Molecular Phylogenetics and Evolution xxx (xxxx) 107954

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Hyper-Cryptic radiation of a tropical montane plant lineage

Ingrid Olivares [a, *], Sergio Tusso [b], María José Sanín [c, d], Marylaure de La Harpe [e], Oriane Loiseau [f], Jonathan Rolland [g], Nicolas Salamin [h], Michael Kessler [i], Kentaro K. Shimizu [a, j, 1], Margot Paris [k, 1]

[a] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland
[b] Department of Genetics, Faculty of Biology, Ludwig Maximilian Universität München, Germany
[c] School of Mathematical and Natural Sciences, Arizona State University. United States
[d] Facultad de Ciencias y Biotecnología, Universidad CES, Colombia
[e] Office for Nature Conservation and Environment of the Canton of Graubünden, Switzerland
[f] School of Geosciences, University of Edinburgh, United Kingdom
[g] CNRS, Laboratoire Evolution et Diversité Biologique, Université Toulouse, France
[h] Department of Computational Biology, University of Lausanne, Switzerland
[i] Department of Systematic and Evolutionary Botany, University of Zurich, Switzerland
[j] Kihara Institute of Biological Research, Yokohama City University, Japan
[k] Department of Biology. University of Fribourg, Switzerland

ARTICLE INFO

Keywords:
Andean diversification
Biodiversity
Genomics
Palms
Population structure
Speciation

ABSTRACT

Species are seen as the fundamental unit of biotic diversity, and thus their delimitation is crucial for defining measures for diversity assessments and studying evolution. Differences between species have traditionally been associated with variation in morphology. And yet, the discovery of cryptic diversity suggests that the evolution of distinct lineages does not necessarily involve morphological differences. Here, we analyze 1,684,987 variant sites and over 4,000 genes for more than 400 samples to show how a tropical montane plant lineage (*Geonoma undata* species complex) is composed of numerous unrecognized genetic groups that are not morphologically distinct. We find that 11 to 14 clades do not correspond to the three currently recognized species. Most clades are genetically different and geographic distance and topography are the most important factors determining this genetic divergence. The genetic structure of this lineage does not match its morphological variation. Instead, this species complex constitutes the first example of a hyper-cryptic plant radiation in tropical mountains.

1. Introduction

Understanding the characteristics of species formation is central to our interpretation of current biodiversity patterns and future predictions (De Aguiar et al., 2009; Butlin et al., 2009). Species diverge by various processes, including geographical isolation between populations or vicariance (Grant 1981; Mayr 1942, 1970, Turelli et al., 2001), ecological specialization (Coyne and Orr 2004; Arnold 2015; Nosil 2012), hybridization, and polyploidization (e.g., Rieseberg and Willis 2007; Shimizu 2022). Because species distributions often span wide ecological and geographical ranges, understanding the tempo and mode of species persistence in space and time amidst such heterogeneity is crucial to understanding what a species itself represents. So far, however, our knowledge about different modes of species diversification has been limited by the paucity of high coverage genetic information needed to assess the boundaries between species.

Mountains are one of the most important arenas for species radiations, as they impose geographical barriers that promote diversification and create elevational gradients that lead to a high diversity of habitats (Hughes and Atchinson 2015; Ebersbach et al., 2017). Geographical barriers, on the one hand, can lead to dispersal limitation (Särkinen et al., 2012) and the geographic diversification of numerous lineages (e.g., Cadena et al., 2012; Londoño et al., 2014; Lagomarsino et al., 2016). Elevational gradients, in turn, open numerous opportunities for divergent selection and local adaptation. This is the case of marginal populations that initially occurred at the upper or lower elevation range limit of the species, and which may adapt to new climatic conditions (Angert et al., 2008). Importantly, the dynamism of mountains leads to fluctuations in

* Corresponding author at: Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.
E-mail address: ilolivaresa@gmail.com (I. Olivares).
[1] Share co-last authorship.

https://doi.org/10.1016/j.ympev.2023.107954
Received 1 May 2023; Received in revised form 23 October 2023; Accepted 25 October 2023
1055-7903/

both dispersal barriers and elevation zones through time (Simpson, 1974; Hooghiemstra and van der Hammen 2004). Such dynamism alters the connectivity between species and populations (Flantua et al., 2014; Sanín et al., 2022a), and therefore also the morphological and genetic disparity between those groups (e.g., Jabaily and Sytsma 2013; Vásquez et al., 2016).

Rapid and recent diversification or evolutionary radiations are an important phenomenon in the evolution of biodiversity (Losos 2010; Hughes et al., 2015). In montane systems, plant radiations result from different speciation modes, including ecological adaptation (e.g., Hughes and Atchison 2015), progenitor-derivate speciation (e.g., Vargas et al., 2020), pollinator isolation (e.g., Lagomarsino et al., 2016), and geographical isolation followed by local adaptation (e.g., Eaton and Ree 2013). Understanding the relative importance of each of these speciation modes remains crucial to unravel the complexity of lineage formation and species maintenance in highly biodiverse regions of the world like the tropical Andes. However, revealing speciation modes in a taxonomic group is complicated by the requirement of having a large-scale sampling that encompasses the full distribution of the group, as well as high-resolution phylogenetic data to distinguish between different populations and eventually species. Currently, only a few montane plant groups have such genome-wide data (e.g., *Costus*: Vargas et al., 2020; *Lupinus*: Contreras et al., 2018; *Pedicularis*: Eaton and Ree 2013).

Montane plant radiations, spanning a large range of phenotypic and ecological conditions, are mostly the result of speciation mediated by island-like ecological opportunities following mountain uplift (Hughes and Eastwood 2006; Hughes and Atchison 2015; Lagomarsino et al., 2016). Little is known, however, about those montane lineages that do not conform with morphologically innovative radiation models. Lineages that are morphologically indistinguishable from each other while showing genetic differences are known as cryptic (Bickford et al., 2007; Struck et al., 2018) and these are important because they may represent unique, overlooked evolutionary trajectories. Without an understanding of the prevalence of such diversity our idea of biodiversity patterns will be strongly skewed towards morphologically observable diversity, perhaps not reflecting overall genetic and phylogenetic diversity. Thus, research on current and future biodiversity patterns must involve the study of cryptic groups (Bálint et al., 2011; Fišer et al., 2018). In fact, Adams et al. (2014) warned about the potential existence of numerous "hyper-cryptic" lineages. Hyper-cryptic groups are those in which –based on the study of multiple, independent, nuclear genes– there is a four-fold or higher increase in species-level diversity. The study of these hyper-cryptic species complexes is paramount to achieve Global Biodiversity Assessments that approximate actual species estimates.

Cryptic diversity has been predominantly studied in animals (e.g., birds: Krabbe et al., 2020; mammals: Nicolas et al., 2012; reptiles: Brown et al., 2012; arthropods: Miller et al., 2013; molluscs: Vriejenhoeck 2009) and more recently in plants from temperate ecosystems (Arctic: Grundt et al., 2006; Europe and Temperate Asia: Theodoridis et al., 2019; North America: Jolles and Wolfe 2012; Mastretta-Yanes et al., 2018; North America and Western Europe: Medina et al., 2012; Boucher et al., 2021). Studies of cryptic tropical plants are very scarce (e.g., Gale et al., 2018; Li et al., 2023; Pillon et al., 2014) and, to our knowledge, no assessment of a cryptic plant lineage has been carried out in the tropical mountains, except for the discovery of a small clade of legumes restricted to dry inter-Andean valleys (Gagnon et al., 2015).

Geonoma is one of the most species rich genera of American palms (Arecaceae) with 68 species and 90 subspecies currently recognized (Henderson et al., 1995; Henderson 2011). Species delimitation within *Geonoma* is challenging because about one fifth of the species exhibit great morphological variation that is not consistently related to geographical or environmental location (Henderson 2011). These species may be considered species complexes or polymorphic species. The Neotropical mountains are home to one such species complex, the *Geonoma undata* lineage (hereafter *undata* complex). Henderson (2011) recognised five species and 16 subspecies within the *undata* complex, distributed from southern Mexico to central Bolivia. *Geonoma* originated ca. 18.5 (11.9–19.5) Ma, and the *undata* complex originated ca. 12.5 Ma with a crown age of about 5.3 Ma (3.8–9.2) (Roncal 2011). Unlike other *Geonoma*, the *undata* complex occupies habitats at elevations at 800–3400 m, and previous studies inferred that this clade originated from lowland lineages through adaptation to cold environments (Roncal et al., 2005; 2011).

