Species rarity has puzzled ecologists for decades owing to its theoretical fundamentals and practical implications. The main concern that rarity raises from a conservation standpoint is that rare species tend to have higher extinction risk. Understanding the underlying phylogenetic patterns of rarity can shed light on the evolutionary processes yielding rarity, contribute to forecasting potential vulnerable species or lineages, and assess the consequences of the loss of rare species. Here, we apply Rabinowitz’s rarity classification scheme and explore phylogenetic patterns in the flora of the Pyrenees, a temperate mountain range. We first categorized species according to their geographic range (endemicity and regional geographic range), habitat specialization and local abundance. Then we analyzed the phylogenetic signal of each rarity type, along with the phylogenetic association between species according to their rarity, and decomposed the variation of rarity among taxonomic levels to better understand its diversification through time. Finally, we estimated the expected loss of phylogenetic diversity (PD) in the regional flora if rare species go extinct. All rarity types showed a certain degree of phylogenetic signal, with endemics, species of limited regional geographic range and habitat specialists forming groups of phylogenetically close species. The decomposition of variation supported these results, as the rarity types with the strongest signal had more variation explained by higher taxonomic levels. The loss of these groups of species had different consequences based on the type of rarity, with the disappearance of habitat specialists having the strongest, negative effect on PD. Our results show a general pattern of phylogenetic association in rarity among the flora of the Pyrenees, which leads to increased vulnerability along certain branches of the Tree of Life. Phylogenetic patterns should be taken into account in conservation planning to effectively protect all facets of biodiversity.

Keywords: endemism, habitat specialization, phylogenetic signal, plant inventories, rare species, threatened
Introduction

The notion of rarity and the mechanisms behind it have interested ecologists for decades (Preston 1948, Stebbins 1980, Rabinowitz 1981) as it is related to one of the central subjects of ecology: the abundance and distribution patterns of species across time and space. There is, however, a more pragmatic aspect of rarity that is of great importance for conservation biology and practice, especially in the current context of global change and biodiversity loss (Pimm et al. 2014, Ceballos et al. 2015): rare species are more likely to go extinct than common species (Mace and Kershaw 1997, Davies et al. 2004).

Rarity is a complex concept that can be approached from several angles, although the classification proposed by Rabinowitz (1981) has been widely adopted (Espeland and Emam 2011, Loza et al. 2017, Choe et al. 2019, Crisfield et al. 2020). Her method classifies species into seven rarity categories based on their geographical range, habitat specificity (HIS) and population size. Although Rabinowitz's ideas were not directly related to vulnerability or extinction risk, later studies have shown that rare species are more susceptible to different drivers of extinction, to the point that some aspects of what we can consider rarity, namely geographic range and population size, are part of the criteria for inclusion in IUCN's Red List (IUCN 2012). First, species with restricted geographical ranges, such as narrow endemics or those with limited spatial distributions, have been shown to have higher extinction risk throughout history (Mckinney 1997, Harnik et al. 2012, Saupe et al. 2015). Second, species limited to certain habitats or with very narrow environmental niches also show increased vulnerability to extinction as they are dependent on the preservation of particular abiotic conditions for their survival, and thus are very susceptible to environmental changes (Saupe et al. 2015, Staude et al. 2020). Finally, species with small populations are more vulnerable to stochastic processes that could lead to their disappearance (Lande 1993, Matthes et al. 2004). Since different kinds of rarity may respond very differently to similar sets of stressors (e.g. global warming, land use change or invasive species), identifying which species are rare and their type of rarity can help to optimize conservation efforts, regarding both the evaluation of vulnerability to different threats and the identification of potentially endangered species.

