopezia jahnii</i>	-71.9208777	8.1087389	Batallón (VEN)	18753847	82201	16156028	10214	388	14.60	10197
<i>Ruilopezia leucactina</i>	-71.9135277	8.1123833	Batallón (VEN)	11290989	61679	9335128	7748	391	12.28	7728
<i>Ruilopezia lindenii</i>	-71.2999444	8.3022512	San José (VEN)	6845497	29844	5873831	5038	409	11.03	5027
<i>Ruilopezia lopez-palacii</i>	-70.1857253	9.2359417	Guaramacal (VEN)	7247985	50186	6234203	4063	456	12.37	4048
<i>Ruilopezia marcescens</i>	-71.2591133	8.3025556	San José (VEN)	4966632	28205	4128547	3301	437	10.72	3284
<i>Ruilopezia paltonioides</i>	-70.1857253	9.2359417	Guaramacal (VEN)	20985038	57386	17634499	11106	375	15.55	11088
<i>Ruilopezia ruizii</i>	-71.5776944	8.2885833	El Molino (VEN)	16367539	36577	14287026	8028	409	13.13	8012
<i>Ruilopezia vergarae</i>	-70.1279166	9.5786111	Cendé (VEN)	14247110	36783	12129235	11116	344	11.00	11099
<i>Ruilopezia viridis</i>	-70.1866272	9.2302028	Guaramacal (VEN)	25079044	151242	21092244	9919	383	17.55	9902
<i>Tamania chardonii</i>	-72.4026666	7.4543888	Tamá (VEN-COL)	19319750	70924	16523043	9545	379	15.06	9528
<i>Rumfordia penninervis</i>	NA	NA	Mexico	17321123	108912	14090591	6848	414	14.62	6838
<i>Smallanthus</i> aff. <i>pyramidalis</i>	NA	NA	Piedras Blancas (VEN)	15267267	62222	12292283	19504	386	9.905	19494

**Supp. Table 2.** Results of the selection of models of molecular evolution of plastid sequences. Loglik: Log-likelihood of the model. AIC: Akaike information criterion. BIC: Bayesian information criterion. df: degrees of freedom.

Model	df	Loglik	AIC	BIC
JC	83	-246937.06	494040.11	494864.10
JC+I	84	-245686.08	491540.16	492374.07
JC+G	84	-246147.07	492462.14	493296.06
JC+G+I	85	-245620.50	491410.99	492254.84
F81	86	-242303.54	484779.08	485632.85
F81+I	87	-241075.18	482324.35	483188.05
F81+G	87	-241517.67	483209.35	484073.05
F81+G+I	88	-241008.39	482192.78	483066.41
K80	84	-246931.10	494030.21	494864.12
K80+I	85	-245682.80	491535.60	492379.44
K80+G	85	-246141.64	492453.28	493297.13
K80+G+I	86	-245617.12	491406.23	492260.00
HKY	87	-242289.05	484752.10	485615.80
HKY+I	88	-241066.86	482309.72	483183.35
HKY+G	88	-241504.41	483184.81	484058.44
HKY+G+I	89	-240999.79	482177.57	483061.13
SYM	88	-246232.75	492641.49	493515.12
SYM+I	89	-244966.23	490110.46	490994.01
SYM+G	89	-245439.49	491056.99	491940.54
SYM+G+I	90	-244898.69	489977.38	490870.86
GTR	91	-241461.32	483104.64	484008.05
GTR+I	92	-240245.22	480674.44	481587.78
GTR+G	92	-240675.42	481534.85	482448.19
GTR+G+I	93	-240173.64	480533.27	481456.54
TrN	88	-242288.67	484753.33	485626.96
TrN+I	89	-241066.54	482311.09	483194.64
TrN+G	89	-241504.03	483186.07	484069.62
TrN+G+I	90	-240999.48	482178.96	483072.44
K81	85	-246239.03	492648.06	493491.91
K81+I	86	-244974.46	490120.92	490974.69
K81+G	86	-245445.92	491063.84	491917.61
K81+G+I	87	-244906.80	489987.61	490851.31
TVM	90	-241461.70	483103.41	483996.89
TVM+I	91	-240245.48	480672.96	481576.38
TVM+G	91	-240675.77	481533.54	482436.95
TVM+G+I	92	-240173.87	480531.73	481445.07

**Supp. Table 3.** Genbank accessions for plastid sequences of taxa used for the analysis of divergence times of Espeletiinae

Species	Tribe	Subtribe	Accession Number
<i>Carramboa trujillensis</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Coespeletia thyrsoformis</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Coespeletia timotensis</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Espeletia brassicoidea</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Espeletia schultzei</i> var. <i>schultzei</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Espeletiopsis pannosa</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Espeletiopsis purpurascens</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Libanothamnus nerifolius</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Paramiflos glandulosus</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Ruilopezia jahni</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Tamania chardonii</i>	Milleriaceae	Espeletiinae	XXXXXX
<i>Rumfordia penninervis</i>	Milleriaceae	Milleriinae	XXXXXX
<i>Smallanthus</i> aff. <i>pyramidalis</i>	Milleriaceae	Milleriinae	XXXXXX
<i>Artemisia frigida</i>	Anthemideae	Artemisiinae	NC_020607
<i>Chrysanthemum indicum</i>	Anthemideae	Artemisiinae	JN867592
<i>Guizotia abyssinica</i>	Milleriaceae	Milleriinae	EU549769
<i>Helianthus annuus</i>	Heliantheae	Helianthiinae	NC_007977
<i>Parthenium argentatum</i>	Heliantheae	Ambrosiinae	NC_013553
<i>Crepis biennis</i>	Cichorieae	Crepidiinae	XXXXXX
<i>Crepis bursifolia</i>	Cichorieae	Crepidiinae	XXXXXX
<i>Crepis capillaris</i>	Cichorieae	Crepidiinae	XXXXXX
<i>Hyoseris radiata</i>	Cichorieae	Hyoseridiinae	XXXXXX
<i>Sonchus arvensis arvensis</i>	Cichorieae	Hyoseridiinae	XXXXXX
<i>Sonchus oleraceus</i>	Cichorieae	Hyoseridiinae	XXXXXX
<i>Sonchus palustris</i>	Cichorieae	Hyoseridiinae	XXXXXX
<i>Taraxacum pacheri</i>	Cichorieae	Crepidiinae	XXXXXX
<i>Scorzonera hirsuta</i>	Cichorieae	Scorzoneriinae	XXXXXX
<i>Scorzonera hispanica</i>	Cichorieae	Scorzoneriinae	XXXXXX
<i>Scorzonera villosa</i>	Cichorieae	Scorzoneriinae	XXXXXX
<i>Tragopogon dubius</i>	Cichorieae	Scorzoneriinae	XXXXXX
<i>Tragopogon porrifolius australis</i>	Cichorieae	Scorzoneriinae	XXXXXX
<i>Tragopogon pratensis orientalis</i>	Cichorieae	Scorzoneriinae	XXXXXX

**Supp. Table 4.** Model selection for the evolution of four morphologic traits according to the Akaike information criterion (AIC). Df: degrees of freedom. ARD: different rates model. ER: equal rate model. Traits used: growth-form (GF, rosette vs. tree), stem type (STEM, branched vs. unbranched), inflorescence position (IPOS, terminal vs. lateral) and inflorescence structure (ISTR, alternate vs. opposite).

Models		Df	lnLik	AIC
GF	ARD	2	-14.545	33.090
	<b>ER</b>	<b>1</b>	<b>-14.546</b>	<b>31.091</b>
STEM	ARD	2	-16.810	37.621
	<b>ER</b>	<b>1</b>	<b>-17.292</b>	<b>36.584</b>
IPOS	<b>ARD</b>	<b>2</b>	<b>-13.080</b>	<b>30.160</b>
	ER	1	-14.486	30.972
ISTR	ARD	2	-15.130	34.259
	<b>ER</b>	<b>1</b>	<b>-15.971</b>	<b>33.942</b>

**Supp. Table 5.** Datasets used for phylogenetic analyses. Plastome and Nuclear refer to whole plastomes and sets of nuclear metacontigs (see text for details). PP: Posterior Probabilities. BPP: Bayesian Posterior Probabilities. sd: standard-deviation.

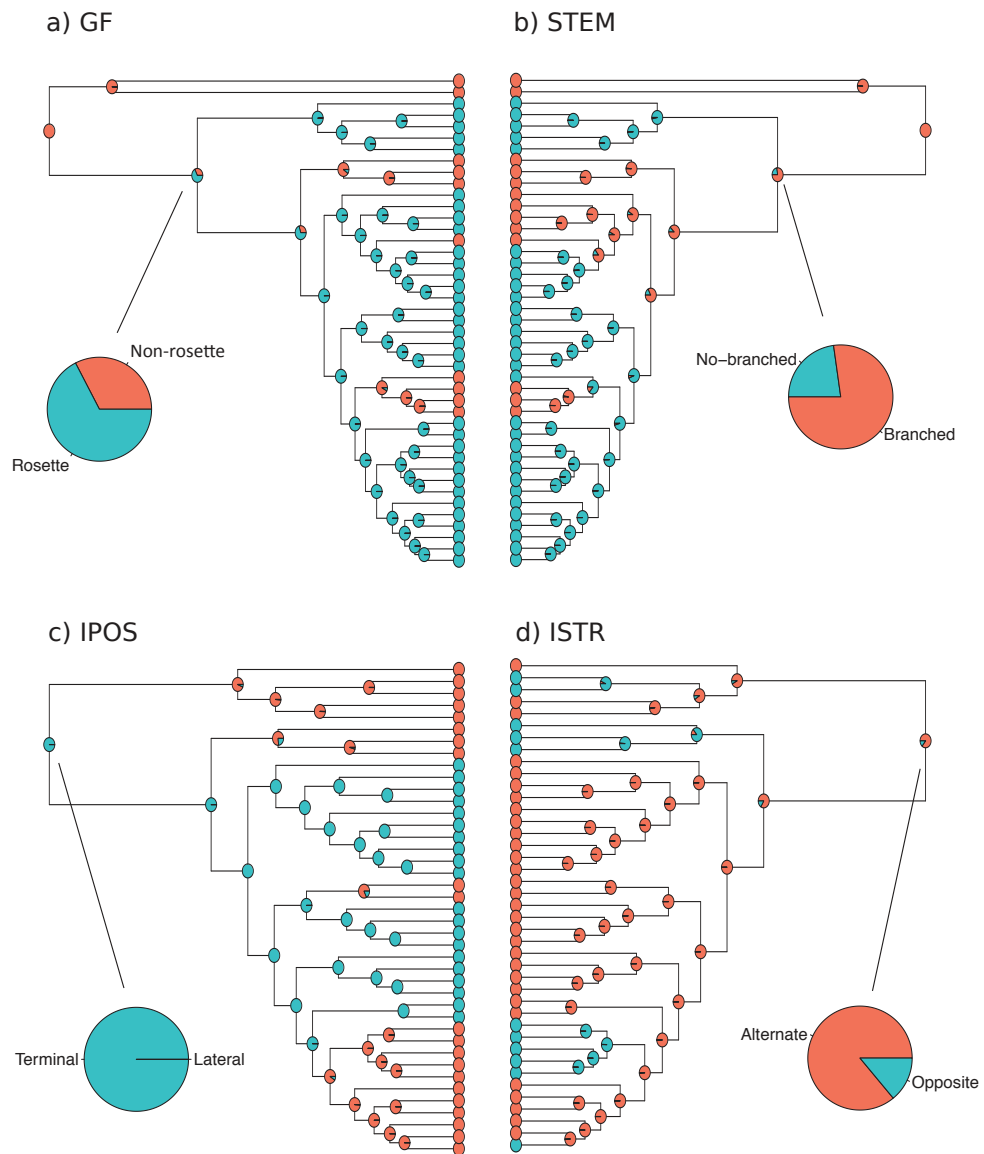
Phylogenetic Datasets	Plastome	Nuclear
Espeletiinae (outgroups)	41 (2)	41 (2)
Alignments	1	1877
Total length of alignment (bp)	151,388	933,394
Ambiguous characters and missing data (%)	0.05	36.34
Polymorphic sites (%)	2.46	7.98
Informative sites (%)	0.60	3.39
Phylogenetic ML inference		
PPmean (sd)	89.12 (19.90)	91.27 (18.70)
PP < 50 (%)	3 (7.50)	2 (5.00)
PP < 90 (%)	9 (22.50)	7 (17.50)
Phylogenetic Bayesian inference		
BPPmean (sd)	0.98 (0.06)	0.99 (0.001)
BPP < 0.50 (%)	1 (2.50)	0 (0.00)
BPP < 0.90 (%)	3 (7.50)	0 (0.00)

**Supp. Table 6.** Analyses of diversification models of Venezuelan Espeletiinae performed with DDD. Node: main clades (as in Fig. 5b). Lambda, mu and K represent rates of speciation, extinction and carrying capacity und, respectively. Loglik, df and AIC give log-likelihood, degrees of freedom and Akaike information criterion values. Models considered: pure-birth (PB), birth-death (BD), density-dependence for speciation (DD) and density dependence for extinction (DDE).

Node	Model	lambda	mu	K	loglik	df	AIC
II	PB	1.121	1e-04	5000	-2.735	1	7.470
	BD	1.121	4e-04	5000	-2.735	2	9.470
	DD	5.033	1e-04	4.975	-1.384	2	6.768
	DDE	5.244	0.021	4.974	-1.388	3	8.776
III	PB	2.060	1e-04	5000	-2.898	1	7.795
	BD	2.052	6e-04	5000	-2.898	2	9.795
	DD	3.560	1e-04	15.85	-2.540	2	9.079
	DDE	9.764	2.225	10.47	-2.067	3	10.13
IV	PB	3.085	1e-04	5000	-2.798	1	7.595
	BD	3.078	0.011	5000	-2.798	2	9.595
	DD	7.320	1e-04	13.82	-1.998	2	7.995
	DDE	8.816	1.458	13.30	-1.970	3	9.939
V	PB	4.370	1e-04	5000	-2.760	1	7.520
	BD	9.442	8.699	5000	-2.552	2	9.104
	DD	10.86	1e-04	12.24	-2.191	2	8.383
VII-VIII	PB	4.863	1e-04	5000	1.942	1	-1.884
	BD	4.854	0.015	5000	1.941	2	0.116
	DD	7.484	1e-04	28.51	2.310	2	-0.062
	DDE	7.643	0.013	27.30	2.313	3	1.373
VI-VIII	PB	4.370	1e-04	5000	1.573	1	-1.146
	BD	5.581	2.252	5000	1.687	2	0.626
	DD	4.894	1e-04	108.5	1.590	2	0.818
	DDE	5.895	0.847	60.96	1.705	3	2.590
V-VIII	PB	4.370	1e-04	5000	0.985	1	0.023
	BD	5.673	2.221	5000	1.151	2	1.697
	DD	4.606	1e-04	477.8	0.982	2	2.036
	DDE	16.00	7.167	32.45	1.817	3	2.367
IV-VIII	PB	3.955	1e-04	5000	-0.174	1	2.347
	BD	4.509	0.957	5000	-0.123	2	4.246
	DD	4.234	1e-04	361.3	-0.156	2	4.313
	DDE	4.482	0.235	242.3	-0.116	3	6.231
III-VIII	PB	3.440	1e-04	5000	-2.014	1	6.030
	BD	3.450	0.066	5000	-2.012	2	8.025
	DD	3.764	1e-04	284.1	-1.951	2	7.902
	DDE	8.925	3.600	62.51	-1.424	3	8.850
II-VIII (VEN)	PB	3.0851	1e-04	5000	-5.9995	1	13.999
	BD	3.2570	0.3710	5000	-5.9739	2	15.969
	DD	3.1708	1e-04	665.84	-5.9847	2	15.947
	DDE	7.0665	2.9567	73.890	-5.5344	3	17.068

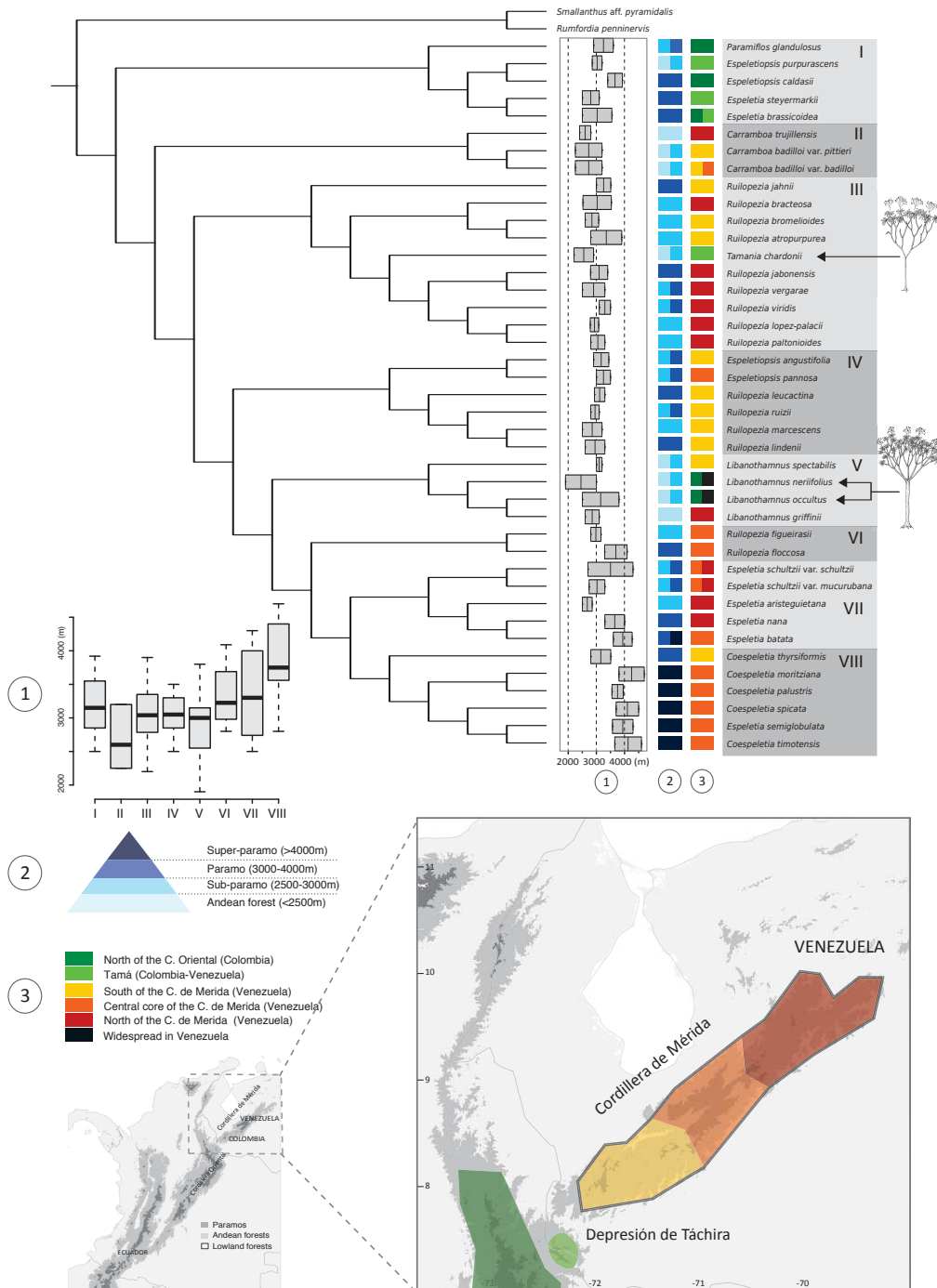


**Supp. Figure 3.** Morphological reconstruction of ancestral state of the Espeletiinae according to the nuclear chronogram for the growth-form (GF), the stem type (STEM), inflorescence position (IPOS) and structure (ISTR) traits. The stem node was only constrained for the GF trait.

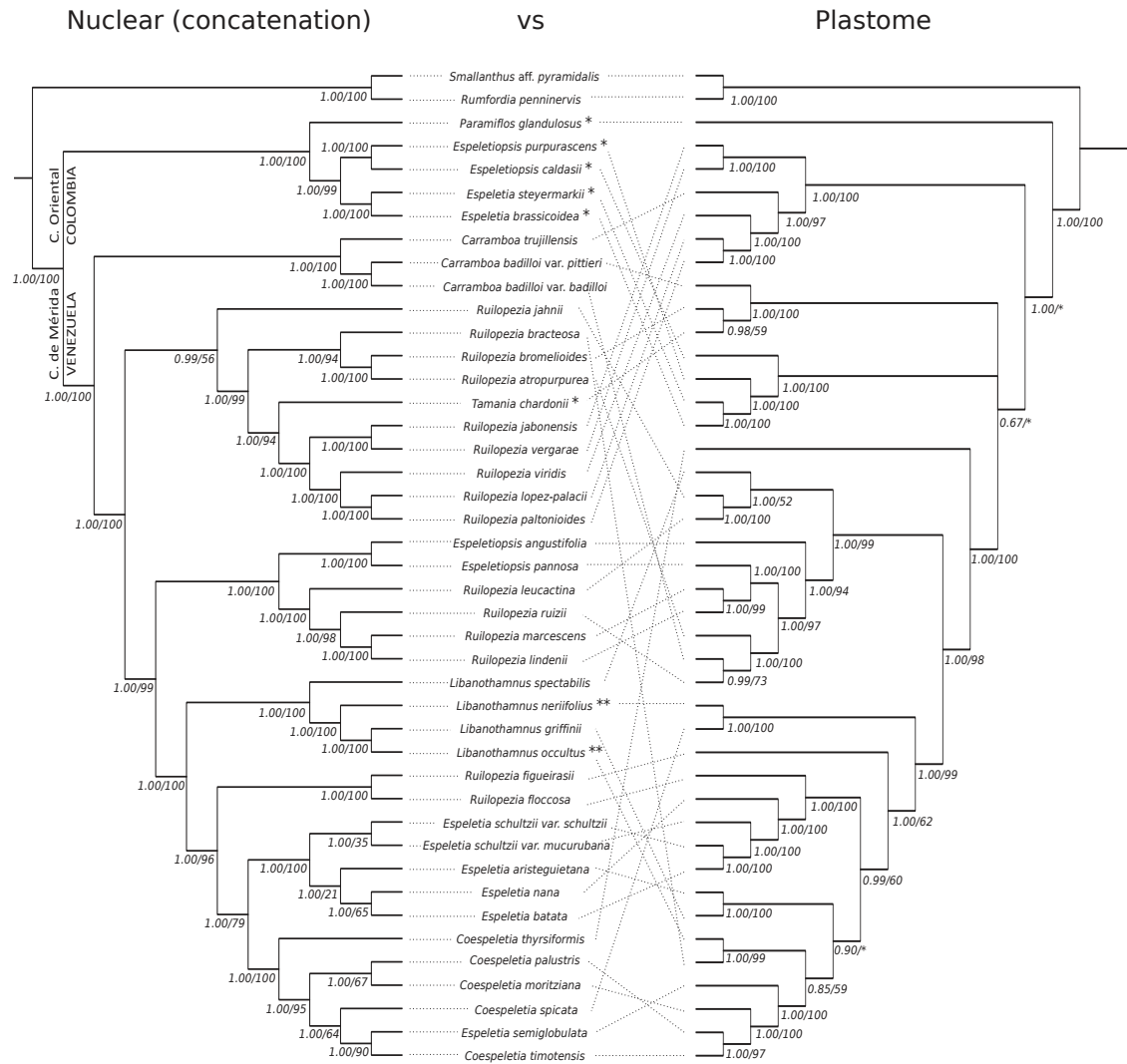




**Supp. Figure 4.** Schematic representation of (1) altitudinal, (2) habitat and (3) geographic distributions in the Colombian and Venezuelan Andes for the taxa of Espeletiinae analysed. Putative dispersals across the Táchira depression have occurred only among low-elevation tree lineages in clades III (*Tamania chardonii*) and V (*Libanothamnus spp.*).



**Supp. Figure 5.** Tanglegram indicating the high degree of discordance between nuclear and plastid phylogenetic trees. Node supports: Bayesian inference (BI/-) and maximum likelihood (-/ML). Species distributions: Venezuelan Andes (no marks), Colombian Andes (\*), both countries (\*\*).



---

## CHAPTER 2 – Specific, Morphological and Ecological Diversification of Espeletiinae

---

*“Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this archipelago, one species has been taken and modified for different ends”*

Darwin, 1845

### 1 | INTRODUCTION

Since Darwin's study of Galapagos finches, understanding the tempo and the mode of diversity changes in species is a major goal of evolutionary biology. Species diversity on Earth does not equivalently appear along different time periods, across organisms or at different ecological or geographic scale (Gaston, 2000). Evolutionary radiation, resulting from the proliferation of species diversity, provides ideal framework to perceive processes at the origin of diversity (Losos and Mahler, 2010). Radiations of species can arise with ecological opportunities providing by biotic or abiotic precursors as the evolution of key innovations allowing colonisation of novel habitats, the release of from competitor's pressures or the emergence of new habitats (Donoghue and Sanderson, 2015). According to the following speciation processes taking place after the colonisation of new environments, different types of radiations have been distinguished depending on their most common speciation mode

Adaptive radiation, based on the concept of adaptive divergence leading to ecological speciation (Givnish, 2015), is an iconic type of evolutionary radiation. It is defined as the

“rise of diversity of ecological roles and attendant adaptation in different species within a lineage” driven by a divergent selection (natural and sexual) in different environments (Schluter, 2000; Simpson, 1953). Most famous examples of ecological radiations are documented in Darwin’s finches (Grant and Grant, 2008), African Rift Lake cichlids (Seehausen, 2006), Caribbean Anolis lizards (Losos, 2009), Hawaiian lobeliads (Givnish et al., 2009) and honeycreepers (Pratt, 2005) or in Andean *Lupinus* (Hughes and Eastwood, 2006). Typically, adaptive radiations are associated with an explosive burst of diversification, but rapid lineage-splitting events are not necessarily a *sine qua non* condition of ecological speciation (Givnish, 2015; Rundell and Price, 2009). For example, Darwin’s finches in the Galapagos diversified no faster than their continental sisters, though exhibited a strong morphologic and ecologic divergence (Sato et al. 2001). According to Givnish (2015), adaptive radiations require a high level of phenotypic disparity and ecological divergence in comparison to sister taxa. However, the pervasive role of ecological speciation in rapid species diversification requires to be fully demonstrated from micro-evolutionary studies of ecological divergence in many groups. In fact, we can legitimately wonder whether any case of plant adaptive radiation has ever been demonstrated.

Unsimilar, ecologically similar species can diverge through geographical separation and limited species dispersal. This could lead to non-adaptive radiation, where ecological selection is not expected to be a primary driver of genetic isolation (Lieberman, 2012; Wiens, 2007). Remote areas (lakes, oceanic islands, sky islands) with low predator or competitor’s pressure allowing lineages to take ecological opportunities have certainly provided an ideal key landscape for both adaptive and non-adaptive radiations. Triggered by the same island-like opportunities, these phenomena may have occurred in almost all continental mountains (considered as “sky-islands”) of the world as in Rocky mountains (Drummond et al., 2012b), New Zealand (Joly et al., 2014), Himalaya (Wen et al., 2014), New Guinea (Brown et al., 2006), in European Alps (Boucher et al., 2012; Roquet et al., 2013) or in the Andes into *páramos* habitats (Hughes and Eastwood, 2006; Madriñán et al., 2013).

*Páramos* are spectacular tropical high-elevation ecosystems occurring in the Northern Andes, between the upper tree-line at *ca.* 3000-3500 m and the permanent snow-line at *ca.* 5000 m. These ecosystems, occupying an area of about 30,000 km<sup>2</sup> (roughly the size of Belgium), extend along the Andes of Venezuela, Colombia and Ecuador, with some outliers in Northern Peru and Panama/Costa Rica (Josee et al., 2009). *Paramos* are also very young environments, because the Northern Andes reached present-day elevations through a rapid final uplift at around 2.7 Ma (late-Pliocene) (Gregory-Wodzicki, 2000; Hooghiemstra et al., 2006). This ecosystem is characterized by low average temperatures, lack of seasonal changes in average daily temperatures coupled with extreme day-night temperature oscillations, high levels of solar radiation and several other features typical of tropical high-elevation ecosystems (Luteyn, 1999). Yet despite its small area, young age and harsh conditions, the *páramo* is extraordinarily diverse, harbouring at least 4700 plant species that is about 10-20% of the Andean floral richness (Rangel-Ch., 2015), and much more than any other tropical high-elevation flora, e.g. East-Africa highlands: 371 species, New Guinea highlands: 1118 species (Sklenář et al., 2011, 2014). Furthermore, the *paramo* is also known as the habitat where some plant clades experienced the World’s fastest rate of species’ diversification, hosting several lineages being both very young (< 4 Ma) and very diverse (> 100 spp)



**Figure 1.** Morphological and ecological diversity among Espeletiinae. (a.) Vegetative traits are composed of rosette forms with varying stem high (1, 2, 4), of tree forms (5, 7), including large and small leaves (3, 6). (b.) Reproductive traits are composed of terminal (1,5,7) and lateral (2, 6) inflorescences, large drooping capitulum (2) or smaller ones (5,6), including different flower colour (3, 4) and length of the inflorescence axis (1, 6). (c.) Ecological distribution of species varies from closed-upper Andean forest (1) to the open grasslands of páramos (2), as well in boggy (3) as rocky (4) areas. Photos by Serge Aubert.

(Hughes and Atchison, 2015; Madriñán et al., 2013). As a consequence, rates of endemism are exceptionally high in páramos, with about 40% of endemic species at a local or country-wise scale (Balslev, 1988; Briceño and Morillo, 2002) and virtually 100% for this habitat as a whole (Madriñán et al., 2013).

The uplift of the Andes played a major role in diversification of Neotropical plant species (Antonelli and Sanmartín, 2011a). Nevertheless, most speciation events highlighted by Madriñán et al. (2013) in páramos clades seem to have occurred during the Pleistocene, thus inviting questions regarding the drivers of speciation. For example, the repeated oscillations and fragmentations of species during climatic cycles of the Pleistocene, causing range expansion and contraction of the vegetation belts, probably explain the exceptional diversification rates of páramos taxa (Madriñán et al., 2013). This ecosystem therefore provides an outstanding framework for evolutionary studies of the origin of tropical plant

biodiversity, with emblematic examples in the *Lupinus* (Drummond et al., 2012b; Nevado et al., 2016), *Bartsia* (Uribe-Convers and Tank, 2015), *Hypericum* (Nürk et al., 2013), *Diplostephium* (Vargas et al., 2017), *Valeriana* (Bell et al., 2015) or in Espeletiinae (Diazgranados and Barber, 2017; Pouchon et al., 2018).

The subtribe Espeletiinae (Asteraceae: Heliantheae) is particularly emblematic of the páramo ecosystem. Numerous morphological combinations characterized this complex of 140 species, as large trees, shrubs, stemless and caulescent rosettes (Fig. 1a) and different reproductive traits (Fig. 1b) (Cuatrecasas, 2013). In addition, this complex has a wide distribution across the páramos from wetlands to open landscape or rocky slopes, and from the upper limit of Andean forests (~2500m) to the edge of glaciers (over 4600m) (Fig. 1c). Through ecophysiological and ecological diversity, Espeletiinae are considered as an emblematic case of adaptive radiation (Cuatrecasas, 1986; Monasterio and Sarmiento, 1991), but this has never been properly established. The understanding of processes at the origin of this subtribe in páramos has been hampered by a lack of phylogenetic resolution (Diazgranados and Barber, 2017; Panero et al., 1999; Rauscher, 2002), which has recently been overcome (Pouchon et al., 2018). According to previous results, Espeletiinae diversified very recently after the final uplift of páramos into two independent radiations in both Colombia and Venezuelan from a branched rosette ancestor. Numerous morphological changes in the complex, usually implying the convergent evolution of growth-forms with frequent losses and gains of various traits, were also highlighted. For example, the unbranched caulescent rosettes evolved independently from a branched-form three times in the subtribe: one in the Colombian radiation and two in the Venezuelan one. This convergence could be the result of colonisation of the same adaptive peaks by distantly related lineages (Schluter, 2000). Repeated evolution of vegetative traits in Espeletiinae could be appeared in response to the cold, dry and high-insolation conditions, which characterize high elevation environments of the páramos.

The influence of traits in diversification appears to be dependant on the geographical, environmental, and temporal context, where both abiotic and biotic forces could act synchronously on different spatial and temporal scales (Bouchenak-Khelladi et al., 2015; Donoghue and Sanderson, 2015). For example, the interaction between geography, environment and biotic traits has recently been highlighted in the diversification of the neotropical Bromeliaceae (Males, 2018). In Espeletiinae, Pouchon et al. (2018) proposed a main hypothesis for the drivers of their radiation into páramos including two processes acting at different time-scales. First, the evolution of rosette probably constituted a key innovation allowing colonization of páramos in both countries, which likely evolved in response to the ecological opportunity allowed by the emergence of novel habitats (Pouchon et al., 2018). Secondly, paleo-environmental changes associated glacial dynamic occurring during the Pleistocene probably promoted diversification of these species (Pouchon et al., 2018). According to theoretical expectations (e.g. in Bouchenak-Khelladi et al., 2015; Donoghue and Sanderson, 2015), shift in vegetative traits to a rosette form would probably trigger radiation of this group by increasing the diversification rate of species, with a modulating effect into such rates from paleo-climatic dynamic during their radiation. Moreover, the subtribe of Espeletiinae also exhibits a high diversity of reproductive traits (Fig. 1b), with altitudinal shift

in the reproductive system from biotic pollinators at low elevation to wind-pollination at high elevations (Berry and Calvo, 1989; Rundel et al., 1994). According to the pollinator-driven model (or Grant-Stebbins model) described in Johnson (2010), shifts in pollination system could accelerate the rate of diversification in species. Examples of such process was recently highlighted for hummingbird-pollinated species in Lobelioideae (Lagomarsino et al., 2016) and in Acanthaceae (Tripp and Tsai, 2017). In addition, Lagomarsino et al. also showed that the combined effect of such pollination shift to the Andean uplift and the paleo-climate shaped the diversification of these plants (Lagomarsino et al., 2016). In Espeletiinae, the differences in reproductive trait in response to the altitudinal pollination shift, hereafter named as pollination syndrome (Fenster et al., 2004), would also probably played an active role in the diversification rate within the species with both key innovation and paleo-climate. Moreover, Grant (1949) hypothesized that diversification for reproductive traits occurred prior to diversification in vegetative traits for lineages depending on specialized biotic pollinators. At the opposite, diversification of vegetative traits preceded reproductive divergence in lineages where pollination depended on unspecialized biotic pollinators and/or abiotic factors (Grant, 1949). As pollination depends principally on generalists' bumblebees and hummingbirds at low elevation and on the wind-pollination at high elevation (Berry and Calvo, 1989; Cuatrecasas, 2013), we hypothesized that vegetative diversification associated to habitat opportunity appeared earlier than pollination syndrome in the diversification of these plants into páramos.

Here, we seek to untangle factors implied in the diversification of Espeletiinae by investigating some of these predictions as some of theoretical assumptions of adaptive radiation. For example, ecological models of radiation predict a slowdown in diversification rates after an initial burst of rapid speciation (called “early burst”). Rates of both morphological and ecological evolution should be inversely correlated to the availability of ecological niches that are progressively filled over time (Losos and Mahler, 2010). In addition, investigating the timing and the mode of evolution for vegetative and reproductive traits provide insights into the dominant mode of speciation during the radiation. We proposed to examine such issues for the Venezuelan radiation of Espeletiinae by using comparative analyses based on our previous phylogenomic results (from the Chapter 1) and using evolutionary models for several morphological and ecological characters. In regard to these expected factors triggering and modulating the diversification of Espeletiinae, we aim to respond to the following questions: (1) did the evolution of both vegetative and reproductive traits correlate to ecological trait and accelerate rate of speciation for this complex? (2) Did climatic dynamic occurring during Pleistocene is associated with a shift in the diversification of Espeletiinae? (3) Did key innovation in vegetative traits happen prior to pollinating syndrome during the radiation of Espeletiinae?

## 2 | METHODS

### Morphological and Ecological Traits Acquisition

Morphological data were retrieved for 136 species of the subtribe from Cuatrecasas (2013)

**Table 1.** Environmental and Morphological data used to estimate divergence among species.

	Environmental	Morphological	
		Vegetative	Reproductive
Continuous	Climate (18 bioclimatic variables of WorldClim); Elevation; Aspect and Slope (Digital Elevation Model)	Stem High (SH, in cm); Lamina length (LL, cm) and width (LW, cm)	Inflorescence length (IL, cm); Capitula diameter (CD, cm); Corolla length (CL, mm); Flowers ratio (FLOR: ray/disc flowers); Length of achene (ALM, mm)
Discrete	Habitat: Andean forest (0-absence/1-presence), subpáramo (0-absence/1-presence), páramo (0-absence/1-presence), superpáramo (0-absence/1-presence)	Growth form (GF, 0-rosette/1-non-rosette); Stem structure (STEM, 0-branched/1-unbranched)	Inflorescence position (IPOS, 0-lateral/1-terminal)

and original diagnoses. We selected a total of 11 continuous or discrete characters that were relevant and could be obtained for each species from a collection of traits, and we classified them into vegetative and reproductive traits (Table 1). We used the mean values provided for each species and performed a log-transformation for continuous traits. Finally, both vegetative and reproductive traits were separately reduced into synthetic axes using the *dudi.hillsmith* R function from the *ade4* package (Dray and Dufour, 2007).

To estimate the climatic niche of species, 6,681 geo-referenced point occurrence points were collected for the same 136 species from our own field surveys or the Paramo Plants Online web database after validation for the location of points (Madriñan et al. 2015). We keep only one occurrence per cell grid (1.0 km x 1.0 km) in order to reduce pseudo-replication of some sampling localities. For each record, we extracted values for elevation, aspect and slope from a digital elevation model and for the 19 climatic variables of the WorldClim dataset (Supp.Table1). Niche positions were then evaluated using the Outlying Mean Index (OMI; Dolédec et al., 2000) following a two-table ordination procedure. First, only uncorrelated environmental variables were kept and used into a principal components analysis (PCA) to estimate environmental gradient using the *dudi.pca* R function of the *ade4* package. Next, the niche position of each species was measured as the distance between the mean environmental conditions used by this species and the mean conditions across the studied localities, using the *niche* function (*ade4*). In addition to the climatic niches, we used the habitat preferences of species in vegetation zonation. A multiple correspondence analysis (MCA) was made in order to keep species association in multiple zones according to their occurrences. Four zones were considered based on overall elevation and vegetation structure: (1) the Andean forest (i.e. the upper part of forest), (2) the shrubby subpáramo (i.e. the ecotone between grass páramos and Andean forest), (3) the treeless open grass páramos, and (4) the superpáramo (the ecotone between páramos and the permanent snow region). Distribution of species in these areas was retrieved from Cuatrecasas (2013) and according to own field observations only for species used in following phylogenies.

## Taxonomic Sampling, DNA Extraction and Shotgun Sequencing

Missing species in a phylogeny can bias pattern of diversification rate through time (FitzJohn et al., 2009). To cope with this issue, sampling was improved compared to the study of



Pouchon et al. (2018) by adding 18 more species. The majority of the samples were collected in the field where leaf tissue was dried in silica gel. Our sampling, focusing here on the Venezuelan radiation, which now includes 94.4% of the 53 Venezuelan recognized species (putative hybrid were not considered from Diazgranados 2012). *Ichthyothere garcia-barrigae* was also added as outgroup according to Diazgranados and Barber (2017), along with *Rumfordia penninervis* and *Smallanthus* aff. *pyramidalis* that were already used in our previous study. In order to investigate the strength of the phylogenetic approach, two individuals of *Espeletia tenorae* were sequenced in addition to two individuals of *Espeletiopsis purpurascens* (including one museum herbarium sample) and one more intraspecific variety of *Carramboa badilloi* (*C. badilloi* var. *littlei*). Genomic DNA was extracted from dried leaves tissue according to the protocol of Pouchon et al. (2018) after shaving the leaves and grinding them with a TissueLyser (Qiagen Corp.). Shotgun libraries were then prepared and sent for shotgun sequencing according to the Illumina HiSeq 2000 protocol from the manufacturer's instructions (Genoscope, France).

### Alignments, Phylogenetic inference and Dating

Sequence reads were filtered based on a Phred score quality value of 20 using the FASTX toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). Cleaned reads were next aligned to the 9,585 nuclear metacontigs loci generated in Pouchon et al. (2018) thanks to BWA 0.7.5 (Li and Durbin, 2009). In order to avoid paralogs biases, only to a single metacontig with maximal MAPQ score of 60 were kept for variant calling using SAMtools v0.1.19 and BCFtools v0.1.19 (Li et al., 2009). SNPs were filtered using the vcfutils script (supplied with SAMtools) according to a root-mean-square of the SNP mapping quality  $\geq 15$  and to a minimum SNP coverage of 10 per taxa. Finally, we excluded loci that were sequenced in less than four taxa from the final dataset.

Phylogenetic reconstructions were made from concatenated alignments of all loci using Maximum Likelihood (ML) and Bayesian approaches. The ML analysis was performed using RAxML-HPC (Stamatakis, 2006) with a GTRCAT substitution model and 500 bootstrap replicates. Bayesian inference was conducted in ExaBayes v.1.4.1 (Aberer et al., 2014) using the default values for temperatures of chain heating and the number of Metropolis-coupled Markov chain Monte Carlo (MCMC) under 500,000 generations. We assessed convergence and mixing using summary statistic of the *postProcParam* function in ExaBayes. We ensured a minimum of 200 for the effective sample sizes of each parameter, and the similarity within and between chain-variance with a potential scale reduction factor close to 1 and below 1.1. A 50% majority-rule consensus phylogram with posterior probabilities was obtained using the ExaBayes *consense* function with a burn-in fraction of 25%.

Calibration of divergence times within Espeletiinae was made according to the approach of Pouchon et al. (2018). The authors estimated the age for the crown node of Espeletiinae at 2.3 (2.0–2.6) Ma. We used this estimates of the complex with relaxed-clock penalised likelihood method in r8s v.1.7.1 (Sanderson, 2003). Uncertainty in the estimation of divergence time was assessed using 200 trees randomly selected from the stationary distribution of the MCMC chain (i.e. after the burn-in phase). Cross-validation was conducted

on each tree replicate to determine the optimal level of rate-smoothing for 100 estimates starting with a log10 value of -5 and increments of 0.4. Finally, mean values and 95% confidence intervals for each node age were estimated from the chronograms using TreeAnnotator (Drummond et al., 2012a).