The lack of genomic data and sampling at the species or subspecies level has precluded our understanding of how much of the diversification of the *undata* complex is the result of adaptation or genetic drift due to geographical isolation. Targeted sequencing to evaluate the phylogenetic relationships of the whole Geonomateae tribe showed that all species within the *undata* complex formed polyphyletic groups not corresponding to the latest taxonomic treatment (Loiseau et al., 2019). This phylogenetic reconstruction suggested lack of genetic structure, leading to few consistent clades each covering a large geographical region. However, a study using the same target sequencing approach, but focused on populations of the *undata* complex in Colombia, identified several genetic clusters, some of them even with sympatric distribution, suggesting reproductive isolation and hence speciation (Sanín et al., 2022b). Thus, the question remains as to how many genetic groups and ultimately species make up the *undata* complex throughout its wide continental distribution.

Here, we aim to understand the genetic structure of a montane lineage that might well correspond to a cryptic evolutionary radiation. Specifically, we tackle this by assessing the diversification of the *undata* complex by conducting a regional analysis that encompasses most of its distribution across the Neotropics. We do this by analysing the most extensive genomic dataset in any species complex of tropical montane plants. This dataset consists of 419 samples from 70 different locations (spanning a gradient of 30° of latitude and 2400 m in elevation) of the three more widespread species within the *undata* complex – *G. lehmannii*, *G. orbignyana and G. undata* – that are all closely related and very difficult to distinguish morphologically from one another; and an extended target sequencing approach covering 1,684,987 variant sites across the genome for the population genetic analyses and over 4,000 genes for the phylogeny. We specifically asked: i) what is the extent of genetic structure within the *undata* complex? ii) are genetic clusters within the *undata* complex so broadly distributed as to be considered part of a single extremely variable species? Or are the same clusters locally restricted? iii) What are the phylogenetic relationships between the resulting clusters?

2. Materials and methods

2.1. The Geonoma undata complex

Geonoma is one of the richest palm genera in the Neotropics. This is a group of mostly slender, understory palms with representatives in almost all forest types. The *undata* complex consists of high elevation species. Based mostly on inflorescence characters, Henderson (2011) recognized five species within the *undata* complex: *G. undata*, *G. orbignyana*, *G. lehmannii*, *G. trigona*, and *G. talamancana.* Using a combination of geographical distribution, stem, and leaf characters, he further distinguished 10 subspecies of *G. undata* and two of each *G. orbignyana* and *G. lehmannii*. Studies on the ecology of the *undata* complex are scarce, but it is known that oilbirds (*Steatornis caripensis*) are important for the dispersal of these species across the Andes (Herzog and Kessler, 1997; Cardenas et al., 2021). The pollination of this group remains to be studied, but in other *Geonoma* species, Euglossine bees and Drosophilidae flies are among the main effective pollinators (Bacon et al., 2021).

2.2. Taxon sampling

We sampled leaf tissue from 419 adult individuals located in ca. 70 sites across three species from the *undata* complex: *G. undata*, *G. orbignyana*, and *G. lehmannii*. *Geonoma undata* and *G. orbignyana* are the most variable, abundant, and widespread species of the group, whereas *G. lehmannii* occurs in small and disjunct populations from Panama to Peru. These distributions explain why our sampling includes only eight samples of *G. lehmannii*, 243 of *G. undata* and 168 of *G. orbignyana*. We were unable to sample the locally endemic species *G. trigona* and *G. talamancana*. We aimed to cover the latitudinal and elevational gradients of the clade, and thus sampled populations from ca. 800 m to 3,400 m in elevation. As outgroups, we used 21 individuals from closely related groups like *G. macrostachys*, and species from 14 different palm genera (Appendix C in Supplementary Material). Overall, 98% of leaf tissues came from field collections, and 2% were obtained from herbarium and botanical garden collections.

2.3. Laboratory and bioinformatics protocol

The detailed protocols used for the laboratory (DNA extraction, library preparation, target capture) and bioinformatics (read trimming, mapping and SNP calling) steps were described in previous publications (de La Harpe et al., 2019; Loiseau et al., 2019). Below, we summarize a few relevant points.

DNA was extracted using the standard instructions in the DNAeasy Plant Mini Kit (Qiagen). Library preparation was done using a KAPA LTP kit (Roche, Basel, Switzerland). Libraries were quantified with a Qubit® Fluorometer v 2.2. For target capture de La Harpe et al. (2019) developed a kit that targets 4,051 genes and 133 non-genic putatively neutral regions. The pooled target capture reactions were sequenced with an Illumina HiSeq3000 sequencer in paired-end 2 × 150 bp mode.

Reads were trimmed using CONDETRI V2.2 (Smeds and Künstner 2011) selecting 20 as the high-quality threshold. We used BOWTIE2 v2.2.5 (Langmead and Salzberg 2012) to map reads to the *G. undata* pseudo reference genome (NCBI project: PRJNA482221). Only reads mapping at a unique location in the genome were kept for analyses. PCR duplicates were masked using PICARD TOOLS v1.119 (https://broadin tatute.github.io/picard). GATK v3.8 (McKenna et al., 2010) was used to base-recalibrate and realign reads around indels. SNPs were called for target regions and their surrounding 1,000 bp using UNIFIEDGENOTYPER also from GATK v3.8 to finally obtain the VCF dataset.

2.4. Filtering of the VCF file

We filtered the VCF file to obtain a reliable dataset in which variants occur in a significant number of individuals and are covered by a suitable number of reads. Sites were filtered with VCFtools v0.1.16 using the options: no indels allowed, minimum quality score of 30, minimum mean depth at 10 × and maximum mean depth at 100 × per site, and a maximum of 50% missing data allowed per site. This resulted in a set of 1,684,987 variant sites with an average depth of 32X and 13.5% missing data. The whole data set containing 445 samples including the outgroup samples were used for phylogenomic analyses, while a subset of only the 422 ingroup samples of *undata* complex were used for downstream population analyses.

2.5. PCA analysis

To explore the distribution of genetic groups in our dataset we performed a Principal Component Analysis (PCA) – We used PLINK 1.9beta6.21 to prune variant sites in linkage disequilibrium by selecting a window size of 50 Kb. We then used a window step size of 10 bp and we pruned any variables that show an r^2 equal or greater than 0.1 within windows. Filtered variants were used to calculate the eigenvectors and eigenvalues and produce a PCA.

2.6. Admixture analysis

We further assessed the population structure and ancestry across the *undata* complex samples using Admixture (v.1.3) (Alexander et al., 2009). We used the pruned output from the previous step which excluded the SNPs in linkage disequilibrium to run Admixture for a number of K clusters varying between 1 and 15. We plotted the results with *Pong* (Behr et al., 2016). We used QGIS 3.4.12 (QGIS.org) to map the locations of each group.

2.7. Phylogenomic analyses

Given that the PCA and Admixture results showed that genomic groups did not correspond to the taxonomic classification and that several independent groups were present, we produced a phylogeny of all samples to understand their genetic clustering. To obtain independent and equally informative genomic regions for the phylogenomic analyses, we used the following conservative approach: we split variant sites into genomic windows of 10 kb, we used these windows to produce independent gene trees that would later be used to build a species tree using multispecies coalescent approaches (see below). Only contigs with over 10 kb in length in the reference genome were considered. We found variation in the number of variant sites between windows suggesting that variant sites are not completely randomly distributed. Thus, based on the distribution of the number of variant sites per window, we selected windows across all sample sequences with an intermediate number of variants to produce the phylogenetic trees. The reason for this is that trees with too few variant sites (genomic regions without markers) will overestimate the uncertainty in the relationship between groups, in particular the relationship between terminal branches/clades. Trees with too many variant sites will have the opposite effect.

From an initial number of 4,718 windows containing SNPs across all samples, we obtained 2,624 windows with 200 – 750 SNPs per window. Variant sites per window were used to produce local alignments by substituting variants sites into the refence genome sequence per samples. This was performed using vcf2phylip.py (Available at https://github.com/edgardomortiz/vcf2phylip). A maximum likelihood tree was produced for each alignment, using IQ-Tree (Minh, et al., 2020). For each tree the option MFP to search for the best suited model and the ultrafast bootstrap option was used. Finally, the 2,624 trees were used in ASTRAL (V. 5.7.8.) (Mirarab et al., 2014; Mirarab and Warnow 2015) to obtain the species tree and support values.

2.8. Divergence and diversity parameters

We calculated population genetic parameters including pairwise nucleotide diversity (π) (Nei and Li 1979), average number of pairwise differences (D_{xy}) (Nei and Li 1979) and pairwise F_{ST} (Reynolds et al., 1983). All parameters were calculated using customised python scripts (from Martin et al., 2019) both per genetic cluster inferred from admixture analyses, as well as for geographical sampling point (populations). Since our approach consist of targeting capture and not whole genome sequencing, relative mean parameters (π, D_{xy} and F_{ST}) were calculated from the average between genomic windows of 500 bp containing at least one variant site. By excluding windows without variant sites, π values might in general be high but this increase is proportional in all measurements which allows the relative comparison between all clusters and populations from this study.