Rarity is also a reflection of the evolutionary history of a species and its ability to establish itself and thrive in different regions and environments (Gaston and Kunin 1997). Under the assumption that traits related to rarity are heritable to some extent (Mouquet et al. 2012), a phylogenetic approach can be helpful to identify potentially rare species (Webb and Gaston 2003). In that case, we would expect rare plants to show some phylogenetic signal; that is, the tendency of phylogenetically close species to resemble one another more than other species (Münkemüller et al. 2012). These phylogenetic patterns vary depending on the spatial scale, the taxonomic group in study and the environmental conditions of the area (Zacaï et al. 2017). However, studying the phylogenetic patterns of rarity can contribute to the identification of potentially rare taxa for which there is no available field information assuming that related species would be similarly rare, which would help us find vulnerable taxa or clades in a phylogeny (Manne and Pimm 2001, Robbins et al. 2006, Toledo et al. 2014). The importance of finding such vulnerable and phylogenetically close taxa lies in that, if these rare species were particularly sensitive to one kind of global driver, the loss in phylogenetic diversity (PD) derived from it would be greater than expected by chance, given the non-random nature of these extinctions (Heard and Mooers 2000, Thuiller et al. 2011).

Here, we apply the rarity framework proposed by Rabinowitz (1981) to the flora of the Pyrenees, a mountain region in southwestern Europe, and investigate the existence of phylogenetic signal associated with geographic rarity (endemicy and regional geographic range (RGR) size), HS and local abundance (LA). In addition, we explore if these rarity types are taxonomically clustered. More specifically, we address the following questions and our expectations about them. 1) Is there any phylogenetic signal for each kind of rarity among the rare plants of the Pyrenees? We expect different phylogenetic patterns for each rarity type because of different underlying mechanisms. For example, the phylogenetic signal of endemicy could depend on the fact that recently diverged species would be closely related (neoeendemisms) and thus show phylogenetic signal, whereas species coming from ancient lineages would be phylogenetically isolated and thus would not show any signal (Mishler et al. 2014). We expect RGR to show some phylogenetic signal related to factors such as limited dispersal ability or niche breadth, which are assumed to be at least partially heritable (Sexton et al. 2017, Saastamoinen et al. 2018). Habitat specialization should also show phylogenetic signal under the assumption of niche conservatism, which has already been observed in plants (Prinzing et al. 2001). Locally scarce species, however, are not expected to show any phylogenetic signal given that many other current factors like resource availability, interspecific interactions, founder effects or environmental filtering have strong influences in the LA of species, dampening any possible phylogenetic patterns. 2) Will the loss of PD be greater than expected by chance if rare plants become extinct? The loss of PD will depend on the degree of phylogenetic relatedness between rare species and the length of the branches in which they are located. If rare species are closely related and located in clades stemming from long branches, which capture more PD, their extinction will lead to a higher loss than expected by chance because that would affect larger and deeper sections of the phylogeny. In contrast, the loss of diversity will be less than randomly expected if rare species are overdispersed in the phylogeny. Finally, to test the relationship between patterns found in our study and practical conservation, we inquire in which way threatened Pyrenean species included in the Pyrenean Red List of vascular plants are associated with phylogenetically close rare plant species. We expect a high degree of overlap between our assessment of rarity and the Red List, given that both share some classification criteria.
Material and methods

Data gathering

We downloaded 18,842 plant inventories carried out over the last 70 years in the Pyrenean area from the Iberian and Macaronesian Vegetation Information System (SIVIM) (Font et al. 2017). This dataset contained around 400,000 plant records of more than 2,300 taxa at species level. Each inventory included information on altitude, some habitat description (phytosociological association, alliance or other), species number and their abundance. The latter was recorded in different scales, depending on the inventories, although most of them follow the classic semi-quantitative scale of Braun–Blanquet that assigns an abundance value to each plant species based on its cover. To properly compare between all species, we transformed all data to the extended Braun–Blanquet scale following van der Maarel (1979), which ranges from 1 to 9. To focus on mountain habitats, we excluded coastal areas and discarded inventories located below 400 m a.s.l. In addition, we removed any inventories containing fewer than five species to ensure proper sampling size and plant representation. Plant names were validated using the Atlas of the Pyrenean Flora (FLORAPYR; <www.atlasflorapyrenaea.eu/src/taxon/index.php?idma=0>), an international project addressing the compilation of all the information available about vascular plants and bryophytes of the Pyrenees and its piedmont. To ensure that the species in our study were representative of the Pyrenean flora, we also excluded all non-native species according to the Atlas of the Pyrenean Flora. Finally, habitats were grouped into one of 14 European Nature Information System (EUNIS) habitats (Table 1; see García et al. 2022 for more information on their geographic distribution in the Pyrenees), a classification for the terrestrial and marine habitat types of the European continent (Moss 2008).