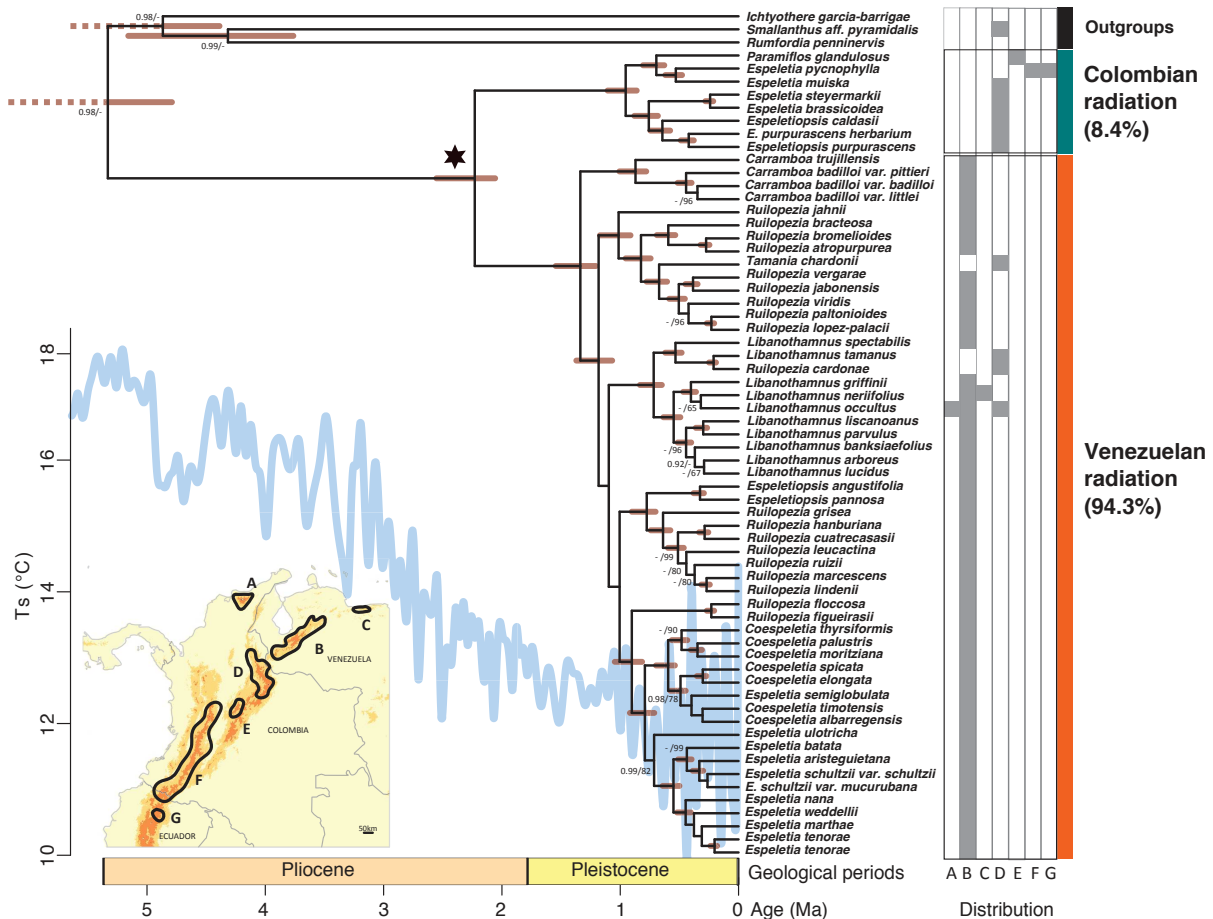
## Morphological and Ecological correlation

We measured the correlation between phenotypes and environments to assess the adaptive values of traits using Phylogenetic generalized least squares (PGLS) regression. PGLS takes phylogenetic relationships into account, assuming that the evolution of residual traits follows a neutral model (Brownian motion). PGLS was conducted using the R package *nlme* (Pinheiro et al., 2018) to test relationship between vegetative or reproductive traits (as phenotype) and the climatic (OMI) or habitat variable (as environment). Phylogenetic relationships used in the PGLS were derived from the time-calibrated Bayesian tree. Regressions were compared to a null model using a likelihood ratio test. As we tested phylogenetic correlations between two dimensions of the vegetative and reproductive PCA, with the two dimensions of the OMI and habitat correlations, we used the method of Benjamini and Hochberg (1995) to adjust the statistical significance of p-values for inflated false positive discovery rates.

## Diversification Analyses

We used the 200 Bayesian chronograms for ML and Bayesian macro-evolutionary models of diversification of lineages and morphological traits for the Venezuelan section of the phylogeny. Trees were first pruned to the species level by removing morphological varieties of *C. badilloi*, *E. schultzi* and keeping only the original species form. Only one individual of *E. tenorae* was kept. In all ML analyses presented below, the fit of various evolutionary models was compared using the Akaike information criterion (AIC).

**Lineage diversification**– In order to investigate shift of lineage diversification during the evolutionary history of Espeletiinae, we used BAMM (Rabosky, 2014). This software estimates speciation and extinction rates through time of lineages by modelling variation of these rates through time but also detecting shifts in diversification dynamics in some subclades that best explain the whole diversification of the clade. We ran BAMM with four reversible jump MCMC chains, each with 10 million of generations and sampled every 10,000 generations. Value of the Poisson process rate prior was set to 1.0 to minimize the number of shifts under the prior alone. Priors for both speciation and extinction rates was determined using the R function *setBAMMpriors*. After a burn-in of 10%, MCMC stationarity and (ESS) were assessed in R using the package CODA (Plummer et al., 2006). BAMM outputs were analysed into the R package ‘BAMMtools’ (Rabosky et al., 2014) to plot the estimated net diversification rates (phylorate), and identify the most credible set of distinct rate shift configurations. The effect of diversity on both speciation and extinction rates were also explored using the method of Etienne et al. (2012) implemented in the R package ‘DDD’. Four models were considered and evaluated: (1) a Pure-Birth (PB, or Yule) model, which assumes no extinction rate (single speciation rate); (2) a Birth-Death (BD) model, allowing extinction and speciation; (3) a Diversity-Dependent Logistic (DDL) model, which assumes a dependence of lineages into the diversification with single speciation rate; (4) a Diversity-

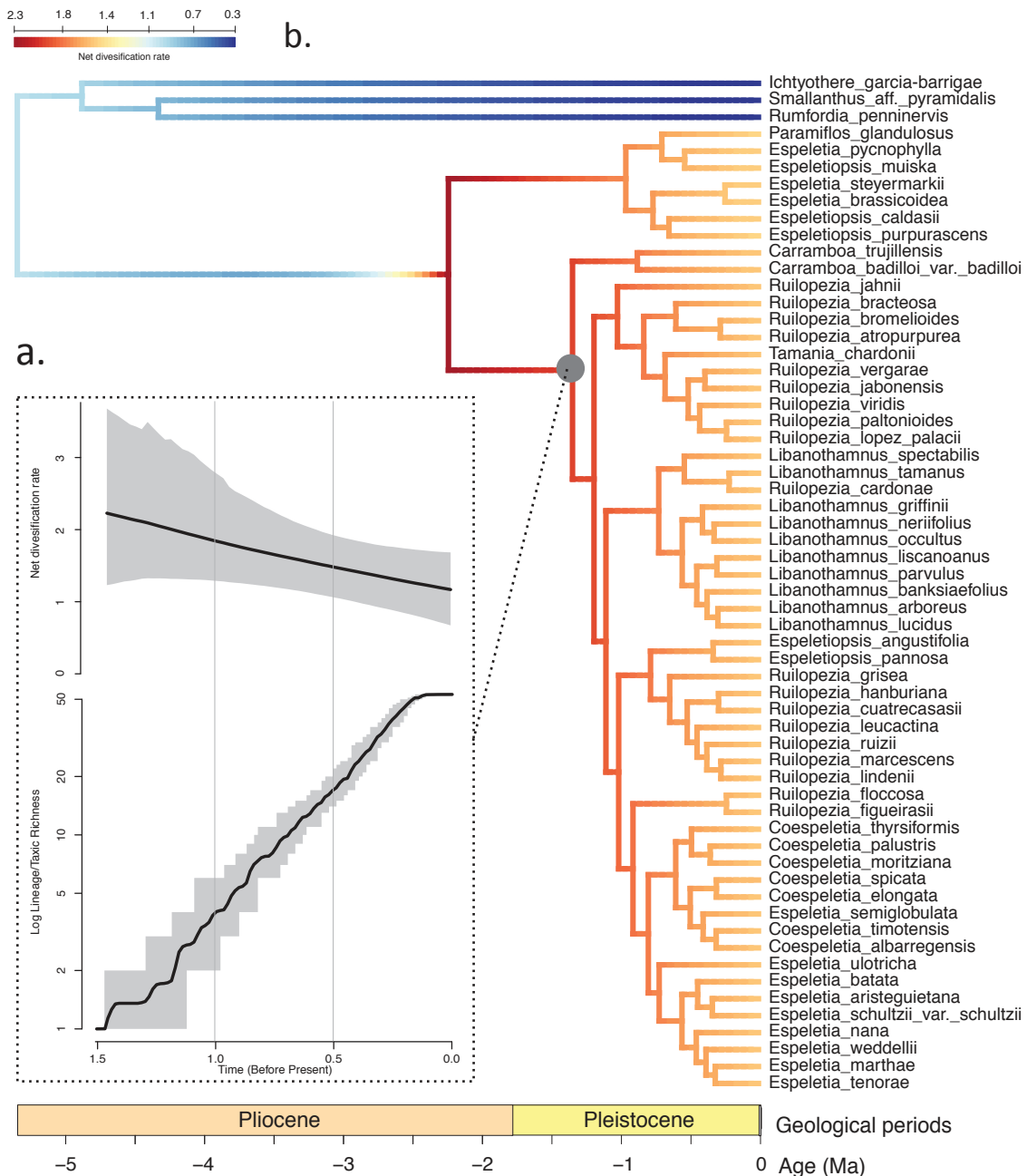


**Figure 2.** Divergence time and phylogenetic relationships of Espeletiinae. Node value indicate Bayesian posterior probabilities (PP) and maximum-likelihood bootstrap percentages (BP). Only not fully supported values are shown (i.e. different from 1.00/100). Bars on nodes represent 95% confidence intervals. Black star indicate node used for calibration according to estimates of Pouchon et al. 2018. Blue lines show sea surface temperatures (Zachos et al. 2001). Distribution of species are indicated for regions in map corresponding to (A) the Sierra Nevada de Santa Marta, (B) the Cordillera de Mérida, (C) Avila, and the paramos (D) of Santander, (E) of Guantiva-La Rusia, (F) of the Cordillera Central Colombian and (G) of Llanganates (Ecuador).

Dependent Logistic (DDL+E) including Extinction rate. The proportion of missing species in the Venezuelan section was incorporated into models.

**Morphological Diversification—** We investigate the influence of both vegetative and reproductive principal components on macroevolutionary lineage diversification. We employed the Quantitative State Speciation and Extinction (QuaSSE) model implemented in the R package ‘diversitree’ (FitzJohn, 2010). This approach fits birth-death models on which speciation and extinction rates vary depending on the value of a continuous trait, which itself evolves along the phylogeny according to a Brownian motion model. The proportion of missing taxa was also incorporated into these models. We fitted models in which the speciation rate was a constant (independent of trait values, null model), linear, sigmoid and hump-shaped function of vegetative and reproductive values. We also fitted the same models by adding a ‘drift’ parameter describing the directional rate of change in the trait to the present time (FitzJohn, 2010).

We also quantified the tempo and the mode of vegetative and reproductive diversification by fitting five models with different evolutionary dynamics: (1) the Brownian



**Figure 3.** Bayesian analysis of macroevolutionary mixtures in Espeletiinae. (a.) Bayesian reconstruction of rate variations in net diversification through time in association with the lineage through time (LTT) plot of the diversification of Venezuelan Espeletiinae for 200 time-calibrated trees (grey) and the median time-calibrated tree (black line). (b.) Maximum a posteriori probability shift configuration represented as a phylorate plot showing variations in speciation rates (cool colours = slow, warm = fast) along each branch of the *Espeletia* phylogeny. Each unique colour section of a branch represents the mean of the marginal posterior density of speciation rates on a localized segment of a phylogenetic tree.

Motion (BM) model, assuming a random walk of trait evolution along the phylogeny; (2) the Ornstein-Uhlenbeck model (OU), assuming adaptive evolution of trait under stabilizing selection toward an adaptive optimum; (3) the Early-Burst (EB) model, assuming exponential decrease of evolutionary rates over time based on “niche filling” assumption; (4) the Pagel’s delta model of time-dependent trait evolution, similar to the EB model, fitting the relative contributions of early versus late evolution in the tree to the covariance of species trait values;

**Table 2.** Lineages diversification of Espeletiinae based on four evolutionary models: constant (1) Pure-Birth (PB) without extinction rate and (2) Birth-Death (BD) model; and density dependent models without (DDL) and with (DDL+E) estimation of extinction rate. Models were estimated from the *diversitree* R package.

	lambda	mu	K	loglik	Df	AIC	P(IChil)
PB	1.823	0.000	INF	-20.71	1	43.43	0.160
BD	1.975	0.143	INF	-19.73	2	43.47	0.157
<b>DDL</b>	5.125	0.000	56.570	-6.657	2	<b>17.31</b>	<b>1.16e-07*</b>
DDL+E	5.210	0.018	55.511	-6.745	3	19.48	8.74e-07*

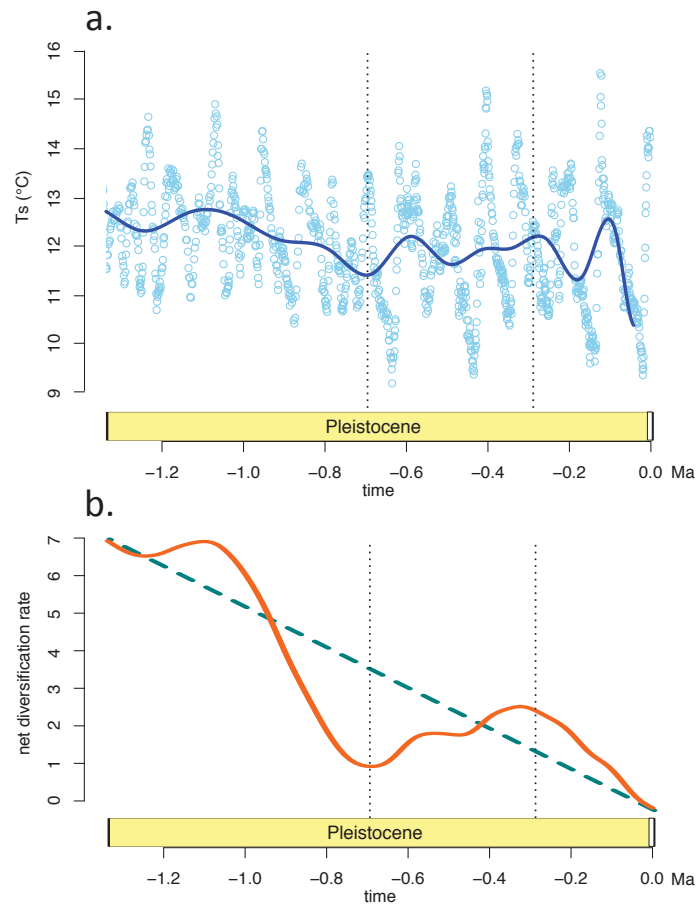
(5) the White-Noise (WN) model, assuming a random walk of trait evolution independent of the phylogeny. Delta values was greater than 1 when recent evolution has been relatively fast and less than 1, when recent evolution has been comparatively slow. Models were fitted by including measurements of the standard errors in trait measurements in the R package ‘geiger’ (Harmon et al., 2008).

**Paleoenvironment-Dependent Diversification**– In order to estimate the effect of Pleistocene climatic dynamic on the diversification of Espeletiinae, we used the derived approach of Condamine et al. (2013), which allows variation of speciation and extinction rates through time according to a paleoenvironmental variable (here global paleotemperature estimated from Zachos et al. (2008), see Fig.2). Extinction rates were set to zero. Following the approach of Morlon et al. (2011), we compared two models of diversification: (1) a constant rate of diversification under a Pure-Birth model (BDCST, null model), and (2) an exponential variation of speciation rate with the temperature (ENVEXP\_L). We also tested (3) linear models with dependence of temperature into diversification (ENVLIN\_L), (4) without effect of temperature (BDLIN), and both linear and exponential models for which the speciation rate was constant and only the extinction rate varied (ENVLIN\_M and ENVEXP\_M). This analysis was performed in R through the package ‘RPANDA’ (Morlon et al., 2016).

### 3 | RESULTS

#### Phylogenetic inference

Phylogenetic alignment comprised 4,025,033 bp, including 5.94% of informative sites and 46.59% of ambiguous or missing data. The concatenated ML and BI phylogenies appeared well resolved, highly congruent, and supported the previous phylogeny of Pouchon et al. (2018) (Fig. 2). Most of the nodes were recovered with a strong support (PP >0.95). Only a main topological difference appeared, consisting in the placement of the genus *Libanothamnus* (clade V in Pouchon et al. 2018), which now appeared more basal within the Venezuelan section in comparison to the clade IV. As for the position of specific taxa, the new sampled species of *Libanothamnus* appeared closely related to others within the same clade, as for the new species of *Espeletia* and *Coespeletia*, which were also closely related respectively to sister taxa in *Espeletia* and *Coespeletia* clades. For *Ruilopezia* species, the new



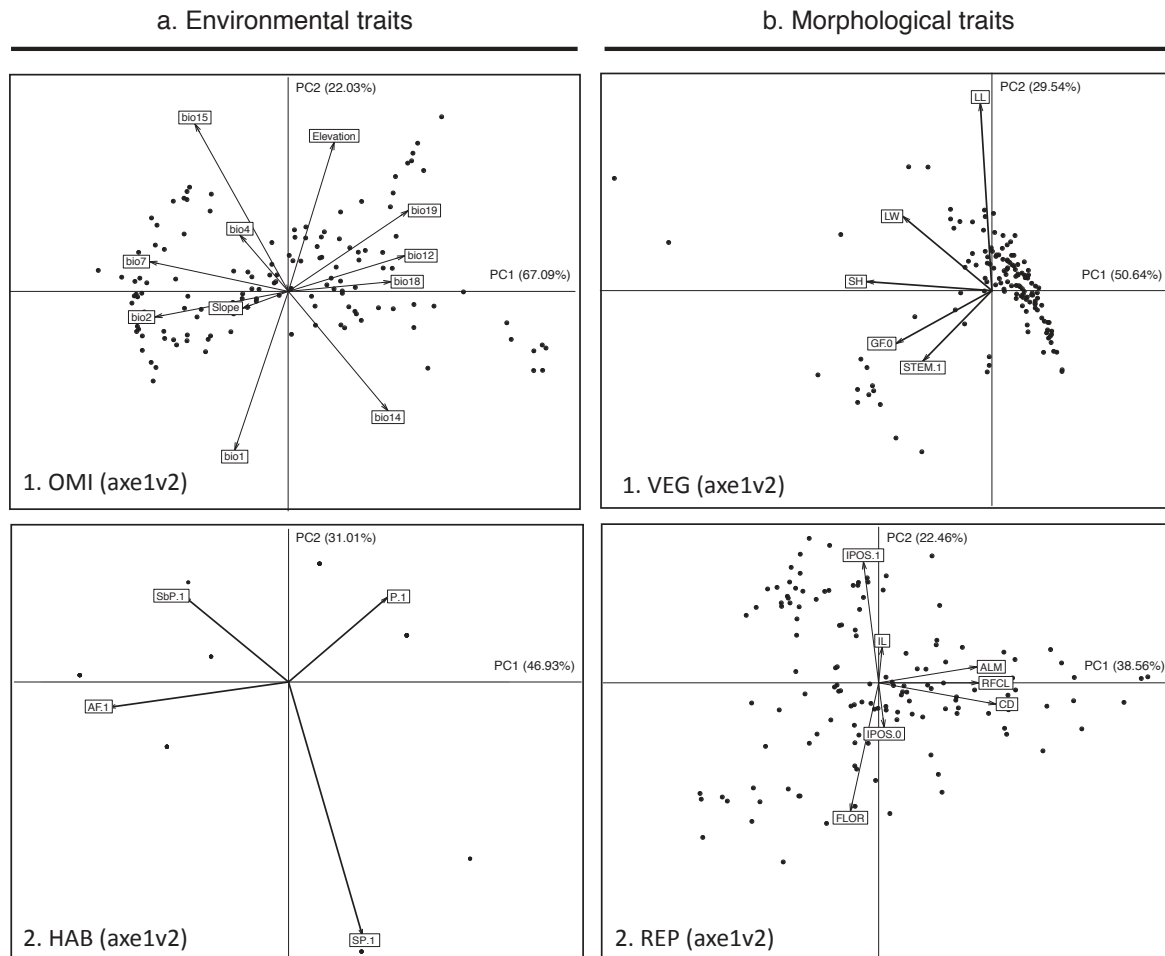
**Figure 4.** Diversification pattern of Espeletiinae. (a.) Past fluctuations of temperatures for the Pleistocene (data retrieved from Zachos et al. 2001). (b.) Speciation rate through time for Espeletiinae obtained from a linear Pure-Birth model (blue dashed line) and from the linear relationship between diversification with paleotemperatures using the approach of Condamine et al. (2013) (models refers to PBLIN and ENVLIN\_L described in Supp. Table 2).

ones formed a clade with the northern *Ruilopezia* (*R. leucactina*, *R. ruizii*, *R. lindenii* and *R. marcescens*). However, *R. cardonae* had an odd position inside the *Libanothamnus* clade, appearing closely related to *L. tamanus*, in contrast to others *Ruilopezia* species. Finally, the intraspecific relationships were fully supported for *Espeletia tenorae* individuals, *Espeletiopsis purpurascens* (including herbarium specimen), and for the morphological varieties of *E. schultzii* and *Carramboia badilloi*, supporting the strength of this approach at recovering phylogenetic relationships between closest relatives.

Divergence time estimations provided narrow age estimates for almost all nodes except for the three most basal ones. The age of the crown node of Espeletiinae was estimated at 2.23 Ma (95% CI: 2.06 – 2.55 Myr ago), which is consistent with the calibration using the plastid-based age estimation of 2.03 – 2.56 Ma of Pouchon et al. (2018).

### Time-dependent diversification

The Bayesian analysis of macroevolutionary mixtures (BAMM) rejected a constant diversification model by showing that diversification rates tended to decline through time but no diversification shift was inferred (Fig. 3). In addition, the ML analysis of lineage



**Figure 5.** Distribution of species into the first two axes of the (a.) environmental and (b.) morphological space. Environmental component is estimated by the occupied climatic niche space (a.1 – OMI) and the habitat (a.2 – HAB) whereas morphological component is estimated by the occupied vegetative space (b.1 – VEG) and the reproductive space (b.2 – REP).

diversification identified the diversity-dependent logistic models (DDL) as the best model of lineage diversification for the Venezuelan radiation of Espeletiinae (Table 2). The DDL model predicts that the diversification of species tended to be stabilized to a carrying capacity (K). Such pattern was concomitant to the LTT plot in Fig. 3, on which the accumulation of the lineages through time tended to slow down recently.

Diversification rates varying as exponential (ENVEXP\_L, ENVEXP\_M) and linear (ENVLIN\_L, ENVLIN\_M) functions of paleotemperature were rejected as best-fitting models of diversification for Espeletiinae (Supp. Table 2). As for previous findings, a model with a linear decrease in diversification rate overtime was preferred (PBLIN, Fig. 4). However, in comparison to the null model, linear model of environmental-dependence for speciation rate (ENVLIN\_L) appeared as the second best-supported model according to the AIC values, on which speciation rate linear decreased with the decrease of the paleotemperature (Fig. 4). As global temperatures tended to cool overtime during the Pleistocene (Zachos et al., 2001), time- and temperature-dependent diversification rates are

**Table 3.** Phylogenetic generalized least-squares (PGLS) models assessing the relationship between the Environmental and the morphological space position occupied by species. Pvalues of the likelihood-ratio test were adjusted by the Benjamini-Hochberg method (1995).

Morphological		Environmental				
		OMI		Habitat		
		PC1	PC2	PC1	PC2	
Vegetative	PC1	coefficient	-0.39746	0.33135	0.33135	0.65758
		likelihood ratio	1.3167	1.8446	16.774	2.6382
		adjusted pvalue	0.456727	0.3488	<b>0.00085*</b>	0.334285
	PC2	coefficient	-0.14236	-0.44245	-0.21208	0.47362
		likelihood ratio	0.265	5.4207	0.61281	1.938
		adjusted pvalue	0.808933	0.0995	0.667230	0.3488
Reproductive	PC1	coefficient	-0.28247	0.58784	0.25071	-0.70283
		likelihood ratio	1.9896	9.3679	1.1616	7.6013
		adjusted pvalue	0.3488	<b>0.02208*</b>	0.4685	<b>0.03888*</b>
	PC2	coefficient	-0.10969	-0.2267	-0.02534	0.039172
		likelihood ratio	0.3067	2.8064	0.019242	0.028181
		adjusted pvalue	0.808933	0.334285	0.905	0.905
	PC3	coefficient	-0.02731	0.264192	0.064983	0.13482
		likelihood ratio	0.014253	2.4576	0.068876	0.20618
		adjusted pvalue	0.905	0.334285	0.905	0.81225

difficult to distinguish.

### Patterns of Ecological divergence

The OMI identified two main gradients explaining over 89% of environmental variation. The first axis (PC1), accounting for 67% of the total variation in environmental variation, could be interpreted as a gradient of continentality: it opposed habitats with highly fluctuating temperatures and precipitations to environments with high total precipitation (Fig. 5.a1, Supp. Fig. 1). The second axis (PC2), explaining 22% of the variance mostly correlated with the altitudinal gradient (Supp. Fig. 2.b). For vegetation zonation two axes explained 78% of the variation among localities, distinguishing between lower ecotones (Andean forest/subpáramos) and upper open páramos in the first axis (Fig. 5.a2, Supp. Fig. 1). Elevated superpáramos were discriminated from páramos along the second axis.

### Patterns of phenotypic divergence

Almost 80% of the variance in vegetative traits could be accounted for by two axes after data reduction. The first principal component (PC1), accounting for 50% of variance, contrasted rosette and tree-form species with intermediated morphotypes (e.g. high stemmed-rosette or branched rosette) (Fig. 5.b1, Supp. Fig. 2). The second axis (PC2) discriminated species with long and wide leaves from species with short and narrow leaves. Data reduction for



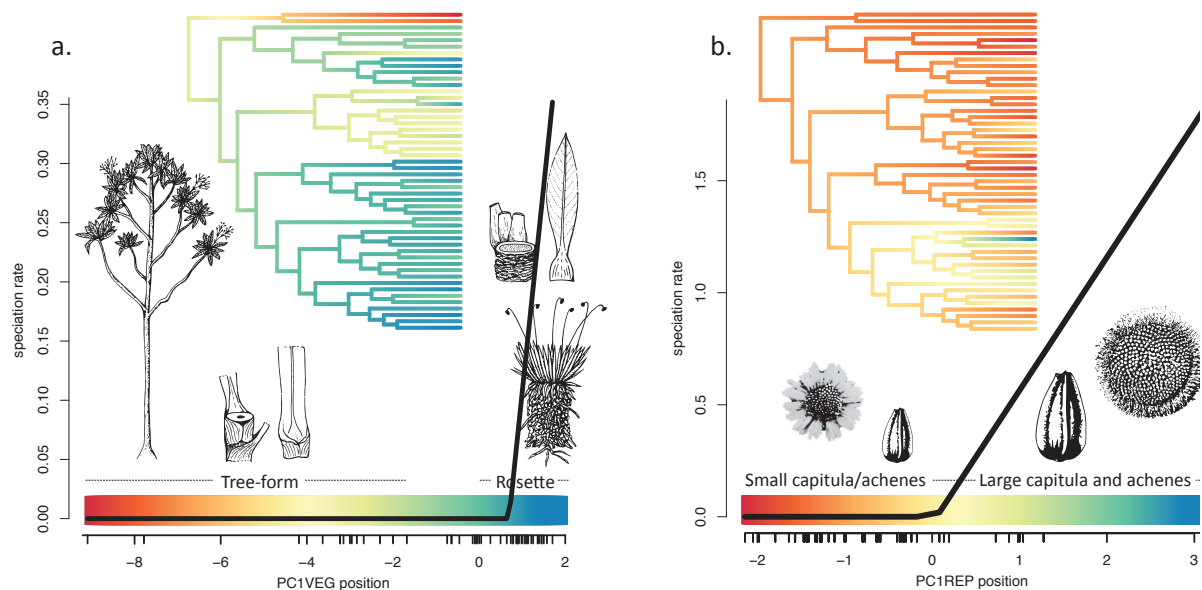
**Table 4.** Models fitting to test for trait-dependent speciation in Espeletiinae.

	df	PC1VEG			PC1REP		
		loglik	AIC	P( Chisql)	loglik	AIC	P( Chisql)
constant	3	-132.48	270.96	NA	-95.198	196.39	NA
linear	4	-132.02	272.04	0.337	-95.020	198.04	0.550
sigmoid	6	-132.48	276.96	0.999	-95.198	202.39	0.999
hump	6	-132.48	276.96	0.999	-95.198	202.39	0.999
<b>linear+drift</b>	5	-118.25	<b>246.50</b>	<b>6.62e-07*</b>	-78.761	<b>167.52</b>	<b>7.27e-08*</b>
sigmoid+drift	7	-132.39	278.78	0.996	-93.690	201.38	0.555
hump+drift	7	-132.39	278.78	0.996	-93.690	201.38	0.555

reproductive traits identified three axes explaining 77% of the observed variance. Species with larger heads (diameter of capitula and length of corolla) and fruits (length of achene) were discriminated from others plants under the first axis (PC1), explaining 38% of the reproductive traits variation (Fig. 5.b2, Supp. Fig. 2). The second component opposed species with terminal inflorescences and a lower ratio of ray/disc flowers to species with higher ratio and lateral inflorescences (PC2). Finally, species with high size inflorescence stem were discriminated from others into the last axis (PC3) in Supp. Fig. 3.

### Traits-dependent diversification

Phylogenetic generalized least-squares (PGLS) regression indicated a strong positive correlation between the PC1 values of vegetative traits and of for habitat variables (PGLS regression:  $P < 0.001$ , Table 3). In others terms, the tree and rosette life forms appeared correlated to the forest-subpáramo and páramo-superpáramo habitats, respectively (Fig. 5, Supp. Fig. 4). Another correlation was found between the PC1 values of reproductive traits and the PC2 values of both OMI and habitat variables ( $P < 0.05$ , Table 3): Capitula presented larger on the plant as well as for seeds were found in the highest elevations and the superpáramo habitats (Fig. 5, Supp. Fig. 5). Others components of vegetative and reproductive traits were not significant correlated to environmental variables. Tests of the influence of traits on diversification rates were performed only on these two-correlated axes (PC1VEG and PC1REP). For both of them, the best-fit QuaSSE model indicated that speciation rate exhibited a positive and linear response to the traits values, evolving directionally (linear+drift model, Table 4). High diversification rates corresponded therefore to high PC1VEG values, i.e. to species with the rosette life-form (Fig. 5.a). Similarly, high diversification rates appeared for high PC1REP values, i.e. to species with larger capitula and longer seeds (Fig. 5.b). Finally, Pagel's delta model best estimated the evolutionary dynamic of the vegetative trait with a delta value of 0.174 (Table 5), which was consistent to the second best fitting EB model. For reproductive traits, no difference could be established between BM, OU and delta models (Table 5). However, in contrast to the vegetative trait,



**Figure 6.** Results of QuaSSE modelling of trait-dependent speciation for the first axis of the (a.) vegetative and the (b.) reproductive space position occupied by species, according to the best fitted models in Table 5. Ancestral state reconstructions were made using the *contMap* function of the *phytools* R package (Revell 2012). The linear function returned by *make.linear.x* put zero values wherever negative for valid likelihood calculations under QuaSSE.

delta value was estimated at 2.543.

## 4 | DISCUSSION

Understanding evolutionary processes leading to plants radiation into páramos has long been complicated by the lack of phylogenetic resolution within clades of interest. Nowadays, the advent of new approaches in the phylogenomic era allow the inference of better-resolved phylogenies, bringing some light on the drivers of diversification in species-rich páramo taxa such as *Lupinus* (Nevado et al., 2016), Lobelioideae (Lagomarsino et al., 2016), *Diplostephium* (Vargas et al., 2017) and Espeletiinae (Pouchon et al., 2018). In this study we have provided a more complete phylogeny for Espeletiinae, reinforcing previous phylogenetic hypotheses (Pouchon et al., 2018) and unraveling the processes underlying the radiation of Espeletiinae into páramos for the Venezuelan section.

### Phylogenetic framework

The new phylogenetic inference corroborated previous relationships made by Pouchon et al. (2018) despite the increase of new sampled taxa (42 to 61 species) and concatenated loci (1,877 to 9,585 loci) into the final supermatrix. Phylogenetic relationships were coherent with the geographical distribution and/or the morphology of species. For example, all the members of the second clade of *Ruilopezia* (clade IV in Pouchon et al. 2018) were distributed to the northern of the Sierra de Merida. Closely geographical species of *Libanothamnus* (clade V) were equally related in the phylogeny as for closely morphological species (Cuatrecasas,

**Table 5.** Evolutionary dynamic of the vegetative and the reproductive position of species into the morphological space. Evolutionary trajectory of trait were modelled under Brownian motion (BM), Ornstein-Uhlenbeck (OU), Early-Burst (EB), Pagel's delta (delta) and White-Noise (WN) models in Geiger R package (Harmon et al. 2008).

	PC1VEG				PC1REP			
	values	loglik	AICc	$\Delta$ AICc	values	loglik	AIC	$\Delta$ AICc
BM	-	-92.02	188.3	7.890	-	-68.68	141.6	<b>1.101</b>
OU (alpha)	0.000	-92.02	190.5	10.15	1.050	-66.99	140.5	<b>0.000</b>
EB (a)	-1.434	-89.35	185.2	4.826	0.000	-68.68	143.8	3.367
delta (d)	0.174	-86.94	180.4	<b>0.000</b>	2.543	-67.03	140.5	<b>0.056</b>
WN	-	-112.9	230.2	49.84	-	-71.33	146.9	6.391

2013). Moreover, the two independent radiations occurring at the origin of the group in both Colombia and Venezuela were recovered, as the main clades decrypting in the previous study. In addition, the same scenarios of morphological change in Espeletiinae of Pouchon et al. (2018) were demonstrated. For example, the crown ancestor of the subtribe for both geographical radiations remained a branched rosette with imbricated open/flat leaf sheaths (Supp. Fig. 6). The unbranched-form of rosettes evolved three times, independently in both countries, likely as a convergent adaptation to life in tropical high-elevation habitats given the strong correlation of this life form with these habitats (Table 3, Supp. Fig. 4). Trees evolved independently three times from the repeated colonization of lower elevations by high-elevation rosette ancestors. Concerning the phylogenetic relationships described in Figure 2, the change in position of trees in clade V into the Venezuelan section had no effect on the morphological evolution scenario that was inferred for this complex. This different placement appeared to be mostly caused by the increase of loci included into the final supermatrix and not the addition of taxa (the same placement was observed when phylogenies were restricted to taxa sampled in Pouchon et al. (2018), results not shown here). An odd position of *Ruilopezia cardonae* (an unbranched rosette) in the genus *Libanothamnus* (trees) was observed with closed relation with *L. tamanus*. The two species are indeed endemic to the páramo of Tama in Colombia and isolated from the Venezuelan sisters by the Táchira depression. Introgression in Espeletiinae is extensive between sympatric and closely-related species (Pouchon et al. 2018, see Chapter 3). A strong level of genetic introgression occurring locally between these species could probably explain such odd position of these species. Finally, this study strengthened the extension of the genome skimming approach described in Pouchon et al. (2018) to a phylogenetic context by showing its good potential to resolve deep nodes as well as intraspecific relationships, from DNA of both fresh or degraded tissues (with the example of the herbarium specimen).

## Ecological opportunity and paleotemperatures led to a burst of lineages diversification

Diversification rate of the Venezuelan Espeletiinae was estimated by Pouchon et al. (2018) at 3.09 spp/Myr, ranking this subtribe among the fastest plant radiations behind *Dianthus* (Valente et al., 2010) and Andean *Lupinus* (Hughes and Atchison, 2015). In this study, our analyses corroborated this explosive burst of diversification for this Venezuelan section by considering an initial speciation rate at 5.12 spp/Myr on the best-fitted model of diversification (DDL model in Table 2). Consistent with predictions from ecological models of adaptive radiation, lineage accumulation over time revealed a density-dependent pattern for this Venezuelan radiation, in which diversification rates decrease as species accumulate in the clade (Table 2). This finding was supported by LTT plots and BAMM analyses (Fig. 3) and by the best-fitting PBLIN model (Supp. Table 2) showing decline of diversification rate through time (Fig. 4). Such a decline in diversification rates as species accumulate in a clade is often interpreted as competition for resources as ecological niches tend to be filled overtime (Losos and Mahler, 2010; Moen and Morlon, 2014). The dynamic of diversification of this Venezuelan section appeared dependent of a carrying capacity close to the extant number of species ( $K=56$  sp. for DDL, Table 2), suggesting that the radiation of these plants was close to its ecological limits (Rabosky, 2009). Consequently, our analyses seemed validate the early ecological opportunities assumption at the origin of the Venezuelan radiation with a probably niche saturation over evolutionary time.

Climate change appears to have played a significant role in the diversification of species, where warmer climates could serve as an evolutionary diversification pump (Erwin, 2009). In Espeletiinae, it has previously been proposed that the radiation has been influenced by paleoenvironmental changes triggered by glacial dynamics during the Pleistocene (Pouchon et al., 2018). In this study, when we applied explicit models described by Condamine et al. (2013), the paleotemperature appeared to have a linear influence in the diversification rate during the radiation of the Venezuelan Espeletiinae in comparison to the null-model (Supp. Table 2). Cooling of temperatures negatively impacted the speciation rate (Fig. 4). Positive effect of paleotemperature in diversification rate was recently found in mammals *Xenarthra* (Gibb et al., 2016) or in plants in Chinese *Primulina* (Kong et al., 2017) and in Andean Lobelioideae (Lagomarsino et al., 2016). In comparison to what was found for Lobelioideae, paleotemperatures seemed to impact the speciation rate but not the extinction rate in Espeletiinae. However, this relationship between diversification rates and paleotemperatures is difficult to distinguish from a relationship between diversification rates and time. Indeed, diversification in Espeletiinae was better explained by time-dependent rather than temperature-dependent models (Fig. 4, Supp. Table 2). Estimates of diversification rate decrease over time (according to the BAMM analysis) with an associated decrease of speciation rate in function of the paleotemperatures (Supp. Table 2). Further investigations could be required to untangle time vs temperature effect on diversification and to evaluate the impact of the climatic variance of temperature increasing during Pleistocene into the diversity of species in contrast to the global temperature. In addition, it would be important to assess if relation between paleotemperatures and diversification rate is direct (e.g. temperature has effects on mutation rate, recombination rate, seed size/dispersal) or indirect (e.g. temperature has an effect on paramo size/connectivity, which in turn has an effect on diversification rate).

## Phenotypic responses to changing ecological opportunities drove speciation

A primary finding of our study is that certain components of Espeletiinae morphological diversity show significant evolutionary correlations with species positions in along climatic and ecological gradients (Table 3), and that this in turn seem to have triggered species diversification.

Regarding vegetative traits, the evolution of the rosette growth-form tended to be associated to the open páramo/superpáramo habitats whereas the branching tree-like forms were associated to the close Andean forest and subpáramos habitats (Fig. 5, Table 4). This trend was not surprising owing to the numerous adaptive values afforded by the rosette into the páramos and the convergent evolution of this growth-form in all tropical alpine ecosystems on Earth. In fact, the rosette confers protection against frost to the apical meristem, reduced water loss through transpiration, increased volume of water stored in the large pith of an insulated stem (Monasterio and Sarmiento, 1991; Smith, 1981). Concerning reproductive traits, the evolution of larger capitula appeared to be associated with the higher elevations and the superpáramo habitat (Fig. 5, Table 3), which is in agreement with the literature (Berry and Calvo, 1989; Cuatrecasas, 2013). In fact, high-elevation species developed fewer but larger capitula by increasing the number of flowers per capitulum (Cuatrecasas, 2013). In addition, species growing in high altitudes had drooping to pendulous capitula with thickly lanate involucres, which can keep the flowers and their pollen dry and facilitates their dissemination (Cuatrecasas, 2013). Such a transition in the capitulum morphology could represent an adaptive shift from insect to wind pollination in Espeletiinae. In particular the low pollinator availability at high elevations in the Andes (Berry and Calvo, 1989), the need to protect capitula from snow or daily frosts, and the extremely long flowering periods of individual heads (Berry and Calvo, 1989; Cuatrecasas, 2013) could have selected for larger capitula containing more flowers. The production of more flowers in wind-pollinated species appears to be cheaper in biomass allocation than those of animal-pollinated species (Cruden and Lyon, 1985), and should maximize fitness of these species (Friedman and Barrett, 2011). In addition, pollen grains of the high-elevated species are more numerous and have shorter spines in comparison to the pollen of the insect-pollinated species, which is a common syndrome of wind pollination in Asteroideae and Asteraceae (Cuatrecasas, 2013).

Both vegetative and reproductive transitions appeared at different moments during the evolution of Espeletiinae. Indeed, Pagel's delta value was  $< 1$  for the vegetative trait and  $> 1$  for the reproductive traits (Table 5). This finding means that evolution of the vegetative trait was concentrated earlier in the phylogenetic tree than the evolution of the reproductive traits. This was consistent with the Grant's theory (1949), for which vegetative diversification appear earlier than reproductive one in the diversification of species pollinating by abiotic (such as wind or water) and biotic generalist pollinators. In addition, such morphological transition was both correlated to speciation rate during the evolution of Espeletiinae. Our study showed that the transition from the branched to the unbranched rosette growth-form increased diversification rate in Espeletiinae (Table 4, Fig. 6). The evolution of rosette probably constituted a "key innovation" at the beginning of the radiation of Espeletiinae, in terms of its adaptive values (Pouchon et al., 2018). Here, according to these results, the evolutionary transition from this form to an unbranched form constituted a secondary

innovation for species, evolving in response to ecological opportunities afforded by the super/páramo habitats, in terms of increasing diversification rates (Donoghue and Sanderson, 2015). In addition, our results showed that the transition from branched rosette to tree forms appeared three times with the colonisation of lower elevations and the forest-páramo ecotones (Fig. 5, Table 4). If such evolution is not correlated to a diversification shift, the tree-form transition constitutes also an inner radiation within the major one, by the mean of a repeated adaptive gain in lower habitats. Slowdown in diversification rates of lineages through time appeared to be synchronized with a slowdown in diversification rate of this vegetative trait enabling Espeletiinae to exploit habitat opportunity, which progressively tended to be filled overtime. Our study also showed that pollinating syndrome related to an abiotic pollinating shift promoted the diversification of the species (Table 4, Fig. 6). This was consistent to the Grant-Stebbins model in which pollinator shifts trigger speciation (Johnson, 2010). It thus seems that the superpáramo habitat constituted a secondary opportunity for rosette species to diversify rapidly with the development of adaptive reproductive traits shaped through the wind-pollination strategy afforded by these high-elevated habitats. The radiation of Espeletiinae seemed consequently to be following two different evolutionary trajectories: an upward and a downward migration both depending to the ecological opportunities.