2.9. Environmental variation

Since the *undata* complex occurs along elevational gradients, we expected that the major environmental differences between populations and clades will be related to differences in temperature. We obtained the values for the mean annual air temperature from the CHELSA V2.1 database (Karger et al., 2017) for each sample and used them to analyze the variance (ANOVA) in temperature between phylogenetic clades. We did the same analysis to test for the differences in elevation between clades.

3. Results

3.1. Genetic structure

Our Principal Component Analysis (PCA) showed that the current taxonomical classification of our samples into the species *G. lehmannii*, *G. orbignyana*, and *G. undata* is not supported by the genomic differentiation between populations (Fig. 1). Therefore, to further understand how individuals and populations cluster and relate to each other, we used Admixture and phylogenomic analyses (see results in the following sections).

To further understand the population structure of our *undata* sampling we examined the outcome from the Admixture analyses. We selected the number of clusters after which there are no significant changes in the cross-validation error (Fig. S2) and loglikelihood values (Fig. S3) (Alexander et al., 2009; Evanno et al., 2005). These correspond to a number between nine and eleven clusters (K = 9–11): nine independent clusters and two with high levels of admixture (Fig. 2a).

3.2. Phylogenomic structure

To understand the clustering patterns within these groups and the relationships between them, we next assessed the phylogenomic results. Based on our maximum likelihood species tree (Fig. 2a), we found 14 well-supported clades in the *undata* complex, although the relationships between some of the clades were not always well-supported. Our phylogenetic tree has three basal and strongly supported splits: the first split corresponds to Clade 1 that groups populations from distant geographical areas; the second split groups populations from Ecuador, Peru, and Bolivia; while the third integrates populations from Colombia, the Caribbean, Central and North America. The limit between the last two clades occurs near the equator. These large clades are furthered divided into smaller clades, some of which contain populations that are geographically narrow, and others in which relatively distant populations group together.

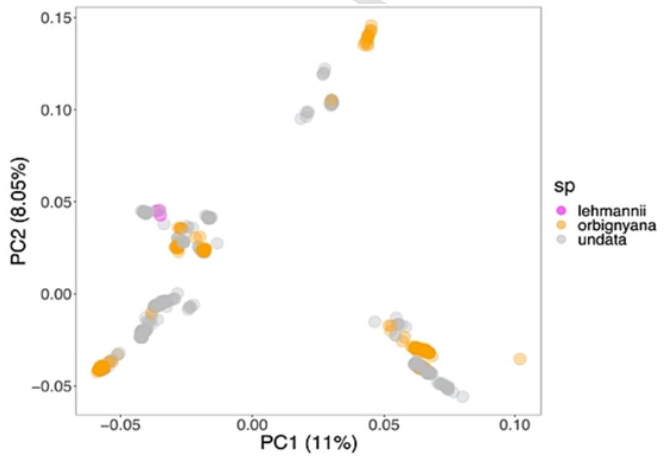

Fig. 1. Population geonomics reveals lack of clustering in the current taxonomic classification of the *undata* complex. Principal Component Analysis (PCA) scatter diagrams show populations scattered in PC1 and PC2 and colored according to the current taxonomical classification.

The clades obtained from the phylogeny were then used to plot the PCA again (Fig. 3a, b). We found that PC1 explained 11% of the genetic variance. Except for one population from Colombia, PC1 separated populations in a north–south gradient. PC2 explained a further 8% of the variance and separated Clade 1 (C1, that combines populations from southern Colombia, southern Ecuador, and northern Bolivia) from the rest. PC3 and PC4 together explained 14 % of the variance and distinguished populations that occur in some of the extremes of the geographical boundaries of our sampling: the Darien region between Colombia and Panama, Tapantí National Park in Costa Rica, La Paz province in northern Bolivia, and the Sierra Nevada de Santa Marta in Colombia (Fig. 4a).

Detailed descriptions of the geographic distribution and populations that integrate each of the 14 clades can be found in the supplementary material. Below, we provide details of the two most distinctive clades.

Clade 1 includes populations in three geographically remote high elevation sites, including specimens morphologically identified as both *G. orbigyana* and *G. undata*. The most divergent population is from northern Bolivia, sister to populations from southern Ecuador and southern Colombia. We confirmed the robustness of this clade by calculating pairwise D_{xy} and F_{ST} for the populations that compose Clade 1 and those populations that occur nearby but that clustered in different clades as this one. Both parameters supported the integrity of Clade 1 by showing less or similar genetic distance between populations within the clade than with populations from other geographically close populations (Fig. S5).

Clade 5 comprises a rheophytic population from southern Ecuador, corresponding to *G. undata* subspecies *pulcherrima* of Henderson (2011). In the Admixture analyses, this population is not separated from Clade 4, whereas in the phylogenetic analysis, Clade 5 is resolved as sister to Clades 1–4. The relationships of this rheophytic population are thus unclear.

In short, our phylogeny and population genetics results demonstrated that the *undata* complex is composed of several genetically independent groups that can occur in close geographical proximity to each other. Three major clusters were strongly supported, one with a disjunct distribution and another two encompassing a north–south gradient. These clusters were further divided into smaller clades occupying relatively small geographical areas or spread along elevation gradients, as described in the following section.

3.3. Divergence and diversity parameters

The parameters D_{xy}, F_{ST} and π confirmed the results obtained previously with the Admixture analysis and later confirmed with the phylogeny. While F_{ST} values showed that the populations within each clade are highly heterogeneous in terms of heterozygosity, D_{xy} showed i) that the biggest differences occur between Clade 1 and all the rest, followed by the differences between southern clades and northern clades. And ii) that northern clades are more divergent between them than those in the south (Fig. S4). The genetic diversity (π) values mirrored the genetic distance patterns by showing that clades spread over large geographical distances (like clades 1, 6 and 13) have the highest diversity whereas clades where individuals belong to a few nearby populations have the lowest genetic diversity. The extreme case of low genetic diversity is clade 7 in the southern Mexico where all individuals come from the same population (Table S1).

We also calculated D_{xy} and F_{ST} to test for the robustness of Clades 1, 8 and 14. We calculated these parameters for i) the populations that compose Clade 1 and populations in Clades 3, 4 and 12 which occur in sites nearby Clade 1; ii) populations in Clades 14 and 8 which also have some range overlap. Both parameters supported the integrity of Clade 1 (Fig. S5), Clade 8 and Clade 14 (Fig. S6) by showing that the divergence