Phylogenetic inference

We used the phylogeny of Roquet and García (2022), a dated genus-level phylogeny built specifically for the Pyrenean flora, using sequences downloaded from GenBank of three chloroplastic regions (rbcL, matK and ndhF), plus the nuclear ribosomal ITS region for some families. It comprises all plant genera in the Pyrenees according to FLORAPYR, except Cytinus, Ptychotis and Xatardia, for which no useful phylogenetic markers were available in GenBank. To be able to work at the species level, we randomly resolved genus-level polytomies following a Yule process. This method randomly resolves polytomies assuming that all taxa have an equal probability of undergoing a speciation event at any moment in time and without following any particular speciation and extinction rates (Gernhard et al. 2008). We repeated this process ten times to produce a distribution of possible evolutionary hypotheses sensu Rangel et al. (2015).

Rarity assessment

We considered four rarity criteria: endemicity and RGR size as complementary components of geographic range; HS; and LA. A species was considered endemic when its global distribution was limited to the Pyrenees. RGR size was measured as the highest number of $10 \times 10$ km UTM cells occupied by the target species within the FLORAPYR grid or the SIVIM database. HS was estimated for each species by combining the frequency of each species and the frequency of the habitats where it occurs using Hurlbert’s $B^\prime$ resource use index as described by Feinsinger et al. (1981): 

$$B^\prime = 1/\sum_i \left( p_i^2 / q_i \right)$$

where $p_i$ represents the proportion of occurrences of the target species in habitat $i$ and $q_i$ the relative abundance of such habitat in the study region. This index ranges from 0 for the rarest species to 1 for the most common species. Its main advantage is that it gives more weight to rare habitats, so that the rarest species are those found in just a few and scarce habitats. To obtain a more in-depth idea of the habitats with which each species is associated, we calculated the IndVal index of

<table>
<thead>
<tr>
<th>EUNIS habitat</th>
<th>n inventories</th>
<th>Species richness</th>
<th>n UTM10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpine and subalpine grasslands</td>
<td>3296</td>
<td>1023</td>
<td>146</td>
</tr>
<tr>
<td>Arctic, alpine and subalpine scrub</td>
<td>446</td>
<td>566</td>
<td>75</td>
</tr>
<tr>
<td>Broadleaved deciduous woodland</td>
<td>3408</td>
<td>1189</td>
<td>310</td>
</tr>
<tr>
<td>Broadleaved evergreen woodland</td>
<td>453</td>
<td>600</td>
<td>94</td>
</tr>
<tr>
<td>Coniferous woodland</td>
<td>1330</td>
<td>875</td>
<td>149</td>
</tr>
<tr>
<td>Dry grasslands</td>
<td>1928</td>
<td>1200</td>
<td>220</td>
</tr>
<tr>
<td>Garrigue</td>
<td>1080</td>
<td>879</td>
<td>188</td>
</tr>
<tr>
<td>Inland cliffs, rock pavements and outcrops</td>
<td>1222</td>
<td>966</td>
<td>191</td>
</tr>
<tr>
<td>Mires and lenses</td>
<td>1209</td>
<td>689</td>
<td>180</td>
</tr>
<tr>
<td>Screes</td>
<td>1178</td>
<td>738</td>
<td>123</td>
</tr>
<tr>
<td>Seasonally wet and wet grasslands</td>
<td>431</td>
<td>716</td>
<td>112</td>
</tr>
<tr>
<td>Surface waters</td>
<td>375</td>
<td>507</td>
<td>123</td>
</tr>
<tr>
<td>Temperate and mediterranean-montane scrub and heathlands</td>
<td>1156</td>
<td>1069</td>
<td>225</td>
</tr>
<tr>
<td>Woodland fringes and clearings, and tall forb stands</td>
<td>1330</td>
<td>1269</td>
<td>239</td>
</tr>
</tbody>
</table>
Dufrene and Legendre (1997). This index gives a degree of association between a single species and each of the habitats in which it is found, considering the abundance of both the species and the habitats. Information on the abundance of habitats, measured as the number of inventories associated with each habitat, and the distribution of species in each of them, was extracted from the SIVIM database. LA for each species was estimated as its average abundance value among all the inventories in SIVIM. Prior to any further analysis, we standardized RGR, HS and LA values to z-scores (mean = 0 and SD = 1) to enable comparisons.