During adaptive radiation, the evolution of key innovation allows species to occupy a new range of environments and the subsequent divergence of other traits for these species evolving in such innovation allows an ecological differentiation in these new environments (Losos and Mahler, 2010; Schluter, 2000). This study seemed confirmed such expectation by the mean of the evolution of the rosette allowing the colonisation of the páramos and the ulterior divergence in reproductive traits for these rosette allowing ecological differentiation in higher elevations. In addition, the repeated evolution of the rosettes to trees allowed species to win the forest-páramo ecotones. Natural selection has clearly favoured the morphological diversification in Espeletiinae however further investigations are necessary to identify signals of ecological speciation during the radiation.

## 5 | CONCLUSION

In a recent debate on the drivers of rapid evolutionary radiations in the high-altitude flora of the Andes, some authors considered the Andean uplift as the main process leading to species diversification by providing new ecological opportunities, promoting habitat heterogeneity, and acting as both an intra and inter Andean vicariant barrier (Antonelli and Sanmartín, 2011a; Rull, 2011a, 2011b). In addition, glacial dynamics of the Pleistocene favouring repeated cycles of connection and isolation of populations was proposed to be responsible for most of the páramo's diversity (Madriñán et al., 2013). Nevertheless, a current consensus views that rapid diversification appeared not depending on single factor but resulting from the synchronous action of biotic and abiotic factors (Antonelli and Sanmartín, 2011a; Bouchenak-Khelladi et al., 2015; Donoghue and Sanderson, 2015; Lagomarsino et al., 2016). Mountain uplift would clearly lead to diversification, but probably simultaneously with shifts between pollination syndromes (Kay et al., 2005; Lagomarsino et al., 2016, 2017) or with the

joint action of allopatric, diversifying selection, adaptation (Drummond et al., 2012b; Nürk et al., 2013) and paleotemperatures (Lagomarsino et al., 2016). This study supported this trend for the Venezuelan radiation of Espeletiinae into páramos, which appeared dependent on ecological opportunities occurring at different temporal scales and appearing progressively filled during their evolution: (1) the habitat opportunity provided by the mountains uplift allowed first Espeletiinae to radiate by the evolution of the rosette and its unbranched form highly adaptive to the páramos; following recently by a (2) pollination syndrome from insect to wind dispersal with the evolution of adaptive floral traits in association to the opportunity of colonizing the highest elevation habitats of the superpáramos. Moreover, in parallel to this upward migration, the radiation of Espeletiinae also arise from a downward migration by the evolution of rosette species (3) repeatedly in tree growth-forms in association to the opportunity of colonizing the lower elevation habitats of the forest-páramo ecotones. Finally, others factors probably acted synchronously during the diversification of Venezuelan Espeletiinae: (4) paleotemperatures cooling seemed negatively influence the speciation rate and (5) geography probably shaped some of speciation events during the radiation according to the geographical consistency of some of the phylogenetic relationships. Further studies are consequently required to untangle the intertwined biotic and abiotic factors at the origin of Espeletiinae, and to determine if same processes may explain the higher specific diversity of the Colombian radiation.

## 6 | SUPPLEMENTARY INFORMATIONS

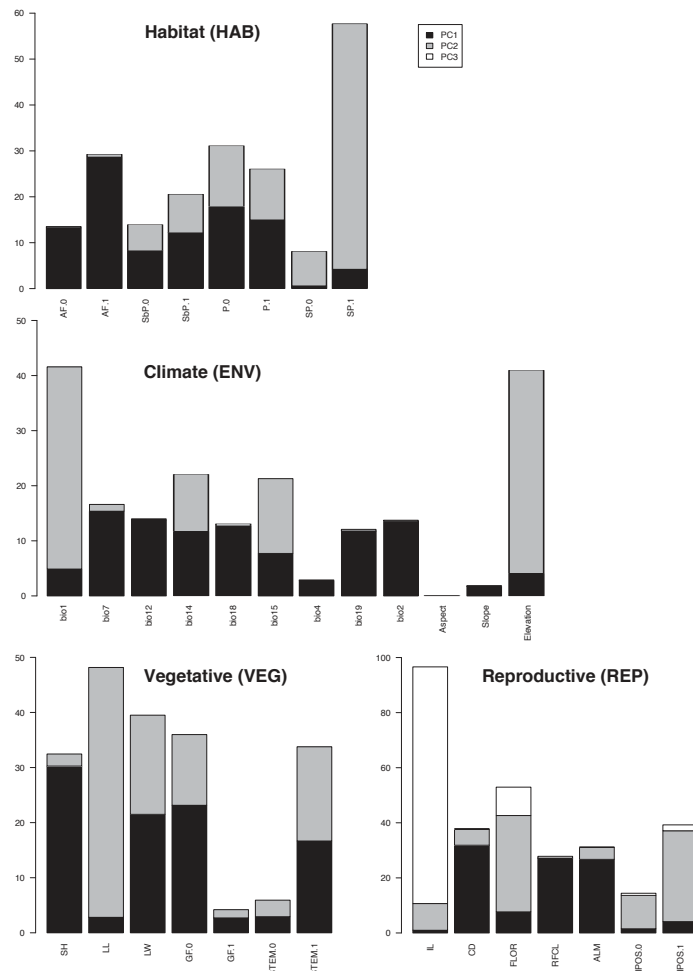
**Supp. Table 1.** Climatic variables used for estimation of climatic niche for species using the Outlying-Mean Index (OMI) method described in Doledec et al. (2000). Variables were extracted from WorldClim database for the region of interest.

WorldClim bioclimatic variables	
bio01	Annual Mean Temperature
bio02	Mean Diurnal Range
bio03	Isothermality
bio04	Temperature Seasonality
bio05	Max Temperature of Warmest Month
bio06	Min Temperature of Coldest Month
bio07	Temperature Annual Range
bio08	Mean Temperature of Wettest Quarter
bio09	Mean Temperature of Driest Quarter
bio10	Mean Temperature of Warmest Quarter
bio11	Mean Temperature of Coldest Quarter
bio12	Annual Precipitation
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality
bio16	Precipitation of Wettest Quarter
bio17	Precipitation of Driest Quarter
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter

**Supp. Table 2.** Diversification of Espeletiinae under time-dependent and environmental-dependent models. Time-dependent models refers to a Pure-Birth (PB) model with constant (CST), linear (LIN) or exponential variation of speciation rate. Environmental-dependent models included exponential and linear variation of speciation rate (ENVLIN\_L and ENVEXP\_L) only and of extinction rate under a constant rate of speciation (ENVLIN\_M and ENVEXP\_M). Models were fitted with the R package RPANDA. Environmental data refer to the global estimation of temperature surface estimated by Zachos et al. (2001).

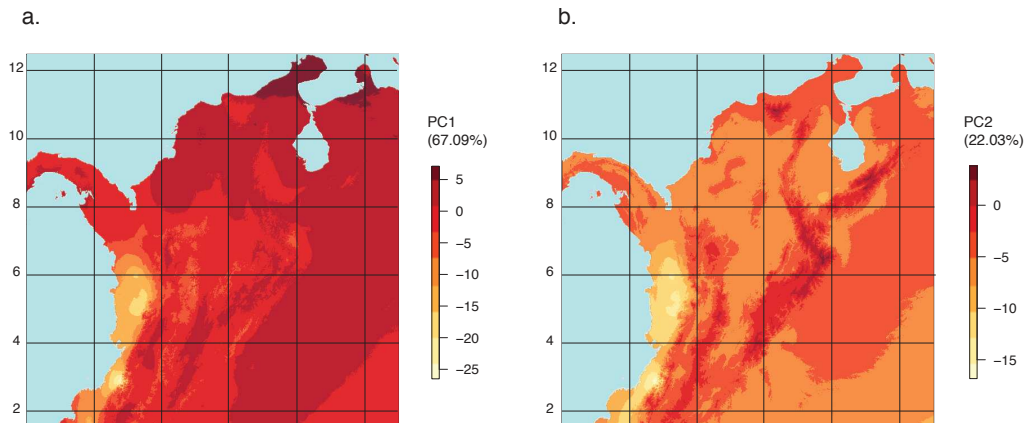
	lambda	mu	loglik	AICc	$\Delta$ AICc
PBCST	1.813	0.000	-20.84	43.77	25.22
PBEXP	0.905; 1.623	0.000	-14.02	32.29	13.75
<b>PBLIN</b>	-0.253; 5.774	0.000	-7.146	18.54	<b>0.000</b>
ENVEXP_L	4.98e-06; 1.075	0.000	-17.22	38.70	20.15
ENVLIN_L	-80.96; 6.979	0.000	-11.62	27.50	8.955
ENVEXP_M	1.813	-0.884; -0.800	-20.84	48.20	29.66
ENVLIN_M	1.819	-2.012; 0.169	-20.92	48.36	29.82

**Supp. Figure 1.** Contribution of both habitat (HAB) and climatic (ENV) variables into the environmental space, and of both vegetative (VEG) and reproductive (REP) trait into the morphological space. Components are represented by black (PC1), grey (PC2) and white (PC3) colours.

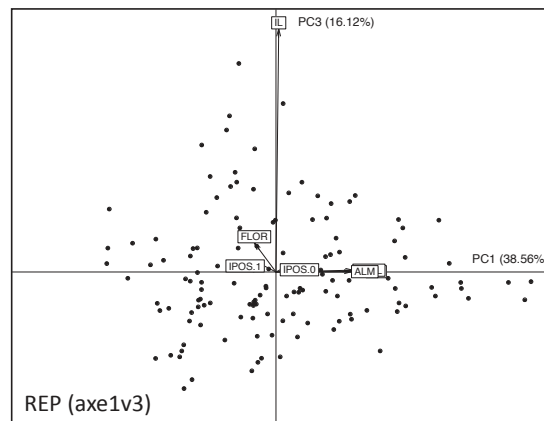




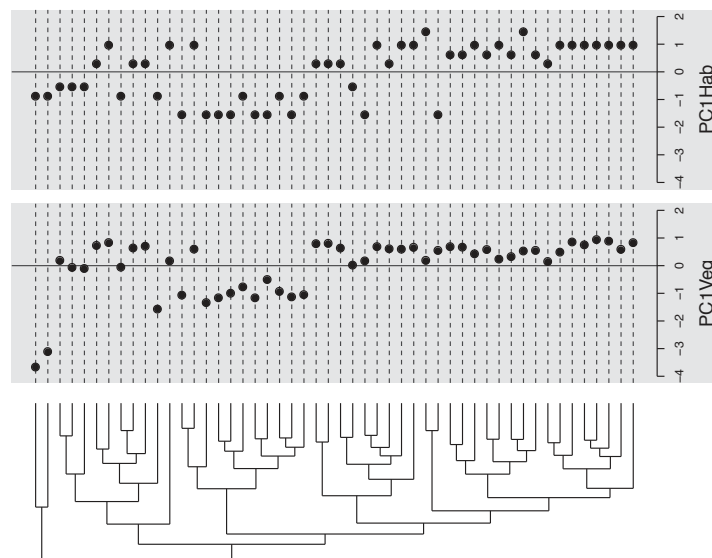
**Supp. Figure 2.** Projections of climatic space estimated by the Outlying-Mean Index approach of Dolédec et al. (2000) for the (a) first (PC1) and the (b.) second (PC2) components.



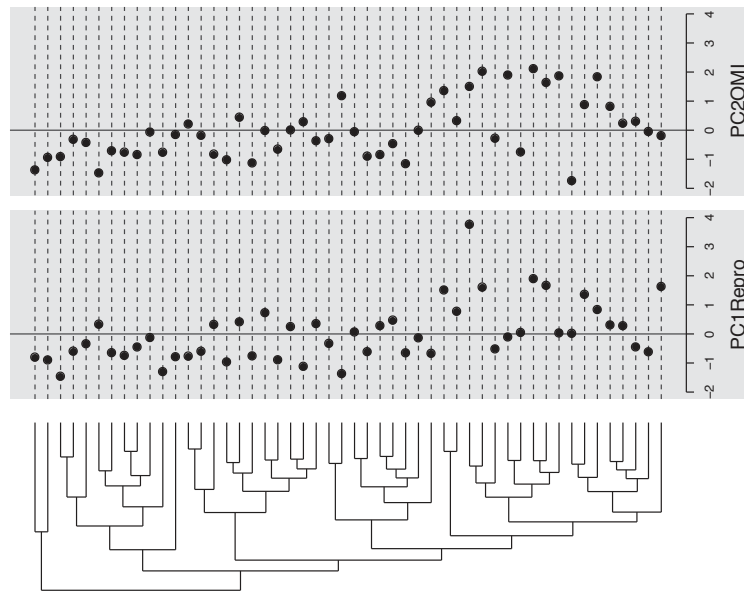
**Supp. Figure 3.** Projection of reproductive space occupied by species for the first and the third components.



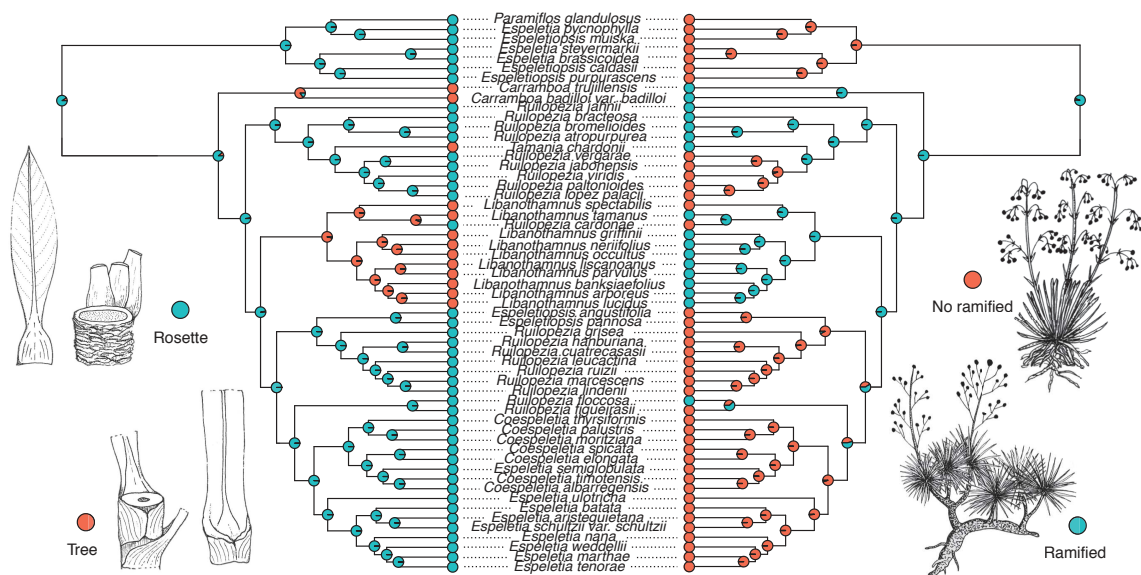
**Supp. Figure 4.** Species level phylogeny of Espeletiinae showing vegetative (PC1VEG) and habitat (PC1HAB) position of species into the morphological and environmental corresponding space.



**Supp. Figure 5.** Species level phylogeny of Espeletiinae showing reproductive (PC1Repro) and climatic (PC2OMI) position of species into the morphological and environmental corresponding space.



**Supp. Figure 6.** Reconstruction of ancestral states of the growth form trait (GF) and the ramification capacity (STEM) in Espeletiinae. Inferences were done using Markov models (Mk) in the R package *Diversitree* (FitzJohn 2012). Two possible discrete Markov models were fitted and compared using AIC values: a full model with different transition rates (ARD) and a constrained model with equal transition rates (ER). The best model (ER for GF and ARD for the STEM traits) was used to run simulations of stochastic character mapping in order to infer most likely scenarios of trait evolution. Illustrations by Christophe Perrier.



---

## CHAPTER 3 – Genomic footprints of hybridisation in the evolution of Espeletiinae

---

*“The importance of the fact that hybrids are very generally sterile, has, I think, been much underrated by some late writers. On the theory of natural selection the case is especially important as the sterility of hybrids could not possibly be of any advantage to them, and therefore could not have been acquired by the continued preservation of successive profitable degrees of sterility.”*

Darwin, 1859

### 1 | INTRODUCTION

There has been a persistent debate in regards with the role of hybridisation during adaptation and evolution of species. Despite the fact that as many as 25% of plant and 10% of animal species are known to hybridize (Mallet, 2005), this process is still a matter of controversy (Abbott et al., 2013; Schwenk et al., 2008). Since Darwin, many biologists have considered hybridisation as an unimportant evolutionary process or even as a dead-end because of the low fitness and sterility of hybrid populations (Schwenk et al., 2008). Nevertheless, during the last twenty years, several advances in biology have highly facilitated the detection of hybridization events and the process is now widely recognized as a major evolutionary force, able to facilitate adaptation of populations through genetic exchanges between populations and to create new species via hybrid speciation (Mallet, 2007; Mavárez and Linares, 2008; Schumer et al., 2015). Possible outcomes of hybridisation will depend on the genetic constitution, spatial distribution, ecology and demographic history of the species involved (Abbott et al., 2013).

One of outcomes of hybridisation is the advent of genetic, phenotypic and adaptive novelties through (1) adaptive introgression, and (2) hybrid recombination. The phenomenon of (1) adaptive introgression corresponds to the transfer of adaptive alleles between two species or only one (i.e. asymmetric introgression) by the hybrid offspring in the parental contact zones (i.e. hybrid zones) (Abbott et al., 2013; Hedrick, 2013). Such alleles providing advantageous traits should introgressed more frequently than alleles leading to genetic incompatibilities and reproductive isolation and could interact in favourable gene combinations in the genetic background of these species accelerating their adaptation (Abbott et al., 2013; Barton, 2001; Hedrick, 2013; Mallet, 2005; Pardo-Diaz et al., 2012). Adaptive introgression has now widely been documented in animals (Figueiró et al., 2017; Hedrick, 2013; Pardo-Diaz et al., 2012; Richards and Martin, 2017) and in plants (Martin et al., 2006; Whitney et al., 2010). For example, gene flow between *Helianthus annuus* and *H. debilis* has increased herbivore resistance in the local introgressed population of *H. annuus* from *H. debilis* (Scascitelli et al., 2010; Whitney et al., 2010), and introgressed alleles of *Iris fulva* in the genome of *I. brevicaulis* have increased the capacity of the latter to survive to flooding (Martin et al., 2006).

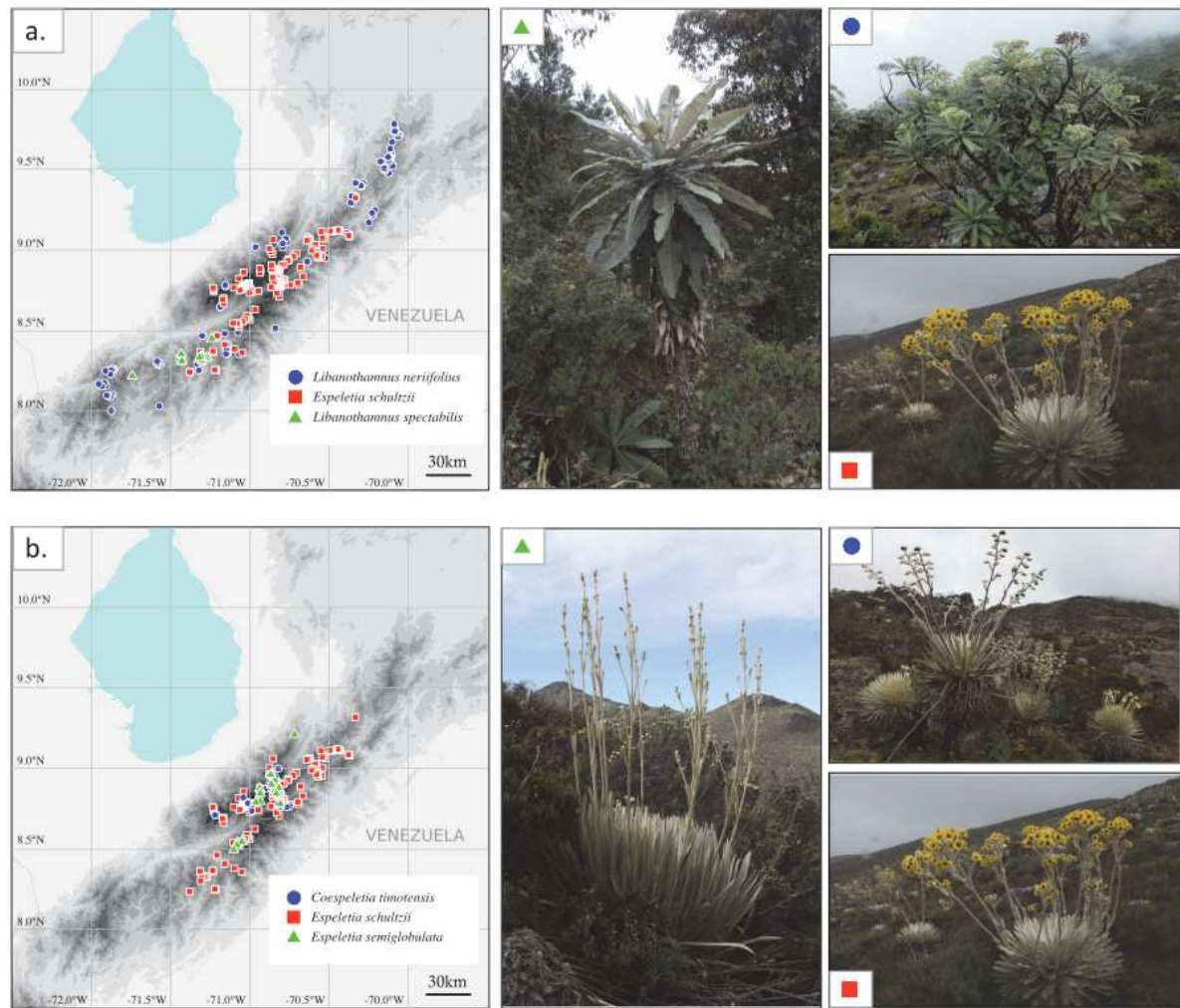
Hybridisation may also generate adaptive novelties in hybrid populations through (2) genomic recombination leading to mosaic of parental traits, intermediate trait or transgressive phenotypes in hybrids (Gross and Rieseberg, 2005). Transgression refers to the production of "extreme" phenotypic traits in hybrid populations exceeding the range of parental phenotypic values (Rieseberg et al., 1999). This process appears mainly as the result of the complementary action of genes (i.e segregation of variance), caused by the interaction of genes with multiple QTL fixed in parents with antagonistic alleles (Mao et al., 2011). Alleles have additive effect during recombination in the genome of the hybrids. In addition, as parental phenotypes are usually restricted to certain values because of the linkage of different QTL alleles for the same trait, hybrids can recombine +/- alleles causing extreme phenotypes (Stelkens and Seehausen, 2009). This process could easily improve fitness in hybrids in comparison to their parents allowing for novel niche adaptation (Mallet, 2007; Seehausen, 2004). Moreover, mixed phenotype and mosaic of parental traits in hybrids could also promote adaptive novelties (Elgvin et al., 2017; Keller et al., 2013; Mavárez et al., 2006). For example, mix of parental wing colored bands involved in mimicry and sexual selection provided new niche adaptation in hybrid lineages of *Heliconius heurippa* (Mavárez et al., 2006). More recently, Elgvin et al. (2017) highlighted that the combination of mosaic parental inheritance and novel divergence within the hybrid lineages has facilitated the origin and maintenance of the Italian sparrow.

Another outcome of the hybridisation is the formation of new species through "hybrid speciation". The emergence of hybrid species could arise from allopolyploid or homoploid hybrid speciation, depending on genomic divergence between hybridizing species (Chapman and Burke, 2007). Under this process, hybrid genomes must be genetically stabilized and reproductively isolated from parents to overcome genetic swamping (i.e. the loss of species identity) and competition (Abbott et al., 2013; Taylor et al., 2006). In allopolyploid speciation, change on ploidy in hybrids, arising for example from the fusion of unreduced gametes, lead to complete and immediate reproductive isolation from their parents through genomic incompatibilities (Mallet, 2007). This process is more common in plants (Barker et

al., 2016), and potentially the more common form of hybrid speciation (Soltis and Soltis, 2009). At the opposite, homoploid hybrids, which are not directly isolated from their parental species, can escape competition by the mean of the natural divergent selection through ecological differentiation into new niches (Buerkle et al., 2000). Indeed, phenotypic novelties, through recombination in hybrid genome, could allow hybrid individuals to colonize and adapt to new ecological niches, unexploited or marginally exploited by parental populations, in which case reproductive isolation results from habitat segregation (Dittrich-Reed and Fitzpatrick, 2013; Seehausen, 2004). For example, different trait mixture in hybrid lineages of cichlids allowed them to use different type of micro-niche slightly different from parental ones (Keller et al., 2013). In addition, hybrids lineages can also become reproductively isolated from parental lineages through geographic isolation without any change of niche preferences, thus conducting to allopatric speciation (James and Abbott, 2005), or by the mean of sexual selection through behaviour processes (Mavárez et al., 2006; Melo et al., 2009; Selz et al., 2014).

As a potentially strong engine of adaptive novelties in species, hybridisation could play a major role at the (1) origin and/or (2) during adaptive radiations (Seehausen, 2004). During the colonisation of new habitats, hybridisation is relatively common through secondary contact between allopatric lineages, and selection against the hybrids appears relatively weak (Abbott et al., 2013). In such context, hybridisation can lead to the fusion of parental lineages into a single lineage in extensive interbreeding zone between hybrids and parents, called “hybrid swarm” (Roy et al., 2015). By elevating rapidly the genetic variance of ecological and functional traits in a new landscape comprising underutilized adaptive peaks, such hybrid swarm can therefore (1) promote the entire adaptive radiation of species (Seehausen, 2004). Furthermore, the “syngameon hypothesis” predicts that hybridisation could also be a determinant process during these radiations (Seehausen, 2004). A syngameon refers to a complex set of hybrid combinations between closely related species (Lotsy, 1925). In such case, few species will dominate the pattern of hybridization. Occasional or locally distributed hybridisation events among species may accelerate rates of diversification in adaptive radiations by providing evolutionary novelties in both hybrid and parental lineages, suitable for the occupancy of untapped niches (Seehausen 2013). However, during these radiations, as theory predicts an increase in the number of species as well as a decrease in available resources over time, the probability that hybrid genomes will find new niches will consequently decrease. According to Seehausen (2004), the diversification of lineages during adaptive radiation will then depend not only on the ecological opportunities and the available genetic variation, but also on the ability of the divergent species to hybridize. Good examples of the role of hybridization during adaptive radiation, including introgression of adaptive traits, transgressive segregation and hybrid speciation, have been demonstrated in Galapagos finches (Lamichhaney et al., 2015), *Heliconius* butterflies (Pardo-Diaz et al., 2012), Hawaiian locusts (Shaw, 2002). The theory of “hybrid swarm origin of radiation” was also demonstrated in the Hawaiian silverswords (Barrier et al., 1999) and mints (Lindqvist et al., 2003), in Alpine whitefish (Hudson et al., 2011) and in African cichlids radiation (Joyce et al., 2011; Meier et al., 2017).

Hybridisation appears relatively prevalent in the high-elevation Andean plant in *Diplostephium* (Vargas et al., 2017), *Hypochaeris* (Tremetsberger et al., 2006), *Lachemilla*



**Figure 1.** Distribution of the two putative hybrids species (a.) *Libanothamnus spectabilis* and (b.) *Espeletia semiglobulata* with their potential parental species *L. neriifolius* and *E. schultzei* for the (a.) and *Coespeletia timotensis* and *E. schultzei* for the (b.). Photos by Serge Aubert.

(Morales-Briones et al., 2018), *Lupinus* (Nevado et al., 2018), *Polylepis* (Schmidt-Lebuhn et al., 2010), *Puya* (Jabaily and Sytsma, 2013), *Senecio* (Dušková et al., 2017) and especially in Espeletiinae (Cuatrecasas, 2013; Diazgranados, 2012; Pouchon et al., 2018). In this last group, if 33 hybrids were recognized (Diazgranados 2012), the frequency of hybridization events is extensive in the field (Diazgranados, 2012; Mavarez personal observations). However, a detailed analysis of the role of hybridisation on patterns of variation for genetic, morphologic and ecologic data during the adaptive radiation of Espeletiinae has never been accomplished. According to Seehausen (2004), evidence that introgression frequently occurred at the base of the radiation could be a proof of “hybrid swarm origin”. Nevertheless, significant evidence in support for gene flow events between closely related and sympatric species in Espeletiinae has been proposed recently by Pouchon et al. (2018). If hybridisation was probably not at the origin of this radiation, this process could play an active role during the radiation in the important morphological variation observed between species known to hybridize frequently, e.g. *E. schultzei* and *C. spicata* (Berry et al., 1988; Cuatrecasas, 2013). Berry et al. (1988) have shown that some inter-specific hybrids in Venezuelan *Espeletia* (e.g.

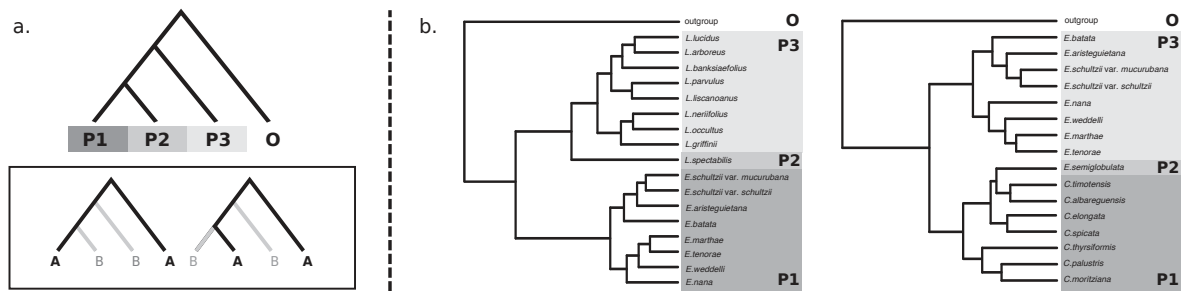
between *E. batata* and *E. schultzei*) closely resemble other well-established and phenotypically stable species (e.g. *E. weddellii*), suggesting that the latter could have had a hybrid origin. Interestingly, some of the species showing significant evidence for genetic introgression also appear to exhibit intermediate morphologies (Pouchon et al., 2018). For example, *Libanothamnus spectabilis* has the characteristic tubular leaf bases and terminal inflorescences of trees in clade V, but it exhibits a caulescent rosette growth-form typical of the rosettes in clade VII (Fig.1a). According to previous findings in phylogenetic analysis performed in SNaQ (see Chapter 1), this species appeared also to share genes between species of the clade V and VII, which could indicate a hybrid origin of this species (Pouchon et al., 2018). Another example is *Espeletia semiglobulata*, which appears phylogenetically close to other rosettes of clade VIII, a group with whom it shares several distinctive morphological traits such as the long and slender leaves, and semiglobose pending/nodding capitula with short ray ligules and small pollen (Cuatrecasas, 2013). Yet *E. semiglobulata* exhibits a clearly dichasial inflorescence, typical of members of rosettes in clade VII.

Here, we attempt to identify molecular signals of hybridisation and introgression in *L. spectabilis* and *E. semiglobulata* using sets of Single Nucleotide Polymorphism (SNP) markers gathered with Restriction-site Associated DNA (RAD) and shotgun sequencing. In light of advances in sequencing techniques, these genome-reduction methods appeared powerful to investigate such patterns. For example, the use of RAD-sequencing has recently allowed investigation of hybrid speciation in *Coenonympha* butterflies (Capblancq et al., 2015), cutthroat trouts (Hohenlohe et al., 2013), cichlid fishes (Meier et al., 2017), *Pedicularis* louseworts (Eaton and Ree, 2013) and páramos plants in the genus *Diplostephium* (Vargas et al., 2017). In Espeletiinae, we used RAD-sequencing to estimate phylogenetic relationships in association with the whole genome shotgun sequencing described in Chapter 1 (see Appendix 1). Both methods appeared relatively consistent at shallower phylogenetic scale (Appendix 1). In this study, as putative hybridisation concerns relatively closed species (Pouchon et al. 2018), we used both of these genome-reduction methods to assess potential footprint of hybridisation based on statistical methods on SNP patterns in taxa alignments. The shotgun approach allows us to increase the marker density on a reduced sample of individuals, while RAD-sequencing, which is less expensive, allows us to increase the sample size of individuals on a reduced number of markers. In addition, the climatic envelopes used by putative hybrid and parental species were also assessed in order to depict the possible impact of hybrid speciation on the ecological divergence of new taxa.

## 2 | METHODS

### Sampling, Sequencing and Data Acquisition

Sampling focused on the two putative hybrid species and their potential parental species according to the previous phylogenetic analyses (Pouchon et al., 2018): polycarpic rosettes with opposite inflorescences in clade VII and trees in clade V for *L. spectabilis*, and rosettes in clades VII and VIII for *E. semiglobulata*. In addition, a new population of *L. spectabilis*



**Figure 2.** a) Schematic description of the Patterson's test. This D-statistic based on a 4-taxon topology ((P1,P2),P3),O) measures incongruent alleles patterns on SNP which are in ABBA and BABA configurations. In the absence of introgression, the counts of ABBA and BABA SNPs are expected to be equals. In opposition, this symmetry deviates from zero in the presence of gene flows between P3 with P1 or P2. b) This test was first made in this study for *Libanothamnus spectabilis* (in P2) by considering all Venezuelan *Libanothamnus* as P3 and all the Venezuelan *Espeletia* as P1. In second, D-statistic was performed for *E. semiglobulata* (in P2) by considering all Venezuelan *Espeletia* (P3) and *Coespeletia* (P1). Phylogenetic a priori was considering according to the previous phylogenies made in the Chapter 2.

was found during our last field mission, located out of the previously known distribution of this species. Individuals in this population differ morphologically from *L. spectabilis* and are treated here as *L. aff. spectabilis*. One individual per species was used in following shotgun approach whereas sampling was improved for RAD-sequencing by adding more individuals (Supp. Table 1).

Genomic DNA was extracted from dried leaves tissue, after shaving the tissue part of interest and grinding them with the TissueLyser (Qiagen Corp.), during 3h for the cell lysis step to preserve the DNA quality by using the DNeasy Plant Mini Kit (Qiagen Corp). However, according to previous experiments, DNA contained high level of inhibitors for enzyme activity (i.e. terpenoids), which prevent restriction digestion of RAD-sequencing protocol (also underlined in Diazgranados and Barber 2017). Following Mastretta-Yanes et al. (2015) recommendations, DNA cleaning was then performed using phenol chloroform before the preparation of RAD libraries to discard any enzymatic inhibitors.

**RAD libraries** – A double-digested RAD (ddRAD) sequencing was performed according to a modified version of the original protocol described in Peterson et al. (2012). Briefly, restriction digestion was done first using 10 units of PstI and MspI restriction endonucleases (NewEngland Biolabs, Inc.) during 3 h at 37°C with 250 ng of purified genomic DNA in a final volume of 36  $\mu$ L using the CutSmart buffer provided with the enzymes. Adapters containing sample-specific barcodes and TruSeq Illumina primers (Illumina Inc.) were next ligated to the DNA thanks to 10 units of T4 DNA ligase (NewEngland Biolabs, Inc), 1  $\mu$ L of 10 mM ribo-ATP (NewEngland Biolabs, Inc). Digestion was continued simultaneous to the ligation step by adding 0.5  $\mu$ L of each digestion restriction enzymes and performed under 60 cycles of 2 min at 16° for the ligation and 37°C for the digestion. Enzymes were next heat-inactivated at 65 °C for 10 min. Samples were then pooled in equivolume mixture and purified with Agencourt AMPure XP beads (Beckman Coulter, France). Fragments in the



250- to 500 bp size range, including adapters, were excised from agarose gel (1.6%) and purified using the QIAquick Gel Extraction Kit (Qiagen Corp.). Libraries were then PCR-amplified in 20  $\mu$ L final reaction volume with 1  $\mu$ L of DNA template, 10 mM of dNTPs, 10  $\mu$ M of each PCR probe (Peterson et al. 2012) and 2 U/  $\mu$ L of Taq Phusion-HF (New England Biolabs Inc.). PCR were performed with an initial denaturation at 98 °C for 10s followed by 15 cycles of 98°C for 10s, 68°C for 30s and 72°C for 1 min; and ended by a final extension at 72 °C for 10 min. Finally, products from each library were purified with with QIAgen MinElute PCR Purification Kit (Qiagen Corp.), controlled in agarose gel migration, quantified using Qubit (Life Technologies) and finally pooled in equimolar amounts. Paired-end 2x150bp sequencing was performed on Illumina HiSeq 2500 rapid run flow cell for each library at Fasteris (Geneva, Switzerland).

Clustering of loci was done using a *de novo* approach in pyRAD (Eaton, 2014), independently for the two putative hybrid taxa and their respective parental species. Reads were first clustered by sample according to a similarity threshold using VSEARCH v2.1.1 program (Rognes et al., 2016) and then aligned with MUSCLE v3.8.31 (Edgar, 2004). Sequencing error rates and heterozygosity were jointly estimated and used to generate consensus sequences for each cluster which were clustered across samples under the same procedure and parameters. We used a clustering thresholds of 85% following results from Appendix 1 and considered a threshold of 75% as the minimum frequency of samples that had to be shared at each locus. Loci sharing heterozygosity at a nucleotide site in 15% minimum of taxa were considered as paralogs and were discarded for the analysis. The minimum depth of coverage required to build a cluster was fixed at 10. Finally, of 12,538 SNPs were recovered for the *L. spectabilis* dataset and of 9,979 SNPs for the *E. semiglobulata* dataset.

**Shotgun libraries** – For the shotgun approach, we used DNA sequences of the taxa studied from previous analyses (Chapter 2). Following the same procedures described in chapters 1 and 2, only one individual per taxon of interest was considered. We also sequenced one individual of *L. aff. spectabilis* on an Illumina HiSeq 2000 at the Genoscope lab (Paris, France) following the manufacturer’s instructions. SNPs were then extracted for each alignment of the reads into the 9,880 metacontigs of reference reconstructed in chapter 1, using BWA (Li and Durbin, 2009). We retained only reads mapping to a single metacontig with a maximal MAPQ score of 60 for the variant calling using SAMtools and BCFtools (Li et al., 2009). SNPs were further filtered with vcfutils using a root-mean-square  $\geq 15$  and a minimum SNP coverage of 10 per marker and per species.

### **Patterson’s D-statistic**

In order to test the molecular evidence of admixture between lineages, we used the Patterson’s D-statistic (Green et al., 2010; Patterson et al., 2012), as implemented in the *evobiR* R package. This method identifies the relative abundances of two phylogenetically incongruent SNPs patterns of ancestral (A) and derived (B) alleles ABBA and BABA, given a 4-taxa topology *a priori* (((P1,P2),P3),O) (Durand et al., 2011) (Fig. 2a). Under a scenario of incomplete lineage sorting (ILS) and no gene flow, these incongruent patterns are expected to occur in equal frequencies. However, introgression could lead these two frequencies to differ

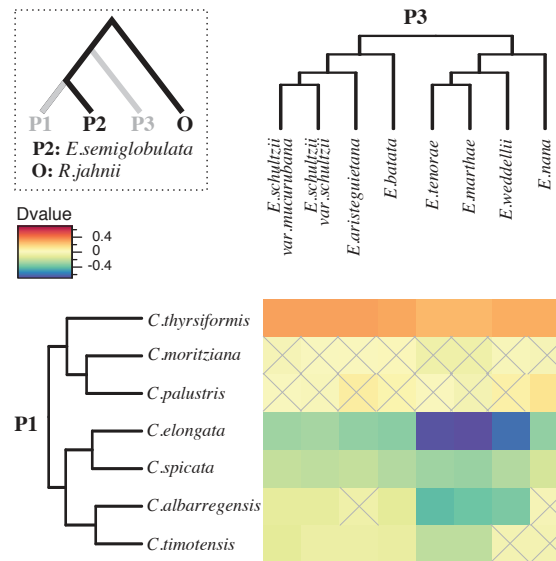
**Table 1.** Summary statistics of Patterson’D-test analyses performed on shotgun SNP for the putative hybrids *Libanothamnus spectabilis* and *Espeletia semiglobulata* according to their respective parents described in Fig. 2, and to three different outgroup species: *Espeletiopsis purpurascens*, *Carramboa trujillensis* and *E.Ruilopezia jahnii*. Mean and standard deviation were estimated from all tests made with each parental combination for the counts of SNP in ABBA and BABA configurations and for the D-values. Significance of tests is shown in Fig. 3 and 4.

P2	Outgroup species	D-value < 0.00			D-value > 0.00		
		ABBA (sd)	BABA (sd)	D (sd)	ABBA (sd)	BABA (sd)	D (sd)
<i>E. semiglobulata</i>	<i>E. purpurascens</i>	115.68 (37.55)	213.92 (51.66)	-0.30 (0.15)	244.22 (45.47)	188.11 (27.90)	0.12 (0.04)
	<i>C. trujillensis</i>	208.32 (52.02)	321.41 (68.56)	-0.20 (0.17)	388.29 (114.33)	286.47 (35.02)	0.13 (0.11)
	<i>R. jahnii</i>	232.51 (53.74)	354.48 (78.26)	-0.20 (0.17)	423.29 (122.60)	291.29 (34.00)	0.16 (0.11)
<i>L. spectabilis</i>	<i>E. purpurascens</i>	NA	NA	NA	453.47 (101.79)	205.33 (61.57)	0.37 (0.19)
	<i>C. trujillensis</i>	NA	NA	NA	821.55 (192.70)	331.14 (78.06)	0.41 (0.16)
	<i>R. jahnii</i>	NA	NA	NA	965.40 (223.95)	389.81 (89.05)	0.41 (0.16)
<i>L. aff. spectabilis</i>	<i>E. purpurascens</i>	NA	NA	NA	451.58 (101.76)	191.38 (43.49)	0.39 (0.16)
	<i>C. trujillensis</i>	NA	NA	NA	780.88 (181.52)	328.51 (60.58)	0.39 (0.14)
	<i>R. jahnii</i>	NA	NA	NA	X	X	X

significantly. Indeed, introgression between P2 and P3 will increase the frequency of the ABBA configuration but not the BABA one. Conversely, introgression between P1 and P3 will favour BABA in comparison to ABBA. This asymmetry between SNP patterns is quantified using the D-statistic by counting the number of SNPs in the genome matching an ABBA or a BABA configurations. D values significantly deviating from zero indicate an excess of shared derived polymorphisms either between P2 and P3 when  $D > 0$  or between P1 and P3 when  $D < 0$ , suggesting introgression between taxa.