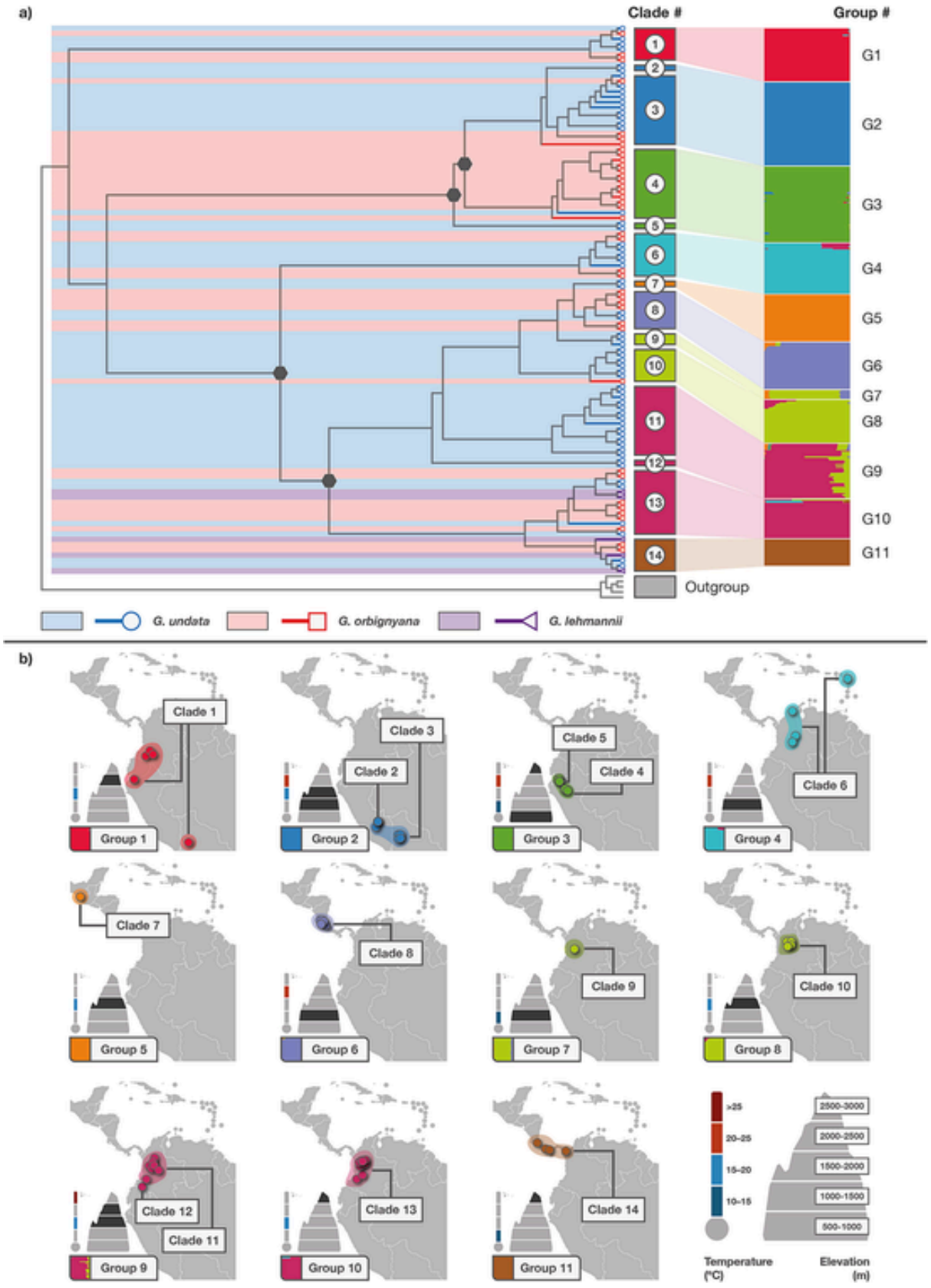

Fig. 2. Phylogenomic and population structure analysis show that the *undata* complex consists of 11 to 14 genetic lineages with various levels of admixture and distinct geographical and elevational distributions. (a) *Left*: Simplified phylogeny including only two samples per population, tip-branches and their background are colored according to their current taxonomic classification, branches with < 0.8 support are indicated with gray dot-hexagons (see full tree and sup-

Fig. 2.—*continued*
port values in Supplementary Material (Appendix B)). *Right*: Admixture results showing the resultant genetic groups. Notice that although nine groups (corresponding to nine different colors) were supported by the cross-validation and loglikelihood values (Fig. S2, S3 in Supplementary Material), these are numbered G1–G11 because we differentiated two groups (G7, G9) that had a high level of admixture. (b) Maps for the distribution of the clades and corresponding genetic groups. Mountain and temperature icons show the elevation and temperature ranges occupied by each clade. See Appendix C to check the precise elevation of each clade.

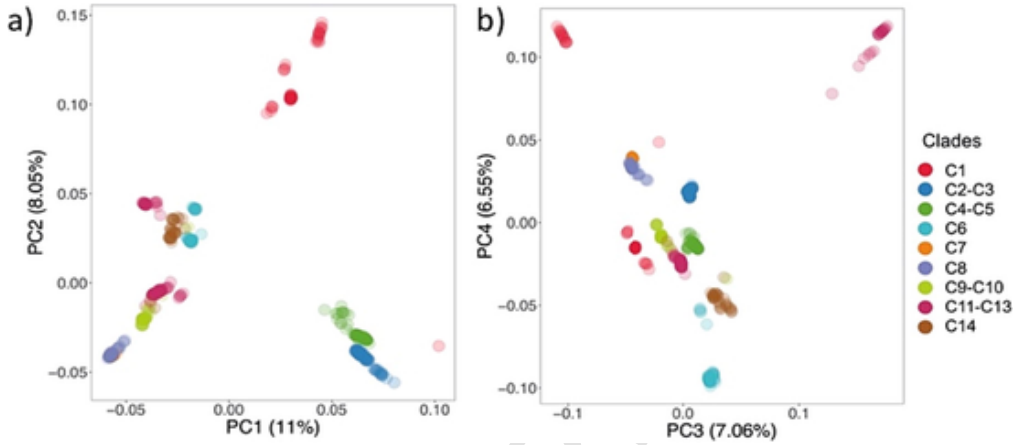

Fig. 3. Populations clustering according to the major clades determined in the phylogenetic analyses. (a) Populations scattered in PC1–PC2 and in (b) PC3–PC4. Some of the clades share their color accordingly with the genetic group they belonged to in the Admixture analysis (see Fig. 2). Most of the variation along the four axes (shown in parentheses) is explained by the geographical location of each cluster (see details in the text).

between the populations that comprise each clade is smaller than with other geographically closer populations. For instance, D_{xy} and F_{ST} between populations 1–3 of clade 1 is lower (or in some cases very similar) than between populations 1–3 of clade 1 and populations 1–4 of clade 3 (Fig. S5). The same is shown for the populations in clades 4 and 12 that occur nearby populations of clade 1.

3.4. Environmental variation

Some of the genetic groups identified by our previous analysis of the *undata* complex occupied habitats with contrasting temperatures due to their non-overlapping distribution (isolation) along elevational gradients. The analysis of variance showed that there are significant differences in the elevation ($R^2 = 0.5757$; $F = 39.45$; p-value < 0.005) and temperature ($R^2 = 0.6014$; $F = 43.87$; p-value < 0.005) values of the sites occupied by the different clades. While some clades occur in premontane habitats (hereafter premontane habitats refer to < 1500 masl), others mostly occur in the cooler highlands (hereafter highlands refer to greater than 1500 masl) (Fig. 2b).

4. Discussion

Here, we demonstrate that the palms of the *undata* complex correspond to a model of hyper-cryptic speciation and that, as a result, about a dozen of species might exist instead of the three species currently recognized and studied here. In other words, the morphological and geographical arguments currently used to delimit species in this complex and probably others in the genus (Henderson 2011) require a complete re-evaluation. In the following sections, we discuss the implications of our results and highlight future research avenues that will contribute to resolve species delimitation in these and other hyper-cryptic mountain lineages.

4.1. Genetic structure within the undata complex

Our population analyses revealed strong population structure between different populations, even among those in the same mountain range. A combination of topography, dispersal limitation, and environmental conditions appear to drive the divergence between members of the *undata* complex.

In contrast to iconic examples of Andean plant radiations, like *Lupinus* (Hughes and Atchinson, 2015) and bellflowers (Lagomarsino et al., 2016) where species diversification is characterised by multiple changes in growth form, habitat preferences, and pollination syndromes, the diversification of the *undata* complex has been much more veiled. The three species included by us are traditionally differentiated only by two subtle traits: the shape and texture of the inflorescences first-bract. Our findings thus suggest that the *undata* complex matches the definition of a hyper-cryptic species complex in which, due to genetic differences, there is a four-fold or higher increase in species number that is largely not evident morphologically (Adams et al., 2014). It is noteworthy that there is quite a lot of morphological variation in the *undata* complex, particularly regarding size variation, that does not correlate to the genetic structure. This means that unlike the traditional lack of variation between cryptic species, the *undata* species have rather a chaotic pattern of variation. However, two clades (Clades 2 and 5) seem to derive phylogenetically from a progenitor-derivate pattern of speciation (Crawford 2010). Individuals in these clades occupy some of the lowest elevations recorded in our sampling and their morphology is the most distinctive of all clades. Still, neither Admixture, D_{xy}, F_{ST}, or π supported a strong divergence or reduced diversity of Clades 2 and 5 in relation to their sister clades (Fig. S4 and Table S1 in Supplementary Material, Appendix A). Therefore, the hypothesis that phenotypic differentiation of these clades confers them evolutionary distinctiveness remains to be tested.

4.2. Morphological and habitat variation

Our results show that the morphological traits used by Henderson (2011) are probably not phylogenetically and taxonomically informative, and that a novel taxonomic classification is needed. Still, some of the clades that we identified appear to correspond to subspecies previously recognized. For example, the distribution of Clade 5 corresponds to G. *undata* subsp. *pulcherrima* and Clade 9 to *G. undata* subsp. *stenothyrsa*, showing that these clades are morphologically distinctive (See Clade 5 pictures in Fig. 4d).