Phylogenetic patterns of rarity

We used two methods to study phylogenetic signal depending on the nature of each rarity type. For RGR, HS and LA we computed Pagel’s $\lambda$ (Pagel 1999), which measures phylogenetic signal in continuous variables and compares it to a Brownian motion model. This index ranges between 0 and 1, with 0 indicating random distribution and 1 evolution under Brownian motion, implying phylogenetic signal. For endemism, which is a binary variable, we computed the D statistic of Fritz and Purvis (2010). This index employs a binomial distribution assuming a latent continuous variable and compares it to a Brownian motion model of evolution. For easier comparison with Pagel’s $\lambda$ we transformed D into $-D+1$, and thus values equal to 0 indicate random distribution and values close to 1 indicate phylogenetic signal (Goberna and Verdú 2016). In addition, we computed pPCA as described in Revell (2009) based on the values of RGR, HS and LA. We applied this method for two reasons: first, it calculates a multivariate $\lambda$ to test for phylogenetic signal in multiple traits at the same time (Ibanez et al. 2016), and second, it allows an easy visualization of the correlation between rarity components and the similarities between species taking into account phylogenetic information (Uyeda et al. 2015).

Common phylogenetic signal metrics like Pagel’s $\lambda$ or Fritz and Purvis’s D give a phylogeny-wide value without identifying the regions of the tree where species that closely resemble one another accumulate. To identify those regions, we computed the local index of phylogenetic association (LIPA) for each species and all rarity types (Keck et al. 2016). This measure is adapted from the local index of spatial association of Anselin (1995), which is a local case of Moran’s autocorrelation index $I$. Positive LIPA values identify species that tend to share similar rarity values with their close relatives. We tested if LIPA values were statistically significant by comparing the observed values to a null model that randomly shuffles the tips of the phylogeny 999 times.

Finally, we tested if the loss of PD in the Pyrenees caused by the extinction of rare species with significant phylogenetic association (i.e. those with positive, significant LIPA values) would be greater than expected by chance. We considered as rare species those with values of RGR, HS and LA lower than average, and all endemics. Then, we followed a procedure similar to Von Euler (2001). First, we calculated the total PD of the phylogeny by summing the length of all branches in the tree and then, for each rarity type separately, we removed rare species from the phylogeny and recalculated PD for the resulting tree. To test if the loss of PD (i.e. the difference in PD before and after the removal of species) was greater than expected by chance, we repeated the process 999 times, removing a set of randomly chosen species of the same size as the number of rare species removed. To compare the loss of PD between rarity types we calculated the standard size effect of the loss (SES$_{PD\_loss}$) for each type, by subtracting the mean of the null values from the observed value and dividing by the standard deviation of the null distribution. Values were considered statistically different from random expectation with a 95% confidence if they were outside the $[-1.96, 1.96]$ interval (Mazel et al. 2016).

Phylogenetic analyses are dependent on the phylogenetic scale used in them (Graham et al. 2018), especially if the phylogeny has deep bifurcations like the one between angiosperms and gymnosperms. Thus, we conducted all analyses using three sets of species: one with all vascular species; one with only angiosperms; and another containing only the oldest groups, gymnosperms and moniliophytes. To consider all possible evolutionary hypotheses among plant species in the Pyrenees we repeated each analysis using our ten phylogenies and averaged the resulting indices.