In this study, we used results from previous phylogenetic studies (Chapters 1 and 2) to formulate hypotheses for putative hybrids. In the case of *L. spectabilis* (P2), we tested all tree species in clade V as P3 taxa and all rosette species in clade VII as P1 taxa (Fig. 2b). For *E. semiglobulata* (P2), we considered all rosette species in clade VIII as P1 taxa and all rosette species in clade VII as P3 (Fig. 2b). To assess the effect of outgroup choice on the analysis, we computed D-statistic tests using three different outgroup taxa (O) sampled at different phylogenetic scale according to previous results (Chapter 2): (O1) *Espeletiopsis purpurascens* (clade I), (O2) *Carramboa trujillensis* (clade II), and (O3) *Ruilopezia jahnii* (clade III). To avoid missing data, only SNPs present in complete quartet alignments and homozygous at the species level were considered. In addition, for ddRAD datasets, a strict consensus was made between intra-specific individuals to consider only one sample per species. The standard deviation of the D-statistic was obtained using 1,000 bootstrap iterations on the SNP sites of each alignment dataset, reduced to selected quartets of taxa. The significant deviations from zero were identified using a Z-score  $> 3$ , corresponding to a p-value of 0.01 after the Benjamini-Hochberg correction (Eaton and Ree, 2013).

## Genetic clustering analyses



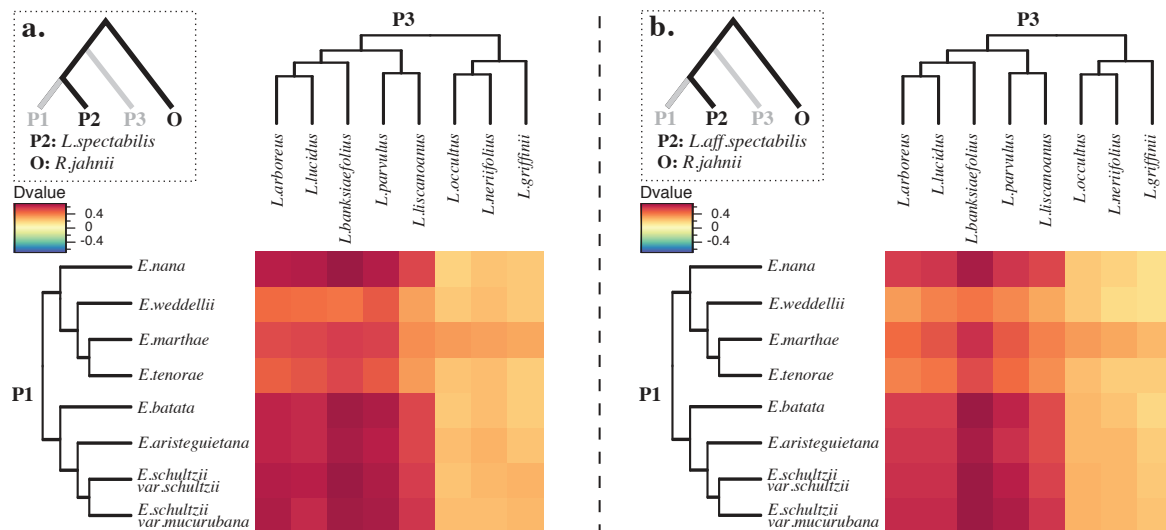
**Figure 3.** Patterson's D-statistic performed on shotgun SNPs considering a four topology  $((P1,P2),P3),O$  with *Ruilopezia jahnii* as O, the Venezuelan rosettes in clade VII as P3, the rosettes in clade VIII species as P3 and *E. semiglobulata* as P2. Crosses highlighted D values which were not statistically significant ( $> 0.05$ ).

Genetic clustering techniques have been shown to be useful to detect population admixture based on geometric multivariate analysis (Novembre and Stephens, 2008; Patterson et al., 2006) or probabilistic approaches (Alexander et al., 2009; Pritchard et al., 2000). In this study, genetic clustering of taxa was assessed using both approaches on the SNP dataset obtained with ddRAD-sequencing owing to population-level sampling for some species.

First, principal component analysis (PCA) was conducted on SNP datasets using *pca* function of the *LEA* R package (Frichot and François, 2015). This geometric method does not make any *a priori* assumptions on population. PCA scores were computed for each genotypic matrix, and eigenvalues of the sample covariance matrix were returned in a screeplot. Each component highlights different aspect of the genetic diversity of individuals. PCA reveals the structure of the population by maximizing the separation between individuals. Genetically close individuals will thus appear close to each other and clustered in plots according to the degree of allele sharing. The number of significant components was evaluated using a graphical method and the Tracy-Widom tests implemented in the *LEA* package.

Admixture analyses were also conducted on each genotypic matrix using the *snmf* function of the same package. In contrast to the PCA approach, this method assumes that K ancestral populations existed. This method estimates individual ancestry proportions by utilizing a sparse non-negative matrix factorisation algorithm to compute least-squares estimates of ancestry coefficients. We conducted analyses using 10 replicates for each number of ancestral populations (K) ranging between 1 and 10, on each genotypic matrix. The quality of fit of the statistical models to our genotypic data was estimated using an entropy criterion with cross-validation technique (Alexander and Lange, 2011; Frichot et al., 2014).

## Ecological divergence analyses



**Figure 4.** Patterson's D-statistic performed on shotgun SNPs considering a four topology (((P1,P2),P3),O) with *R. jahnii* as O, tree species in clade V as P3, rosette species in clade VII as P1 and (a.) *L. spectabilis* or (b.) *L. aff. spectabilis* as P2. All tests were statistical significant ( $p < 0.01$ ).

Occurrence data were collected for the two sets of putative hybrid and parental species from field observations and the Paramo Plants Online web database (<http://paramo.uniandes.edu.co>). We kept only one occurrence per grid cell in order to reduce over-representation of sampling localities. These occurrences were next associated with the 19 bioclimatic variables available in the WORLDCLIM database (<http://worldclim.org>; Hijmans et al., 2005) to evaluate environmental niches of species by using the 'Outlying Mean Index' (OMI; Dolédec et al., 2000). This approach gives the position and the breadth of each species niche along principal component axes calculated from a set of environmental variables. The niche position corresponds to the mean location of the species in the environmental space, and the breadth represents the variability of the used environment (SD of the species values in the OMI axes; Thuiller et al., 2004). The analysis was performed in R using the 'ade4' package (Dray and Dufour, 2007).

### 3 | RESULTS

#### D-statistic inferences

The 4-taxon test performed for *E. semiglobulata* using SNPs data recovered with the shotgun alignments did not show significant signals of introgression among examined species, regardless of the choice of the outgroup taxon (Table 1). Here, only results involving *Ruilopezia jahnii* as outgroup are shown (Fig. 3). Only the tests considering *C. thyrsoformis* as P1 showed a significant signal of introgression ( $p < 0.001$ ) with  $D > 0$ . In contrast, significant introgression was found between clade VIII (P1) and clade VII (P3) species (in particular for dwarf species *E. tenorae*, *E. marthae* and *E. weddellii*), with D values significantly  $< 0$ .

Patterson's D analyses performed on SNPs data recovered with ddRAD alignments were not congruent with the shotgun alignments due to a very low number of SNPs in ABBA or BABA configuration (Supp. Fig. 1, Supp. Table 2).

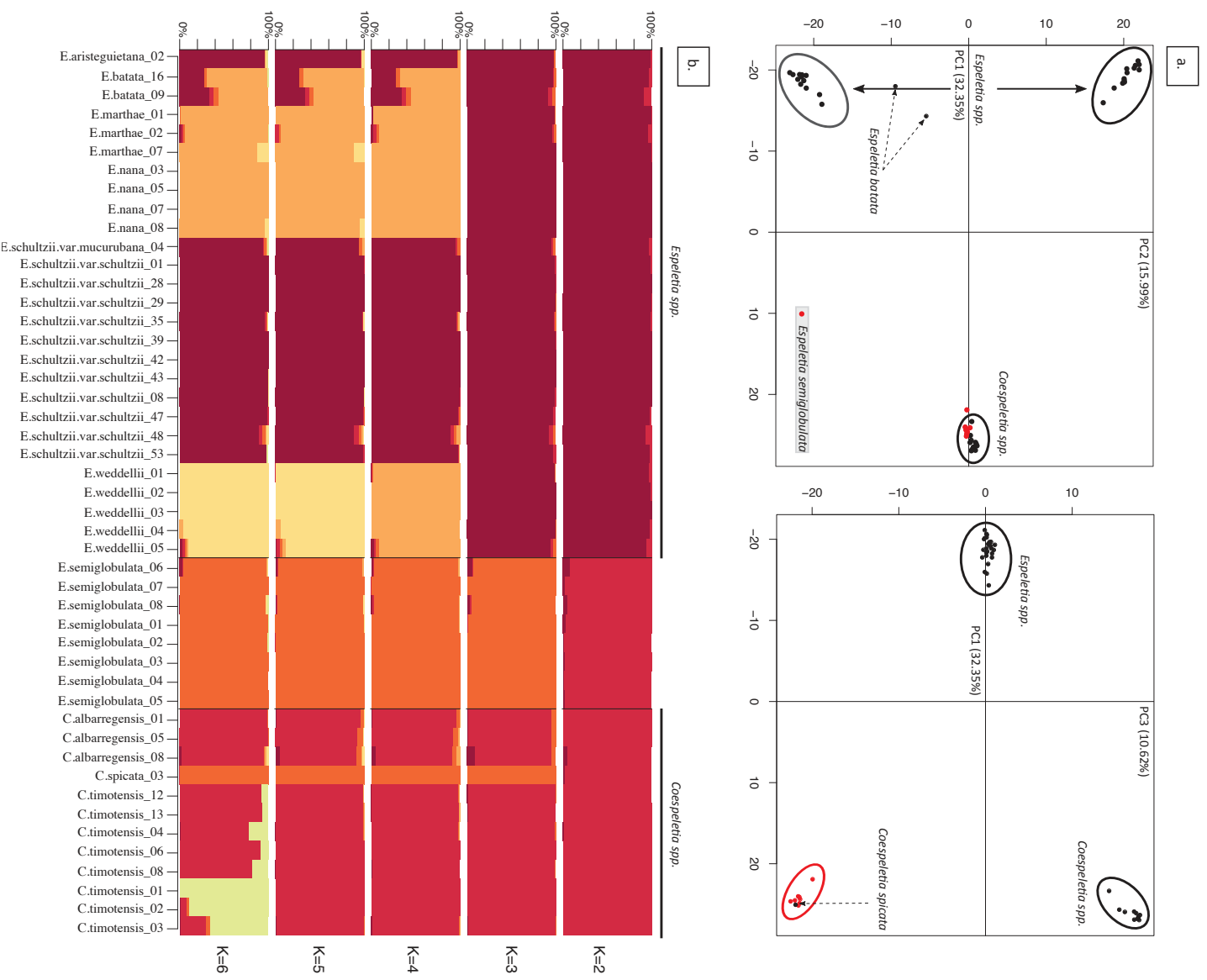
In the case of *L. spectabilis*, Patterson's D-statistic based on SNPs data recovered with the shotgun alignments showed significant levels of introgression among species regardless of the choice of outgroup, P1 or P3 species (Table 1). Significant positive values of D were in fact observed in all combinations of tree members in clade V and rosette members of clade VII as parental species, with some differences in the magnitude of D values depending on the species considered (Fig. 4a). The same pattern was observed for *L. aff. spectabilis* (Fig. 4b. and Table 1). Unfortunately, this analysis does not allow reaching a conclusion in regards with the parental species at the origin of either *L. spectabilis* or *L. aff. spectabilis*. Analyses of the SNPs data recovered with the ddRAD alignments showed a similar pattern in comparison with data from shotgun alignments for both putative hybrid species (Supp. Fig. 2, Supp. Table 2).

### Genetic structuration

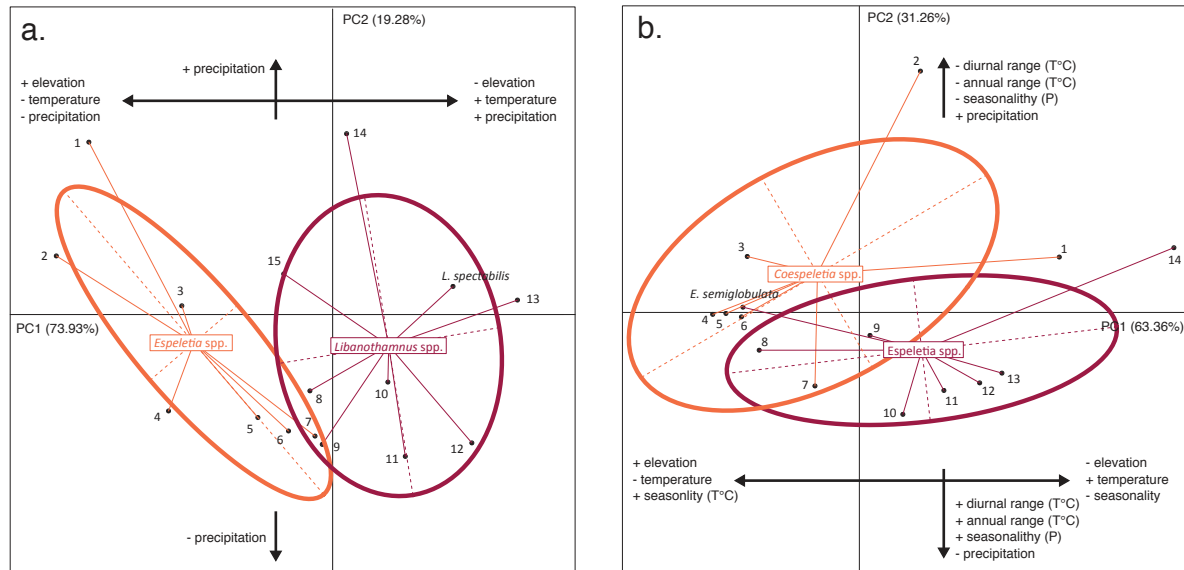
Genetic PCA made on ddRAD markers allowed the identification of three well-differentiated genetic clusters: (1) tree species in clade V, (2) rosette species in clade VII and (3) *L. spectabilis* and *L. aff. spectabilis* individuals (Fig. 5a.). The highest interspecific genetic differentiation appeared on the first principal component (PC1-34.97%) between trees in cluster 1 and rosettes in cluster 2, while the putative hybrid taxa in cluster 3 showed somewhat intermediate score. The second PCA component (PC2-11.29%) differentiated clusters 1 and 3, while the third PCA component (PC3-9.35%) discriminated large and dwarf rosettes in cluster 2. Admixture analyses showed highly similar patterns (Fig. 5b.). For instance, at K=2 the putative hybrid taxa appear admixed for rosettes and trees clusters, with a greater genetic contribution for the latter. When increasing to K=3, the rosette cluster is split into dwarf and large rosettes. For K=4 the putative hybrid taxa are split from the remaining tree species. At higher values of K, individual species appear (e.g. *E. weddellii* at K=5, *L. nerifolius* at K=6). Finally, no differentiation was found between the two populations of *L. spectabilis*. As for the PCA method, no conclusive assignation could be made on the specific identification of parental species contributing to the hybrid origin of *L. spectabilis*. Finally, the two individuals of *E. batata* appeared genetically admixed according to both multivariate and clustering analyses.

In the case of *E. semiglobulata*, genetic PCA discriminated between species in clade VII and VIII in the first axis (PC1-32.35%) (Fig. 6a.). As above, the two other axes differentiated respectively between dwarf and large rosettes in clade VII (PC2-15.99%) and between *E. semiglobulata* and other rosettes in clade VIII (on PC3-10.62%). These results of the admixture analyses reinforced these views; by showing that *E. semiglobulata* did not appear admixed for genetic polymorphisms associated with rosettes in clades VII and VIII (Fig. 6b.). Indeed, at K=2, rosettes in clades VII and VIII are well separated, with *E. semiglobulata* nested within the latter. At K=3, *E. semiglobulata* is split from other member of clade VIII. At K=4, the rosettes in clade VII are split into large (*E. schultzii* and *E. aristeguietana*) and dwarf forms (*E. nana*, *E. marthae* and *E. weddellii*), while the dwarf





## Ecological divergence



**Figure 7.** Ecological space between putative hybrids (a.) *Libanothamnus spectabilis* and (b) *Espeletia semiglobulata* with their respective parental species. Analyses were performed using the Outlying Mean Index (OMI) approach (Dolédéc et al., 2010). Only the mean niche position of each taxon is plotted. Species for the first analysis (a.) are encoded as: 1. *E. aristeiguetana*; 2. *E. batata*; 3. *E. schultzii* var. *schultzii*; 4. *E. marthae*; 5. *E. weddellii*; 6. *E. tenorae*; 7. *E. nana*; 8. *L. arboreus*; 9. *L. liscanoanus*; 10. *L. banksiaefolius*; 11. *L. parvulus*; 12. *L. griffinii*; 13. *L. neriifolius*; 14. *L. occultus*; 15. *L. lucidus*. For the second one (b.): 1. *Coespeletia elongata*; 2. *C. thyriformis*; 3. *C. albarregensis*; 4. *C. timotensis*; 5. *C. moritziana*; 6. *C. spicata*; 7. *C. palustris*; 8. *E. batata*; 9. *E. schultzii* var. *schultzii*; 10. *E. marthae*; 11. *E. weddellii*; 12. *E. tenorae*; 13. *E. nana*; 14. *E. aristeiguetana*.

The analysis of climatic niches of *L. spectabilis* and their putative parental species is shown in the Figure 7a. The first component, accounting for 73.93% of the ecological variance, discriminated between tree species in clade V and rosette species in clade VII according to a gradient of elevation, temperature and precipitation. As expected, the climatic niches of trees were associated with lower elevations and higher temperatures and precipitations than those of rosettes, corresponding to habitats of the upper of Andean forests and páramos, respectively. The second component (19.28% of the ecological variance), opposed species with climatic niches characterized by differences in precipitation. The position of *L. spectabilis* into the climatic space appeared nested within the space defined by climatic niches of trees, and quite similar to that of *L. neriifolius*.

Climatic niches of *E. semiglobulata* and their putative parental species are shown in the Figure 7b. The first axis (63.36% of the variance) opposed low- and high-elevation habitats, whereas the second one discriminated niches according to fluctuations in temperatures and precipitations. The climatic niche position of *E. semiglobulata* was clearly similar to these estimated for other rosettes in clade VIII (e.g. *C. timotensis*, *C. spicata*, *C. moritziana*), characterized by high elevations, and low temperatures and precipitations. Such conditions are typical of the superpáramo habitat in which these four species are almost exclusively found; sometimes in sympatry with *E. batata* and *E. schultzii* var. *schultzii* from clade VII.



## 4 | DISCUSSION

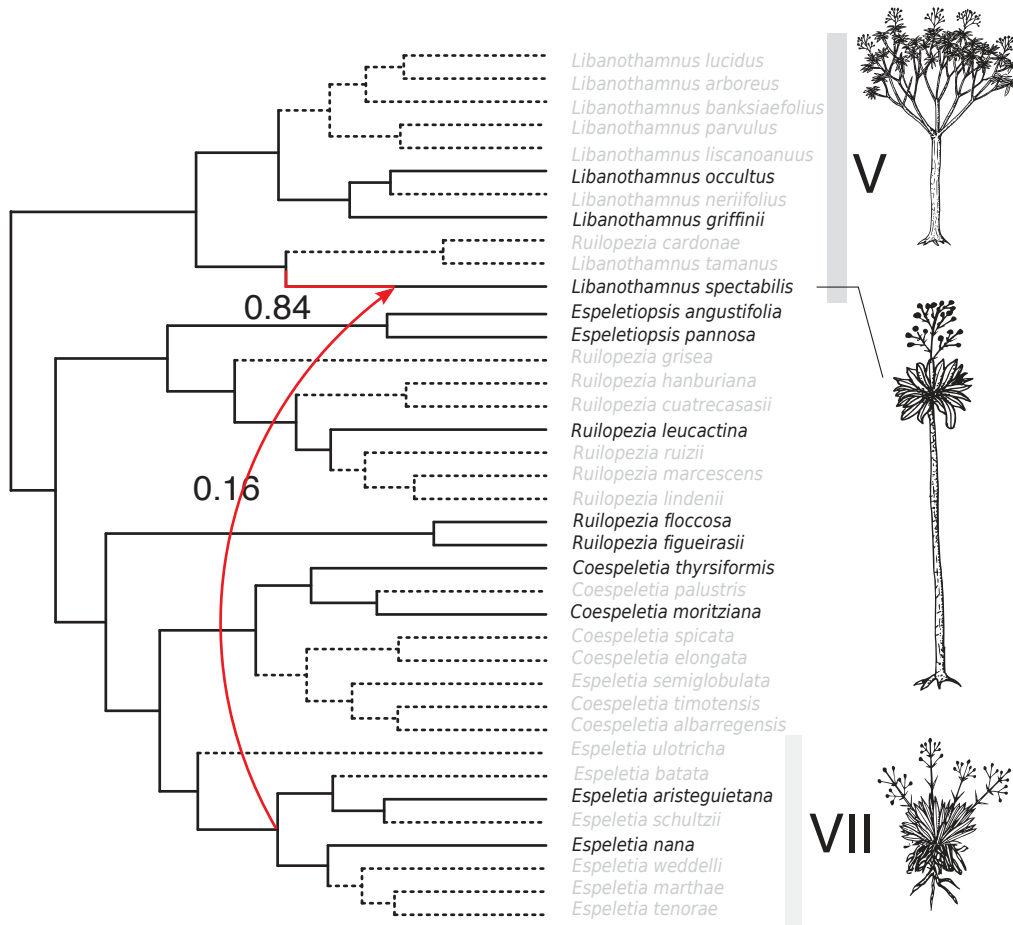
Since 1760 and its first description by Linnaeus until today, hybridisation has challenged the biological species concept (Mayr, 1942) and its role in evolution has been controversial (Mallet, 2005, 2007). The current consensus, however, views hybridisation as a force acting both positively and negatively in the evolution of species. Indeed, depending on the genetic, demographic and ecological backgrounds, hybridisation could erode genetic diversity and lead to de-speciation (Abbott et al., 2013), but could also promote genetic novelty, enhancing in some case adaptation and speciation (Abbott et al., 2013; Mallet, 2005; Mavárez and Linares, 2008). Such findings have been mainly facilitated by the advent of both technical improvements such as next generation sequencing, and new statistical methods that use genomic data to detect gene flow events between taxa (Goulet et al., 2017). In this chapter, by combining genomic data generated with both ddRADs and the shotgun sequencing described in Pouchon et al. (2018), we describe hybridisation and introgression processes in Espeletiinae that most certainly contributed to the evolution and diversity of this group.

### Hybrid speciation hypothesis in *L. spectabilis*

**Genomic footprints evidences** – Pouchon et al. (2018) hypothesized based on their results a putative hybrid origin of *Libanothamnus spectabilis*, which exhibits a unique rosette species with leaf sheaths typical of trees, according to its phylogenetic and morphological relationships and with trees in clade V and rosettes in clade VII.

In this study, significant Patterson's D values were obtained for both *L. spectabilis* and *L. aff. spectabilis* ( $D > 0.37$ ;  $p < 0.001$ ), suggesting gene flow events between trees in clade V and rosettes in clade VII, regardless of the outgroup choice and the genomic data used (Fig. 4, Table 1 and Supp. Table 2). This finding suggested that both parental groups had contributed significant amounts of genetic variation to these two populations of *L. spectabilis*. In addition, the genetic clustering performed on the SNPs data from ddRAD alignments corroborated this result by showing admixtures of genetic clusters of species in clades V and VII in the genomic composition of *L. spectabilis* individuals (Fig. 5). These population-level results, together with the species-level results presented in Chapter 1 (Pouchon et al., 2018), suggest that *L. spectabilis* probably originated from a hybridisation event involving trees in clade V and rosettes in clade VII. Results also indicated that both *L. spectabilis* and *L. aff. spectabilis* could not be differentiated in term of hybrid ancestry. A single event of hybridisation would probably lead to this species before a geographical divergence of populations. In addition, neither the Patterson's D-statistic nor the genetic clustering approaches allowed determining the parental species at the origin of this hybrid speciation event of both *L. spectabilis*. Such findings corroborated the results from the Chapter 1. Indeed, SNaQ analysis performed on a subset of taxa highlighted an ancient reticulation contributing a proportion  $\gamma = 0.16$  of genes from ancestral lineages of rosettes in clade VII to this species (Fig. 8). As for our genetic clustering analyses, a greater genetic contribution of ancestral genome from the trees in clade V instead of rosettes in clade VII was found in the genome of *L. spectabilis*.

Hybrid speciation is a relatively common process in plants and animals (Abbott et al., 2013; Yakimowski and Rieseberg, 2014). Numerous hybrid species have been confirmed in



**Figure 8.** Reticulate evolution estimated with SNaQ from 1870 genes for a subset of species. This analysis was performed in Pouchon et al. (2018). Dashed lines represents missing taxa for these clades that were not included. One reticulation was found in *L. spectabilis* with a probability 0.16 of having genes from the rosette species of the clade VII.

plants; e.g. in *Helianthus anomalus*, *H. deserticola* and *H. paradoxus* (Rieseberg et al., 2003); *Scaevola procera* (Howarth and Baum, 2005); *Senecio squalidus* (Brennan et al., 2012); *Iris nelsonii* (Taylor et al., 2013); *Pinus densata* (Gao et al., 2012) and *Paeonia anomala* (Pan et al., 2007). Moreover, in past decade, the detection of such forms of speciation has also widely been facilitated by genomic approaches allowing identification of hybrid ancestry in genome of species (Capblancq et al., 2015; Malinsky et al., 2017; Marcussen et al., 2014; Meier et al., 2017; Schumer et al., 2016; Tung and Barreiro, 2017; Turissini and Matute, 2017). In this study, such genomic hybrid ancestry would probably imply hybrid origin to this species from ancestral rosette and tree-form species, constituting a striking example of hybrid speciation during the evolution of the Espeletiinae between highly morphological differentiated species. Nevertheless, linking the genomic hybrid ancestry to the process of hybrid speciation requires that hybrid lineages were persistent, and morphologically and ecologically distinct from its parental species through reproductive isolation (Feliner et al., 2017). The following section proposed to verify these conditions in *L. spectabilis*.

**Insights on Reproductive Isolation** – In plants, reproductive isolation could arise from pre- and post-pollination barriers. Niche differentiation (e.g. ecogeographic isolation with colonization of transgressive habitat), phenological isolation (in timing of reproduction) and pollinator specialization prevent gene flow before pollination of species. For example, phenotypic novelties in floral traits providing by hybridisation drive pollinator shifts in orchids and reproductive isolation in highly specific plant-pollinator interactions (Vereecken et al., 2010). At opposite, pollen-stigma interactions, pollen competition, hybrid seed formation or hybrid fertility and viability isolate species after pollination (Baack et al., 2015).

**Pre-pollination barriers:** we showed in this study that *L. spectabilis* shared a somewhat similar climatic niche than *L. neriifolius* (Fig. 7a). This finding was not surprising due to the sympatric distribution of these species locally in the upper section of the Venezuelan Andean forests (Cuatrecasas, 2013). Climatic niches inferred on this study for tree species were coherent with their distribution at lower elevations in comparison with rosette species in clade VII, which are mostly restricted to páramo habitats at higher elevations. Consequently, reproductive isolation between *L. spectabilis* and other trees can hardly be explained by a novel climatic niche or habitat preferences. Nevertheless, species could diverge in terms of their microhabitats, which you could not measure using climatic variables.

Pollinator's specialization and phenology do not seem to be strong sources of isolation since flowers in all these taxa are mainly visited by generalist pollinators, such as bees and bumblebees (Berry and Calvo, 1989; Cuatrecasas, 2013), with a reproductive synchronization (Mavarez, pers. obs.). However, reproductive isolation could be arising from pollinator's attraction. Indeed, *L. spectabilis* exhibits a terminal inflorescence typical of trees in clade V on an unbranched rosette growth-form (Cuatrecasas, 2013). The recombination of parental trait led to a shift in the reproductive strategy; a terminal inflorescence without branching implies a monocarpic (semelparous) species that reproduces only once and then dies (Fig. 9b.), in contrast to polycarpic (iteroparous) species that reproduce multiple times during their lifetime (Cuatrecasas, 2013) (Fig. 9a). The monocarpic syndrome in insect-pollinated species could be of selective advantage with a more than proportional increase in fitness with a size in reproductive outputs and traits for example (Stearns, 1992), leading to a mass-flowering strategy for pollinators. This can be achieved by larger and colourful ligules, more nectar or scents and, of course, the largest one of the sub-tribe in terms of number of capitula observed in *L. spectabilis* (Cuatrecasas, 2013). Mass-flowering strategy for pollinator assemblage, arising from hybrid recombination, would increase the efficiency of pollination in this species (Kakishima et al., 2011; Kelly, 1994); and make it more competitive with parents. Finally, by overcoming inbreeding depression effects, a reproductive transition in flowering plants to self-mating could provide another selective advantage in hybrids (Wright et al., 2013).

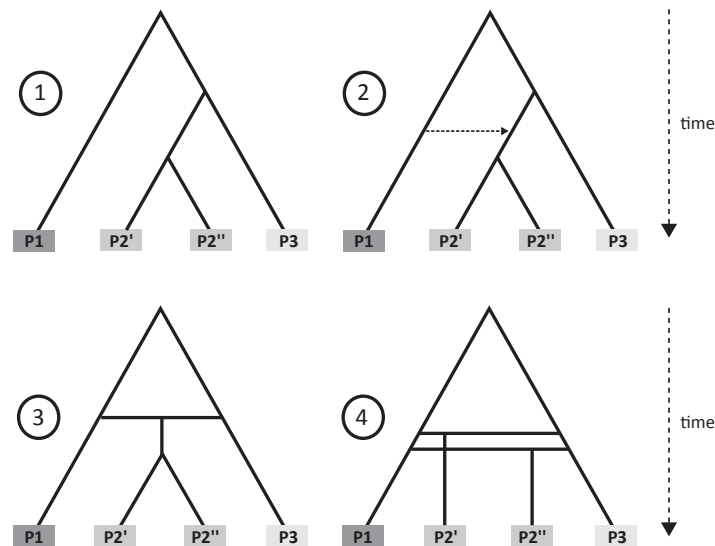
**Post-pollination barriers:** Hybrids between *L. spectabilis* and other sympatric trees from clade V (*L. neriifolius*) or rosettes from clade VII (*E. schultzii*) have not been observed (Mavarez, pers. obs.), which suggests that unviability could be a source of reproductive isolation between *L. spectabilis* and its putative parental species. If that is the case, such post-pollination barriers must have evolved after the speciation event of *L. spectabilis* since F1



**Figure 9.** Differences in reproductive strategies between (a.) polycarpic trees in clade V (here *L. neriifolius*), which reproduce repeatedly through adult life and (b) monocarpic hybrid species *L. spectabilis*, which dies after reproduction. Photos by Serge Aubert

hybrids between trees in clade V and rosettes in clade VII are viable and produce normal-shaped inflorescences and capitula (Mavarez, in prep). Unfortunately, no studies of inter-specific pollen-stigma compatibility, seed germination or seedling survival have ever been performed in *L. spectabilis*, which provide information necessary for the understanding of post-pollination reproductive isolation in plants (Johnston et al., 2003). On the other hand, the absence or rarity of hybrids between *L. spectabilis* and other sympatric *Espeletia* could indeed be more the consequence of strong pre-pollination barriers (e.g. such hybrids are not produced) than of post-pollination barriers (e.g. such hybrids have low survival). A posterior confirmation of this scenario would thus fit a growing body of empirical evidence suggesting that pre-pollination tends to contribute more than post-pollination to reproductive isolation in plants (Lowry et al., 2008).

**Perspectives** – Further investigations will then be needed to clarify the hybrid species status of *L. spectabilis*. If hybrid ancestry was shown on this study, reproductive mechanisms need to be validated with field experiments (e.g. in Carrió and Güemes, 2014; Murúa et al., 2017). Furthermore, evolutionary history scenarios of hybrid speciation could also be tested under coalescent models as performed in Capblancq et al. (2015) and Barrera-Guzmán et al. (2017). These methods simulated a large number of SNPs and estimated posterior probabilities under each scenario from the similarity between the summary statistics of the simulated and the



**Figure 10.** Evolutionary scenarios that could be tested in the coalescent analysis to clarify the hybrid status of *L. spectabilis* with models involving only cladogenesis events (1) and hybridisation events (2-4). P1 corresponds to rosettes in clade VII, P3 to trees in clade V, P2' to *L. spectabilis* lineages and P2'' to *L. aff. spectabilis* lineages.

observed SNPs (e.g.  $F_{st}$  or Patterson's  $D$  values), enabling the identification of the most probable evolutionary history (Cornuet et al., 2014). In this study, no distinction in hybrid ancestry could be made between the two geographical isolated populations of *L. spectabilis* and *L. aff. spectabilis* (Fig. 4 and 5). In such case, multiple evolutionary scenarios could be tested and compared in further study to confirm or invalidate hybrid speciation hypothesis: (1) a cladogenesis event of ancestral lineage from the clade V followed by a geographic divergence of *L. spectabilis* and *L. aff. spectabilis*, (2) a same scenario with ancestral introgression of rosette after the cladogenesis, (3) a single event of hybrid speciation from ancestral trees and rosettes followed by a geographic divergence, and (4) two hybrid speciation events leading respectively to *L. spectabilis* and *L. aff. spectabilis* (Fig. 10).

### Introgression between superpáramos species

The hypothesis of a hybrid origin for *E. semiglobulata* was addressed owing to its combination of morphologic traits from rosettes of clade VII, *i.e.* an opposite inflorescence; and from rosettes of clade VIII, *i.e.* long linear-lanceolate leaves, capitulum with pending/nodding disposition, short ligules and small pollen (Cuatrecasas, 2013).

Here, the  $D$ -statistic values for *E. semiglobulata* performed on SNPs from the shotgun alignments showed significant results when placing *C. thyriformis* as P1 and rosette species from clade VII as P3, although with slightly lower values compared to *L. spectabilis* (Fig. 3, Table 1). On the other hand, genetic clustering analyses showed that *E. semiglobulata* appears to be closely related to other rosettes in clade VIII, without evidence of admixture with rosettes in clade VII (Fig. 6). Unfortunately, it is difficult at this time to tease apart these conflicting results since individuals of *C. thyriformis* could not be included in the genetic clustering analyses due to the low number of SNPs recovered in this species with the ddRAD

approach. In view of these results, no firm conclusions could be reached in regards with hybrid speciation in *E. semiglobulata*.

Furthermore, according to Pouchon et al. (2018), introgression events was found between *C. thyrsoformis* not only with *E. semiglobulata*, but with all members of clade VIII and between *C. spicata* (clade VIII) and rosettes in clade VII. Such patterns were highlighted in this study by significant levels of introgression between other members of clades VII and VIII ( $D < 0$ ,  $p < 0.001$  in Fig. 3 and in Supp. Fig. 3). Our clustering analyses performed on SNPs from the ddRAD alignments also highlighted a genetically mixture in the genome of *E. batata*, between dwarf (*E. marthae* and *E. nana*) and large rosettes species in clade VII (*E. aristeguietana* and *E. schultzei*) (Fig. 5 and 6). Natural hybridisation and introgression are more frequent in closely related species because of their higher genetic compatibility (Mallet 2005; Hochkirch 2013). The divergence between clades VII and VIII is very recent, estimation to less than 0.80 Ma (Pouchon et al., 2018, Chapter 2). In addition, most of these species are sympatric in the Venezuelan páramo of the Piedras Blancas (Cuatrecasas, 2013), and appeared climatically close according to our niche reconstruction within high páramos and superpáramos (Fig. 7b). High levels of introgression between species in these two clades are consequently not surprising. In particular, some of these species are known to hybridize frequently, e.g. *E. schultzei* var. *schultzei*, *E. schultzei* var. *mucurubana*, *C. spicata* (Berry et al., 1988; Cuatrecasas, 2013). Hybrid populations between *E. batata* and *E. schultzei* are also often observed on field (J. Mavarez personal observation) and in the literature (Álvarez et al., 1990). Sequenced individuals of *E. batata* in this study could probably have been sampled in such hybrid zone. Pervasive introgressions among members of these two clades prevent any conclusions regarding hybrid speciation assumption. Further investigations will thus be necessary to discriminate between hybrid origin of *E. semiglobulata* and recent introgressions of these rosette species of the clades VII and VIII into superpáramos. Such gene flows could be in part responsible of the important phenotypic diversity observed in reproductive traits within these species, e.g. the structure of inflorescences, number of capitula per inflorescence, diameter and shape of capitula, number of flowers and seeds etc (Cuatrecasas, 2013). Divergence in reproductive traits, such as the size of the capitula and the seeds, seemed correlated to adaptation to the high elevation habitat of superpáramos during the evolution of Espeletiinae (see Chapter 2). Adaptive introgression, appearing as an important source of genomic variation that could accelerate adaptation (Barton, 2001), is common in plants (Suarez-Gonzalez et al., 2018). For example, this process has facilitated adaptive divergence in a recent radiation of monkeyflowers (Stankowski and Streisfeld, 2015). In Espeletiinae, gene flows in relation to reproductive traits correlated to pollination shift between these species could have accelerated the diversification rate of these species in these habitats (Chapter 2). Nevertheless, showing such adaptive evidence requires identifying introgressed regions in genome, genomic signature of selection, adaptively relevant phenotypic consequences of introgressed alleles and correlation between resultant phenotypes and fitness (Suarez-Gonzalez et al., 2018). With an increase in both sampling and sequencing depth, such of these points could be tested between these species in further studies.

## Hybridisation and Evolution in Espeletiinae

Through phenotypic novelties for selection to act upon by genetic recombination and adaptive introgression, hybridisation could play an important role during adaptive radiations (Kagawa and Takimoto, 2017; Seehausen, 2004). Such process was recently highlighted as a major driver of the radiation of cichlid fishes in East-African lakes (Meier et al., 2017), as for adaptive introgression facilitating radiation of the *Heliconius* (Consortium et al., 2012) and Darwin's finches (Lamichhane et al., 2015). According to Seehausen (2004), testing the impact of hybridisation in generation of adaptive radiation requires to verify that extensive introgression occurred on the onset of the radiation. Indeed, frequent hybridisation events observed in rapid radiation could be a consequence of recently diverged species (Wiens et al., 2006). Hybrid origin of adaptive radiation was for example shown in Malawi's cichlid fish radiation (Genner and Turner, 2012) or in Hawaiian silverswords plants (Barrier et al., 1999). Moreover, hybridisation appears to be a key factor in the adaptive radiation of the clownfishes, occurring throughout the evolutionary history of the group, which is still happening (Litsios and Salamin, 2014). In Espeletiinae, hybridisation was detected relatively recently in their evolution between closely related species (Pouchon et al., 2018; in this study), and is still happening (Diazgranados, 2012; Mavarez personal observations). No strong evidence of extensive introgression was highlighted at the origin of the group (Chapter 1). However, this process seemed play an active role in species (i.e. *L. spectabilis*) and phenotypic diversity (introgression clade VII/clade VIII), enhancing of both adaptive and specific evolution of this complex in the páramos. This study confirmed the important role of hybridisation in the evolution of species as it was shown in others Andean plant radiations (Dušková et al., 2017; Nevado et al., 2018; Vargas et al., 2017).

## 5 | SUPPLEMENTARY INFORMATIONS

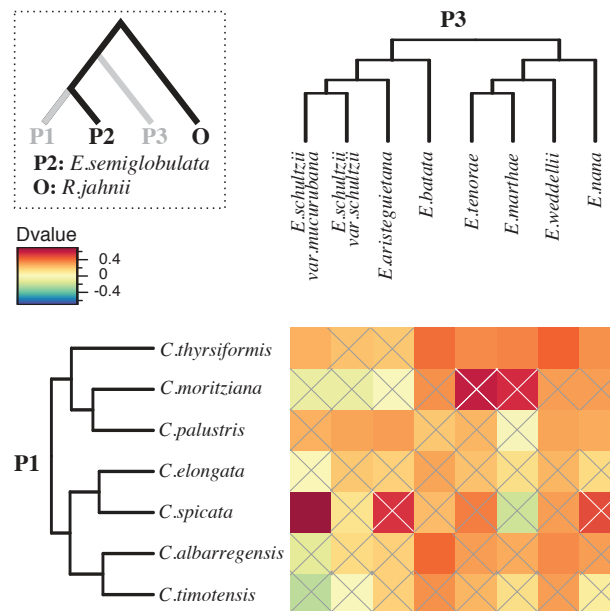
**Supp. Table 1.** Species sampled in this study.

	Species (samples)
Clade I	<i>Espeletiopsis purpurascens</i> (2)
Clade II	<i>Carramboa trujillensis</i> (2)
Clade III	<i>Ruilopezia jahnii</i> (2)
Clade V	<i>Libanothamnus arboreus</i> (4), <i>L. banksiaefolius</i> (1), <i>L. griffinii</i> (2), <i>L. liscanoanus</i> (2), <i>L. lucidus</i> (1), <i>L. neriifolius</i> (9), <i>L. occultus</i> (2), <i>L. parvulus</i> (3), <i>L. spectabilis</i> (5), <i>L. aff. spectabilis</i> (3)
Clade VII	<i>Espeletia aristeguietana</i> (1), <i>E. batata</i> (2), <i>E. marthae</i> (3), <i>E. nana</i> (4), <i>E. schultzi</i> (12), <i>E. tenorae</i> (1), <i>E. weddellii</i> (5)
Clade VIII	<i>Coespeletia albarregensis</i> (3), <i>C. elongata</i> (2), <i>C. moritziana</i> (2), <i>C. palustris</i> (1), <i>C. spicata</i> (1), <i>C. thyriformis</i> (2), <i>Espeletia semiglobulata</i> (8)

**Supp. Table 2.** Summary statistics of Patterson's D-test analyses performed on ddRAD SNP for the putative hybrids *Libanothamnus spectabilis* and *Espeletia semiglobulata* according to their respective parents described in Fig. 2, and to three different outgroup species: *Espeletiopsis purpurascens*, *Carramboia trujillensis* and *Ruilopezia jahnii*. Mean and standard deviation were estimated from all tests made with each parental combination for the counts of SNP in ABBA and BABA configurations and for the D-values.