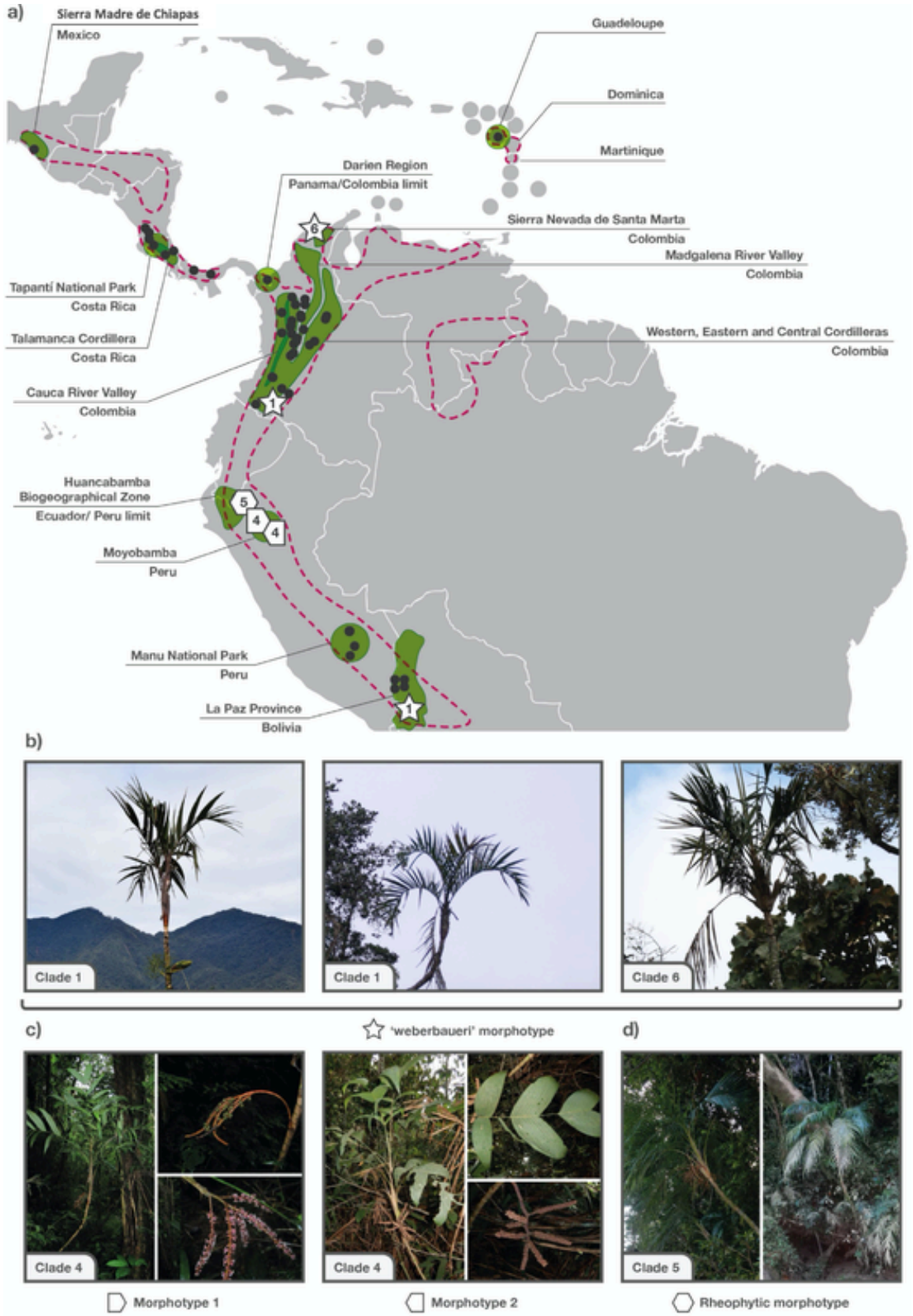

Fig. 4. a) Geographical distribution of the undata complex populations sampled in this study. Location and names of specific places and topographical features mentioned in the text are indicated. Dotted-red areas indicate the whole distribution of the undata complex based on botanical records (Henderson, 2011). Although there is phenotypic convergence and differences among some of the clades, our results showed that these differences might not necessarily confer evolutionary dis-

Fig. 4.—*continued*
tinctiveness. b) Morphological convergence of populations from high elevations but located in distant places. c) Differentiation in leaf division and inflorescence thickness that can be typically observed in nearby populations of the undata complex. d) Distinct morphology of a rheophythic population from the eastern Ecuadorian premontane forests. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Our phylogenomic results showed that geographically close populations occurring at different elevations are often genetically isolated. This elevational isolation likely affects the adaptation of the species in the *undata* complex to the environment, which might manifest in some contrasting morphologies. For example, we found that the distinct high-elevation plants corresponding to the morphotype *weberbaueri* of *G. undata* subsp. *undata* (Henderson 2011) occur at distant geographic sites and fall into distinct genetic clades. Thus, *weberbaueri*-type plants from the Sierra Nevada de Santa Marta in northern Colombia were recovered in Clade 6, those from southern Colombia in Clade 1, and those from northern Bolivia in Clade 1 (Fig. 4b). As already suggested by Henderson (2011), this suggests convergent morphological evolution between populations as the result of adaptive selection.

Nevertheless, the limited morphological variation in the *undata* complex is not consistent across topographic or environmental gradients. Thus, while there are cases like the example above in which there seems to be convergence of high elevation populations or those in which a specific morphotype can be expected under certain environmental conditions (e.g., Clade 5: rheophytes with tall and stout stems, and leaves with numerous narrow pinnae, Fig. 4d), there are other cases in which populations exhibit little morphological variation —relative to their sister groups— other than in the size of different organs (e.g., Clade 2, a premontane population with plant size, leaf, and inflorescence division above the average within the *undata* complex). Also, the variation along the elevational gradient was not consistent. We found nearby clades at different elevations that were sometimes grouped in the same cluster, e.g., a premontane (Clade 2) with a high elevation (Clade 3) clade; or placed in different ones, e.g., Clades 8 and 14. More studies will be needed to assess whether morphological differences actually relate to species boundaries, and how the genetic differentiation relates to environmental factors.

4.3. Phylogenetic relationships between clades

In the phylogeny, the most divergent clade (Clade 1) has a peculiar distribution (with disjunct populations in Colombia, Ecuador, and Bolivia) and is sister to all other clades. Although we cannot exclude the possibility that further unsampled populations of this clade occur in between our records, the fact that numerous populations assigned to other clades exist at geographically intermediate locations support the notion that this is a genetically distinct, but geographically disjunct clade. The distribution of Clade 1 is peculiar because, as detailed below, the rest of the clades are consistently separated in northern and southern groups. Future research that includes more populations and perhaps a whole genome sequencing approach would help to elucidate the origin of the relationship between these distant populations.

Regarding the remaining clades, the primary geographical division is between a northern and southern group. This boundary coincides with the Huancabamba biogeographical zone located between southern Ecuador and northern Peru (ca. latitude 5° S) (Fig. 4a). Several studies suggest that the distinct climatic conditions in this area, characterized by extreme precipitation gradients, strong winds, and an atypically low treeline elevation, act as a dispersal barrier for montane plants (Simpson 1975; Ayers 1999; Cosacov et al., 2009; Jabaily and Sytsma 2013; Vargas et al., 2017, Contreras et al., 2018) and animals (Bonaccorso 2009; Chaves and Smith 2011; Gutiérrez-Pinto et al., 2012). However, it has also been demonstrated that the area offers a diversity of habitats that contribute to the diversification of several plant groups (Weigend 2002). We found clades to the north (Clades 6–14) and south of the Huancabamba zone (Clades 2–3) as well as clades that overlap its range (Clades 4–5). Our findings suggest that the diversification of the *undata* complex has been significantly influenced by the Huancabamba biogeographical zone, which probably acted both as a dispersal barrier and as an arena for species divergence.

Further subdivisions resulted in the recognition of several major genetic clusters largely found in geographically distinct regions, and often corresponding to separate mountain regions. These groups thus correspond to geographically isolated populations in distinct mountain ranges, suggesting a role of geographical (allopatric) speciation. In the Colombian Andes, the genetic structure is more complex. Here, three parallel mountain ranges are inhabited by six clades. There is evidence that the clades found in Colombia have some degree of evolutionary independence even if they are geographically close, which is likely the result of reduced gene flow caused by the dissected topographical conditions of the Andean mountains (Sanín et al., 2022b). In Colombia, two taxonomically recognized groups turned into four genetically independent groups, a pattern that was to some extent echoed at the continental scale in this study. Despite there is evidence that some of the genetically independent groups in the Western and Central Cordilleras of Colombia are sympatric (Sanín et al., 2022b), further research shall reveal whether populations are in syntopy. Our results provide a wider understanding of the taxonomic versus genetic variation of this and other species complex with similar variation in the Neotropics.

The fact that some of the *undata* clades occur sympatrically in some of the distribution areas implies that the diversification of the group is not only the result of pure random drift caused by geographical isolation. Previous studies with other species complexes within *Geonoma* have found evidence of reproductive differences between sympatric species (Listabarth 1993; Knudsen 1999; Borchsenius et al., 2016). Thus, future studies will need to assess whether reproductive or any other differences exist but also evaluate the environmental and biotic pressures that might drive those differences to ultimately understand the evolution of sympatric clades within the *undata* complex.