Taxonomic patterns of rarity

To better understand how rarity is distributed among different taxonomic levels, we fitted a Bayesian random effect model for each rarity type (binomial for endemism and Gaussian for scaled and centered RGR, HS and LA) with rarity as dependent variable and a random effect consisting of genus nested within family and both nested within order. This method partitions the variation (variance for RGR, HS and LA, and deviance for endemism) of each rarity type among taxonomic levels and an unexplained residual component. While considering the nested nature of taxonomic classification (Asner and Martin 2016, Oliveras et al. 2020, Martinelli et al. 2021), Models were fitted using uninformative priors and four Markov chains with 4000 iterations each, a thinning interval of 10 and a burn-in period of 1000 iterations. We calculated the proportion of variation explained by each taxonomic level by comparing it to the total variation explained by the random effects, including the random residual component.

Rarity and conservation status

The conservation status of the plants in our dataset was obtained from the Red List of the Pyrenean vascular flora published by the FLORAPYR project (<www.opcc-ctp.org/en/florapyr>). To explore how rarity relates to threatened species in the Red List (critically endangered: CR, endangered: EN and vulnerable: VU), each species of our dataset was plotted along the first two components of our phylogenetic principal component analysis (pPCA).
All analyses were conducted in R ver. 4.1.2 (<www.r-project.org>). Pagel's $\lambda$ and LIPA values were computed with the phylosignal package (Keck et al. 2016) while Purvis's $D$ was calculated with the caper package (Orme et al. 2018). IndVals were calculated with the labdsv package (Roberts 2019). Random models were fitted using the rstanarm package with the uninformative priors and parameters provided by the `stan_glmer` function (Goodrich et al. 2020). The variance or deviance of the random effects of each random model was assessed with the insight package (Lüdecke et al. 2019). The variance or deviance of the random effects of each random model was assessed with the insight package (Lüdecke et al. 2019). Package phytools was used for the phylogenetic PCA (Revell 2012).

**Results**

According to the SIVIM database, plant species in the Pyrenees occupied, on average, 133.02 (SD = 114.87) $10 \times 10$ km UTM cells, had an average HS of 0.2 (SD = 0.12) and an average LA of 2.78 (SD = 0.87) in the Braun–Blanquet extended scale. Only 78 (3.31%) of the 2351 species in our dataset were endemic to the region. Analysis found that 568 species (24.16% of the total) had below-average values of RGR, HS and LA at the same time. In addition, 28 of these species were also endemic to the region (Supporting information).

**Phylogenetic patterns of rarity and their consequences**

We observed very similar and consistent patterns in phylogenetic signal between all three datasets of major plant groups (Fig. 1 and Supporting information) and thus only the results from the complete dataset are reported in the main text. Every rarity type showed statistically significant phylogenetic signal ($p < 0.05$), but the strength varied between types: LA exhibited the strongest signal, followed by HS, RGR and endemism (Fig. 1). The pPCA indicated a certain degree of signal for RGR, HS and LA together (mean $\lambda = 0.43$, SD = 0.01).

A total of 595 of 2351 species (25.3%) had significant, positive LIPA values (Fig. 2, Supporting information). Although we observed differences between each aspect of rarity type in how these species were distributed in the phylogeny, we detected that endemism, RGR and HS presented clearly defined groups of species that contributed more to their phylogenetic signal, both for higher and lower values than the average (Fig. 2). LA, on the other hand, had significant LIPA values spread across the whole phylogeny, mostly caused by species with LA values lower than the average.

In contrast to the loss of other kinds of rare species with significant LIPA values greater than 0, the removal of endemics did not result in statistically significant changes in PD ($\text{SES}_{PD,\text{Loss}} = -0.76$). However, the decrease in PD differed between the other rarity types (Fig. 1b): the loss of habitat specialists ($\text{SES}_{PD,\text{Loss}} = 45.47$) and species with low LA ($\text{SES}_{PD,\text{Loss}} = 10.16$) resulted in a much higher PD loss than expected under the random loss, whereas the loss of species with limited RGR led to a lower PD loss than expected ($\text{SES}_{PD,\text{Loss}} = -8.89$).