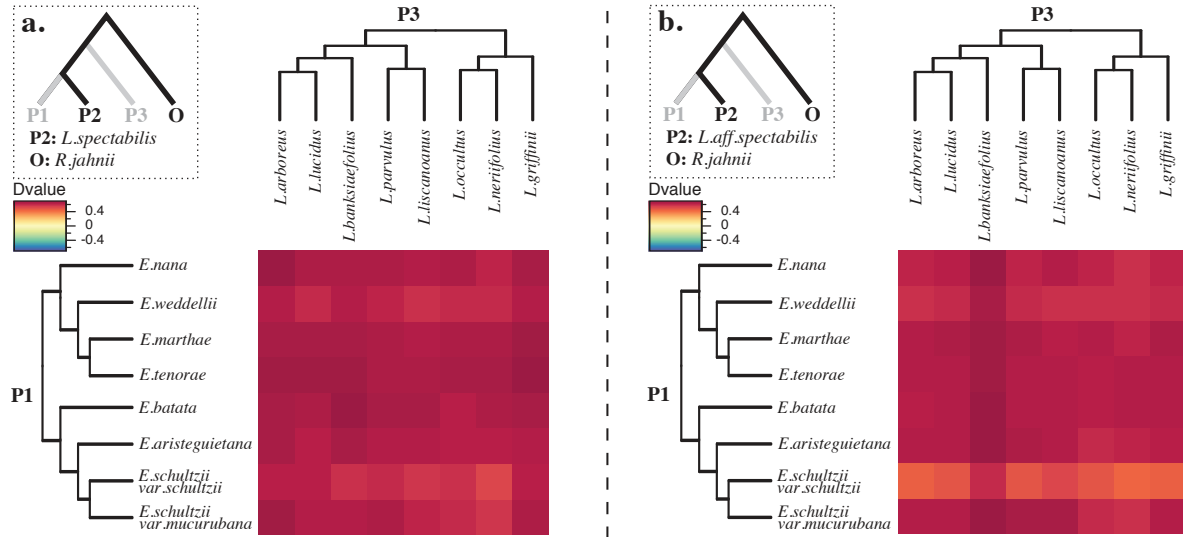
P2	Outgroup species	D-value < 0.00			D-value > 0.00		
		ABBA (sd)	BABA (sd)	D (sd)	ABBA (sd)	BABA (sd)	D (sd)
<i>E. purpurascens</i>							
<i>E. semiglobulata</i>	<i>C. trujillensis</i>	23.8 (5.06)	32.8 (2.86)	-0.16 (0.08)	33.24 (19.85)	17.84 (9.96)	0.30 (0.15)
	<i>R. jahnii</i>						
	<i>E. purpurascens</i>	NA	NA	NA	279.51 (61.50)	23.56 (5.19)	0.83 (0.04)
<i>L. spectabilis</i>	<i>C. trujillensis</i>	NA	NA	NA	309.76 (69.46)	42.88 (11.65)	0.75 (0.03)
	<i>R. jahnii</i>	NA	NA	NA	310.13 (70.76)	36.36 (9.45)	0.78 (0.04)
	<i>E. purpurascens</i>	NA	NA	NA	257.29 (58.83)	28.88 (6.04)	0.79 (0.04)
<i>L. aff. spectabilis</i>	<i>C. trujillensis</i>	NA	NA	NA	280.38 (60.10)	47.97 (12.57)	0.70 (0.04)
	<i>R. jahnii</i>	NA	NA	NA	279.29 (66.87)	45.23 (8.74)	0.71 (0.06)

**Supp. Figure 1.** Patterson's D-statistic performed on ddRAD SNPs considering a four topology (((P1,P2),P3),O) with *Ruilopezia jahnii* as O, the Venezuelan *Espeletia* as P3, *Coespeletia* species as P1, and *E. semiglobulata* as P2. Crosses highlighted D values which were not statistically significant.





**Supp. Figure 2.** Patterson's D-statistic performed on *ddRAD* SNPs considering a four topology (((P1,P2),P3),O) with *Ruilopezia jahnii* as O, *Libanothamnus* species as P3, the Venezuelan *Espeletia* as P1 and (a.) *L. spectabilis* or (b.) *L. aff. spectabilis* as P2. All gene flows events involving  $D > 0$  were statistical significant.



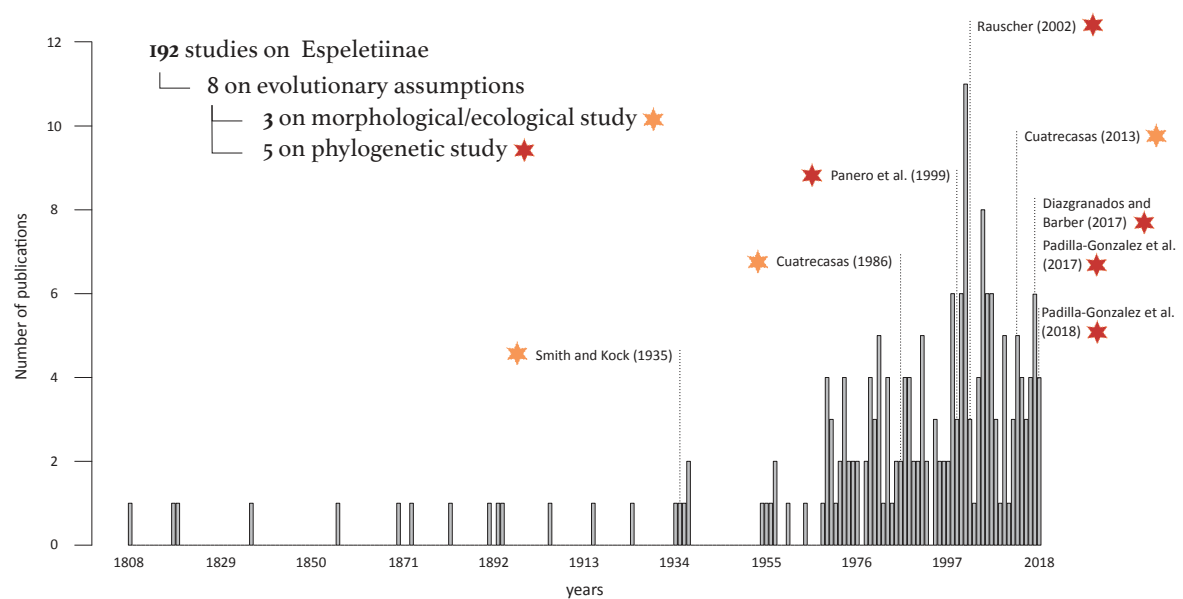
## DISCUSSION – Adaptive radiation of Espeletiinae in tropical alpine ecosystems

---

*“Quand on cherche, on trouve généralement quelque chose, mais ce n'est pas toujours exactement ce qu'on voulait. Ce fut ce qui se passa en l'occurrence.”*

Gandalf, 1000-3021 du 3ème Age

Les Espeletiinae (Asteraceae ; Heliantheae) ont fasciné un grand nombre de botanistes au cours d'expéditions dans les Andes, depuis le 18<sup>ème</sup> siècle jusqu'à aujourd'hui, en raison de leur forme de vie remarquable en rosette caulescente géante, secouant ainsi le paradigme selon lequel les plantes alpines sont généralement plus petites dans les milieux extrêmes de haute altitude. Depuis leur première découverte par Mutis, et leur description établie par Humboldt & Bonpland (1809), nombreux scientifiques se sont intéressés à la systématique, à la physiologie et à l'écologie des Espeletiinae dans les páramos pendant près de 210 ans, conduisant à l'aboutissement de 190 études (Fig. 1). Néanmoins, parmi celles-là, seulement six études ont proposé des scénarios quand à l'origine évolutive et la diversification de ces plantes dans les páramos, et seulement trois de ces études se basent sur des phylogénies moléculaires (Fig. 1) (Diazgranados and Barber, 2017; Padilla-González et al., 2017, 2018). Cependant, face à l'absence de signal phylogénétique robuste, la compréhension des mécanismes évolutifs à l'origine de leur radiation a longtemps été tenue mystérieuse. Cette dernière partie du manuscrit résume les principaux résultats obtenus durant ce travail de thèse, en montrant comment, à partir de l'utilisation de données issues séquençage haut-débit, ils



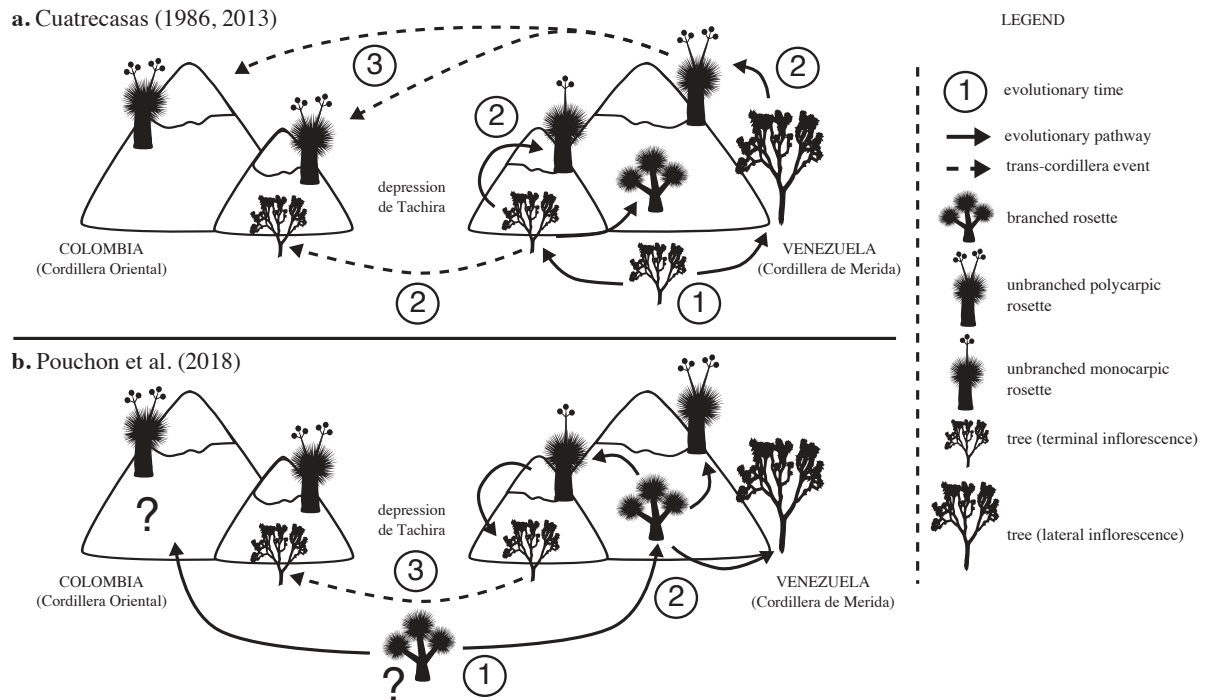
**Figure 1.** Trends in research on Espeletiinae as reflected in the number of articles published annually since their first discovery in 1808. The literature search was conducted using the Google Scholar database on May 20, 2018.

contribuent à avancer nos connaissances sur les facteurs explicatifs de la diversification de Espeletiinae dans les páramos, et plus généralement sur les scénarios évolutifs ayant permis la diversification explosive de certains groupes de plantes.

## 1 | RADIATION DES ESPELETIINAE DANS LES PARAMOS

### Patrons Globaux

**Deux radiations indépendantes** – Selon les premières théories évolutives, les Espeletiinae aurait évolué il y a probablement moins de 5 millions d’années (Ma) à partir d’un ancêtre commun arbustif issus des forêts andines néo tropicale dans les páramos (Cuatrecasas, 1986, 2013; Smith and Koch, 1935). Cette radiation aurait débuté au Venezuela, en raison d’une plus grande diversité morphologique retrouvée dans cette région, par plusieurs colonisations verticales successives des páramos suite à des évolutions répétées de la forme en rosette monocarpique ou polycarpique à partir de cette forme ancestrale, suivies de nombreuses colonisations horizontales secondaires entre les cordillères vénézuéliennes et colombiennes (Cuatrecasas, 1986, 2013) (Fig. 2a). Selon ce scénario évolutif, les arbres à inflorescences terminales issues des forêts andines du Venezuela (les *Libanothamnus*), qui seraient une évolution primitive de l’ancêtre commun à ces plantes, auraient colonisé la partie supérieure des forêts colombiennes à partir de la dépression de Táchira connectant les deux cordillères. De plus, en dépit de leur faible capacité de dispersion (Cuatrecasas, 2013), les rosettes polycarpiques de l’étage des páramos vénézuélien des genres *Espeletia* et *Espeletiopsis*, évoluant à partir ces formes arbustives, auraient également traversé deux fois la dépression



**Figure 2.** Evolutionary hypotheses of the Espeletiinae into the páramos habitats made by (a) Cuatrecasas (1986, 2013) and (b) established in this thesis manuscript.

(de plus de 30km) à environ 2,000m plus bas en plaine, afin de rallier la cordillère orientale Colombienne. Une telle dispersion répétée semblerait toutefois peu plausible pour ces plantes alpines, malgré les connexions possibles des páramos pendant les intenses périodes glaciaires du Pléistocène (Flantua et al., 2014). Contrairement à cette théorie, les phylogénies moléculaires réalisées dans ce manuscrit ont pu mettre en évidence un évènement important de vicariance à la base de la radiation de ces plantes dans ces écosystèmes (Chapitre 1 et 2). Les résultats obtenus soutiennent une évolution unique de la rosette en une forme ramifiée ancestrale qui aurait permis de coloniser indépendamment les páramos colombien et vénézuélien (étape 1 Fig. 2b). Cette nouvelle théorie évolutive de vicariance ancestrale converge avec celle établie par Diazgranados et Barber (2017), et supporte l'avènement de deux radiations indépendantes de ces plantes en Colombie et au Venezuela. Les mouvements trans-cordillères seraient ainsi limités aux espèces arbustives de basses altitudes comme les *Libanothamnus* et non plus aux rosettes de hautes altitudes (étape 3, Fig. 2b).

**Une radiation explosive durant le Pléistocène** – Les forts taux de diversification des plantes retrouvés dans les páramos, considérés comme les plus rapides de tous les biomes (Madríñan et al. 2013), ont principalement été attribués au phénomène géologique d'une surrection récente de ces nouveaux habitats mais également au phénomène climatique d'une dynamique glaciaire importante pendant le Pléistocène (Antonelli and Sanmartín, 2011a). Les Andes du Nord ont atteint 40% de leur élévation actuelle entre 10.0-5.0 Ma (Miocène), jusqu'à leur élévation finale au cours des 5.0-2.7 Ma pendant le Pliocène (Gregory-Wodzicki, 2000; Mora et al., 2011). L'élévation actuelle de la limite supérieure de la forêt, correspondant à la limite inférieure de l'habitat des páramos, serait apparue peu de temps après à 2.558 Ma (van der

Hammen and Hooghiemstra, 2000). Cette période correspondrait alors à l'apparition de conditions favorable à l'établissement d'une végétation des páramos. Une végétation arbustive proto-paraméenne aurait occupé ces habitats, à la fin du Pliocène/début Pléistocène (1.8-2.5 Ma), au dessus de la limite supérieure des forêts entre 2000 et 3000m d'altitudes, soit plus bas que le páramo actuel (Hooghiemstra and Van der Hammen, 2004b). Les traces de pollen retrouvé dans les sédiments confirment la présence de taxon que l'on retrouve actuellement dans le páramo. Par exemple, l'augmentation de pollen des taxons *Aragoa*, *Ericaceae* et *Hypericum*, retrouvée dans les sédiments, indique qu'une végétation arbustive s'est développée il y a ~2.0 Ma dans un écotone de type subpáramo et s'est intensifiée après il y a ~1.7-1.8 Ma (Torres et al., 2013). D'après Madriñan et al. (2013), la divergence des Espeletiinae coïnciderait avec cette surrection Andine et l'émergence suivante des páramos, en raison d'une origine à ces plantes estimée à 4.0 Ma (5.9-2.4). De plus, la majorité des évènements de spéciation se serait passé pendant le Pléistocène (Madriñan et al., 2013). Ces deux facteurs géologique et climatiques sembleraient expliquer l'origine de la radiation des Espeletiinae dans ces habitats. Les datations moléculaires effectuées durant ce travail de thèse ont permis de mieux estimer l'âge des Espeletiinae, à partir du génome chloroplastique de ces plantes et de calibrations fossiles, à 2.3 Ma (2.0-2.6) (Chapitre 1) avec une radiation Vénézuélienne et Colombienne qui aurait débuté respectivement il y a 1.3 Ma (1.5-1.2) et 0.9 Ma (1.1-0.8) (Chapitre 2). Par conséquent, avec pas moins de 140 espèces apparues en 2.3 MA, le taux de diversification de ces plantes est encore plus rapide que précédemment estimé par Madriñan et al. (2013), avec une estimation à 5.12 espèces/Myr pour la radiation vénézuélienne (Chapitre 2). De plus, la dynamique de diversification des Espeletiinae se déroulerait entièrement durant le Pléistocène, soit 300 KyA après l'émergence potentielle d'un proto-páramo il y a ~2.6 Ma (van der Hammen and Hooghiemstra, 2000; Hooghiemstra and Van der Hammen, 2004b). L'innovation clé de la forme en rosette ramifiée apparaîtrait ainsi comme le facteur explicatif ayant permis à ces plantes de coloniser de nouvelles niches écologiques dans cet habitat en place, qui ne pouvaient pas être exploités par les lignées ancestrales des Espeletiinae. En connaissant les fortes valeurs adaptatives liées à la forme en rosette (Beck et al., 1982; Monasterio and Sarmiento, 1991; Smith, 1981), cette forme ancestrale ramifiée était potentiellement adaptée à ces proto-páramos, et à la d'une végétation arbustive de subpáramos il y a ~2.0 Ma (Torres et al., 2013). Finalement, la spéciation des Espeletiinae semblerait s'intensifier plus récemment lors d'un changement de cycles climatiques caractérisé par des plus larges amplitudes thermiques pendant le Pléistocène, et par des évolutions répétées de la rosette en une forme non ramifiée au Venezuela et en Colombie (étape 1 et 2, Fig. 2b ; voir Chapitre 1). Si cette apparition semblerait être moteur à la radiation des Espeletiinae, le paléoclimat pourrait également jouer un rôle majeur dans les mécanismes de spéciation chez ces plantes. En raison d'un échantillonnage presque complet, ces hypothèses portant sur l'importance de l'innovation clé et du climat dans la diversification de ces plantes ont pu être vérifiées pour la section vénézuélienne de cette radiation.

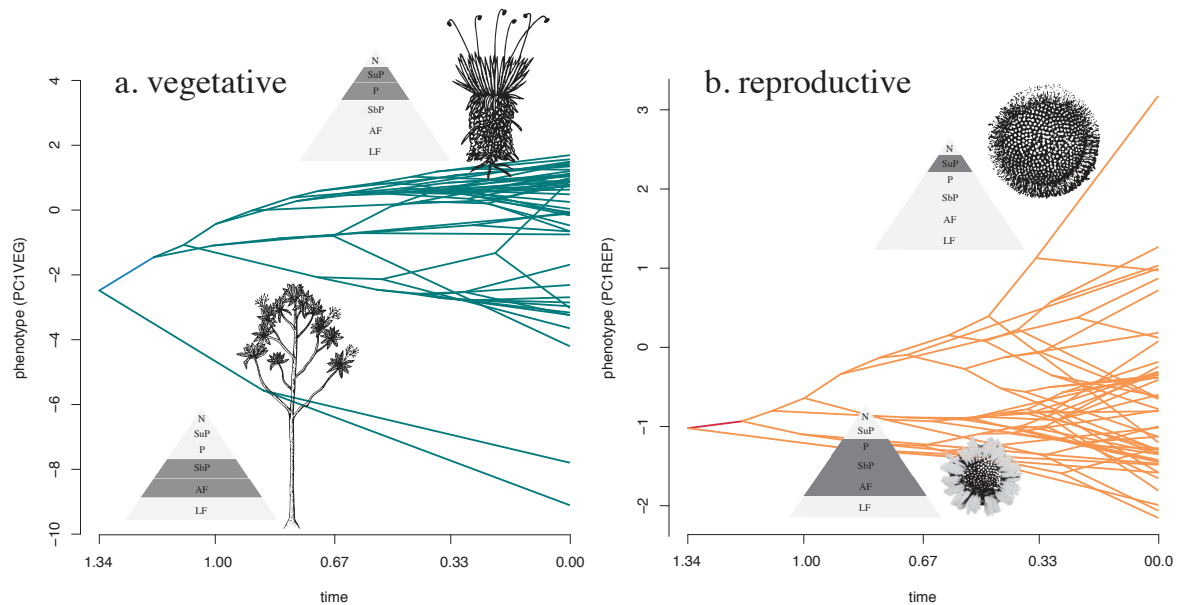
## Diversification au sein de la radiation vénézuélienne

**Des opportunités écologiques** – Les innovations clés, à l'origine de nouvelles opportunités écologiques pour les espèces, ont souvent été proposées pour les plantes et les animaux

(Losos, 2009; Matschiner et al., 2011; Rainford et al., 2014) mais il s'est avéré difficile pour trouver ces traits uniques dans la plupart des groupes malgré un grand intérêt à le faire (Glor, 2010). Chez les plantes, ce processus a été démontré à l'origine de la radiation des *Lupinus* dans les páramos où l'apparition d'une forme de vie pérenne aurait permis à ces plantes de coloniser les habitats de montagnes récemment apparus, chose impossible pour les lignées ancestrales annuelles, et d'influencer le taux de diversification de ces plantes avec d'autres évolutions dérivées de cette forme (Drummond, 2008; Drummond et al., 2012b). L'apparition d'un métabolisme de type acide crassulacéen (CAM) et d'un réservoir de stockage de l'eau constituent également des innovations clés ayant augmentées le taux de diversification des plantes de la sous-famille des Bromelioideae à l'origine de leur radiation dans les forêts néotropicales (Silvestro et al., 2014). Concernant les Espeletiinae, ce manuscrit a permis de mieux caractériser l'influence de traits dans la radiation vénézuélienne de ces plantes, qui aurait suivie deux trajectoires évolutives différentes: une évolution vers le haut et une vers le bas, mais toute deux dépendantes des opportunités écologiques du milieu.

Les analyses de diversification effectuées dans le Chapitre 2 ont démontré que l'évolution des traits végétatifs chez les Espeletiinae vers la rosette non-ramifiée aurait accéléré le taux de spéciation au sein de ce groupe. Cette forme de vie apparaît être hautement adaptée à l'écosystème des páramos du fait par exemple d'une protection thermique des bourgeons au sein de la couronne de feuille et de l'eau au sein d'un unique tronc par des feuilles marcescentes, d'une protection contre la sécheresse par une capacitance élevée ou encore de la présence d'une moelle importante pour la stockage hydrique au sein de ce tronc (Meinzer and Goldstein, 1985; Monasterio and Sarmiento, 1991; Smith, 1981). Chez les Espeletiinae, cette forme en rosette monocaule apparaît être liée au cours de l'évolution de ces plantes à la colonisation de l'étage du páramo (voir Table 3 et Supp.Fig.4 du chapitre 2). Si l'évolution de la forme ancestrale aurait permis de coloniser un subpáramo (~2.3 Ma), l'évolution secondaire au Venezuela de cette forme en rosette caulescente monocaule (< 1.0 Ma) constituerait une innovation clé additionnelle chez ces plantes, leur ayant permis de coloniser l'étage propre du páramo et d'augmenter les taux de diversification chez ces plantes. La même affirmation pourrait être établie en Colombie où cette forme monocaule chez les espèces colombienne aurait évoluée également il y a moins de 1 Ma (Chapitre 1 et 2). En parallèle à cette évolution convergente vers de plus hautes élévations, l'évolution répétée des arbres au Venezuela a permis une recolonisation des forêts Andine à de plus basse altitude (Chapitre 2). Cette différenciation de traits végétatifs est apparue rapidement dans l'évolution de ces plantes (Fig. 3a) et aurait ainsi permis aux Espeletiinae de coloniser rapidement de nouvelles zones adaptatives dans le páramo et en forêt de plus basse altitude, constituant des opportunités écologiques importantes à la radiation ces espèces.

Une troisième opportunité écologique est également constatée plus récemment dans l'évolution des Espeletiinae, celle de la colonisation du superpáramos. En effet, les modèles de diversification ont montré que ces plantes ont pu coloniser ces habitats de plus haute élévation par l'émergence de capitules plus larges et de graines plus grosses, adaptés à une pollinisation par le vent dans ces nouveaux milieux (Berry and Calvo, 1989; Cuatrecasas, 2013). Cette différenciation de traits reproductifs, ou syndrome de pollinisation, également associé à une accélération récente du taux de diversification de ces plantes, leur a permis récemment de coloniser une nouvelle zone adaptative (Fig. 3b). Chez les Angiospermes, une



**Figure 3.** Phenotypic divergence through the evolution of the Espeletiinae for both (a) vegetative and (b) reproductive traits. The vegetative divergence occurred earlier in the evolution between tree-form (linked to the Andean forest and the subpáramo habitat) and unbranched rosette-form (associated to the páramo and superpáramo habitats). In opposite, the reproductive divergence appeared more recently with the emergence of higher floral trait, associated to the colonization of the superpáramos. Vegetation zonation corresponds to the lowland forests (LF), the Andean forest (AF), the subpáramo (SbP), the proper páramo (P), the superpáramo (SuP) and the nival zone (N).

étude récente menée à large échelle a démontré que les variations phénotypiques dans la taille et la masse des graines étaient fortement liés aux taux de spéciation dans l'évolution des plantes à fleurs (Igea et al., 2017). Les plantes à petites graines ont diversifié plus rapidement en raison de leur capacité de colonisation plus importante. Dans cette étude, nous fournissons un exemple opposé de ce processus à une échelle phylogénétique plus fine où cette fois-ci les plantes à graine plus larges ont diversifié très rapidement en raison de leur capacité d'adaptation à ces nouveaux milieux de hautes élévations et non de dispersion (Berry and Calvo, 1989).

Cette transition altitudinale vers de nouvelles niches écologiques plus récemment témoigne d'une nouvelle exploration de l'espace écologique chez ces espèces. D'après les modèles écologiques de radiations adaptatives, le taux de diversification des espèces apparaîtrait inversement corrélé à la disponibilité des niches écologiques fournies par cette opportunité (Losos and Mahler, 2010). Celles-ci tendraient à saturer au cours du temps au travers de compétition pour l'espace écologique diminuant ainsi la diversification des espèces (Etienne and Haegeman, 2012; Rabosky and Hurlbert, 2015). En considérant ces nombreuses opportunités écologiques développées chez les Espeletiinae, on s'attendrait à un tel phénomène dans l'évolution de ces plantes. Dans ce manuscrit, les modèles de diversification utilisés valident bien cette prédiction en démontrant que la diversification de ces plantes tend à diminuer jusqu'à un nombre d'espèce limitant dans le milieu étant proche du nombre d'espèces existantes (Chapitre 2). En supposant que la diversité des traits végétatifs et reproductifs qui est corrélée à la diversité d'habitat soit un proxy des niches écologiques

(Chapitre 2), la radiation vénézuélienne serait proche de sa limite écologique dans les páramos (Rabosky, 2009). Toutefois, des analyses complémentaires sont nécessaires afin de nuancer l'effet d'une saturation de l'espace géographique à une saturation écologique de l'archipel des páramos.

**Une influence du paléoclimat** – Les changements climatiques qui se sont produit à large échelle ont modifié la distribution des niches disponibles pour les espèces (Ohlemüller, 2011) et ainsi influencé la diversification des lignées (Kolář et al., 2016) et leur distribution actuelle (Svenning et al., 2015). Ces changements climatiques favorisent également le contact secondaire entre des communautés par la migration d'espèces allopatriques en sympatrie, qui entraînant de nouvelles interactions écologiques entre ces espèces peut être moteur aux radiations adaptatives (Gavrilets and Losos, 2009; Losos and Ricklefs, 2009). De nombreuses études ont récemment soulignées l'effet direct ou indirect des fluctuations climatiques passés sur la diversification à différentes échelles et sur différents groupes taxonomiques comme chez les mammifères (Gibb et al., 2016), les oiseaux (Claramunt and Cracraft, 2015), les insectes (Condamine et al., 2016), les serpents (Cyriac and Kodandaramaiah, 2017) ou encore chez les plantes (Kong et al., 2017; Lehtonen et al., 2017). Au niveau des páramos, les oscillations climatiques durant le Pléistocène auraient modelé indirectement les taux de diversification chez les plantes en connectant et isolant ces habitats de façon répétées, favorisant ainsi la spéciation (Antonelli and Sanmartín, 2011a; Madriñán et al., 2013). Aguilée et al. (2012) ont démontré qu'une alternation répétée d'allopatrie et de sympatrie chez les taxons génère une plus grande diversité que ces mécanismes de spéciation seuls. Ce phénomène de répétitions d'isolement spatiale et de contact secondaire, appelés "pompe à espèces", a notamment contribué à la radiation des cichlides dans les lacs Malawi et Tanganyika (Schwarzer et al., 2012; Sturmbauer, 1998), et pourrait être responsable de la forte diversification des plantes dans les páramos (Madriñán et al., 2013). Cependant, à ce jour, l'effet des paleotemperatures et du climat sur les taux de diversification dans les páramos a seulement été démontré chez un groupe de Campanulacées, par un lien direct entre taux de mutation et températures (Lagomarsino et al., 2016). Dans ce manuscrit, le taux de diversification de la section vénézuélienne des Espeletiinae apparaît lié de façon linéaire aux températures passées (Chapitre 2). Le refroidissement de ces dernières pendant le Pléistocène aurait négativement influencé le taux de spéciation chez ces plantes validant l'importance du climat dans la diversification des Espeletiinae et des plantes des páramos de façon plus générale (Chapitre 1; Madriñán et al. 2013). Néanmoins, ces températures passées apparaissent dépendantes du facteur temps et le taux de diversification des Espeletiinae vénézuéliens tend à décliner au cours du temps (voir section précédente). Cette corrélation pourrait ainsi refléter un lien direct entre la diversification et les températures ou être la conséquence de la relation que tous les deux ont avec le facteur temps. Par exemple, la diversification pourrait diminuer au cours du temps en raison d'une saturation de l'espace écologique ou géographique, comme démontré ici, alors que la température moyenne diminue à cause des phénomènes climatiques du Pléistocène. Une autre possibilité est que la spéciation soit plus élevée à fortes températures, comme démontré dans ce travail, quand les páramos sont fragmentés et isolés, et plus faible à des températures basses quand ils sont interconnectés. Cette hypothèse, étant reliés à la théorie de pompe à espèces de Madriñán et



al. (2013), implique que l'essentiel des événements de spéciations soient arrivés en allopatrie. Des analyses complémentaires sont ainsi nécessaires afin de démêler l'effet de la saturation des niches des páramos et l'effet propre du climat dans la diversification des Espeletiinae.

**Rôle de l'Hybridation** – L'hybridation est reconnue aujourd'hui comme un processus évolutif à l'origine des espèces (Mallet, 2007), et par la génération de variations génétiques et phénotypique utilisable pour la sélection, à l'origine des radiations adaptatives (Seehausen, 2004, 2013). Ces nouveautés adaptatives arrivent par des phénomènes d'introgession adaptative d'allèles favorables à la sélection chez les espèces parentales (i.e. introgession adaptative), et/ou par des recombinaisons de traits parentaux dans le génome des lignées hybrides leur pouvant être favorables dans de nouveaux milieux (Abbott et al., 2013; Gross and Rieseberg, 2005). Une hybridation intense entre des lignées lors de la colonisation de nouveaux habitats peut ainsi rapidement permettre la colonisation de nombreuses niches adaptatives vacantes, et promouvoir la radiation des lignées (Kagawa and Takimoto, 2017; Seehausen, 2004). D'une autre manière, des événements ponctuels d'hybridation localisés chez certaines espèces ayant diversifiées peuvent également moduler cette radiation en leur permettant d'exploiter des ressources encore non utilisées (Seehausen, 2004). Le rôle de l'hybridation, pouvant être ainsi précurseur ou acteur aux radiations adaptatives a récemment été démontré chez les animaux dans la radiation des cichlides africains (Herder et al., 2006; Meier et al., 2017), des *Heliconius* (Kozak, 2016; Pardo-Diaz et al., 2012), des pinsons des Galápagos (Lamichhaney et al., 2015), des criquets Hawaïen (Shaw, 2002) ou encore des salmonidés (Hudson et al., 2011). Peu d'études se sont toutefois intéressées à l'impact de ce phénomène dans la radiation des plantes. Quelques exemples figurent parmi les radiations insulaires des Crassulaceae des îles des Canaries (Jorgensen and Olesen, 2001) et de l'alliance des *silverswords* des îles d'Hawaï (Barrier et al., 1999; García-Verdugo et al., 2013), ou des îles continentales des páramos chez les Astereae (Vargas et al., 2017), les Asteraceae (Dušková et al., 2017) les Fabaceae (Nevado et al., 2018) et les Rosaceae (Morales-Briones et al., 2018).

Chez les Espeletiinae, l'hybridation est particulièrement répandue dans les páramos (Cuatrecasas, 2013; Diazgranados, 2012) et pourrait être un mécanisme important à la radiation de ces plantes (Diazgranados and Barber, 2017). Néanmoins, une hybridation importante peut survenir comme la conséquence d'un isolement incomplet des lignées plutôt qu'une causalité à leur radiation en raison de la diversification rapide et récente de ces taxons (< 2.30 Ma, Chapitre 1 et 2) (Wiens et al., 2006). Un moyen de vérifier cette prédiction selon Seehausen (2004) est de savoir si l'hybridation chez ces plantes est apparue au début de la radiation (i.e. rôle de précurseur) et si l'hybridation a joué un rôle dans la spéciation et la diversification adaptative durant la radiation (i.e. rôle d'acteur). Dans ce manuscrit, des événements d'hybridation ont pu être identifiés relativement récemment dans l'évolution des Espeletiinae entre des espèces phylogénétiquement et géographiquement proches (Chapitre 1 et 3). Ainsi la plupart de ces événements apparaît après la prolifération rapides des espèces et ne semble pas être un précurseur à la radiation comme attendue sous l'hypothèse de Seehausen (2004). Cependant ce processus, qui est toujours omniprésent chez les Espeletiinae (Diazgranados, 2012; Mavarez personal observations), semble jouer un rôle actif dans la diversité spécifiques et phénotypique ayant pu promouvoir l'évolution et l'adaptation de ces

plantes durant leur radiation (Chapitre 3). Des évidences génomiques fortes ont en effet démontré la présence d'une admixture ancienne chez *L. spectabilis*, qui couplé à une recombinaison de traits morphologiques parentaux chez cette espèce nous laissent supposer à une hypothèse de spéciation hybride à l'origine de cette plante entre des rosettes et des arbres (Chapitre 3). Des événements importants d'introgression génétiques ont également été démontrés chez les rosettes de plus hautes altitudes à la limite haute du páramo/superpáramo (Chapitre 1 et 3). Ces flux de gènes pourraient être en partie responsables d'une grande diversité phénotypique observée chez ces espèces, notamment au niveau des traits reproductifs (Cuatrecasas, 2013). Dans le chapitre 2, nous avons également vu qu'une divergence de certains de ces traits, comme la taille des capitules et des graines, a permis à ces espèces de s'adapter rapidement à une pollinisation par le vent dans ces milieux. Une introgression adaptative entre ces espèces aurait ainsi pu accélérer leur adaptation à ces habitats de hautes montagnes (Barton 2001). Des analyses complémentaires sont toutefois requises pour affirmer le statut d'espèce hybride chez *L. spectabilis*, et pour identifier chez les espèces des superpáramos des régions introgressées dans le génome codant pour des phénotypes adaptatifs à ces milieux (Suarez-Gonzalez et al., 2018).

## 2 | UNE RADIATION ADAPTATIVE ?

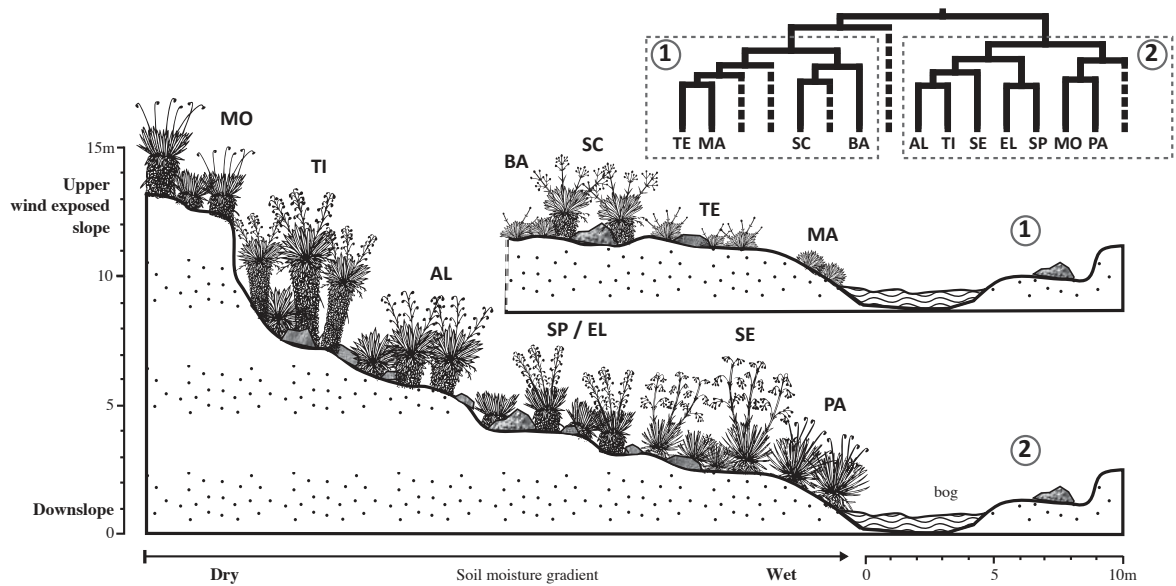
Depuis les travaux de Darwin sur les pinsons des Galápagos, les radiations adaptatives, associant l'écologie et l'adaptation des espèces, ont longtemps sollicité l'intérêt de nombreux chercheurs en biologie évolutive. Néanmoins, malgré un tel engouement, la suggestion d'une radiation adaptative comme un processus évolutif à la diversification d'un groupe figure comme une hypothèse qui doit être testée et non pas comme un a priori (Glor, 2010). Or peu d'études empiriques ont directement testé les prérequis des radiations adaptatives (Gavrilets and Losos, 2009), et certains aspects même de la définition ne sont pas toujours vérifiés en dépit d'une divergence écologique des espèces liées à des adaptations morphologiques (Givnish, 2015).

Les radiations adaptatives ont été définies comme une diversification rapide des espèces à partir d'une lignée ancestrale, associée à une forte diversification de stratégies écologiques (Rundell and Price, 2009; Schluter, 2000; Simpson, 1953). Or cette notion de spéciation "explosive" au cours du temps, traduisant une accélération de la diversification des espèces par rapports à leurs lignées sœurs, n'apparaît pas être vérifiée chez les pinsons de Darwin (Grant and Grant, 2008) ou chez les broméliacées (Givnish et al., 2014), figurant parmi des exemples majeurs de radiation adaptatives (Givnish, 2015). De plus, l'émergence d'opportunités écologiques, pouvant survenir à partir de l'apparition de nouveaux habitats, d'innovation clés ou encore à partir d'extinction massive (Stroud and Losos, 2016), a souvent été décrite comme une composante fondamentale à ce processus (Burress and Tan, 2017; Schluter, 2000; Simpson, 1953; Yoder et al., 2010). Or bien qu'elles puissent fournir des conditions adéquates à la spéciation aux travers d'adaptations à différentes niches écologiques, ces opportunités ne conduisent pas forcément à une diversification ou même à une radiation adaptative des espèces (Fortuna et al., 2017; Stroud and Losos, 2016). Par

exemple, l'évolution de la forme en coussin constitue une innovation clé chez le genre *Androsace* ayant permis à ces plantes de s'adapter aux conditions alpines (Boucher et al., 2012), bien que leur radiation soit expliquée par des phénomènes géographiques (Roquet et al., 2013).

Dans ce contexte, Givnish (2015) définit une radiation adaptative comme “la propagation d'une diversité de rôles écologiques et d'adaptations spécifiques associés à ces rôles chez les lignées” et donne une série de points essentiels pour tester cette prédiction. En raison d'une forte diversification morphologique, spécifique et écologique, les Espeletiinae ont longtemps été décrit comme un exemple majeur de radiation adaptative (Cuatrecasas, 2013; Monasterio and Sarmiento, 1991). Peu d'études ont réellement démontré de façon macro-évolutive ou micro-évolutive la présence de radiations adaptatives chez les plantes comme chez les Asclepiadinae Américain (Agrawal et al., 2009), les lobelioids Hawaïen (Givnish et al., 2009), l'alliance de *silveswords* d'Hawaï (Baldwin and Wagner, 2010; Blonder et al., 2016) ou encore les lupins (Drummond et al., 2012b). Chez ces plantes, des évidences récentes de traces de sélection dans l'expression des gènes ont récemment apporté un supplément à l'hypothèse d'une radiation adaptative des lupins (Nevado et al., 2016). Cependant aucuns des attendus théoriques de radiation adaptative n'ont pas été directement testés chez les Espeletiinae. Cette partie de la discussion propose de vérifier chacun des points de définition de Givnish (2015) permettant d'affirmer que la radiation des Espeletiinae dans les páramos vénézuélien constitue bien un exemple majeur de radiation adaptative.