4.4. Pervasiveness of hyper-cryptic radiations

Given the similarity in variation patterns between *G. undata* and several other palm species complexes in *Geonoma* (e.g., *G. macrostachys,* Bacon et al., 2021) and other genera (e.g., *Aiphanes,* Sanín et al., 2022c), it seems likely that a similar diversification pattern exists in those other groups even when they are predominantly from the lowlands. However, mechanisms for range fragmentation could be different for them, linked e.g., to Pleistocene forest-cover dynamics, changing riverbeds or others. This makes the hyper-cryptic radiation observed in *G. undata* a relevant pattern to understand plant diversification in the Neotropics and likely in many other regions. If hyper-cryptic radiations are so widespread then in addition to the taxonomy, the criteria used for the conservation status of many of these species might also need a re-evaluation.

4.5. Future directions for species delimitation

We propose that the *undata* complex consists of several clades that are genetically so distinct even in close geographical proximity to other clades (e.g., Clades 9 to 13) that they are likely independent. However, there is also evidence of gene flow between some of the clades, suggesting that species boundaries are not yet fully formed. We suggest that the resulting clades likely correspond to roughly one dozen biological species. Considering that our sampling did not include some geographically highly restricted and morphologically distinct populations recog-

nized as subspecies by Henderson (2011), the actual number of species in the *undata* complex may even be higher. As highlighted in *section 4.4* several other groups in the Neotropics share the pattern of variation of the *undata* complex which compels us to find new ways to resolve the taxonomy of cryptic species.

Singhal et al. (2018) in their *framework to resolving cryptic species* provided four steps to diagnose species boundaries across cryptic lineages: 1) statistical species delimitation, 2) *post hoc* discovery of phenotypic differences between clades, 3) indirect or direct estimates of evolutionary isolation between clades, and 4) calibration-based approaches. The present study contributes to step 1 and can be considered as the first step to determine which lineages are unique at the regional scale. To tackle step 2, future studies will have to assess morphological or other phenotypic (e.g., physiological) characters that might relate to species distinctiveness. Step 3 will require future research of genetic structure and reproductive biology across different elevational or climatic gradients. Finally, calibration approaches (step 4) combine information on genetics, morphology, and reproductive isolation directly from contact zones to delimit species boundaries. The latter is an appealing approach that will, however, require the compilation of considerable information and which could constitute a long-term objective in the study of hyper-cryptic lineages.

5. Conclusions

Our study shows the *undata* complex is an example of an active hyper-cryptic radiation. Current morphological classification of the *undata* complex was not supported by the genetic groups identified here. Instead of three species, the samples studied here might better correspond to about dozen species. Our results support a strong population structure with little gene flow between most of the genetic clusters. We found a combination of early divergent clades that might preserve the ancestral genetic background and other genetic groups that likely represent well-defined species. A few groups showed evidence of gene flow which might later lead to further species differentiation or to the unification of current incipient species. We propose that topographical features, dispersal limitation, and environmental changes along elevational gradients are the main factors driving the diversification of the *undata* species complex. Since several species complexes in the Neotropics show a similar variation-pattern to that found in *G. undata*, we suggest that hyper-cryptic radiations might constitute a common feature of Neotropical plant diversification. This would in turn impact the Global Biodiversity Assessments by significantly increasing the estimations of tropical biodiversity.

6. Data accessibility

Sequences for 162 samples used in this study were publicly available on NCBI (NCBI bioprojects PRJNA482221, PRJNA541164, PRJNA707300 and PRJNA689999 from de La Harpe et al., 2019; Loiseau et al., 2019; Sanín et al., 2022b; and Sanín et al., 2022c). Sequences for the other 278 samples are available under XXX (These data are currently being uploaded and the project numbers and corresponding links to NCBI will be added when we receive the manuscript proofs).

Funding

This research was funded by a Forschungskredit Postdoc grant from the University of Zurich to Ingrid Olivares (FK-21–133). Sample collection, identification and sequencing were funded by a Swiss National Science Foundation Grant to Nicolas Salamin, Christian Lexer and Michael Kessler (SNSF, Grant Number: CRSII3_14763).

Uncited references

Flantua et al. (2019).

CRediT authorship contribution statement

Ingrid Olivares : Conceptualization, Methodology, Funding acquisition, Data curation, Investigation, Writing – original draft, Writing – review & editing. **Sergio Tusso :** Methodology. **María José Sanín :** Data curation, Investigation, Writing – review & editing. **Marylaure de La Harpe :** Data curation, Investigation. **Oriane Loiseau :** Data curation, Investigation. **Jonathan Rolland :** Data curation, Investigation, Writing – review & editing. **Nicolas Salamin :** Conceptualization, Funding acquisition, Writing – review & editing. **Michael Kessler :** Conceptualization, Funding acquisition, Writing – review & editing. **Kentaro K. Shimizu :** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Margot Paris :** Conceptualization, Methodology, Data curation, Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Thelma Barbara, Frédérique Fardin, Libertad Silva, Benoit Villain, Anna Weigand, and all collaborators of the POPCORN project in Colombia, Ecuador, Peru, and Bolivia for assistance during fieldwork, permit applications, and lab work, Henrik Balslev, Rodrigo Bernal, Angela Cano, Gloria Galeano and Julissa Roncal for sharing samples. To Andrew Henderson for interesting discussions, his encouragement to proceed with this study, and for the confirmation of a few sample identifications. To all members of the Shimizu group for their collegiality, Epifanía Arango, Masaomi Hatakeyama and Stefan Milosavljevic for bioinformatics support. We especially thank the late Christian Lexer, Professor at the University of Fribourg, who conceptualised and initiated this population genomics study together with Michael Kessler and Nicolas Salamin. To the handling editor Dr. Guillermo Orti and two anonymous reviewers for their suggestions which significantly improved the final version of this manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107954.

References

Adams, M., Raadik, T.A., Burridge, C.P., Georges, A., 2014. Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? Syst. Biol. 63 (4), 518–533. https://doi.org/10.1093/sysbio/syu017.

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664. http://www.genome.org/cgi/doi/10.1101/gr.094052.109.

Angert, A.L., Bradshaw, H.D., Schemske, D.W., 2008. Using experimental evolution to investigate geographic range limits in monkeyflowers. Evolution 62, 2660–2675. https://doi.org/10.1111/j.1558-5646.2008.00471.x.

Arnold, M.L., 2015. Divergence with genetic exchange. Oxford University Press, Oxford.

Ayers, T.J., 1999. Biogeography of *Lysipomia* (Campanulaceae), a high elevation endemic: an illustration of species richness at the Huancabamba Depression, Peru. Arnaldoa 6, 13–28.

Bacon, C.D., Roncal, J., Andermann, T., Barnes, C.J., Balslev, H., Gutiérrez-Pinto, N., Morales, H., Núñez-Avellaneda, L., Tunarosa, N., Antonelli, A., 2021. Genomic and niche divergence in an Amazonian palm species complex. Bot. J. Linn. Soc. 197 (4), 498–512. https://doi.org/10.1093/botlinnean/boab012.

Bálint, M., Domisch, S., Engelhardt, C.H.M., Haase, P., Lehrian, S., Sauer, J., Theissinger, K., Pauls, S.U., Nowak, C., 2011. Cryptic biodiversity loss linked to global climate