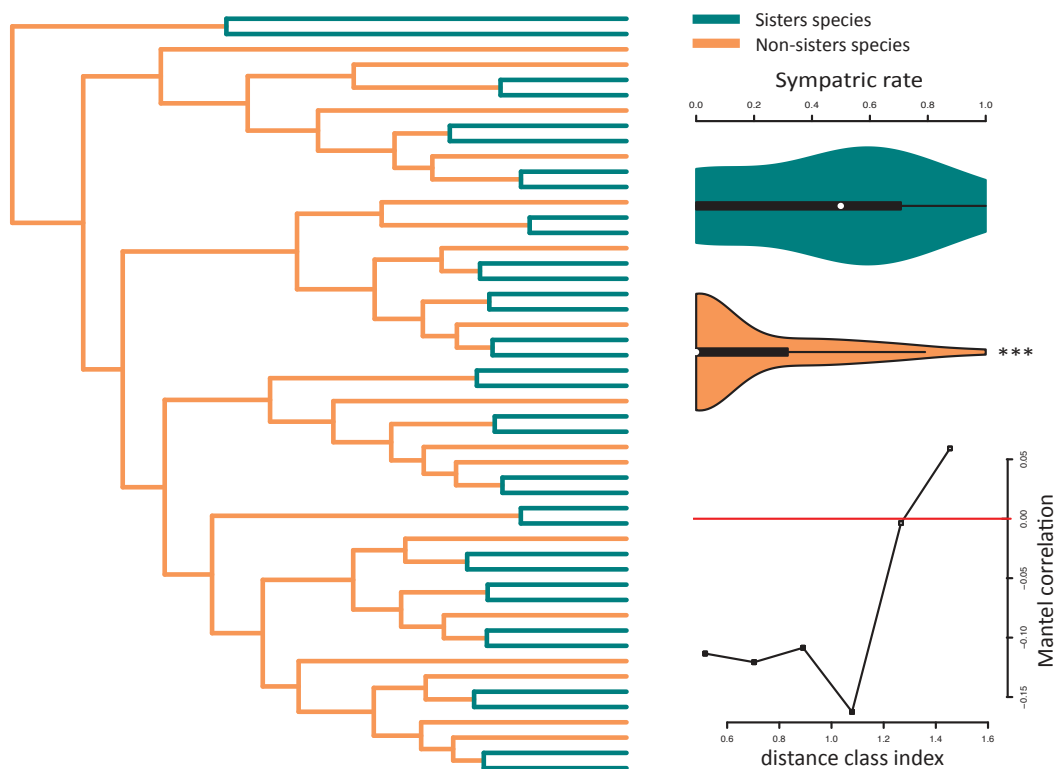
- (1) *Une augmentation des rôles écologiques et de la diversité de traits associés à ces rôles* – La diversité écologique chez les Espeletiinae a augmenté durant la diversification du groupe en association avec un grand nombre de traits morphologiques et écologiques. Par exemple les habitats occupés par ces espèces sont associés à la forme de vie de ces dernières (Cuatrecasas, 1986, 2013), la distribution altitudinale est liée à la capacité d'évitement du gel (Goldstein et al., 1985), et les vecteurs de pollinisations sont liées à la morphologie des fleurs (Berry and Calvo, 1989). Les résultats obtenus dans ce manuscrit soutiennent ces observations en démontrant l'association des rosettes au páramos, des formes arbustives aux forêts andines et des capitules plus larges aux superpáramos au cours de l'évolution de ces plantes (Chapitre 2).
- (2) *Les traits associés à ces rôles sont adaptatifs* – Des évidences intra- et interspécifiques confirme les fortes valeurs adaptatives de ces traits en particulier pour la rosette caulescente dans les páramos: une pubescence foliaire dense (Baruch and Smith, 1979; Meinzer and Goldstein, 1985), un volume de moelle important dans le tronc (Goldstein et al., 1984; Meinzer et al., 1985), une dense couche de feuille mortes recouvrant le tronc (Monasterio and Sarmiento 1991). Les capitules des espèces des superpáramos apparaissent également adaptées à ces milieux et à la pollinisation par le vent avec un nombre réduits mais une forme plus large et compacte liée à l'augmentation du nombre de fleurs, des capitules tombant, des grains de pollen à épines raccourcies, un involucre fortement recouvert de poils pour maintenir au sec le pollen afin de faciliter sa dispersion (Cuatrecasas, 2013).
- (3) *Une augmentation de l'espace morphologique et écologique des espèces par des innovations clés* – Concernant ce point, l'innovation clé à l'origine d'opportunités



**Figure 4.** Topographic and vegetation sequence for the Espeletiinae distribution for sympatric and closely related species according to moist soil and slope gradient. Profile was made according to Perez (1995) and field observations in the páramo of Piedras Blancas. Phylogeny represents clade VII (here 1) and VIII (here 2) estimated in Chapter 2. Only sympatric species of these clades are shown in the profile and in the phylogenetic tree by solid branches. Acronyms represent species names with: AL-*Coespeletia albarregensis*, BA-*Espeletia batata*, EL-*C. elongata*, MA-*E. marthae*, MO-*C. moritziana*, PA-*C. palustris*, SC-*E. schultzei*, SE-*E. semiglobulata*, SP-*C. spicata*, TE-*E. tenorae* and TI-*C. timotensis*.

écologiques ne me paraît pas être une condition sine qua non aux radiations adaptatives en raison de la survenue de radiation non-adaptatives suivant l'émergence d'une telle évolution (Fortuna et al., 2017; Stroud and Losos, 2016). Toute fois en considérant cette définition selon Givnish (2015), l'évolution de la rosette peut être considérée comme une innovation morphologique ayant permis la colonisation et la diversification initiale des Espeletiinae dans l'habitat des páramos au Venezuela et en Colombie (Chapitre 1). De plus, l'établissement et la diversification suivante de cette sous-tribu dans les páramos Vénézuélien a été achevé par l'évolution d'une innovation supplémentaire une forme en rosette marcescente non ramifiée, et d'une pollinisation par le vent à plus haute altitude par l'évolution de traits reproductif, ayant tous deux accéléré la diversification de ces plantes dans ces habitats (Chapitre 2).

- (4) *Une diminution de l'espace morphologique au cours du temps* – Le taux de changement morphologique chez les Espeletiinae apparaît augmenter rapidement durant une courte période entre 1.75 and 1.25 Ma, et tend ensuite à décroître vers le présent (Chapitre 1). Ce résultat est également retranscrit dans le Chapitre 2 où l'évolution de la diversité phénotypique des traits végétatifs apparaît être rapide au début de la radiation de ces plantes (Chapitre 2, Fig. 3a).
- (5) *Un chevauchement géographique d'espèces écologiquement divergentes* – La divergence écologique semble avoir joué un rôle important dans la structuration des communautés des Espeletiinae (Berry and Calvo, 1989; Cuatrecasas, 2013). Par exemple, les espèces retrouvées en sympatrie dans le páramos de Piedras Blancas (Venezuela) sont composées :



**Figure 5.** Sympatric rate of species estimated for sisters and non-sisters taxa shown on our phylogenetic reconstruction of the Venezuelan radiation (see Chapter 2). Geographic overlap was estimated using the Pianka's niche overlap metric in the *spaa* R package from the occurrence of species into páramos. Sisters-taxa are significantly more sympatric than non-sisters species according to the Kruskal-Wallis test shown in top right panel (\*\*\*)  $p < 0.001$ ). Overlap distance appeared also significantly correlated to the phylogenetic distance according to the Mantel correlogram (bottom right panel). This test, estimating relationship between phylogenetic and geographic distances across time shows an early negative correlation meaning that sympatry tends to decrease over time.

de grandes rosettes polycarpiques pollinisées par le vent (*Coespeletia timontensis* and *C. spicata*), de rosette sessiles polycarpique pollinisées par le vent en milieu humide (*C. palustris* and *Espeletia semiglobulata*) ou sec (*C. moritziana*), de rosettes polycarpiques pollinisées par des insectes en milieu sec (*E. schultzei*) ou humide (*E. batata*) et des rosettes monocarpiques pollinisées par les insectes en milieu sec (*Ruilopezia floccosa*) (Fig. 4).

- (6) *Compétition et sélection divergente entre espèces phylogénétiquement proches* – En estimant les taux de sympatrie pour les espèces échantillonnées dans ce travail, on peut s'apercevoir que les espèces sœurs (soit les espèces phylogénétiquement proches) se retrouvent généralement plus en sympatrie (Fig. 5). Ce résultat figure comme un attendu des radiations adaptatives (Losos and Mahler, 2010). De plus, les taxons sympatriques chez les Espeletiinae présentent fréquemment d'importantes différences au niveau de leurs préférences écologiques, qui ont été interprétées comme la conséquence de compétition interspécifique (Smith, 1981). Par exemple, les nombreuses espèces sympatriques dans le páramo de Piedras Blancas, qui sont proches phylogénétiquement, sont clairement

distribuées le long d'un gradient environnementale au vue de l'humidité dans le sol, la taille des particules et la pente par exemple (Fig. 4 ; voir Pérez, 1996).

### 3 | CONCLUSION

#### Une diversification dictée par des facteurs biotiques et abiotiques

Deux principaux mécanismes liés à la sélection naturelle ont été proposés et opposés afin d'expliquer l'apparition et le maintien de la diversification des espèces (Barnosky, 2001; Benton, 2009). La théorie de la "Reine Rouge" (ou *Red Queen*), basée sur les observations de Darwin et de Wallace, propose que la dynamique de cette diversité soit liée à des facteurs biotiques (intrinsèques) régulés par les interactions entre espèces et les traits d'histoires de vie de ces dernières (Benton, 2009). Cette théorie met en avant l'apparition de nouveautés chez des taxons utilisables pour la sélection naturelle, tout en sachant que d'autres espèces également puissent être soumises à leurs propres nouveautés. À l'opposée, le modèle du "Fou du Roi" (ou *Court Jester*) propose que cette dynamique de diversification soit expliquée par des facteurs abiotiques (extrinsèques), où des changements géologiques et climatiques ont régi les taux de spéciation et d'extinction (Barnosky, 2001). La divergence entre ces deux modèles est généralement considérée comme étant dépendante de l'échelles spatiale et temporelle (Benton 2009). Par exemple, les facteurs biotiques agirait plutôt à une échelle locale sur des pas de temps court tandis que les facteurs abiotiques agirait sur des échelles régionales au niveau d'échelle temps plus globales.

Si ces théories évolutives de la reine rouge et du fou du roi ont longtemps été opposées (Benton, 2009), des évidences récentes ont soulignées l'importance de l'interaction des deux facteurs biotiques et abiotiques dans la diversification des espèces (Bouchenak-Khelladi et al., 2015; Condamine et al., 2018; Donoghue and Sanderson, 2015). De plus, ces prédictions spatiales et temporelles ne sont pas toujours vérifiées (Liow et al., 2015; Silvestro et al., 2015). Par exemple, la dynamique de glaciation (composante abiotique) pendant le Pléistocène pourrait expliquer les taux de diversification élevés chez les plantes des páramos (Madriñán et al., 2013). Or chez les *Lupinus*, l'apparition de nouveaux traits de vie (composante biotique) pendant le Miocène/Pliocène semblerait être moteur de la diversification de ces espèces (Drummond et al., 2012b; Nevado et al., 2016). La surrection des Andes aurait également joué un rôle important dans la diversification des espèces en association avec des shifts dans le système de pollinisation (composante biotique) (Kay et al., 2005; Lagomarsino et al., 2016, 2017), ou encore en association avec le paléoclimat (Lagomarsino et al., 2016) ou l'action combinée d'innovation biotiques et de spéciations géographiques (Drummond et al., 2012b; Nürk et al., 2013). L'ensemble des résultats de ce manuscrit souligne cette interaction entre ces deux forces abiotiques et biotiques qui agirait de manière synchrone dans la diversification des Espeletiinae : par (1) l'apparition d'opportunités écologiques liées à l'émergence d'innovation clés et par (2) l'effet du paléoclimat sur les taux de diversifications au Venezuela. De plus, (3) l'hybridation, à

l'origine d'interactions biotiques entre espèces et qui est très présente chez ces plantes, aurait également activement participé plus récemment à la forte diversité morphologique et écologiques observées chez les Espeletiinae durant leur évolution.

## Perspectives

Ce travail fournit une vue d'ensemble de certains des processus évolutifs ayant agi dans la diversification des Espeletiinae, et constitue un exemple supplémentaire remarquable de radiations adaptatives, qui je l'espère viendra aider quant à la compréhension générale de ce processus. De plus, il soulève également de nombreuses autres questions et perspectives chez ces plantes et à plus large échelle.

Une continuité à ce travail serait de pouvoir considérablement augmenter l'échantillonnage de la section colombienne afin de déterminer si cette section sœur à la radiation vénézuélienne, serait une radiation de type adaptative ou non. En effet, les espèces de la radiation colombienne, plus diversifiées spécifiquement mais moins diverses morphologiquement (Cuatrecasas, 2013), ont une distribution latitudinale et longitudinale plus étendue au niveau des cordillères orientale, centrale et occidentales de la Colombie par rapport à leur groupe frère restreint à la cordillère de Mérida (Cuatrecasas, 2013). En raison de la faible capacité de dispersion de ces espèces, on pourrait s'attendre à ce que ces phénomènes de spéciation géographiques soient en partie plus importants dans la radiation colombienne que vénézuélienne. Il serait ainsi intéressant de pouvoir étudier l'effet d'évènements biogéographiques au niveau de ces deux radiations. De plus, ces espèces sont majoritairement représentées par une forme de vie en rosette caulescente non ramifiées, qui constitue une innovation clé à l'origine d'opportunités écologiques dans la section vénézuélienne. Il serait intéressant de voir si les mêmes opportunités écologiques liées à cette forme, et par exemple aux changements de morphologie des capitules en haute élévation, seraient également retrouvées dans ces habitats.

De nombreux autres traits, variant avec l'élévation chez les Espeletiinae, pourraient également être mesurés et intégrés aux modèles de diversification afin d'apporter un regard nouveau quant à l'association de plusieurs traits dans la diversification des espèces. Par exemple le degré de pubescence foliaire et de marcescence, hautement adaptées à ces habitats (Meinzer and Goldstein, 1985; Monasterio and Sarmiento, 1991), pourrait agir de manière synchrone à la structure en rosette caulescente non-ramifiée sur les taux de diversification des espèces. Blonder et al. (2016) ont démontré l'importance de la diversification fonctionnelle de traits foliaires, dans la radiation adaptative des *silverswords* d'Hawaii, en lien avec les niches climatiques. Un certain nombre de ces traits fonctionnels pourrait être mesuré afin de tester si de telles corrélations apparaissent également chez les Espeletiinae. Enfin la suite de ce travail serait de pouvoir augmenter l'effort de séquençage et d'échantillonnage du Chapitre 3 afin de déterminer et de démêler les cas d'hybridation à l'origine de la spéciation, des cas introgressions adaptatives durant l'évolution de ces plantes mais également de pouvoir déterminer des traces de sélection dans le génome de ces plantes, par des méthodes de transcriptomique comme utilisées par Nevado et al. (2016), afin de supporter l'hypothèse d'une radiation adaptative.

Finale­ment, ce manuscrit a égale­ment ouvert de nom­breuses perspectives quant à l’uti­lisation gé­nérale de données de séquençage haut-débit dans la réso­lution de radiations récentes (Fernández-Mazuecos et al., 2018; Léveillé-Bourret et al., 2018; Nevado et al., 2016; Vargas et al., 2017). Ces méthodes pour­raient consi­déra­blement amé­liorer les phylogé­nies et les datations des lignées dans les páramos et dans les autres biomes terrestres à différentes échelles phylogé­néti­ques, spatiales et temporelles. Au travers de ces méthodes associées à des analyses comparatives variées comme utilisées au cours de ce travail ou dans d’autres études récentes (Condamine et al., 2018; Lagomarsino et al., 2016), nous pourrions déter­miner si la dynamique de glacia­tion impor­tante durant le Pléistocène serait un facteur extrinsèque commun à la diversification des plantes dans ces biomes tropicaux alpins, comme pour les Espeletiinae et les Campanulacées (Lagomarsino et al., 2016), ou un facteur commun à la diversification des espèces dans les biomes alpins tropicaux et alpins tempérés en général. Cela nous permettrait d’élargir notre compréhension des mécanismes macro-évolutifs dans la diversification des espèces en précisant l’effet joint de plusieurs facteurs biogéographique, morphologiques, et environnementaux dans les patrons de diversité actuels.



# APPENDIX – Shotgun versus RAD- sequencing for phylogenetic issues in Espeletiinae

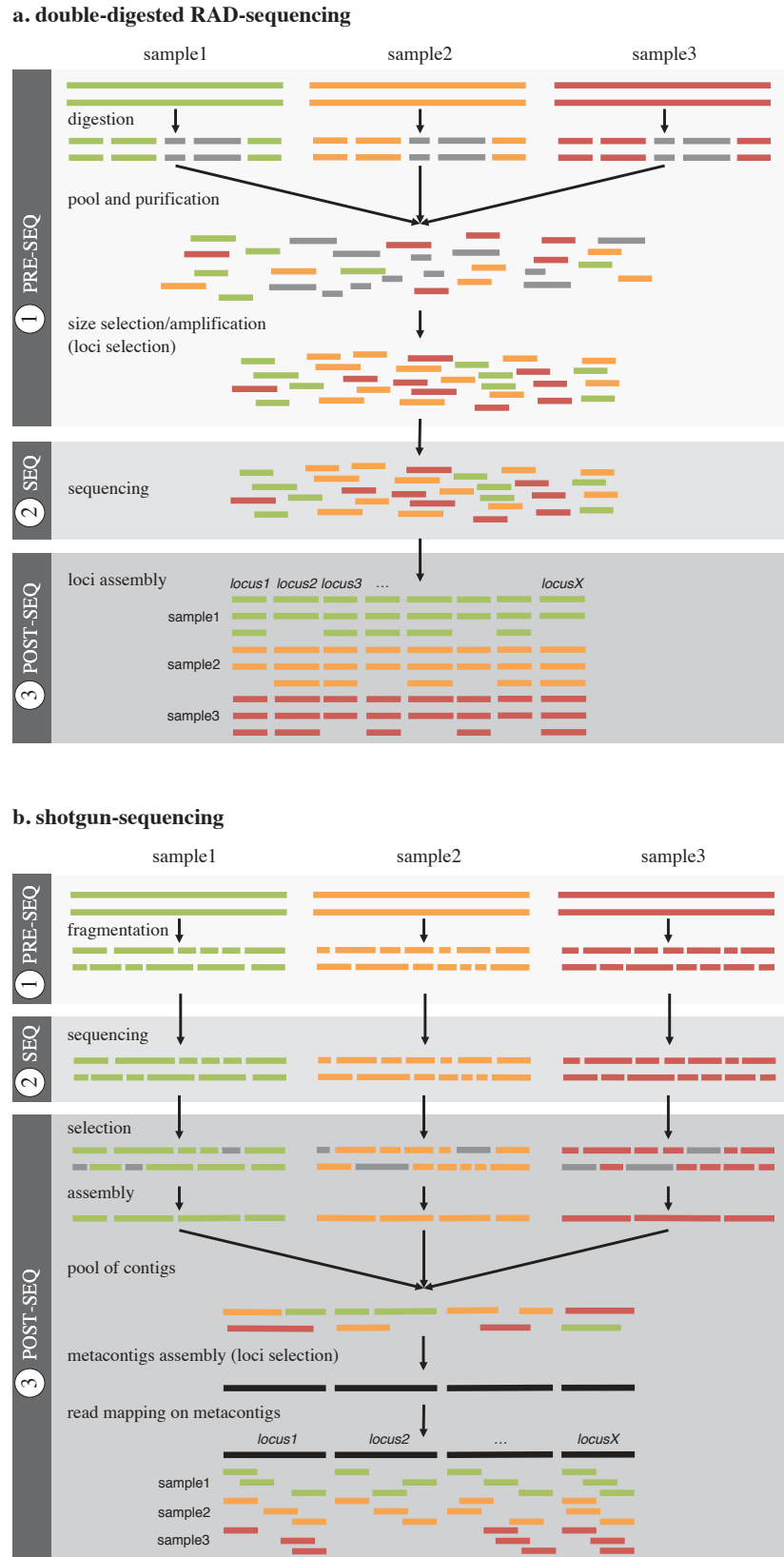
---

## 1 | INTRODUCTION

Capturing sufficient polymorphic signal from recent radiated and closely related species or from non-model organisms is a major challenge in the phylogenetic area owing to the lack of phylogenetic signal provided by the traditional markers and to the conflicting signal caused by short internodes (Whitfield and Lockhart, 2007). Within the last decade, the advent and the development of next-generation sequencing (NGS) offered unprecedented possibilities in phylogenomics by increasing greatly both taxonomic coverage and the amount of sequence data, while remaining at a relatively low cost. According to these technical developments, new bioinformatics advances such as the coding of fast and efficient genome assemblers, e.g. Velvet (Zerbino and Birney, 2008), ABySS (Simpson et al., 2009), have highly increased our ability to produce large sets of assembled contigs and even genomes (Moore et al., 2006; Schwartz et al., 2010; Zimmer and Wen, 2015) or to discover and genotype thousands of single nucleotide polymorphism (SNP) across taxa by mapping the sequenced reads against reference genomes (Zhang et al., 2011). For example, with the combined use of genome shotgun sequence data and bioinformatics tools we can now accurately hope to resolve phylogenetic relationships, even among closely related taxa (Herrera and Shank, 2016; Lemmon and Lemmon, 2013; Malé et al., 2014; McCormack et al., 2012; Straub et al., 2012).

Whole genome re-sequencing (WGR) approaches are powerful for addressing evolutionary biology issues by providing the highest genetic marker density of the current genomic methods (Fuentes-Pardo and Ruzzante, 2017). However, such techniques are still expensive in sequencing and computing resources and require the availability of reference genomes. Consequently, uses of WGR are limited in non-model species, for which reduced-representation sequencing methods, not depending on a reference genome, are preferred. According to Fuentes-Pardo and Ruzzante in their recent review (2017), these approaches can be classified in three groups: (i) Restriction-site Associated DNA sequencing (RAD-sequencing), (ii) Sequencing of transcribed DNA from the mRNA (RNA-sequencing) and (iii) Whole-Exome Sequencing (WES). These reduced-representation sequencing methods supply enough depth of coverage per locus with a greater number of taxa at a lower budget and do not require any prior genomic knowledge. Restriction site associated DNA sequencing (RAD-sequencing) has become the most used reduced-representation sequencing approach in the literature to generate thousands of SNPs in non-model species (Andrews et al., 2016). This method, including the “genotyping by sequencing” (GBS) technique, targets homologous sequences using specific restriction enzyme cutting sites through the genome (Baird et al., 2008; Peterson et al., 2012; Seeb et al., 2011). A huge number of RAD loci can then be obtained and if each of them is relatively short (~ 100-300 base-pair) with few SNPs involving limited phylogenetic information, when taken together, the phylogenetic signal reconstructed can be strong (Eaton and Ree, 2013; Vargas et al., 2017). Original RAD protocol used one restriction enzyme digestion followed by a physical shearing to reduce the length of fragments for sequencing (Baird et al., 2008). However, RAD-sequencing-related techniques have been developed more recently to simplify the original protocol with the use of two restriction enzymes to eliminate the mechanical shearing step (Peterson et al., 2012). This new protocol, called ddRAD-sequencing, for double digest RAD-sequencing, involves lots of pre-sequencing efforts to generate homologous loci, and consists of the following steps: double restriction enzymes digestion, adapter ligation using individual barcodes, size selection and amplification of specific fragments (Fig. 1a).

Pouchon et al. (2018) used a new phylogenomic approach, minimising the pre-sequencing sample preparation. In contrast to the reduced-representation sequencing methods using restriction enzymes, this method, based on shotgun sequencing, requires more sequencing effort to keep reliable read coverage and a more intensive post-sequencing bioinformatics step to discover homologous loci (Fig. 1b.). This method, which is close to that of Schwartz et al. (2015) and Fu et al. (2016), can be seen as an extending of genome skimming (Straub et al., 2012). Whole genomic DNA is directly individually sequenced after a fragmentation step to recover high-copy DNA regions. If genome skimming allows then access to organelle genomes or nuclear ribosomal DNA regions (i.e. internal transcribe spacer) using single sequencing run, only 10% of reads are usually used. In addition, the reconstructed phylogeny with chloroplast DNA or ribosomal clusters could still be non-informative in recent radiated species (Panero et al., 1999; Pouchon et al., 2018; Rauscher, 2002). The “extending” approach then allows the use of remaining reads to access to nuclear genome by assembling individual nuclear contigs which are then pooled and assembled into meta-contigs in order to discover homologous regions, used as chimeric reference for the read mapping (Pouchon et al., 2018).



**Figure 1.** Diagrams of different laboratory and bioinformatics steps for assembly of homologous loci for (a) ddRAD-seq and for (b) shotgun sequencing as performed in Pouchon et al. 2018. Steps were organised into (1) pre-sequencing, (2) sequencing and (3) post-sequencing efforts.

This shotgun procedure allowed to describe the phylogeny of the entire morphological

groups in the complex of Espeletiinae, a rapid radiation, providing fully resolved phylogenies (see Chapter 1 and 2) where traditional markers failed (Diazgranados and Barber, 2017; Panero et al., 1999; Rauscher, 2002). This approach used shotgun sequencing to discover random homologous loci (Pouchon et al., 2018). By contrast, RAD-sequencing targets specific loci using restriction sites. If phylogenies are difficult to estimate in rapid radiation, RAD-sequencing also appeared strongly efficient to puzzle out poorly resolved relationships (Wagner et al., 2013), especially within Andean taxa in the *Diplostephium* radiation (Vargas et al. 2017). These new techniques providing high-resolution genomic data were often compared with plastid or traditional nuclear markers (Cruaud et al., 2014; Fu et al., 2016; Pouchon et al., 2018; Vargas et al., 2017). However, only few studies compared them with nuclear markers obtained from reduced-representation methods on phylogenomic context (Harvey et al., 2016; Rodríguez et al., 2017). Here, we propose to evaluate the power of both shotgun- and ddRAD-sequencing methods to infer the phylogeny of Espeletiinae.

## 2 | METHODS

### Taxonomic Sampling

Sampling was focused on selected species from the shotgun-based phylogenies mentioned in the previous chapter, by using two individuals from each species and by adding some Colombian species. The majority of the samples were collected in the field where leaf tissue was dried in silica gel. Sampling accounts for 24.2% of the Colombian and 94.4% of the Venezuelan recognized species (Cuatrecasas, 2013; Diazgranados, 2012).

DNA extraction in Espeletiinae is relatively arduous due to the heavy indumentum of the leaves. Genomic DNA was extracted from dried leaves tissue, after shaving the tissue part of interest and grinding them with the TissueLyser (Qiagen Corp.), during 3h for the cell lysis step to preserve the DNA quality by using the DNeasy Plant Mini Kit (Qiagen Corp). According to previous experiments, DNA of Espeletiinae contained high level of inhibitors for enzyme activity (i.e. terpenoids), which prevent restriction digestion of ddRAD-sequencing protocol (also underlined in (Diazgranados and Barber, 2017)). Following Mastretta-Yanes et al. (2015) recommendations, DNA cleaning was then performed using phenol chloroform before the preparation of libraries to discard any enzymatic inhibitors.

### ddRAD-sequencing Library preparation

Double-digested RAD-sequencing was performed according to a modified version of the original protocol described in Peterson et al. (2012). Briefly, restriction digestion was done first using 10 units of PstI and MspI restriction endonucleases (NewEngland Biolabs, Inc.) during 3 h at 37°C with 250 ng of purified genomic DNA in a final volume of 36 µL using the CutSmart buffer provided with the enzymes. Adapters containing sample-specific barcodes and TruSeq Illumina primers (Illumina Inc.) were next ligated to the DNA thanks to 10 units of T4 DNA ligase (NewEngland Biolabs, Inc), 1 µL of 10 mM ribo-ATP (NewEngland Biolabs, Inc). Digestion was continued simultaneous to the ligation step by adding 0.5 µL of

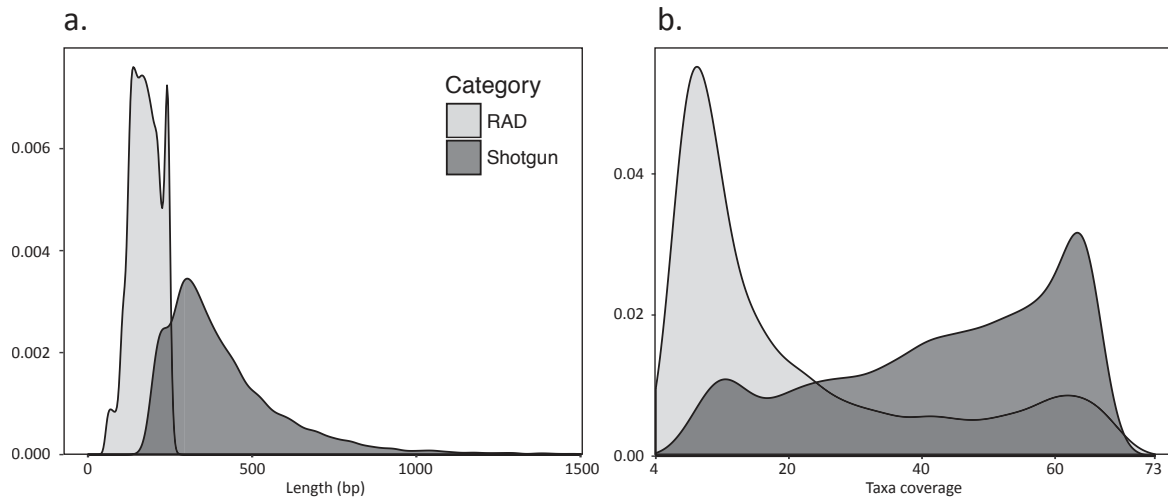
**Table 1.** Resuming alignments and phylogenetic characteristics of different matrices. Three alignments were made for RAD-sequencing data according to the clustering thresholds values in pyRAD. Shotgun alignment was recovered from the Chapter 3. For all alignments, a minimum of four taxa was set to consider locus in the final matrix.

Phylogenetic Datasets	ddRAD-clust80	ddRAD-clust85	ddRAD-clust90	Shotgun
Samples (species)	73 (70)	73 (70)	73 (70)	65 (61)
Assembled loci	8,056	8,429	8,903	9,585
Length of alignment (bp)	1,402,181	1,465,761	1,549,823	4,025,033
Ambiguous characters and missing data (%)	70.23	70.52	71.48	46.59
Polymorphic sites (%)	5.61	5.28	4.80	11.68
Informative sites (%)	2.17	2.02	1.78	5.94
Phylogenetic ML inference				
PPmean (sd)	92.44 (13.17)	92.60 (11.92)	99.55 (1.26)	96.43 (9.65)
PP < 50 (%)	1 (1.43)	0 (0.00)	0 (0.00)	0 (0.00)
PP < 90 (%)	18 (22.50)	19 (27.14)	20 (28.57)	7 (11.29)

each digestion restriction enzymes and performed under 60 cycles of 2 min at 16° for the ligation and 37°C for the digestion. Enzymes were next heat-inactivated at 65 °C for 10 min. Samples were then pooled in equivolume mixture and purified with Agencourt AMPure XP beads (Beckman Coulter, France). Fragments in the 250- to 500 bp size range, including adapters, were excised from agarose gel (1.6%) and purified using the QIAquick Gel Extraction Kit (Qiagen Corp.). Libraries were then PCR-amplified in 20 µL final reaction volume with 1 µL of DNA template, 10 mM of dNTPs, 10 µM of each PCR probe (Peterson et al. 2012) and 2 U/ µL of Taq Phusion-HF (New England Biolabs Inc.). PCR were performed with an initial denaturation at 98 °C for 10s followed by 15 cycles of 98°C for 10s, 68°C for 30s and 72°C for 1 min; and ended by a final extension at 72 °C for 10 min. Finally, products from each library were purified with with QIAgen MinElute PCR Purification Kit (Qiagen Corp.), controlled in agarose gel migration, quantified using Qubit (Life Technologies) and finally pooled in equimolar amounts. Paired-end 2x150bp sequencing was performed on Illumina HiSeq 2500 rapid run flow cell for each library at FASTERIS (Geneva, Switzerland).

## Post-sequencing processing

Sequencing adapters were identified and trimmed from reads, then reads of unexpected sizes (fewer than 150bp) were removed, using cutadapt v1.13 (Martin, 2011). The *process\_radtags* function of STACKS (Catchen et al., 2013) was run on each library to demultiplex and filter the cleaned reads based on their quality with a sliding window of 15% of the read length and a Phred score limit of 20. Reads with low quality or uncalled bases were removed. Paired-end reads were next merged using FLASH v1.0.2 (Magoč and Salzberg, 2011) or concatenated when they could not be aligned. *De novo* clustering of RAD loci was done using pyRAD (Eaton, 2014) because of its capacity to deal with insertion-deletion variation among loci, which improved identification of homology across divergent samples. This was accomplished using an alignment-clustering method to perform both within- and across-sample clustering. In addition to the Long Branch Attraction (LBA) effect generated by missing data on the concatenated supermatrix (Liu et al., 2015; Xi et al., 2016), incomplete characters in close

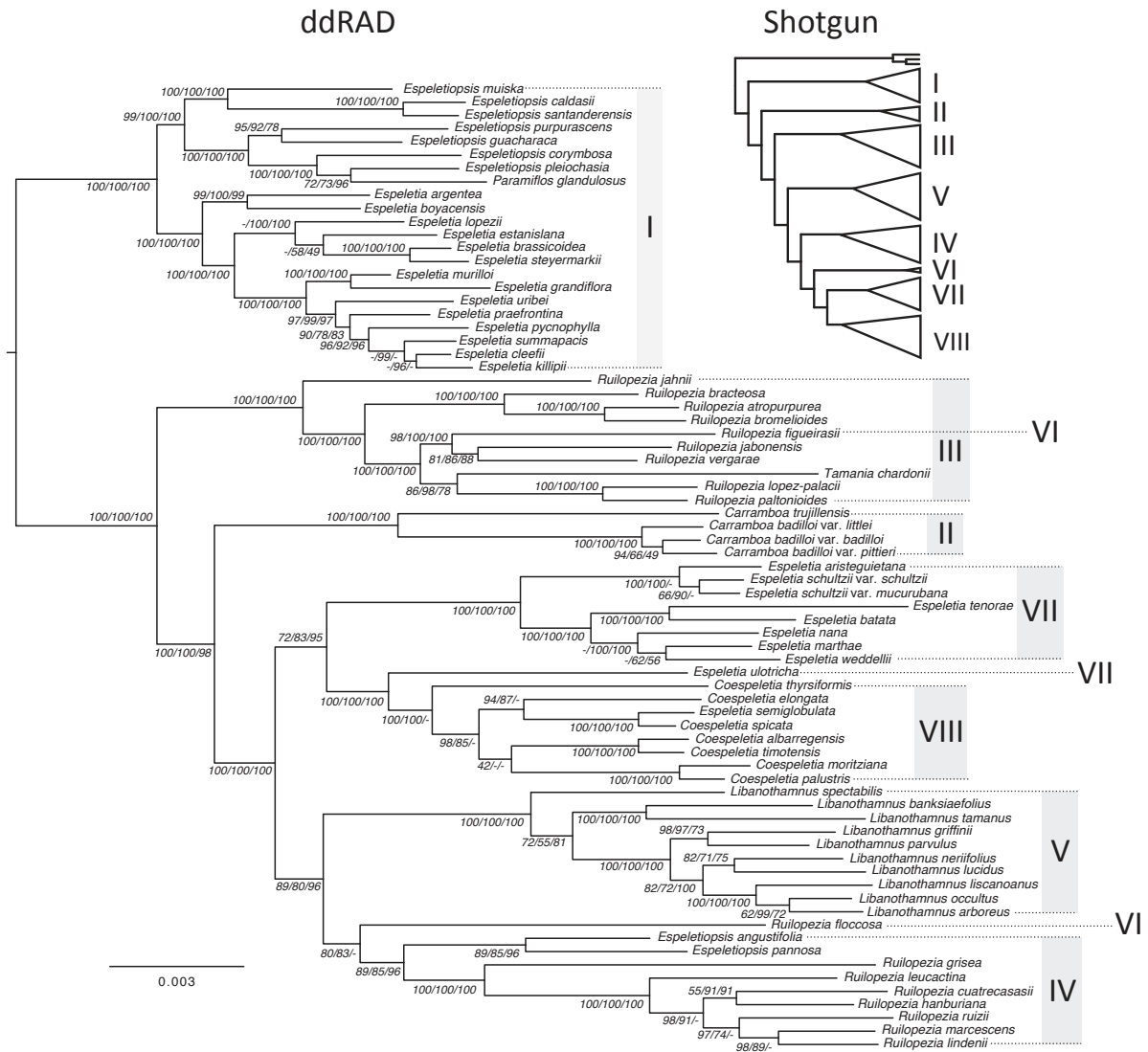


**Figure 2.** Density of the (a) length and the (b) taxa coverage of assembled loci from RAD and shotgun approaches.

outgroups can have an impact on phylogenetic accuracy (Roure et al., 2013). Because of the absence of taxa recovery between ingroup and outgroup species in this study, we did not consider outgroup sequences for the further processes. The reads from ingroup species were first clustered by sample according to a similarity threshold using VSEARCH v2.1.1 program (Rognes et al., 2016) and then aligned with MUSCLE v3.8.31 (Edgar, 2004). Sequencing error rate and heterozygosity were jointly estimated and used to generate consensus sequences for each cluster, which were clustered across samples under the same procedure and parameters. We tested a range of clustering thresholds (clust-80, clust-85 and clust-90%) with a minimum of four taxa that had to be shared at each locus. Loci sharing heterozygosity in at least 15% of taxa at a given site were considered paralogs and were discarded. The minimum depth of coverage required to build a cluster was fixed at 10. Finally, three aligned data matrices were generated.

## Data Exploration and Phylogenetic Inference

For all datasets, the proportion of loci shared among individuals was estimated in R using *RADami* package (Hipp et al., 2014). Phylogenetic inference was calculated using the Maximum Likelihood (ML) approach implemented in RAxML v8.1.3 (Stamatakis, 2006) on the concatenated loci. ML searches were conducted under the GTRCAT substitution model and according to rapid bootstrapping analysis using 500 bootstraps replicates. Moreover, in order to assess for the sensitivity of the concatenated approach to conflicts among gene trees due to ILS, phylogenetic reconstructions were also estimated by a coalescent-based multispecies tree approach as implemented in ASTRAL II (Mirarab and Warnow, 2015), as made for shotgun phylogenies in Pouchon et al. (2018). Here, 8,429 gene trees were first estimated using RAxML under the GTRCAT model with 200 bootstrap replicates performed on the alignment of each RAD-loci (corresponding to the clust85 dataset in Table 1). To assess for gene conflicts among them, we used the multilocus bootstrapping procedure from ASTRAL II with 100 replicates combining each RAxML best tree and its bootstrap replicates.



**Figure 3.** Phylogeny obtained from the ddRAD-sequencing concatenated alignments under Maximum Likelihood inference performed in RAxML with 500 bootstrap replicates. Bootstrap support was shown for nodes for the three matrices according to their clustering thresholds clust80/clust85/clust90. Backbone topology of the shotgun phylogeny was extracted from the Chapter 3 and compared to the one obtained with the clustering threshold of 85%.

## 3 | RESULTS

### Datasets characteristics

Assemblies of RAD-seq loci according to the three different clustering thresholds comprised between 1,402,181 and 1,549,823 nucleotides (Nt) from the concatenation of respectively 8,056 to 8,903 loci (Table 1). The number of loci included in the shotgun analysis (from Chapter 3) was somewhat similar (with 9,585 loci). However, the size of the alignments was reduced in the RAD-seq datasets due to the shorter assembled loci provided by the read sequencing size of target regions (Fig. 2). Missing data was consistent among the RAD-seq

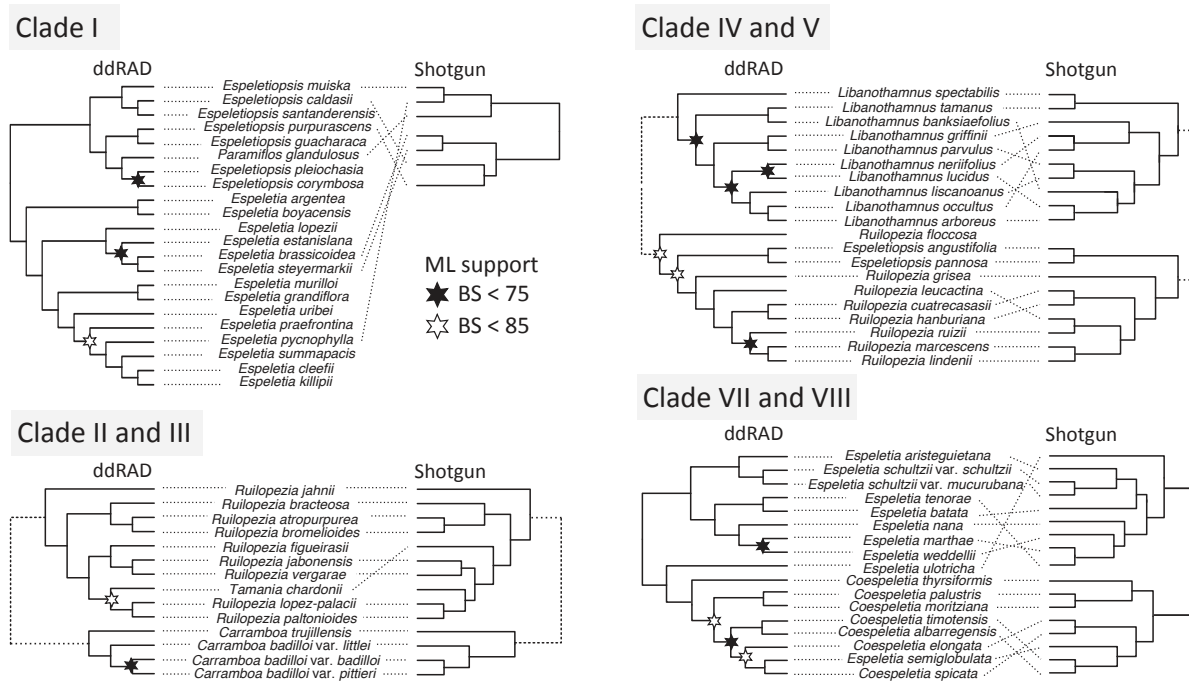
datasets but appeared more important than for shotgun analysis. Shotgun loci comprised higher taxa coverage than RAD-seq loci (Fig. 2). Finally, the RAD-seq matrices, with 27,586 to 30,427 Nt, included about 100 times less informative sites in the recovered alignment than the shotgun one, which comprised nearly 239,086 informative sites (Table 1).

## Phylogenetic inferences and Loci Representation

ML phylogenies made on the three RAD-seq concatenated alignments were highly congruent and well supported (Table 1, Fig. 3). A single topology was therefore shown in the Figure 3. Divergence of taxa was first highlighted between the Colombian and the Venezuelan species, corroborating the pattern observed in Pouchon et al. (2018) and Diazgranados and Barber (2017) (clade I in Fig. 3). In contrast to the shotgun phylogeny, the monophyletic clade of the Colombian species was more represented with 22/91 sampled species (instead of 7/91), allowing new phylogenetic interpretations. The topology showed that *Espeletia* and *Espeletiopsis* formed monophyletic groups in Colombia with *Paramiflos glandulosus* closely related to the *Espeletiopsis* (Fig. 3). For the Venezuelan species, with the exception of the shotgun clade VI (represented only by *Ruilopezia figueirasii* and *R. floccosa*), we found the same other clades but incongruences in the backbone topology. In fact, clade III of *Ruilopezia* including *Tamania chardonii*, appeared at the basal position of the Venezuelan radiation instead of the following diverged *Carramboa* species (clade II). In addition, *R. figueirasii* took place in this *Ruilopezia* clade. Afterwards, in contrast to the shotgun topology, the four others clades were arranged in two groups of two sister clades (Fig. 3). Clade V of *Libanothamnus* that was closely related to clade IV of *Espeletiopsis/Ruilopezia* on the shotgun topology was sister of this clade on the RAD-seq phylogenies. Change occurred with the position of *R. floccosa*, now appearing at the root of clade IV. Finally, close phylogenetic relationships between clade VII and VIII found on shotgun analysis were corroborated in this topology. Intra-clade comparisons were illustrated in Figure 4. Omitting the Colombian clade (I), which was not sufficiently represented for shotgun analysis, we showed that only relationships in clades II and III were stable between the two approaches (Fig. 4). Regarding unstable positions, only *R. leucactina* was unstable in clade IV, and *E. semiglobulata* in clade VIII. For others, most of the differences could be explained by the low supports recovered in the RAD-sequencing phylogeny (Fig. 3 and 4). Furthermore, species tree topology inferred in ASTRAL had lower node support values in comparison to the concatenated topology (Supp. Fig. 1). This topology involved mostly intra-clade changes, and principally one inter-clade change at the base of the Venezuelan radiation with an unresolved position of *Carramboa* and the first clade of *Ruilopezia* (BS of 63).

Finally, by examining the loci shared among species across both phylogenies, we showed great differences between shotgun and RAD-sequencing methods. For the RAD-sequencing approach, a small number of loci (10-20%) was shared between species (Fig. 5a). Such pattern was recovered regardless of the clustering threshold choice (Supp. Fig. 2). On the other hand, for the shotgun approach, species shared in average between 40 to 60% of homologous loci (Fig. 5b). Outgroup species shared very few loci with ingroup taxa and some clusters of loci were shared mostly between Venezuelan or Colombian species. This result



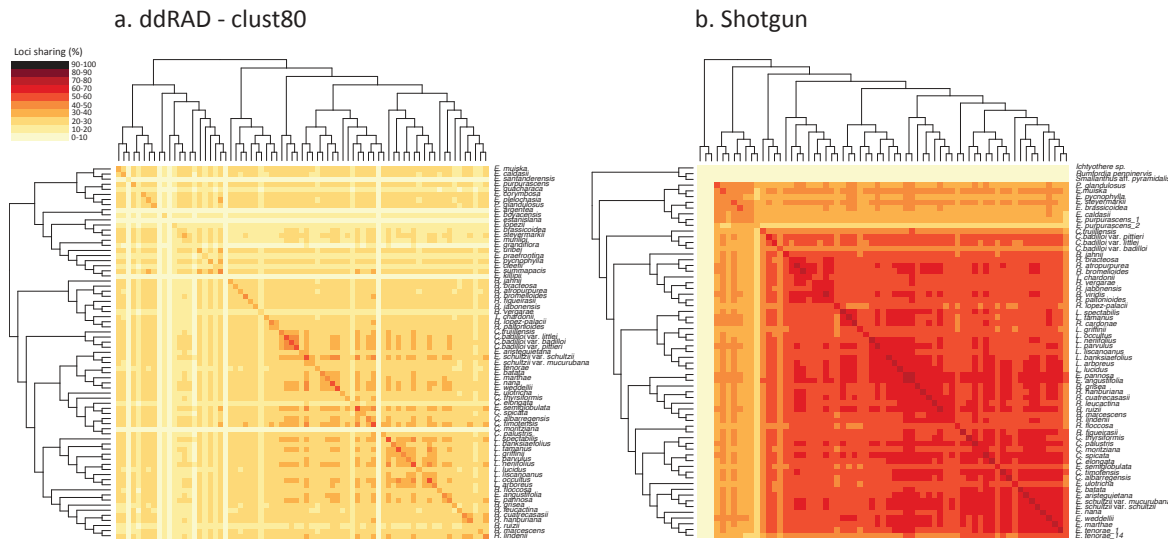


**Figure 4.** Intraclade comparisons of ddRAD and shotgun based phylogenies. Topology for the ddRAD phylogeny corresponds to the one obtained with a clustering threshold of 85%.

corroborates the higher taxa coverage highlighted by the shotgun approach in Fig. 2b leading to fewer missing data into the final supermatrix (Table 1).

## 4 | DISCUSSION

The advent of NGS technologies has revolutionized the field of phylogenomics, with reduced-representation methods providing high-resolution genomic data for non-model organisms at a reasonable cost. These techniques provide unprecedented opportunities to resolve phylogenetic relationships between species using RAD-sequencing (Eaton and Ree, 2013; Vargas et al., 2017; Wagner et al., 2013), RNA-sequencing (Rodríguez et al., 2017), sequences capture (Leaché et al., 2014; McCormack et al., 2012) or multiplexed-shotgun sequencing (Fu et al., 2016; Pouchon et al., 2018; Wessinger et al., 2016). However, this increase in genomic loci data also brings some biases with conflicting branching patterns on these loci due to incomplete lineage sorting (ILS), gene flow or gene duplication/loss, especially when assessing node support, through bootstrapping and Bayesian posterior probabilities, which tend to be very high with large amounts of data. Artefacts could lead to an increase of node probabilities because of missing data rather than real phylogenetic signal (Dell’Ampio et al., 2014). Wrong phylogenies could thus be highly supported. Here, this study seemed to support this trend.



**Figure 5.** Proportion of loci shared among species for all the pairwise comparison for (a) ddRAD-sequencing loci built with a clustering threshold of 80% similarity on pyRAD (Eaton 2014) and for (b) shotgun-sequencing loci from the Chapter 3.

## Phylogenetic conflicting relationships

Phylogeny made on RAD-sequencing approach appeared well supported (Fig. 3) but contradicted the backbone topology inferred on shotgun sequencing (Chapter 1 and 2). However, RAD-loci were very few and less shared between taxa (Fig. 2b and Fig. 5), generating an important level of missing data (Table 1), and were less informative than shotgun-loci (roughly with a factor of 100, respectively  $\sim 30,000$  to 239,086 Nt; Table 1). Closely related species shared more loci than others (Fig. 4). Differences in topology could probably be artifactual due to the few informative sites represented on the backbone topology in spite of node support probabilities. To evaluate this “big data” effect, (Irisarri et al., 2017) and Rodríguez et al. (2017) proposed resampling loci rather than individual nucleotide sites in a concatenated analysis, which could be more informative than traditional bootstrap or posterior probability analyses, using loci jackknifing. This approach could be used here to untangle artifactual from true phylogenetic signal.

Furthermore, the loci concatenation method was recently debated in the literature (Edwards et al., 2016; Springer and Gatesy, 2016). This approach may be biased in presence of incomplete lineage sorting (ILS), gene flow or gene duplication/loss (Springer and Gatesy, 2016), which could be especially important with the increase in loci supplied by NGS. To cope with this issue, methods based on multispecies coalescence have been recently developed to reconstruct the evolutionary history of species in the presence of gene-tree discordance (Mirarab et al., 2014; Solís-Lemus et al., 2016). Most of the phylogenomic studies now compare traditional concatenation to multispecies coalescent approaches (Nevado et al., 2016; Pouchon et al., 2018; Vargas et al., 2017). Here, by investigating species tree inferences accounting for ILS, we showed that ASTRAL-tree highlighted a polytomy on the basal position of the Venezuelan lineages in comparison to the concatenated tree involving the *Carramboa* and *Ruilopezia* clades (Supp. Fig. 1). This validated a lack of phylogenetic

signal and a contrasting pattern on RAD-loci supplied by strong ILS on this backbone position. On the other hand, the concatenated approach appeared more similar to the multispecies coalescent approach in the shotgun dataset (Pouchon et al., 2018). Regarding the phylogenetic relationships, the basal difference in the Venezuelan section of these plants highlighted by the RAD-concatenation tree was consequently not supported. The remaining main backbone differences with the shotgun approach concerned the *Espeletiopsis/Ruilopezia* (clade IV) and the *Libanothamnus* (clade V), which appeared more closely related in contrast to Pouchon et al. 2018 (Fig. 3, Supp. Fig. 1). However, this relationship was not fully supported by the RAD-concatenated tree (BS < 85), and the *Libanothamnus* appeared unstable in shotgun phylogenies with the increase of loci in the phylogenetic inference compared to clade IV (Chapter 2). A doubt could be issued as to the exact position of the *Libanothamnus* within this Venezuelan section. Other changes concerned intra-clade relations involving low node supports (Fig. 3 and 4), and the position of *R. figueirasii*/*R. floccosa*. These two species, endemic of central part of the cordillera de Merida, were monophyletic in the shotgun study (clade VI; Pouchon et al., 2018). Here, *R. figueirasii* appeared close to the northern *Ruilopezia* (*R. vergarae* and *R. jabonensis* of clade III), whereas *R. floccosa* was close to the southern *Ruilopezia* (clade IV; BS < 85). Further investigations will consequently be necessary to assess the position of these species, by adding more individuals into the sampling. Finally, some phylogenetic differences were highlighted in clades VII and VIII. These species, genetically close, are also geographically close (Cuatrecasas, 2013). In addition, important events of introgression were shown within and between these two clades (Pouchon et al. 2018, Chapter 3). Gene flows, expected to generate phylogenetic conflicts, could probably explain such changes. As applied in Pouchon et al. (2018), the SNaQ approach, developed by Solís-Lemus et al. (2017), could be used in order to identify conflicting positions in the RAD-based phylogeny due to hybridisation.

## Evolutionary reconstructions agreement

In spite of these topological differences, both RAD- and shotgun-sequencing approaches corroborated the same scenarios of geographical and morphological evolution previously hypothesised (Chapter 1 and 2).

Regarding the geographical component, the two independent radiations occurring at the origin of the group in both Colombia and Venezuela were recovered. This finding corroborates the results of Diazgranados and Barber (2017), and of Pouchon et al. (2018). The main clades decrypted in the previous studies (except for clade VI) were also recovered, although inter-clade relationships were not totally similar. Shallow relationships were consequently better estimated than deeper nodes by RAD-sequencing methods. In addition, the similarities observed support the assumptions regarding the morphological evolution described in Chapter 1 and 2. For example, by investigating ancestral state reconstructions as performed in Pouchon et al. (2018), we showed that the crown ancestor of the subtribe for both radiations remained a rosette with imbricated open/flat leaf sheaths (Supp. Fig. 3). Only the evolution of the ramification capacity changed from the previous reconstructions (Supp. Fig. 3). The Colombian radiation, which is now more completely understood, is exclusively composed of unbranched rosettes species (at the exception of *Espeletia pleiochasia*). Such

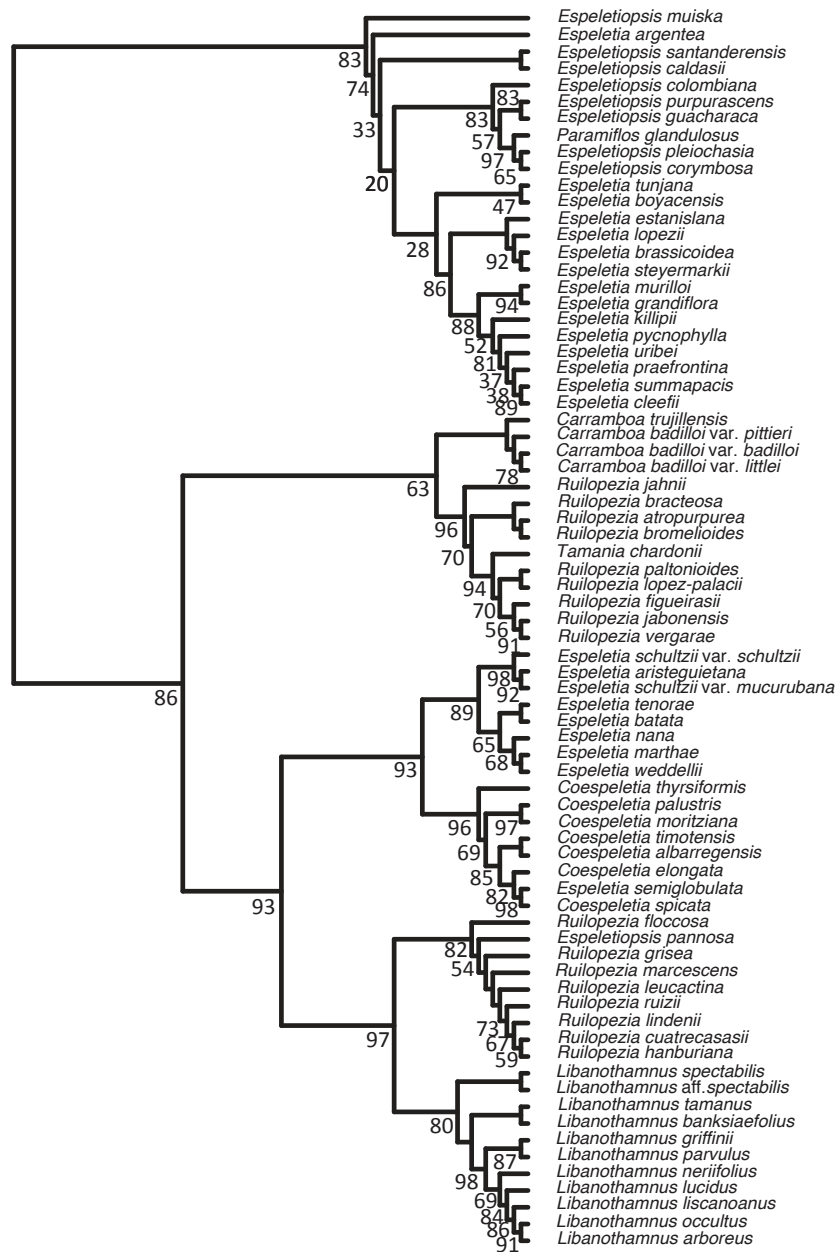
pattern could probably explain the change in ancestral state reconstruction for the crown ancestor. Nevertheless, as previously published, the tree growth-form evolved independently three times from the repeated colonization of lower elevations by high-elevation rosette ancestors in the Venezuelan section.

## 5 | CONCLUSION

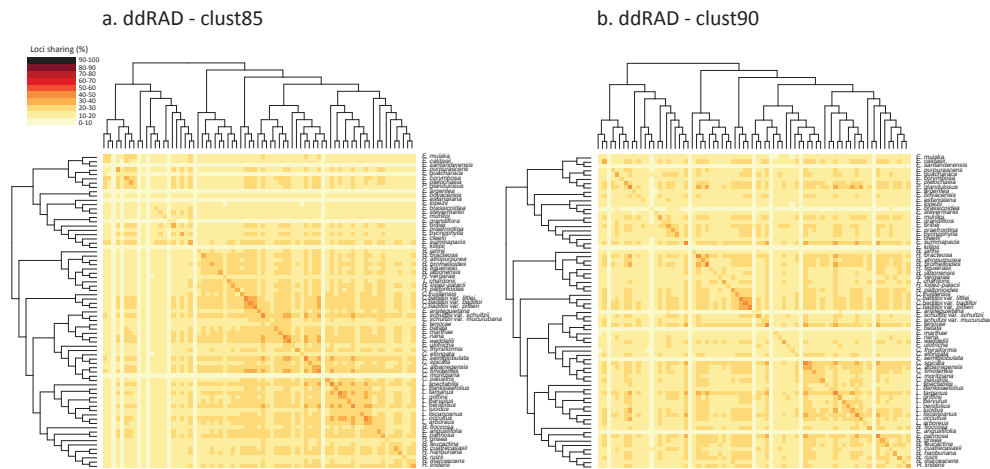
To conclude, the multiplexed shotgun-sequencing methods described in Pouchon et al. (2018) appeared more effective to infer phylogenetic relationships within the radiation of Espeletiinae. Although it was more expensive, it allowed to recover both deep and shallow resolutions of the phylogeny of these species (Pouchon et al. 2018), as well as using degraded tissues, as opposed to the ddRAD-sequencing approach (Chapter 2). As DNA extraction of Espeletiinae species was arduous (see Methods section), such a benefit is not negligible. Moreover, shotgun sequencing allowed us to access the organelle genomes (Chapter 1), which was very useful in molecular dating of deep divergence. On the other hand, RAD-sequencing seemed not to provide enough phylogenetic signal to resolve all the radiation, especially on deep nodes, even though the same evolutionary scenarios for morphological trait evolution were recovered with both methods. This lack of signal, in contrast to the shotgun method, was also demonstrated in Chapter 3 when testing SNPs configuration in Paterson's D-test to detect hybridisation footprints in *Espeletia semiglobulata*. Nevertheless, this method was very useful at population levels to estimate the structuration of populations, where more individuals could be included in the sequencing runs (Chapter 3). Additional sequencing effort in RAD libraries will probably overcome this issue in the phylogenetic inferences regarding the radiation of the Espeletiinae.

## 6 | SUPPLEMENTARY INFORMATIONS

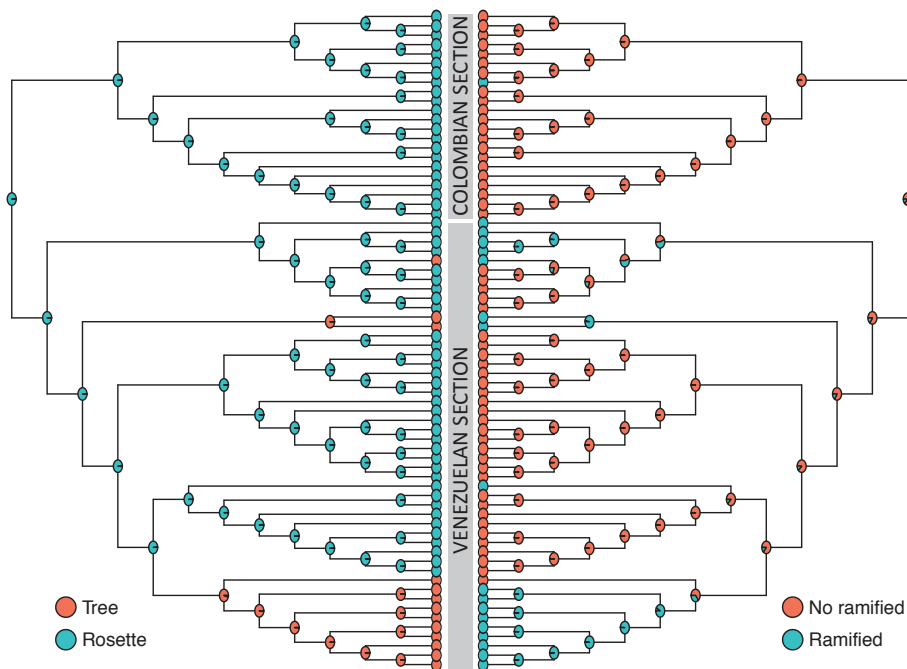
**Supp. Figure 1.** Species tree reconciliation based on the 8,429 gene trees inferred for the RAD-loci. Gene trees were estimated from RAxML (Stamatakis 2006) and the species tree was estimated using ASTRAL-II (Mirarab and Warnow 2015) using a bootstrapping procedure. Only bootstrap probabilities < 100 were drawn for the nodes.



**Supp. Figure 2.** Proportion of loci shared among species for all the pairwise comparisons for ddRAD-sequencing loci built with a clustering threshold of (a) 85% similarity and (b) 90% similarity on pyRAD (Eaton 2014).



**Supp. Figure 3.** Reconstruction of ancestral states of the growth form trait (GF) and the ramification capacity (STEM) in Espeletiinae. Inferences were done using Markov models (Mk) with the R package Diversitree (FitzJohn 2012). Two possible discrete Markov models were fitted and compared using AIC values: a full model with different transition rates (ARD) and a constrained model with equal transition rates (ER). The best model (ER for both traits) was used to run simulations of stochastic character mapping in order to infer the most likely scenarios of trait evolution.



## Literature cited

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., et al. (2013). Hybridization and speciation. *J. Evol. Biol.* *26*, 229–246.

Aberer, A.J., Kobert, K., and Stamatakis, A. (2014). ExaBayes: massively parallel bayesian tree inference for the whole-genome era. *Mol. Biol. Evol.* *31*, 2553–2556.

Abrahamczyk, S., Souto-Vilarós, D., and Renner, S.S. (2014). Escape from extreme specialization: passionflowers, bats and the sword-billed hummingbird. *Proc R Soc B* *281*, 20140888.

Acosta, M.C., and Premoli, A.C. (2010). Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, *Nothofagaceae*). *Mol. Phylogenet. Evol.* *54*, 235–242.

Agrawal, A.A., Fishbein, M., Halitschke, R., Hastings, A.P., Rabosky, D.L., and Rasmann, S. (2009). Evidence for adaptive radiation from a phylogenetic study of plant defenses. *Proc. Natl. Acad. Sci.* *106*, 18067–18072.

Aguilée, R., Claessen, D., and Lambert, A. (2012). Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. *Evol. Int. J. Org. Evol.* *67*, 1291–1306.

Alexander, D.H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* *12*, 246.

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* *19*, 1655–1664.

Almén, M.S., Lamichhaney, S., Berglund, J., Grant, B.R., Grant, P.R., Webster, M.T., and Andersson, L. (2016). Adaptive radiation of Darwin's finches revisited using whole genome sequencing. *BioEssays News Rev. Mol. Cell. Dev. Biol.* *38*, 14–20.

Álvarez, M.J.M., Hurtado, J., Armas, J.F.S., G, M.R.F., and Azócar, A. (1990). Algunas evidencias poblacionales de formación de híbridos entre *espeletia batata* cuartr. y *E. schultzi* Wedd. en el altiano venezolano. *Ecotrópicos* *3*, 57–66.

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., and Hohenlohe, P.A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* *17*, 81–92.

- Antonelli, A., and Sanmartín, I. (2011a). Why are there so many plant species in the Neotropics? *Taxon* *60*, 403–414.
- Antonelli, A., and Sanmartín, I. (2011b). Mass Extinction, Gradual Cooling, or Rapid Radiation? Reconstructing the Spatiotemporal Evolution of the Ancient Angiosperm Genus *Hedyosmum* (Chloranthaceae) Using Empirical and Simulated Approaches. *Syst. Biol.* *60*, 596–615.
- Antonelli, A., Nylander, J.A.A., Persson, C., and Sanmartín, I. (2009). Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proc. Natl. Acad. Sci.* *106*, 9749–9754.
- Aubert, S., Boucher, F., Lavergne, S., Renaud, J., and Choler, P. (2014). 1914–2014: A revised worldwide catalogue of cushion plants 100 years after Hauri and Schröter. *Alp. Bot.* *124*, 59–70.
- Baack, E., Melo, M.C., Rieseberg, L.H., and Ortiz-Barrientos, D. (2015). The origins of reproductive isolation in plants. *New Phytol.* *207*, 968–984.
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., and Johnson, E.A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLOS ONE* *3*, e3376.
- Baldwin, B.G., and Wagner, W.L. (2010). Hawaiian angiosperm radiations of North American origin. *Ann. Bot.* *105*, 849–879.
- Balslev, H. (1988). Distribution Patterns of Ecuadorean Plant Species. *Taxon* *37*, 567–577.
- Barker, M.S., Arrigo, N., Baniaga, A.E., Li, Z., and Levin, D.A. (2016). On the relative abundance of autopolyploids and allopolyploids. *New Phytol.* *210*, 391–398.
- Barnosky, A.D. (2001). Distinguishing the Effects of the Red Queen Court Jester on Miocene Mammal Evolution in the Northern Rocky Mountains. *J. Vertebr. Paleontol.* *21*, 172–185.
- Barrera-Guzmán, A.O., Aleixo, A., Shawkey, M.D., and Weir, J.T. (2017). Hybrid speciation leads to novel male secondary sexual ornamentation of an Amazonian bird. *Proc. Natl. Acad. Sci.* 201717319.
- Barrier, M., Baldwin, B.G., Robichaux, R.H., and Purugganan, M.D. (1999). Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Mol. Biol. Evol.* *16*, 1105–1113.
- Barton, N.H. (2001). The role of hybridization in evolution. *Mol. Ecol.* *10*, 551–568.
- Baruch, Z., and Smith, A.P. (1979). Morphological and physiological correlates of niche breadth in two species of *Espeletia* (Compositae) in the Venezuelan Andes. *Oecologia* *38*, 71–82.
- Bayzid, M.S., and Warnow, T. (2013). Naive binning improves phylogenomic analyses. *Bioinforma. Oxf. Engl.* *29*, 2277–2284.
- Beck, E., Senser, M., Scheibe, R., Steiger, H.-M., and Pongratz, P. (1982). Frost avoidance and freezing tolerance in Afroalpine “giant rosette” plants. *Plant Cell Environ.* *5*, 215–222.
- Bell, G. (2016). Experimental macroevolution. *Proc. R. Soc. B Biol. Sci.* *283*.
- Bell, C.D., Calderon, G., Gonzalez, L., Scholz, A., and Liede-Schumann, S. (2015). Resolving Relationships within Valerianaceae (Dipsacales): New Insights and Hypotheses from Low-Copy Nuclear Regions. *Syst. Bot.* *40*, 327–335.



- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Benton, M.J. (2009). The Red Queen and the Court Jester: Species Diversity and the Role of Biotic and Abiotic Factors Through Time. *Science* 323, 728–732.
- Bermúdez, M.A., Kohn, B.P., Beek, P.A. van der, Bernet, M., O’Sullivan, P.B., and Shagam, R. (2010). Spatial and temporal patterns of exhumation across the Venezuelan Andes: Implications for Cenozoic Caribbean geodynamics. *Tectonics* 29.
- Berry, P.E., and Calvo, R.N. (1989). Wind Pollination, Self-Incompatibility, and Altitudinal Shifts in Pollination Systems in the High Andean genus *Espeletia* (Asteraceae). *Am. J. Bot.* 76, 1602–1614.
- Berry, P., Z, S.B., and M, R.C. (1988). La hibridización en la evolución de los frailejones ( *espeletia*, *astiraceae* ). *Ecotrópicos* 1, 11–24.
- Blonder, B., Baldwin, B.G., Enquist, B.J., and Robichaux, R.H. (2016). Variation and macroevolution in leaf functional traits in the Hawaiian silversword alliance (Asteraceae). *J. Ecol.* 104, 219–228.
- Bouchenak-Khelladi, Y., Onstein, R.E., Xing, Y., Schwery, O., and Linder, H.P. (2015). On the complexity of triggering evolutionary radiations. *New Phytol.* 207, 313–326.
- Boucher, F.C., Thuiller, W., Roquet, C., Douzet, R., Aubert, S., Alvarez, N., and Lavergne, S. (2012). Reconstructing the origins of high-alpine niches and cushion life form in the genus *Androsace* S.L. (Primulaceae). *Evol. Int. J. Org. Evol.* 66, 1255–1268.
- Bragg, J.G., Potter, S., Bi, K., and Moritz, C. (2016). Exon capture phylogenomics: efficacy across scales of divergence. *Mol. Ecol. Resour.* 16, 1059–1068.
- Brawand, D., Wagner, C.E., Li, Y.I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A.Y., Lim, Z.W., Bezault, E., et al. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513, 375–381.
- Brennan, A.C., Barker, D., Hiscock, S.J., and Abbott, R.J. (2012). Molecular genetic and quantitative trait divergence associated with recent homoploid hybrid speciation: a study of *Senecio squalidus* (Asteraceae). *Heredity* 108, 87–95.
- Briceño, B., and Morillo, G. (2002). CATÁLOGO ABREVIADO DE LAS PLANTAS CON FLORES DE LOS PÁRAMOS DE VENEZUELA. PARTE I. DICOTILEDÓNEAS (MAGNOLIOPSIDA). *Acta Botánica Venezolánica* 25, 1–46.
- Brown, G.K., Nelson, G., and Ladiges, P.Y. (2006). Historical biogeography of *Rhododendron* section *Vireya* and the Malesian Archipelago. *J. Biogeogr.* 33, 1929–1944.
- Bruun-Lund, S., Clement, W.L., Kjellberg, F., and Rønsted, N. (2017). First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). *Mol. Phylogenet. Evol.* 109, 93–104.
- Buerki, S., Forest, F., Alvarez, N., Nylander, J.A.A., Arrigo, N., and Sanmartín, I. (2010). An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *J. Biogeogr.* 38, 531–550.
- Buerkle, C.A., Morris, R.J., Asmussen, M.A., and Rieseberg, L.H. (2000). The likelihood of homoploid hybrid speciation. *Heredity (Edinb)* 84 ( Pt 4), 441–451.

- Burns, K.J., Shultz, A.J., Title, P.O., Mason, N.A., Barker, F.K., Klicka, J., Lanyon, S.M., and Lovette, I.J. (2014). Phylogenetics and diversification of tanagers (Passeriformes: Thraupidae), the largest radiation of Neotropical songbirds. *Mol. Phylogenet. Evol.* 75, 41–77.
- Burress, E.D., and Tan, M. (2017). Ecological opportunity alters the timing and shape of adaptive radiation. *Evol. Int. J. Org. Evol.* 71, 2650–2660.
- Capblancq, T., Després, L., Rioux, D., and Mavárez, J. (2015). Hybridization promotes speciation in *Coenonympha* butterflies. *Mol. Ecol.* 24, 6209–6222.
- Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinforma. Oxf. Engl.* 25, 1972–1973.
- Carlquist, Sherwin, Baldwin, Bruce, Carr, and Gerald (2003). *Tarweeds & Silverswords: Evolution of the Madiinae* (St. Louis, Missouri: Missouri Botanical Garden Press).
- Carrió, E., and Güemes, J. (2014). The effectiveness of pre- and post-zygotic barriers in avoiding hybridization between two snapdragons (*Antirrhinum* L.: Plantaginaceae). *Bot. J. Linn. Soc.* 176, 159–172.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. (2013). Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140.
- Chapman, M.A., and Burke, J.M. (2007). Genetic divergence and hybrid speciation. *Evol. Int. J. Org. Evol.* 61, 1773–1780.
- Chaves, J.A., Cooper, E.A., Hendry, A.P., Podos, J., León, L.F.D., Raeymaekers, J.A.M., MacMillan, W.O., and Uy, J.A.C. (2016). Genomic variation at the tips of the adaptive radiation of Darwin's finches. *Mol. Ecol.* 25, 5282–5295.
- Claramunt, S., and Cracraft, J. (2015). A new time tree reveals Earth history's imprint on the evolution of modern birds. *Sci. Adv.* 1, e1501005.
- Condamine, F.L., Rolland, J., and Morlon, H. (2013). Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16, 72–85.
- Condamine, F.L., Clapham, M.E., and Kergoat, G.J. (2016). Global patterns of insect diversification: towards a reconciliation of fossil and molecular evidence? *Sci. Reports* 6, 19208.
- Condamine, F.L., Rolland, J., Höhna, S., Sperling, F.A.H., and Sanmartín, I. (2018). Testing the Role of the Red Queen and Court Jester as Drivers of the Macroevolution of Apollo Butterflies. *Syst. Biol.*
- Consortium, T.H.G., Dasmahapatra, K.K., Walters, J.R., Briscoe, A.D., Davey, J.W., Whibley, A., Nadeau, N.J., Zimin, A.V., Hughes, D.S.T., Ferguson, L.C., et al. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487, 94–98.
- Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., and Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189.
- Cronn, R., Knaus, B.J., Liston, A., Maughan, P.J., Parks, M., Syring, J.V., and Udall, J. (2012). Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.* 99, 291–311.

- Cruaud, A., Gautier, M., Galan, M., Foucaud, J., Sauné, L., Genson, G., Dubois, E., Nidelet, S., Deuve, T., and Rasplus, J.-Y. (2014). Empirical assessment of RAD sequencing for interspecific phylogeny. *Mol. Biol. Evol.* *31*, 1272–1274.
- Cruden, R.W., and Lyon, D.L. (1985). Patterns of biomass allocation to male and female functions in plants with different mating systems. *Oecologia* *66*, 299–306.
- Cuatrecasas, J. (1976). A new subtribe in the Heliantheae (Compositae): Espeletiinae. *Phytologia*. *35*, 43–61.
- Cuatrecasas, J. (1986). *Speciation and Radiation of the Espeletiinae in the Andes* (Oxford University Press).
- Cuatrecasas, J. (2013). *A Systematic Study of the Subtribe Espeletiinae Heliantheae Asteraceae Memoirs of The New York Botanical Garden Volume 107*.
- Cyriac, V.P., and Kodandaramaiah, U. (2017). Paleoclimate determines diversification patterns in the fossorial snake family Uropeltidae Cuvier, 1829. *Mol. Phylogenet. Evol.* *116*, 97–107.
- Dalquen, D.A., Altenhoff, A.M., Gonnet, G.H., and Dessimoz, C. (2013). The Impact of Gene Duplication, Insertion, Deletion, Lateral Gene Transfer and Sequencing Error on Orthology Inference: A Simulation Study. *PLOS ONE* *8*, e56925.
- Darwin, C. (1859). *On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life* / (London :: John Murray, Albemarle Street.).
- Degnan, J.H., and Rosenberg, N.A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* *24*, 332–340.
- Dell’Ampio, E., Meusemann, K., Szucsich, N.U., Peters, R.S., Meyer, B., Borner, J., Petersen, M., Aberer, A.J., Stamatakis, A., Walz, M.G., et al. (2014). Decisive data sets in phylogenomics: lessons from studies on the phylogenetic relationships of primarily wingless insects. *Mol. Biol. Evol.* *31*, 239–249.
- Diazgranados, M. (2012). A nomenclator for the frailejones (Espeletiinae Cuatrec., Asteraceae). *PhytoKeys* *16*, 1–52.
- Diazgranados, M., and Barber, J.C. (2017). Geography shapes the phylogeny of frailejones (Espeletiinae Cuatrec., Asteraceae): a remarkable example of recent rapid radiation in sky islands. *PeerJ* *5*.
- Dittrich-Reed, D.R., and Fitzpatrick, B.M. (2013). Transgressive Hybrids as Hopeful Monsters. *Evol. Biol.* *40*, 310–315.
- Dolédec, S., Chessel, D., and Gimaret-Carpentier, C. (2000). Niche Separation in Community Analysis: A New Method. *Ecology* *81*, 2914–2927.
- Donoghue, M.J., and Sanderson, M.J. (2015). Confluence, synnovation, and depauperons in plant diversification. *New Phytol.* *207*, 260–274.
- Dray, S., and Dufour, A.-B. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. *J. Stat. Softw.* *022*.

- Drummond, C.S. (2008). Diversification of *Lupinus* (Leguminosae) in the western New World: Derived evolution of perennial life history and colonization of montane habitats. *Mol. Phylogenet. Evol.* *48*, 408–421.
- Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012a). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* *29*, 1969–1973.
- Drummond, C.S., Eastwood, R.J., Miotto, S.T.S., and Hughes, C.E. (2012b). Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): testing for key innovation with incomplete taxon sampling. *Syst. Biol.* *61*, 443–460.
- Durand, E.Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* *28*, 2239–2252.
- Dušková, E., Sklenář, P., Kolář, F., Vásquez, D.L.A., Romoleroux, K., Fér, T., and Marhold, K. (2017). Growth form evolution and hybridization in *Senecio* (Asteraceae) from the high equatorial Andes. *Ecol. Evol.* *7*, 6455–6468.
- Eaton, D.A.R. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinforma. Oxf. Engl.* *30*, 1844–1849.
- Eaton, D.A.R., and Ree, R.H. (2013). Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Syst. Biol.* *62*, 689–706.
- Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* *32*, 1792–1797.
- Edwards, S.V., Xi, Z., Janke, A., Faircloth, B.C., McCormack, J.E., Glenn, T.C., Zhong, B., Wu, S., Lemmon, E.M., Lemmon, A.R., et al. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Mol. Phylogenet. Evol.* *94*, 447–462.
- Elgvin, T.O., Trier, C.N., Tørresen, O.K., Hagen, I.J., Lien, S., Nederbragt, A.J., Ravinet, M., Jensen, H., and Sætre, G.-P. (2017). The genomic mosaicism of hybrid speciation. *Sci. Adv.* *3*, e1602996.
- Erwin, D.H. (2009). Climate as a Driver of Evolutionary Change. *Curr. Biol.* *19*, R575–R583.
- Etienne, R.S., and Haegeman, B. (2012). A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *Am. Nat.* *180*, E75–89.
- Etienne, R.S., Morlon, H., and Lambert, A. (2012). Estimating the Duration of Speciation from Phylogenies. *Evolution* *68*, 2430–2440.
- Fagua, J.C., and Gonzalez, V.H. (2007). Growth rates, reproductive phenology, and pollination ecology of *Espeletia grandiflora* (Asteraceae), a giant Andean caulescent rosette. *Plant Biol. Stuttg. Ger.* *9*, 127–135.
- Feliner, G.N., Álvarez, I., Fuertes-Aguilar, J., Heuertz, M., Marques, I., Moharrek, F., Piñeiro, R., Riina, R., Rosselló, J.A., Soltis, P.S., et al. (2017). Is homoploid hybrid speciation that rare? An empiricist's view.
- Fenster, C.B., Armbruster, W.S., Wilson, P., Dudash, M.R., and Thomson, J.D. (2004). Pollination Syndromes and Floral Specialization. *Annu. Rev. Ecol. Evol. Syst.* *35*, 375–403.
- Fernández-Mazuecos, M., Mellers, G., Vigalondo, B., Sáez, L., Vargas, P., and Glover, B.J. (2018). Resolving Recent Plant Radiations: Power and Robustness of Genotyping-by-Sequencing. *Syst. Biol.* *67*, 250–268.

- Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., Santos, S.H.D., Hughes, G.M., Komissarov, A., Antunes, A., et al. (2017). Genome-wide signatures of complex introgression and adaptive evolution in the big cats.
- FitzJohn, R.G. (2010). Quantitative traits and diversification. *Syst. Biol.* 59, 619–633.
- FitzJohn, R.G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* 3, 1084–1092.
- FitzJohn, R.G., Maddison, W.P., and Otto, S.P. (2009). Estimating Trait-Dependent Speciation and Extinction Rates from Incompletely Resolved Phylogenies. *Syst. Biol.* 58, 595–611.
- Flantua, SGA, Hooghiemstra, H., van Boxel, J.H., Cabrera, M., González-Carranza, Z., González-Arango, C., Stevens, W.D., Montiel, O.M., Raven, P.H., and Paleocology and Landscape Ecology (IBED, FNWI) (2014). Connectivity dynamics since the Last Glacial Maximum in the northern Andes: a pollen-driven framework to assess potential migration. In *Monographs in Systematic Botany from the Missouri Botanical Garden*, (Missouri Botanical Garden Press),.
- Folk, R.A., Mandel, J.R., and Freudenstein, J.V. (2017). Ancestral Gene Flow and Parallel Organellar Genome Capture Result in Extreme Phylogenomic Discord in a Lineage of Angiosperms. *Syst. Biol.* 66, 320–337.
- Fortuna, M.A., Zaman, L., Wagner, A., and Bascompte, J. (2017). Non-adaptive origins of evolutionary innovations increase network complexity in interacting digital organisms. *Phil Trans R Soc B* 372, 20160431.
- Frichot, E., and François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* 6, 925–929.
- Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and François, O. (2014). Fast and Efficient Estimation of Individual Ancestry Coefficients. *Genetics* 196, 973–983.
- Friedman, J., and Barrett, S.C.H. (2011). The Evolution of Ovule Number and Flower Size in Wind-Pollinated Plants. *Am. Nat.* 177, 246–257.
- Fu, Y.-B., Dong, Y., and Yang, M.-H. (2016). Multiplexed shotgun sequencing reveals congruent three-genome phylogenetic signals for four botanical sections of the flax genus *Linum*. *Mol. Phylogenet. Evol.* 101, 122–132.
- Fuentes-Pardo, A.P., and Ruzzante, D.E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Mol. Ecol.* 26, 5369–5406.
- Gabaldón, T., and Koonin, E.V. (2013). Functional and evolutionary implications of gene orthology. *Nat. Rev. Genet.* 14, 360–366.
- Gao, J., Wang, B., Mao, J.-F., Ingvarsson, P., Zeng, Q.-Y., and Wang, X.-R. (2012). Demography and speciation history of the homoploid hybrid pine *Pinus densata* on the Tibetan Plateau. *Mol. Ecol.* 21, 4811–4827.
- García-Verdugo, C., Friar, E., and Santiago, L.S. (2013). Ecological Role of Hybridization in Adaptive Radiations: A Case Study in the *Dubautia arborea* – *Dubautia ciliolata* (Asteraceae) Complex. *Int. J. Plant Sci.* 174, 749–759.

- Garzzone, C.N., Hoke, G.D., Libarkin, J.C., Withers, S., MacFadden, B., Eiler, J., Ghosh, P., and Mulch, A. (2008). Rise of the Andes. *Science* 320, 1304–1307.
- Gaston, K.J. (2000). Global patterns in biodiversity. *Nature*.
- Gavrilets, S., and Losos, J.B. (2009). Adaptive radiation: contrasting theory with data. *Science* 323, 732–737.
- Genner, M.J., and Turner, G.F. (2012). Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. *Mol. Biol. Evol.* 29, 195–206.
- Gibb, G.C., Condamine, F.L., Kuch, M., Enk, J., Moraes-Barros, N., Superina, M., Poinar, H.N., and Delsuc, F. (2016). Shotgun Mitogenomics Provides a Reference Phylogenetic Framework and Timescale for Living Xenarthrans. *Mol. Biol. Evol.* 33, 621–642.
- Givnish, T.J. (2010). Giant lobelias exemplify convergent evolution. *BMC Biol.* 8, 3.
- Givnish, T.J. (2015). Adaptive radiation versus “radiation” and “explosive diversification”: why conceptual distinctions are fundamental to understanding evolution. *New Phytol.* 207, 297–303.
- Givnish, T.J., Millam, K.C., Mast, A.R., Paterson, T.B., Theim, T.J., Hipp, A.L., Henss, J.M., Smith, J.F., Wood, K.R., and Sytsma, K.J. (2009). Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proc. R. Soc. Lond. B Biol. Sci.* 276, 407–416.
- Givnish, T.J., Barfuss, M.H.J., Van Ee, B., Riina, R., Schulte, K., Horres, R., Gonsiska, P.A., Jabaily, R.S., Crayn, D.M., Smith, J.A.C., et al. (2014). Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Mol. Phylogenet. Evol.* 71, 55–78.
- Glor, R.E. (2010). Phylogenetic Insights on Adaptive Radiation. *Annu. Rev. Ecol. Evol. Syst.* 41, 251–270.
- Goldstein, G., Meinzer, F., and Monasterio, M. (1984). The role of capacitance in the water balance of Andean giant rosette species. *Plant Cell Environ.* 7, 179–186.
- Goldstein, G., Rada, F., and Azocar, A. (1985). Cold Hardiness and Supercooling along an Altitudinal Gradient in Andean Giant Rosette Species. *Oecologia* 68, 147–152.
- Goslee, and Urban (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data.
- Goulet, B.E., Roda, F., and Hopkins, R. (2017). Hybridization in Plants: Old Ideas, New Techniques. *Plant Physiol.* 173, 65–78.
- Grant, V. (1949). Pollination Systems as Isolating Mechanisms in Angiosperms. *Evolution* 3, 82–97.
- Grant, B.R., and Grant, P.R. (2008). Fission and fusion of Darwin's finches populations. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 2821–2829.
- Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., et al. (2010). A Draft Sequence of the Neandertal Genome. *Science* 328, 710–722.
- Gregory-Wodzicki, K.M. (2000). Uplift history of the Central and Northern Andes: A review. *GSA Bull.* 112, 1091–1105.
- Gross, B.L., and Rieseberg, L.H. (2005). The ecological genetics of homoploid hybrid speciation. *J. Hered.* 96, 241–252.