change. Nat. Clim. Chang. 1 (6), 313–318. https://doi.org/10.1038/nclimate1191.
Behr, A.A., Liu, K.Z., Liu-Fang, G., Nakka, P., Ramachandran, S., 2016. *Pong*: fast analysis and visualization of latent clusters in population genetic data. Bioinformatics 32 (18), 2817–2823. https://doi.org/10.1093/bioinformatics/btw327.
Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22, 148–155. https://doi.org/10.1016/j.tree.2006.11.004.
Bonaccorso, E., 2009. Historical biogeography and speciation in the neotropical highlands: molecular phylogenetics of the jay genus *Cyanolyca*. Mol. Phylogenetics Evol. 50, 618–632. https://doi.org/10.1016/j.ympev.2008.12.012.
Borchsenius, F., Lozada, T., Knudsen, J.T., 2016. Reproductive isolation of sympatric forms of the understorey palm *Geonoma macrostachys* in western Amazonia. Bot. J. Linn. Soc. 182, 398–410. https://doi.org/10.1111/boj.12428.
Boucher, F.C., Dentant, C., Ibanez, S., Capblancq, T., Boleda, M., Boulangeat, L., Smyčka, J., Roquet, C.C., Lavergne, S., 2021. Discovery of cryptic plant diversity on the rooftops of the Alps. Sci. Rep. 11 (1), 1–10. https://doi.org/10.1038/s41598-021-90612-w.
Brown, R.M., Siler, C.D., Grismer, L.L., Das, I., McGuire, J.A., 2012. Phylogeny and cryptic diversification in Southeast Asian flying geckos. Mol. Phylogenetics Evol. 65, 351–361. https://doi.org/10.1016/j.ympev.2012.06.009.
Butlin, R., Bridle, J., Schluter, D. (Eds.). 2009. Speciation and patterns of diversity. Cambridge University Press.
Cadena, C.D., Kozak, H.K., Gómez, J.P., Parra, J.L., McCain, C.M., Bowie, R.C.K., Carnaval, A.C., Moritz, C., Rahbek, C., Roberts, T.E., Sanders, N.J., Schneider, C.J., VanDerWal, J., Zamudio, K.R., Graham, C.H., 2012. Latitude, elevational climatic zonation and speciation in New World vertebrates. Proc. R. Soc. B Biol. Sci. 279, 194–201. https://doi.org/10.1098/rspb.2011.0720.
Cárdenas, S., Echeverry-Galvis, M.A., Stevenson, P.R., 2021. Seed dispersal effectiveness by oilbirds (*Steatornis caripensis*) in the Southern Andes of Colombia. Biotropica 53, 671–680. https://doi.org/10.1111/btp.12908.
Chaves, J.A., Smith, T.B., 2011. Evolutionary patterns of diversification in the Andean hummingbird genus *Adelomyia*. Mol. Phylogenetics Evol. 60, 207–218. https://doi.org/10.1016/j.ympev.2011.04.007.
Contreras-Ortiz, N., Atchison, G.W., Hughes, C.E., Madriñán, S., 2018. Convergent evolution of high elevation plant growth forms and geographically structured variation in Andean *Lupinus* (Fabaceae). Bot. J. Linn. Soc. 187 (1), 118–136. https://doi.org/10.1093/botlinnean/box095.
Cosacov, A., Sérsic, A.N., Sosa, V., De-Nova, J.A., Nylinder, S., Cocucci, A.A., 2009. New insights into the phylogenetic relationships, character evolution, and phytogeographic patterns of *Calceolaria* (Calceolariaceae). Am. J. Bot. 96, 2240–2255. https://doi.org/10.3732/ajb.0900165.
Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associates, Sunderland, MA.
Crawford, D.J., 2010. Progenitor-derivative species pairs and plant speciation. Taxon 59, 1413–1423. https://doi.org/10.1002/tax.595008.
De Aguiar, M.A.M., Baranger, M., Baptestini, E.M., Kaufman, L., Bar-Yam, Y., 2009. Global patterns of speciation and diversity. Nature 460 (7253), 384–387. https://doi.org/10.1038/nature08168.
de La Harpe, M., Hess, J., Loiseau, O., Salamin, N., Lexer, C., Paris, M., 2019. A dedicated target capture approach reveals variable genetic markers across micro-and macro-evolutionary time scales in palms. Mol. Ecol. Resour. 19, 221–234. https://doi.org/10.1111/1755-0998.12945.
Eaton, D.A.R., Ree, R.H., 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). Syst. Biol. 65, 689–706. https://doi.org/10.1093/sysbio/syt032.
Ebersbach, J., Schnitzler, J., Favre, A., Muellner-Riehl, A.N., 2017. Evolutionary radiations in the species-rich mountain genus Saxifraga L. BMC Evol. Biol. 17, 1–13. https://doi.org/10.1186/s12862-017-0967-2.
Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.
Fišer, C., Robinson, C.T., Malard, F., 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27, 613–635. https://doi.org/10.1111/mec.14486.
Flantua, S.G.A., O'Dea, A., Onstein, R.E., Giraldo, C., Hooghiemstra, H., 2019. The flickering connectivity system of the North Andean paramos. J. Biogeogr. 46, 1808–1825. https://doi.org/10.1111/jbi.13607.
Gagnon, E., Hughes, C.E., Lewis, G.P., Bruneau, A., 2015. A new cryptic species in a new cryptic genus in the Caesalpinia group (Leguminosae) from the seasonally dry inter-Andean valleys of South America. Taxon 64 (3), 468–490. https://doi.org/10.12705/643.6.
Gale, S.W., Duangjai, S., Li, J., Ito, Y., Watthana, S., Termwutthipreecha, P., Cheuk, M.L., Suddee, S., 2018. Integrative analyses of *Nervilia* (Orchidaceae) section *Linervia* reveal further undescribed cryptic diversity in Thailand. Syst. Biodivers. 16 (4), 377–396. https://doi.org/10.1080/14772000.2017.1415233.
Grant, V., 1981. Plant speciation, 2nd ed. Columbia Univ. Press, New York.
Grundt, H.H., Kjolner, S., Borgen, L., Rieseberg, L.H., Brochmann, C., 2006. High biological species diversity in the arctic flora. PNAS 103, 972–975. https://doi.org/10.1073/pnas.0510270103.
Henderson, A., 2011. A revision of *Geonoma* (Arecaceae). Phytotaxa 17, 1–271. https://doi.org/10.11646/phytotaxa.17.1.1.
Henderson, A., Galeano, G., Bernal, R., 1995. A field guide to the palms of the Americas. Princeton University Press, Princeton, New Jersey.
Herzog, S.K., Kessler, M., 1997. Dieta de una colonia de guácharos (Steatornis caripensis) en el Parque Nacional Carrasco, Cochabamba, Bolivia. Ecología en Bolivia 30, 69–73.
Hooghiemstra, H., van der Hammen, T., 2004. Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. Philos. Trans. R. Soc. B. 359, 173–180. https://doi.org/10.1098/rstb.2003.1420.
Hughes, C.E., Atchison, G.W., 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. New Phytol. 207, 275–282. https://doi.org/10.1111/nph.13230.
Hughes, C., Eastwood, R., 2006. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. PNAS 103 (27), 10334–10339. https://doi.org/10.1073/pnas.060192810.
Hughes, C.E., Nyffeler, R., Linder, H.P., 2015. Evolutionary plant radiations: where, when, why and how? New Phytol. 207, 249–253. . https://www.jstor.org/stable/newphytologist.207.2.249.
Jabaily, R.S., Sytsma, K.J., 2013. Historical biogeography and life-history evolution of Andean *Puya* (Bromeliaceae). Bot. J. Linn. Soc. 171, 201–224. https://doi.org/10.1111/j.1095-8339.2012.01307.x.
Jolles, D.D., Wolfe, A.D., 2012. Genetic differentiation and crypsis among members of the myco-heterotrophic *Pyrola picta* species complex (Pyroleae: Monotropoideae: Ericaceae). Syst. Bot. 37 (2), 468–477. https://doi.org/10.1600/036364412X635511.
Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, P., Kessler, M., 2017. Climatologies at high resolution for the Earth land surface areas. Sci. Data 4, 170122. https://doi.org/10.1038/sdata.2017.122.
Knudsen, J., 1999. Floral scent differentiation among coflowering, sympatric species of *Geonoma* (Arecaceae). Plant Species Biol. 14, 137–142. https://doi.org/10.1046/j.1442-1984.1999.00017.x.
Krabbe, N.K., Schulenberg, T.S., Hosner, P.A., Rosenberg, K.V., Davis, T.J., Rosenberg, G.H., Lane, D.F., Andersen, M.J., Robbins, M.K., Cadena, C.D., Valqui, T., Salter, J.F., Spencer, A.J., Angulo, F., Fjeldså, J., 2020. Untangling cryptic diversity in the High Andes: Revision of the *Scytalopus* [magellanicus] complex (Rhinocryptidae) in Peru reveals three new species. Auk 137 (2), ukaa003. https://doi.org/10.1093/auk/ukaa003.
Lagomarsino, L.P., Condamine, F.L., Antonelli, A., Mulch, A., Davis, C.C., 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). New Phytol. 210, 1430–1442. https://doi.org/10.1111/nph.13920.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.
Li, Z., Peng, S., Wang, Q., Li, W., Liang, S., Chen, J., 2023. Cryptic diversity of the genus *Ottelia* in China. Biodivers. Sci. 31 (2), 22394. . https://www.biodiversity-science.net/EN/10.17520/biods.2022394.
Listabarth, C., 1993. Pollination *in Geonoma macrostachys* and three congeners *G. acaulis, G. gracilis*, and *G. interrupta*. Botanica Acta 106, 496–506. https://doi.org/10.1111/j.1438-8677.1993.tb00779.x.
Loiseau, O., Olivares, I., Paris, M., de La Harpe, M., Weigand, A., Koubínová, D., Rolland, J., Bacon, C.D., Balslev, H., Borchsenius, F., Cano, A., Couvreur, T.L.P., Delnatte, C., Fardin, F., Gayot, M., Mejía, F., Mota-Machado, T., Perret, M., Roncal, J., Sanín, M.J., Stauffer, F., Lexer, C., Kessler, M., Salamin, N., 2019. Targeted capture of hundreds of nuclear genes unravels phylogenetic relationships of the diverse Neotropical palm tribe Geonomateae. Front. Plant Sci. 12 (10), 864. https://doi.org/10.3389/fpls.2019.00864.
Londoño, C., Cleef, A.M., Madriñán, S., 2014. Angiosperm flora and biogeography of the paramo region of Colombia, northern Andes. Flora: Morphol. Distrib. Funct. Ecol. Plants 209, 81–87. https://doi.org/10.1016/j.flora.2013.11.006.
Losos, J.B., 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. Am. Nat. 175, 623–639. https://doi.org/10.1086/652433.
Martin, S.H., Davey, J.W., Salazar, C., Jiggins, C.D., 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 17, e2006288.
Mastretta-Yanes, A., Xue, A.T., Moreno-Letelier, A., Jorgensen, T.H., Alvarez, N., Pinero, D., Emerson, B.C., 2018. Long-term in situ persistence of biodiversity in tropical sky islands revealed by landscape genomics. Mol. Ecol. 27 (2), 432–448. https://doi.org/10.1111/mec.14461.
Mayr, E., 1942. Systematics and the Origin of Species. Columbia Univ. Press, New York.
Mayr, E., 1970. Populations, species, and evolution. Harvard Univ. Press, Cambridge.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. . http://www.genome.org/cgi/doi/10.1101/gr.107524.110.
Medina, R., Lara, F., Goffinet, B., Garilleti, R., Mazimpaka, V., 2012. Integrative taxonomy successfully resolves the pseudo-cryptic complex of the disjunct epiphytic moss *Orthotrichum consimile* sl (Orthotrichaceae). Taxon 61 (6), 1180–1198. https://doi.org/10.1002/tax.616002.
Miller, A.D., Skoracka, A., Navia, D., de Mendonca, R.S., Szydło, W., Schultz, M.B., Smith, C.M., Truol, G., Hoffmann, A.A., 2013. Phylogenetic analyses reveal extensive cryptic speciation and host specialization in an economically important mite taxon. Molec. Phylogenet. Evol. 66, 928–940. https://doi.org/10.1016/j.ympev.2012.11.021.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548. https://doi.org/10.1093/bioinformatics/btu462.
Mirarab, S., Warnow, T., 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44–i52. https://doi.org/10.1093/bioinformatics/btv234.
Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of