- Von Hagen, K.B., and Kadereit, J.W. (2003). The Diversification of *Halenia* (Gentianaceae): Ecological Opportunity Versus Key Innovation. *Evolution* 57, 2507–2518.
- Van der Hammen, T., and Hooghiemstra, H. (2000). Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quat. Sci. Rev.* 19, 725–742.
- Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., and Challenger, W. (2008). GEIGER: investigating evolutionary radiations. *Bioinformatics* 24, 129–131.
- Harvey, M.G., Smith, B.T., Glenn, T.C., Faircloth, B.C., and Brumfield, R.T. (2016). Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. *Syst. Biol.* 65, 910–924.
- Hedberg, O. (1964). *Features of afroalpine plant ecology* (Uppsala: Almqvist and Wiksells Boktryckeri).
- Hedrick, P.W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* 22, 4606–4618.
- Hejase, H.A., and Liu, K.J. (2016). A scalability study of phylogenetic network inference methods using empirical datasets and simulations involving a single reticulation. *BMC Bioinformatics* 17.
- Herder, F., Nolte, A.W., Pfaender, J., Schwarzer, J., Hadiaty, R.K., and Schliewen, U.K. (2006). Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc. R. Soc. B Biol. Sci.* 273, 2209–2217.
- Herrera, S., and Shank, T.M. (2016). RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Mol. Phylogenet. Evol.* 100, 70–79.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Clim.* 25, 1965–1978.
- Hipp, A.L., Eaton, D.A.R., Cavender-Bares, J., Fitzek, E., Nipper, R., and Manos, P.S. (2014). A Framework Phylogeny of the American Oak Clade Based on Sequenced RAD Data. *PLOS ONE* 9, e93975.
- Hohenlohe, P.A., Day, M.D., Amish, S.J., Miller, M.R., Kamps-Hughes, N., Boyer, M.C., Muhlfeld, C.C., Allendorf, F.W., Johnson, E.A., and Luikart, G. (2013). Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Mol. Ecol.* 22, 3002–3013.
- Holzman, R., and Hulsey, C.D. (2017). Mechanical Transgressive Segregation and the Rapid Origin of Trophic Novelty. *Sci. Reports* 7, 40306.
- Hönisch, B., Hemming, N.G., Archer, D., Siddall, M., and McManus, J.F. (2009). Atmospheric Carbon Dioxide Concentration Across the Mid-Pleistocene Transition. *Science* 324, 1551–1554.
- Hooghiemstra, H., and Van der Hammen, T. (2004a). Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359, 173–180; discussion 180–181.
- Hooghiemstra, H., and Van der Hammen, T. (2004b). Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. *Philos. Trans. R. Soc. B Biol. Sci.* 359, 173–181.
- Hooghiemstra, H., Wijninga, V.M., and Cleef, A.M. (2006). The paleobotanical record of colombia: implications for biogeography and biodiversity. *Ann. Mo. Bot. Gard.* 93, 297–325.

- Hoorn, C., Wesselingh, F.P., Steege, H. ter, Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., et al. (2010). Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. *Science* 330, 927–931.
- Hoorn, C., Wesselingh, F.P., Steege, H.T., Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., et al. (2011). Origins of Biodiversity—Response. *Science* 331, 399–400.
- Hoorn, C., Mosbrugger, V., Mulch, A., and Antonelli, A. (2013). Biodiversity from mountain building. *Nat. Geosci.* 6, 154.
- Howarth, D.G., and Baum, D.A. (2005). Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaevola* (goodeniaceae) in the Hawaiian Islands. *Evol. Int. J. Org. Evol.* 59, 948–961.
- Hudson, A.G., Vonlanthen, P., and Seehausen, O. (2011). Rapid parallel adaptive radiations from a single hybridogenic ancestral population. *Proc. Biol. Sci.* 278, 58–66.
- Hughes, C., and Eastwood, R. (2006). Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proc. Natl. Acad. Sci.* 103, 10334–10339.
- Hughes, C.E., and Atchison, G.W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.* 207, 275–282.
- Humboldt, A. von A. du texte, and Bonpland, A. (1773-1858) A. du texte (1809). *Alexandri de Humboldt et Amati Bonpland Plantae aequinoctiales : per regnum Mexici, in provinciis Caracarum et novae Andalusiae, in Peruvianorum, Quitensium, Novae Granatae Andibus, ad Oronoci, fluvii Nigri, fluminis Amazonum ripas nascentes.* Tome 2 (Paris: Schoell).
- Igea, J., Miller, E.F., Papadopulos, A.S.T., and Tanentzap, A.J. (2017). Seed size and its rate of evolution correlate with species diversification across angiosperms. *PLOS Biol.* 15, e2002792.
- Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.-Y., Kupfer, A., Petersen, J., Jarek, M., Meyer, A., Vences, M., et al. (2017). Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nat. Ecol. Evol.* 1, 1370–1378.
- Ivory, S.J., Blome, M.W., King, J.W., McGlue, M.M., Cole, J.E., and Cohen, A.S. (2016). Environmental change explains cichlid adaptive radiation at Lake Malawi over the past 1.2 million years. *Proc. Natl. Acad. Sci.* 113, 11895–11900.
- Jabaily, R.S., and Sytsma, K.J. (2013). Historical biogeography and life-history evolution of Andean Puya (Bromeliaceae). *Bot. J. Linn. Soc.* 171, 201–224.
- James, J.K., and Abbott, R.J. (2005). Recent, allopatric, homoploid hybrid speciation: the origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evol. Int. J. Org. Evol.* 59, 2533–2547.
- Johnson, S.D. (2010). The pollination niche and its role in the diversification and maintenance of the southern African flora. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 499–516.
- Johnston, J.A., Arnold, M.L., and Donovan, L.A. (2003). High hybrid fitness at seed and seedling life history stages in Louisiana irises. *J. Ecol.* 91, 438–446.
- Joly, S., Heenan, P.B., and Lockhart, P.J. (2014). Species radiation by niche shifts in New Zealand's rockcresses (*Pachycladon*, Brassicaceae). *Syst. Biol.* 63, 192–202.



- Jorgensen, T.H., and Olesen, J.M. (2001). Adaptive radiation of island plants: evidence from *Aeonium* (Crassulaceae) of the Canary Islands. *Perspect. Plant Ecol. Evol. Syst.* 4, 29–42.
- Josee, C., Cuesta, F., Navarro, G., Cabrera, E., Chacón Moreno, E., Ferreira, W., Peralvo, M., Saito, J., and Tovar, A. (2009). Ecosistemas de Los Andes del Norte y Centro. Bolivia, Colombia, Ecuador, Perú y Venezuela.
- Joyce, D.A., Lunt, D.H., Genner, M.J., Turner, G.F., Bills, R., and Seehausen, O. (2011). Repeated colonization and hybridization in Lake Malawi cichlids. *Curr. Biol.* 21, R108–R109.
- Kagawa, K., and Takimoto, G. (2017). Hybridization can promote adaptive radiation by means of transgressive segregation. *Ecol. Lett.* 21, 264–274.
- Kakishima, S., Yoshimura, J., Murata, H., and Murata, J. (2011). 6-year periodicity and variable synchronicity in a mass-flowering plant. *PLoS ONE* 6, e28140.
- Kay, K.M., Reeves, P.A., Olmstead, R.G., and Schemske, D.W. (2005). Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. *Am. J. Bot.* 92, 1899–1910.
- Keller, I., Wagner, C.E., Greuter, L., Mwaiko, S., Selz, O.M., Sivasundar, A., Wittwer, S., and Seehausen, O. (2013). Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol. Ecol.* 22, 2848–2863.
- Kelly, D. (1994). The evolutionary ecology of mast seeding. *Trends Ecol. Evol.* 9, 465–470.
- Kolář, F., Dušková, E., and Sklenář, P. (2016). Niche shifts and range expansions along cordilleras drove diversification in a high-elevation endemic plant genus in the tropical Andes. *Mol. Ecol.* 25, 4593–4610.
- Kong, H., Condamine, F.L., Harris, A.J., Chen, J., Pan, B., Möller, M., Hoang, V.S., and Kang, M. (2017). Both temperature fluctuations and East Asian monsoons have driven plant diversification in the karst ecosystems from southern China. *Mol. Ecol.* 26, 6414–6429.
- Koonin, E.V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annu. Rev. Genet.* 39, 309–338.
- Körner, C. (2003). *Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems* (Berlin Heidelberg: Springer-Verlag).
- Kozak, K.M. (2016). *Macroevolution and phylogenomics in the adaptive radiation of Heliconiini butterflies*. Thesis. University of Cambridge.
- Kubatko, L.S. (2009). Identifying hybridization events in the presence of coalescence via model selection. *Syst. Biol.* 58, 478–488.
- Kubatko, L.S., and Degnan, J.H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56, 17–24.
- Lagomarsino, L.P., Condamine, F.L., Antonelli, A., Mulch, A., and Davis, C.C. (2016). The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* 210, 1430–1442.
- Lagomarsino, L.P., Forrestel, E.J., Muchhala, N., and Davis, C.C. (2017). Repeated evolution of vertebrate pollination syndromes in a recently diverged Andean plant clade. *Evolution* 71, 1970–1985.

- Lamarck, J.-B. (1809). *Philosophie Zoologique, ou Exposition des Considérations Relatives à l'Histoire Naturelle des Animaux*.
- Lamichhaney, S., Berglund, J., Almén, M.S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., et al. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518, 371–375.
- Leaché, A.D., Fujita, M.K., Minin, V.N., and Bouckaert, R.R. (2014). Species Delimitation using Genome-Wide SNP Data. *Syst. Biol.* 63, 534–542.
- Lehtonen, S., Silvestro, D., Karger, D.N., Scotese, C., Tuomisto, H., Kessler, M., Peña, C., Wahlberg, N., and Antonelli, A. (2017). Environmentally driven extinction and opportunistic origination explain fern diversification patterns. *Sci. Reports* 7, 4831.
- Lemmon, E.M., and Lemmon, A.R. (2013). High-Throughput Genomic Data in Systematics and Phylogenetics. *Annu. Rev. Ecol. Evol. Syst.* 44, 99–121.
- Lemmon, A.R., Emme, S.A., and Lemmon, E.M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* 61, 727–744.
- Léveillé-Bourret, É., Starr, J.R., Ford, B.A., Moriarty Lemmon, E., and Lemmon, A.R. (2018). Resolving Rapid Radiations within Angiosperm Families Using Anchored Phylogenomics. *Syst. Biol.* 67, 94–112.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.* 25, 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* 25, 2078–2079.
- Lieberman, B.S. (2012). Adaptive Radiations in the Context of Macroevolutionary Theory: A Paleontological Perspective. *Evol. Biol.* 39, 181–191.
- Lindqvist, C., Motley, T.J., Jeffrey, J.J., and Albert, V.A. (2003). Cladogenesis and reticulation in the Hawaiian endemic mints (Lamiaceae). *Cladistics* 19, 480–495.
- Linné, C. von (1758). *Caroli Linnaei...Systema naturae per regna tria naturae :secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. (Holmiae :: Impensis Direct. Laurentii Salvii.)*.
- Liow, L.H., Reitan, T., and Harnik, P.G. (2015). Ecological interactions on macroevolutionary time scales: clams and brachiopods are more than ships that pass in the night. *Ecol. Lett.* 18, 1030–1039.
- Litsios, G., and Salamin, N. (2014). Hybridisation and diversification in the adaptive radiation of clownfishes. *BMC Evol. Biol.* 14, 245.
- Liu, L., Xi, Z., Wu, S., Davis, C.C., and Edwards, S.V. (2015). Estimating phylogenetic trees from genome-scale data. *Ann. N. Y. Acad. Sci.* 1360, 36–53.
- Liu, Y., Li, D., Zhang, Q., Song, C., Zhong, C., Zhang, X., Wang, Y., Yao, X., Wang, Z., Zeng, S., et al. (2017). Rapid radiations of both kiwifruit hybrid lineages and their parents shed light on a two-layer mode of species diversification. *New Phytol.* 215, 877–890.

- Loeuille, B., Semir, J., Lohmann, L.G., and Pirani, J.R. (2015). A Phylogenetic Analysis of Lychnophorinae (Asteraceae: Vernonieae) Based on Molecular and Morphological Data. *Syst. Bot.* *40*, 299–315.
- Londoño, C., Cleef, A., and Madriñán, S. (2014). Angiosperm flora and biogeography of the páramo region of Colombia, Northern Andes. *Flora - Morphol. Distrib. Funct. Ecol. Plants* *209*, 81–87.
- Losos, J. (2009). *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles* (University of California Press).
- Losos, J.B., and Mahler, D.L. (2010). Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. *Evol. Darwin First* *150*, 381–420.
- Losos, J.B., and Ricklefs, R.E. (2009). Adaptation and diversification on islands. *Nature*.
- Lotsy, J.P. (1925). Species or lineage. *Genetica* *7*, 487–506.
- Lowry, D.B., Modliszewski, J.L., Wright, K.M., Wu, C.A., and Willis, J.H. (2008). The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philos. Trans. R. Soc. B Biol. Sci.* *363*, 3009–3021.
- Luebert, F., and Weigend, M. (2014). Phylogenetic insights into Andean plant diversification. *Front. Ecol. Evol.* *2*.
- Luteyn, J.L. (1999). *Páramos: a checklist of plant diversity, geographical distribution, and botanical literature* (Bronx, N.Y.: New York Botanical Garden Press).
- Maddison, W.P., and Knowles, L.L. (2006). Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* *55*, 21–30.
- Madriñán, S., Cortés, A.J., and Richardson, J.E. (2013). Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Front. Genet.* *4*.
- Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinforma. Oxf. Engl.* *27*, 2957–2963.
- Malé, P.-J.G., Bardon, L., Besnard, G., Coissac, E., Delsuc, F., Engel, J., Lhuillier, E., Scotti-Saintagne, C., Tinaut, A., and Chave, J. (2014). Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. *Mol. Ecol. Resour.* *14*, 966–975.
- Males, J. (2018). Geography, environment and organismal traits in the diversification of a major tropical herbaceous angiosperm radiation. *AoB PLANTS* *10*.
- Malinsky, M., Svoldal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F., and Durbin, R. (2017). Whole Genome Sequences Of Malawi Cichlids Reveal Multiple Radiations Interconnected By Gene Flow. *bioRxiv* 143859.
- Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends Ecol. Evol.* *20*, 229–237.
- Mallet, J. (2007). Hybrid speciation. *Nature* *446*, 279–283.
- Mao, D., Liu, T., Xu, C., Li, X., and Xing, Y. (2011). Epistasis and complementary gene action adequately account for the genetic bases of transgressive segregation of kilo-grain weight in rice. *Euphytica* *180*, 261–271.

- Marcussen, T., Sandve, S.R., Heier, L., Spannagl, M., Pfeifer, M., International Wheat Genome Sequencing Consortium., Jakobsen, K.S., Wulff, B.B.H., Steuernagel, B., Mayer, K.F.X., et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345, 1250092.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, pp. 10–12.
- Martin, N.H., Bouck, A.C., and Arnold, M.L. (2006). Detecting Adaptive Trait Introgression Between *Iris fulva* and *I. brevicaulis* in Highly Selective Field Conditions. *Genetics* 172, 2481–2489.
- Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Piñero, D., and Emerson, B.C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol. Ecol. Resour.* 15, 28–41.
- Matschiner, M., Hanel, R., and Salzburger, W. (2011). On the Origin and Trigger of the Notothenioid Adaptive Radiation. *PLOS ONE* 6, e18911.
- Mavárez, J., and Linares, M. (2008). Homoploid hybrid speciation in animals. *Mol. Ecol.* 17, 4181–4185.
- Mavárez, J., Salazar, C.A., Bermingham, E., Salcedo, C., Jiggins, C.D., and Linares, M. (2006). Speciation by hybridization in *Heliconius* butterflies. *Nature* 441, 868–871.
- Mayr, E. (1942). *Systematics and the Origin of Species, from the Viewpoint of a Zoologist* (Harvard University Press).
- McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., and Glenn, T.C. (2012). Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Res.* 22, 746–754.
- Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L., and Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* 8, 14363.
- Meinzer, F., and Goldstein, G. (1985). Some Consequences of Leaf Pubescence in the Andean Giant Rosette Plant *Espeletia Timotensis*. *Ecology* 66, 512–520.
- Meinzer, F.C., Goldstein, G.H., and Rundel, P.W. (1985). Morphological changes along an altitude gradient and their consequences for an andean giant rosette plant. *Oecologia* 65, 278–283.
- Melo, M.C., Salazar, C., Jiggins, C.D., and Linares, M. (2009). Assortative mating preferences among hybrids offers a route to hybrid speciation. *Evol. Int. J. Org. Evol.* 63, 1660–1665.
- Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinforma. Oxf. Engl.* 31, i44–52.
- Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., and Warnow, T. (2014). ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinforma. Oxf. Engl.* 30, i541–548.
- Moen, D., and Morlon, H. (2014). Why does diversification slow down? *Trends Ecol. Evol.* 29, 190–197.
- Monasterio, M., and Sarmiento, L. (1991). Adaptive radiation of *Espeletia* in the cold andean tropics. *Trends Ecol. Evol.* 6, 387–391.

- Moore, M.J., Dhingra, A., Soltis, P.S., Shaw, R., Farmerie, W.G., Foltá, K.M., and Soltis, D.E. (2006). Rapid and accurate pyrosequencing of angiosperm plastid genomes. *BMC Plant Biol.* 6, 17.
- Mora, A., Baby, P., Roddaz, M., Parra, M., Brusset, S., Hermoza, W., and Espurt, N. (2011). Tectonic History of the Andes and Sub-Andean Zones: Implications for the Development of the Amazon Drainage Basin. In *Amazonia: Landscape and Species Evolution*, (Wiley-Blackwell), pp. 38–60.
- Morales-Briones, D.F., Romoleroux, K., Kolář, F., and Tank, D.C. (2018). Phylogeny and Evolution of the Neotropical Radiation of *Lachemilla* (Rosaceae): Uncovering a History of Reticulate Evolution and Implications for Infrageneric Classification. *Syst. Bot.* 43, 17–34.
- Morillo, G., and Briceño, B. (2007). Estudio sobre *Carramboa tachirensis* (Aristeg.) Cuatrec. (Asteraceae) y sus afines. *Rev. Fac. Agron.* 24.
- Morlon, H., Parsons, T.L., and Plotkin, J.B. (2011). Reconciling molecular phylogenies with the fossil record. *Proc. Natl. Acad. Sci.* 108, 16327–16332.
- Morlon, H., Lewitus, E., Condamine, F.L., Manceau, M., Clavel, J., and Drury, J. (2016). RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* 7, 589–597.
- Murúa, M., Espíndola, A., González, A., and Medel, R. (2017). Pollinators and crossability as reproductive isolation barriers in two sympatric oil-rewarding <Emphasis Type="Italic">Calceolaria</Emphasis> (Calceolariaceae) species. *Evol. Ecol.* 31, 421–434.
- Muschick, M., Indermaur, A., and Salzburger, W. (2012). Convergent Evolution within an Adaptive Radiation of Cichlid Fishes. *Curr. Biol.* 22, 2362–2368.
- Nevado, B., Atchison, G.W., Hughes, C.E., and Filatov, D.A. (2016). Widespread adaptive evolution during repeated evolutionary radiations in New World lupins. *Nat. Commun.* 7, 12384.
- Nevado, B., Contreras-Ortiz, N., Hughes, C., and Filatov, D.A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytol.* 0.
- Nilsson, M.A., Churakov, G., Sommer, M., Tran, N.V., Zemann, A., Brosius, J., and Schmitz, J. (2010). Tracking Marsupial Evolution Using Archaic Genomic Retroposon Insertions. *PLOS Biol.* 8, e1000436.
- Novembre, J., and Stephens, M. (2008). Interpreting principal component analyses of spatial population genetic variation. *Nat. Genet.* 40, 646–649.
- Nürk, N.M., Scheriau, C., and Madriñán, S. (2013). Explosive radiation in high Andean *Hypericum*—rates of diversification among New World lineages. *Front. Genet.* 4.
- Ohlemüller, R. (2011). Running Out of Climate Space. *Science* 334, 613–614.
- Padilla-González, G.F., Diazgranados, M., and Costa, F.B.D. (2017). Biogeography shaped the metabolome of the genus *Espeletia*: a phytochemical perspective on an Andean adaptive radiation. *Sci. Reports* 7, 8835.
- Padilla-González, G.F., Diazgranados, M., Oliveira, T.B., Chagas-Paula, D.A., Costa, D., and B, F. (2018). Chemistry of the subtribe *Espeletiinae* (Asteraceae) and its correlation with phylogenetic data: an in silico chemosystematic approach. *Bot. J. Linn. Soc.* 186, 18–46.
- Pan, J., Zhang, D., and Sang, T. (2007). Molecular phylogenetic evidence for the origin of a diploid hybrid of *Paeonia* (Paeoniaceae). *Am. J. Bot.* 94, 400–408.

- Panero, J.L., Jansen, R.K., and Clevinger, J.A. (1999). Phylogenetic Relationships of Subtribe Ecliptinae (Asteraceae: Heliantheae) Based on Chloroplast DNA Restriction Site Data. *Am. J. Bot.* 86, 413–427.
- Pardo-Diaz, C., Salazar, C., Baxter, S.W., Merot, C., Figueiredo-Ready, W., Joron, M., McMillan, W.O., and Jiggins, C.D. (2012). Adaptive Introgression across Species Boundaries in Heliconius Butterflies. *PLOS Genet.* 8, e1002752.
- Patterson, N., Price, A.L., and Reich, D. (2006). Population Structure and Eigenanalysis. *PLOS Genet.* 2, e190.
- Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient Admixture in Human History. *Genetics* 192, 1065–1093.
- Pease, J.B., Haak, D.C., Hahn, M.W., and Moyle, L.C. (2016). Phylogenomics Reveals Three Sources of Adaptive Variation during a Rapid Radiation. *PLOS Biol.* 14, e1002379.
- Pérez, F.L. (1996). The effects of giant Andean rosettes on surface soils along a high paramo toposequence. *GeoJournal* 40, 283–298.
- Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLOS ONE* 7, e37135.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., and R-core (2018). nlme: Linear and Nonlinear Mixed Effects Models.
- Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News* 6, 7–11.
- Popescu, A.-A., Huber, K.T., and Paradis, E. (2012). ape 3.0: New tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinforma. Oxf. Engl.* 28, 1536–1537.
- Pouchon, C., Fernández, A., Nassar, J.M., Boyer, F., Aubert, S., Lavergne, S., and Mavárez, J. (2018). Phylogenomic Analysis of the Explosive Adaptive Radiation of the Espeletia Complex (Asteraceae) in the Tropical Andes. *Syst. Biol.*
- Pratt, H.D. (2005). *The Hawaiian Honeycreepers: Drepanidinae* (Oxford, New York: Oxford University Press).
- Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.
- Rabosky, D.L. (2009). Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.* 12, 735–743.
- Rabosky, D.L. (2014). Automatic Detection of Key Innovations, Rate Shifts, and Diversity-Dependence on Phylogenetic Trees. *PLOS ONE* 9, e89543.
- Rabosky, D.L., and Hurlbert, A.H. (2015). Species richness at continental scales is dominated by ecological limits. *Am. Nat.* 185, 572–583.

- Rabosky, D.L., Grudler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W., Huang, H., and Larson, J.G. (2014). BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* 5, 701–707.
- Rada, F., Goldstein, G., Azocar, A., and Meinzer, F. (1985). Freezing avoidance in Andean giant rosette plants. *Plant Cell Environ.* 8, 501–507.
- Rainford, J.L., Hofreiter, M., Nicholson, D.B., and Mayhew, P.J. (2014). Phylogenetic Distribution of Extant Richness Suggests Metamorphosis Is a Key Innovation Driving Diversification in Insects. *PLOS ONE* 9, e109085.
- Rangel-Ch., J.O. (2015). La biodiversidad de Colombia: significado y distribución regional. *Rev. Acad. Colomb. Ciencias Exactas Físicas Nat.* 39, 176–200.
- Rauscher, J.T. (2000). Molecular Systematics of the Espeletia Complex: Evidence from NrITS Sequence on the Evolution of an Andean Adaptive Radiation (Washington University).
- Rauscher, J.T. (2002). Molecular phylogenetics of the Espeletia complex (Asteraceae): evidence from nrDNA ITS sequences on the closest relatives of an Andean adaptive radiation. *Am. J. Bot.* 89, 1074–1084.
- Reznick, D.N., and Ricklefs, R.E. (2009). Darwin's bridge between microevolution and macroevolution. *Nature* 457, 837–842.
- Richards, E.J., and Martin, C.H. (2017). Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. *PLOS Genet.* 13, e1006919.
- Rieseberg, L.H., and Soltis, D.E. (1991). Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Plants* 5, 65–84.
- Rieseberg, L.H., Archer, M.A., and Wayne, R.K. (1999). Transgressive segregation, adaptation and speciation. *Heredity (Edinb)* 83 ( Pt 4), 363–372.
- Rieseberg, L.H., Raymond, O., Rosenthal, D.M., Lai, Z., Livingstone, K., Nakazato, T., Durphy, J.L., Schwarzbach, A.E., Donovan, L.A., and Lexer, C. (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301, 1211–1216.
- Robinson, H.E. (1981). A Revision of the Tribal and Subtribal Limits of the Heliantheae (Asteraceae).
- Roch, S., and Steel, M. (2015). Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. *Theor. Popul. Biol.* 100C, 56–62.
- Rodríguez, A., Burgon, J.D., Lyra, M., Irisarri, I., Baurain, D., Blaustein, L., Göçmen, B., Künzel, S., Mable, B.K., Nolte, A.W., et al. (2017). Inferring the shallow phylogeny of true salamanders (*Salamandra*) by multiple phylogenomic approaches. *Mol. Phylogenet. Evol.* 115, 16–26.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4.
- Rokas, A., and Carroll, S.B. (2006). Bushes in the Tree of Life. *PLOS Biol.* 4, e352.
- Roquet, C., Boucher, F.C., Thuiller, W., and Lavergne, S. (2013). Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* 40, 1874–1886.

- Roure, B., Baurain, D., and Philippe, H. (2013). Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Mol. Biol. Evol.* *30*, 197–214.
- Roy, D., Lucek, K., Walter, R.P., and Seehausen, O. (2015). Hybrid “superswarm” leads to rapid divergence and establishment of populations during a biological invasion. *Mol. Ecol.* *24*, 5394–5411.
- Rubin, B.E.R., Ree, R.H., and Moreau, C.S. (2012). Inferring Phylogenies from RAD Sequence Data. *PLOS ONE* *7*, e33394.
- Rull, V. (2011a). Neotropical biodiversity: timing and potential drivers. *Trends Ecol. Evol.* *26*, 508–513.
- Rull, V. (2011b). Origins of biodiversity. *Science* *331*, 398–399; author reply 399–400.
- Rundel, P.W., Smith, A.P., and Meinzer, F.C. (1994). *Tropical Alpine Environments: Plant Form and Function* (Cambridge University Press).
- Rundell, R.J., and Price, T.D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evol.* *24*, 394–399.
- Sánchez-Baracaldo, P., and Thomas, G.H. (2014). Adaptation and Convergent Evolution within the *Jamesonia-Eriosorus* Complex in High-Elevation Biodiverse Andean Hotspots. *PLOS ONE* *9*, e110618.
- Sanderson, M.J. (2003). r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinforma. Oxf. Engl.* *19*, 301–302.
- Scascitelli, M., Whitney, K.D., Randell, R.A., King, M., Buerkle, C.A., and Rieseberg, L.H. (2010). Genome scan of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric patterns of introgression and small islands of genomic differentiation. *Mol. Ecol.* *19*, 521–541.
- Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. *Bioinformatics* *27*, 592–593.
- Schluter, D. (2000). *The Ecology of Adaptive Radiation* (OUP Oxford).
- Schmickl, R., Liston, A., Zeisek, V., Oberlander, K., Weitemier, K., Straub, S.C.K., Cronn, R.C., Dreyer, L.L., and Suda, J. (2016). Phylogenetic marker development for target enrichment from transcriptome and genome skim data: the pipeline and its application in southern African *Oxalis* (Oxalidaceae). *Mol. Ecol. Resour.* *16*, 1124–1135.
- Schmidt-Lebuhn, A.N., Kessler, M., and Hensen, I. (2007). Hummingbirds as drivers of plant speciation? *Trends Plant Sci.* *12*, 329–331.
- Schmidt-Lebuhn, A.N., Fuchs, J., Hertel, D., Hirsch, H., Toivonen, J., and Kessler, M. (2010). An Andean radiation: polyploidy in the tree genus *Polylepis* (Rosaceae, Sanguisorbeae). *Plant Biol.* *12*, 917–926.
- Schumer, M., Cui, R., Rosenthal, G.G., and Andolfatto, P. (2015). Reproductive Isolation of Hybrid Populations Driven by Genetic Incompatibilities. *PLOS Genet.* *11*, e1005041.
- Schumer, M., Cui, R., Powell, D.L., Rosenthal, G.G., and Andolfatto, P. (2016). Ancient hybridization and genomic stabilization in a swordtail fish. *Mol. Ecol.* *25*, 2661–2679.



- Schwartz, R.S., Harkins, K.M., Stone, A.C., and Cartwright, R.A. (2015). A composite genome approach to identify phylogenetically informative data from next-generation sequencing. *BMC Bioinformatics* 16, 193.
- Schwartz, T.S., Tae, H., Yang, Y., Mockaitis, K., Van Hemert, J.L., Proulx, S.R., Choi, J.-H., and Bronikowski, A.M. (2010). A garter snake transcriptome: pyrosequencing, de novo assembly, and sex-specific differences. *BMC Genomics* 11, 694.
- Schwarzer, J., Swartz, E.R., Vreven, E., Snoeks, J., Cotterill, F.P.D., Misof, B., and Schlieven, U.K. (2012). Repeated trans-watershed hybridization among haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution. *Proc. R. Soc. B Biol. Sci.* 279, 4389–4398.
- Schwenk, K., Brede, N., and Streit, B. (2008). Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 2805–2811.
- Seeb, J.E., Pascal, C.E., Grau, E.D., Seeb, L.W., Templin, W.D., Harkins, T., and Roberts, S.B. (2011). Transcriptome sequencing and high-resolution melt analysis advance single nucleotide polymorphism discovery in duplicated salmonids. *Mol. Ecol. Resour.* 11, 335–348.
- Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207.
- Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proc. Biol. Sci.* 273, 1987–1998.
- Seehausen, O. (2013). Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* 26, 279–281.
- Selz, O.M., Thommen, R., Maan, M.E., and Seehausen, O. (2014). Behavioural isolation may facilitate homoploid hybrid speciation in cichlid fish. *J. Evol. Biol.* 27, 275–289.
- Shaw, K.L. (2002). Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16122–16127.
- Silvestro, D., Zizka, G., and Schulte, K. (2014). Disentangling the effects of key innovations on the diversification of Bromelioideae (bromeliaceae). *Evol. Int. J. Org. Evol.* 68, 163–175.
- Silvestro, D., Antonelli, A., Salamin, N., and Quental, T.B. (2015). The role of clade competition in the diversification of North American canids. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8684–8689.
- Simpson, G.G. (1953). *The Major Features of Evolution* (Columbia University Press).
- Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19, 1117–1123.
- Sklenář, P., Dušková, E., and Balslev, H. (2011). Tropical and Temperate: Evolutionary History of Páramo Flora. *Bot. Rev.* 77, 71–108.
- Sklenář, P., Hedberg, I., and Cleef, A.M. (2014). Island biogeography of tropical alpine floras. *J. Biogeogr.* 41, 287–297.
- Smith, A.P. (1974). Bud Temperature in Relation to Nyctinastic Leaf Movement in an Andean Giant Rosette Plant. *Biotropica* 6, 263–266.
- Smith, A.P. (1981). Growth and Population Dynamics of Espeletia (Compositae) in the Venezuelan Andes.

- Smith, A.C., and Koch, M.F. (1935). The Genus *Espeletia*: A Study in Phylogenetic Taxonomy. *Brittonia* 1, 479–530.
- Smith, S.A., and Pease, J.B. (2017). Heterogeneous molecular processes among the causes of how sequence similarity scores can fail to recapitulate phylogeny. *Brief. Bioinform.* 18, 451–457.
- Solís-Lemus, C., Yang, M., and Ané, C. (2016). Inconsistency of Species Tree Methods under Gene Flow. *Syst. Biol.* 65, 843–851.
- Solís-Lemus, C., Bastide, P., and Ané, C. (2017). PhyloNetworks: A Package for Phylogenetic Networks. *Mol. Biol. Evol.* 34, 3292–3298.
- Soltis, P.S., and Soltis, D.E. (2009). The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* 60, 561–588.
- Springer, M.S., and Gatesy, J. (2016). The gene tree delusion. *Mol. Phylogenet. Evol.* 94, 1–33.
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.
- Stankowski, S., and Streisfeld, M.A. (2015). Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proc R Soc B* 282, 20151666.
- Stearns, S.C. (1992). *The Evolution of Life Histories* (Oxford ; New York: OUP Oxford).
- Stelkens, R., and Seehausen, O. (2009). Genetic distance between species predicts novel trait expression in their hybrids. *Evol. Int. J. Org. Evol.* 63, 884–897.
- Stephens, J.D., Rogers, W.L., Heyduk, K., Cruse-Sanders, J.M., Determann, R.O., Glenn, T.C., and Malmberg, R.L. (2015). Resolving phylogenetic relationships of the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. *Mol. Phylogenet. Evol.* 85, 76–87.
- Straub, S.C.K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C., and Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *Am. J. Bot.* 99, 349–364.
- Stroud, J.T., and Losos, J.B. (2016). Ecological Opportunity and Adaptive Radiation. *Annu. Rev. Ecol. Evol. Syst.* 47, 507–532.
- Struwe, L., Haag, S., Heiberg, E., and Grant, J.R. (2009). Andean speciation and vicariance in neotropical macrocarpaea (gentianaceae–helieae). *Ann. Mo. Bot. Gard.* 96, 450–469.
- Stuessy, T.F. (1973). A SYSTEMATIC REVIEW OF THE SUBTRIBE MELAMPODIINAE (COMPOSITAE, HELIANTHEAE). *Contrib. Gray Herb. Harv. Univ.* 65–80.
- Sturmbauer, C. (1998). Explosive speciation in cichlid fishes of the African Great Lakes: a dynamic model of adaptive radiation. *J. Fish Biol.* 53, 18–36.
- Suarez-Gonzalez, A., Lexer, C., and Cronk, Q.C.B. (2018). Adaptive introgression: a plant perspective. *Biol. Lett.* 14.
- Svenning, J.-C., Eiserhardt, W.L., Normand, S., Ordonez, A., and Sandel, B. (2015). The Influence of Paleoclimate on Present-Day Patterns in Biodiversity and Ecosystems. *Annu. Rev. Ecol. Evol. Syst.* 46, 551–572.

- Takahashi, T., and Koblmüller, S. (2011). The Adaptive Radiation of Cichlid Fish in Lake Tanganyika: A Morphological Perspective. *Int. J. Evol. Biol.* 2011.
- Taylor, E.B., Boughman, J.W., Groenenboom, M., Sniatynski, M., Schluter, D., and Gow, J.L. (2006). Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Mol. Ecol.* 15, 343–355.
- Taylor, S.J., Rojas, L.D., Ho, S.W., and Martin, N.H. (2013). Genomic collinearity and the genetic architecture of floral differences between the homoploid hybrid species *Iris nelsonii* and one of its progenitors, *Iris hexagona*. *Heredity* 110, 63–70.
- Thuiller, W., Lavorel, S., Midgley, G., Lavergne, S., and Rebelo, T. (2004). Relating Plant Traits and Species Distributions Along Bioclimatic Gradients for 88 Leucadendron Taxa. *Ecology* 85, 1688–1699.
- Torres, V., Hooghiemstra, H., Lourens, L., and Tzedakis, P.C. (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. *Quat. Sci. Rev.* 63, 59–72.
- Tremetsberger, K., Stuessy, T.F., Kadlec, G., Urtubey, E., Baeza, C.M., Beck, S.G., Valdebenito, H.A., Ruas, C. de F., and Matzenbacher, N.I. (2006). AFLP Phylogeny of South American Species of *Hypochaeris* (Asteraceae, Lactuceae). *Syst. Bot.* 31, 610–626.
- Tremetsberger, K., Gemeinholzer, B., Zetsche, H., Blackmore, S., Kilian, N., and Talavera, S. (2013). Divergence time estimation in Cichorieae (Asteraceae) using a fossil-calibrated relaxed molecular clock. *Org. Divers. Evol.* 13, 1–13.
- Tripp, E.A., and Tsai, Y.-H.E. (2017). Disentangling geographical, biotic, and abiotic drivers of plant diversity in neotropical *Ruellia* (Acanthaceae). *PLOS ONE* 12, e0176021.
- Tsitrone, A., Kirkpatrick, M., Levin, D.A., and Morgan, M. (2003). A model for chloroplast capture. *Evolution* 57, 1776–1782.
- Tung, J., and Barreiro, L.B. (2017). The contribution of admixture to primate evolution. *Curr. Opin. Genet. Dev.* 47, 61–68.
- Turissini, D.A., and Matute, D.R. (2017). Fine scale mapping of genomic introgressions within the *Drosophila yakuba* clade. *PLOS Genet.* 13, e1006971.
- Uribe-Convers, S., and Tank, D.C. (2015). Shifts in diversification rates linked to biogeographic movement into new areas: An example of a recent radiation in the Andes. *Am. J. Bot.* 102, 1854–1869.
- Valente, L.M., Savolainen, V., and Vargas, P. (2010). Unparalleled rates of species diversification in Europe. *Proc. R. Soc. Lond. B Biol. Sci.* rspb20092163.
- Vargas, O.M., and Madriñán, S. (2012). Preliminary Phylogeny of *Diplostephium* (Asteraceae): Speciation Rate and Character Evolution. *Lundellia* 15, 1–15.
- Vargas, O.M., Ortiz, E.M., and Simpson, B.B. (2017). Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytol.* 214, 1736–1750.
- Vereecken, N.J., Cozzolino, S., and Schiestl, F.P. (2010). Hybrid floral scent novelty drives pollinator shift in sexually deceptive orchids. *BMC Evol. Biol.* 10, 103.

- Vuilleumier, F., and Monasterio, M. (1986). *High Altitude Tropical Biogeography* (Oxford University Press).
- Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., and Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Mol. Ecol.* 22, 787–798.
- Wen, J., Zhang, J.-Q., Nie, Z.-L., Zhong, Y., and Sun, H. (2014). Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front. Genet.* 5.
- Wessinger, C.A., Freeman, C.C., Mort, M.E., Rausher, M.D., and Hileman, L.C. (2016). Multiplexed shotgun genotyping resolves species relationships within the North American genus *Penstemon*. *Am. J. Bot.* 103, 912–922.
- Whitfield, J.B., and Lockhart, P.J. (2007). Deciphering ancient rapid radiations. *Trends Ecol. Evol.* 22, 258–265.
- Whitney, K.D., Ahern, J.R., Campbell, L.G., Albert, L.P., and King, M.S. (2010). Patterns of hybridization in plants. *Perspect. Plant Ecol. Evol. Syst.* 12, 175–182.
- Wiens, J.J. (2007). Speciation and Ecology Revisited: Phylogenetic Niche Conservatism and the Origin of Species. *Evolution* 58, 193–197.
- Wiens, J.J., Engstrom, T.N., Chippindale, P.T., and Schwenk, K. (2006). Rapid diversification, incomplete isolation, and the “speciation clock” in north american salamanders (genus *plethodon*): testing the hybrid swarm hypothesis of rapid radiation. *Evolution* 60, 2585–2603.
- Wiens, J.J., Kuczynski, C.A., Smith, S.A., Mulcahy, D.G., Sites, J.W., Townsend, T.M., Reeder, T.W., and Zamudio, K. (2008). Branch Lengths, Support, and Congruence: Testing the Phylogenomic Approach with 20 Nuclear Loci in Snakes. *Syst. Biol.* 57, 420–431.
- Wollenberg, K.C., Wang, I.J., Glor, R.E., and Losos, J.B. (2013). Determinism in the Diversification of Hispaniolan Trunk-Ground Anoles (*anolis Cybotes* Species Complex). *Evolution* 67, 3175–3190.
- Wright, S.I., Kalisz, S., and Slotte, T. (2013). Evolutionary consequences of self-fertilization in plants. *Proc. R. Soc. B Biol. Sci.* 280.
- Xi, Z., Liu, L., and Davis, C.C. (2016). The Impact of Missing Data on Species Tree Estimation. *Mol. Biol. Evol.* 33, 838–860.
- Yakimowski, S.B., and Rieseberg, L.H. (2014). The role of homoploid hybridization in evolution: a century of studies synthesizing genetics and ecology. *Am. J. Bot.* 101, 1247–1258.
- Yoder, J.B., Clancey, E., Des Roches, S., Eastman, J.M., Gentry, L., Godsoe, W., Hagey, T.J., Jochimsen, D., Oswald, B.P., Robertson, J., et al. (2010). Ecological opportunity and the origin of adaptive radiations. *J. Evol. Biol.* 23, 1581–1596.
- Young, T.P., and Augspurger, C.K. (1991). Ecology and evolution of long-lived semelparous plants. *Trends Ecol. Evol.* 6, 285–289.
- Zachos, J., Pagani, M., Sloan, L., Thomas, E., and Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292, 686–693.
- Zachos, J.C., Dickens, G.R., and Zeebe, R.E. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451, 279–283.

Zerbino, D.R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829.

Zhang, J., Chiodini, R., Badr, A., and Zhang, G. (2011). The impact of next-generation sequencing on genomics. *J. Genet. Genomics Yi Chuan Xue Bao* 38, 95–109.

Zimmer, E.A., and Wen, J. (2015). Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches. *J. Syst. Evol.* 53, 371–379.