restriction endonucleases. Proc. Natl. Acad. Sci. 76, 5269–5273. https://doi.org/10.1073/pnas.76.10.5269.
Nicolas, V., Herbreteau, V., Couloux, A., Keovichit, K., Douangboupha, B., Hugot, J.P., 2012. A remarkable case of micro-endemism in *Laonastes aenigmamus* (Diatomyidae, Rodentia) revealed by nuclear and mitochondrial DNA sequence data. PLoS One 7, e48145.
Nosil, P., 2012. Ecological speciation. Oxford University Press, Oxford, United Kingdom.
Pillon, Y., Hopkins, H.C., Rigault, F., Jaffré, T., Stacy, E.A., 2014. Cryptic adaptive radiation in tropical forest trees in New Caledonia. New Phytol. 202 (2), 521–530. https://doi.org/10.1111/nph.12677.
QGIS.org. QGIS Geographic Information System. QGIS Association. http://www.qgis.org.
Reynolds, J., Weir, B.S., Cockerham, C.C., 1983. Estimation of the coancestry coefficient: Basis for a short-term genetic distance. Genetics 105 (3), 767–779. https://doi.org/10.1093/genetics/105.3.767.
Rieseberg, L.H., Willis, J.H., 2007. Plant speciation. Science 317, 910–914. https://doi.org/10.1126/science.1137729.
Roncal, J., Francisco-Ortega, J., Asmussen, C.B., Lewis, C.E., 2005. Molecular phylogenetics of tribe Geonomeae (Arecaceae) using nuclear DNA sequences of Phosphoribulokinase and RNA Polymerase II. Syst. Bot. 30, 275–283. https://doi.org/10.1600/0363644054223620.
Roncal, J., Blach-Overgaard, A., Borchsenius, F., Balslev, H., Svenning, J.C., 2011. A dated phylogeny complements macroecological analysis to explain the diversity patterns in *Geonoma* (Arecaceae). Biotropica 43, 324–334. https://doi.org/10.1111/j.1744-7429.2010.00696.x.
Sanín, M.J., Cardona, A., Valencia-Montoya, W.A., Jiménez, M.F.T., Carvalho-Madrigal, S., Gómez, A.C., Bacon, C.D., Tangarife, T.R., Jaramillo, J.S., Zapata, S., Valencia, V., 2022a. Volcanic events coincide with plant dispersal across the Northern Andes. Glob. Planet. Change 210, 103757. https://doi.org/10.1016/j.gloplacha.2022.103757.
Sanín, M.J., Mejía-Franco, F.G., Paris, M., Valencia-Montoya, W.A., Salamin, N., Kessler, M., Olivares, I., Jaramillo, J.S., Cardona, A., 2022b. Geogenomics of montane palms points to Miocene-Pliocene Andean segmentation related to strike-slip tectonics. J. Biogeogr. 49 (9), 1711–1725. https://doi.org/10.1111/jbi.14327.
Sanín, M.J., Borchsenius, F., Paris, M., Carvalho-Madrigal, S., Hoyos, A.C.G., Cardona, A., Arcila-Marín, N., Ospina, Y., Hoyos-Gómez, S.E., Manrique, E.F., Bernal, R., 2022c. The tracking of moist habitats allowed *Aiphanes* (Arecaceae) to cover the elevation gradient of the northern Andes. Front. Plant Sci. 13, 881879. https://doi.org/10.3389/fpls.2022.881879.
Särkinen, T., Pennington, R.T., Lavin, M., Simon, M.F., Hughes, C.E., 2012. Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. J. Biogeogr. 39, 884–900. https://doi.org/10.1111/j.1365-2699.2011.02644.x.
Shimizu, K.K., 2022. Robustness and the generalist niche of polyploid species: Genome shock or gradual evolution? Curr. Opin. Plant Biol. 69, 102292. https://doi.org/10.1016/j.pbi.2022.102292.
Simpson, B.B., 1974. Glacial migrations of plants: island biogeographical evidence. Science 185, 698–700. https://doi.org/10.1126/science.185.4152.698.
Simpson, B.B., 1975. Pleistocene changes in the flora of the high tropical Andes. Paleobiology 1, 273–294. https://doi.org/10.1017/S0094837300002530.
Singhal, S., Hoskin, C.J., Couper, P., Potter, S., Moritz, C., 2018. A framework for resolving cryptic species: a case study from the lizards of the Australian wet tropics. Syst. Biol. 67 (6), 1061–1075. https://doi.org/10.1093/sysbio/syy026.
Smeds, L., Künstner, A., 2011. ConDeTri-a content dependent read trimmer for Illumina data. PLoS One 6, e26314.
Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K.H., Liow, L.H., Nowak, M.D., Stedje, B., 2018. Finding evolutionary processes hidden in cryptic species. Trends Ecol. Evol. 33 (3), 153–163. https://doi.org/10.1016/j.tree.2017.11.007.
Theodoridis, S., Nogués-Bravo, D., Conti, E., 2019. The role of cryptic diversity and its environmental correlates in global conservation status assessments: Insights from the threatened bird's-eye primrose (*Primula farinosa* L.). Divers. Distrib. 25 (9), 1457–1471. https://doi.org/10.1111/ddi.12953.
Turelli, M., Barton, N.H., Coyne, J.A., 2001. Theory and speciation. Trends Ecol. Evol. 16, 330–343. https://doi.org/10.1016/S0169-5347(01)02177-2.
Vargas, O.M., Ortiz, E.M., 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. https://doi.org/10.1111/nph.14530.
Vargas, O.M., Goldston, B., Grossenbacher, D.L., Kay, K.M., 2020. Patterns of speciation are similar across mountainous and lowland regions for a Neotropical plant radiation (Costaceae: *Costus*). Evolution 74 (12), 2644–2661. https://doi.org/10.1111/evo.14108.
Vrijenhoek, R.C., 2009. Cryptic species, phenotypic plasticity, and complex life history: assessing deep-sea fauna diversity with molecular markers. Deep-Sea Res Part II - Topical Stud. Oceanogr. 56, 1713–1723. https://doi.org/10.1016/j.dsr2.2009.05.016.
Weigend, M., 2002. Observations on the biogeography of the Amotape-Huancabamba Zone in northern Peru. Bot. Rev. 68, 38–54. https://doi.org/10.1663/0006-8101(2002)068[0038:OOTBOT]2.0.CO;2.