**Taxonomic patterns of rarity**

The partitioning of variation (variance or deviance) of each rarity type among taxonomic levels using random models indicated that between 10 and 42% of variation was explained by taxonomy (Fig. 3a). Endemism and LA had the highest proportion of variation explained by all taxonomic levels together (42.5% and 36.4%, respectively). Between taxonomic levels, the highest variation was found at the genus level, except for LA, where family accounted for the highest proportion of variation. These results were congruent with the analysis of LIPA. The species with significant LIPA values for each rarity type belonged to families with a higher proportion of endemics and lower values of RGR, HS and LA, according to the random model (Supporting information).

---

**Figure 1.** (a) Distribution of the estimated values of phylogenetic signal (Pagel’s $\lambda$ for regional geographic range (RGR), habitat specificity (HS) and local abundance (LA) and Purvis’s $D+1$ for endemism) using ten versions of the phylogeny for each rarity type. (b) Standard effect size of the loss of phylogenetic diversity (PD) after removing rare species with significant local index of phylogenetic association (LIPA) values. Shaded area indicates the 95% confidence interval $[-1.96$ to $1.96]$. 

---

Page 5 of 11
A total of 11 threatened species was present in our data-
set, representing a small fraction of the 64 threatened taxa
included in the Red List of the Pyrenean vascular flora. Half
of these species (1 critically endangered and 5 vulnerable)
had significant LIPA values and were endemic or had below-
average values of RGR, HS or LA. In addition, all threatened
species tended to have small RGR and be habitat specialists,
but did not necessarily show low LA (Fig. 3b).

Conservation status of rare plants in the Pyrenees

A total of 11 threatened species was present in our data-
set, representing a small fraction of the 64 threatened taxa
included in the Red List of the Pyrenean vascular flora. Half
of these species (1 critically endangered and 5 vulnerable)
had significant LIPA values and were endemic or had below-
average values of RGR, HS or LA. In addition, all threatened
species tended to have small RGR and be habitat specialists,
but did not necessarily show low LA (Fig. 3b).

Discussion

We present the first study of the taxonomic and phyloge-
netic patterns of four kinds of rarity in the flora of a temper-
ate and diverse mountain region, the Pyrenees. We detected
significant phylogenetic signal of varying intensity for all
aspects of rarity: endemism, RGR, HS and LA. In addition,
our analysis of the local contribution of each species to phylo-
genetic signal revealed distinct groups of closely related
species that were similar in different aspects of rarity, espe-
cially RGR and HS. The taxonomic analysis was congruent
with the observed phylogenetic signal. These results support
our hypothesis that rarity in plants is conserved to different
degrees through phylogenies. Thus, rare species tend to be
closely related, leading to phylogenetic clustering of species
more vulnerable to extinction.

Phylogenetic and taxonomic aspects of species rarity

Rarity is a multifaceted and complex phenomenon whose
phylogenetic patterns can vary depending on the spatial and
phylogenetic scope of the analysis. These patterns can also
vary depending on the group studied and type of rarity, due to

Figure 2. Phylogenetic tree of the Pyrenean species included in
Iberian and Macaronesican Vegetation Information System
(SIVIM) inventories (one of the ten versions produced), with dots
and bars in color depicting those species with significant local index
of phylogenetic association (LIPA) values (p ≤ 0.05). Rings from
the inside out: endemic species (white circles), scaled regional geo-
graphic range (RGR) values (red bars), scaled habitat specialization
values (blue bars), scaled local abundance (yellow bars).

Figure 3. (a) Proportion of variation explained by each taxonomic
level in the random models fitted for each rarity type. (b) First two
components of the phylogenetic principal component analysis
(pPCA) based on regional geographic range (RGR), habitat speci-
ficity (HS) and local abundance (LA). Squares indicate endemic
species and colored points threatened species in the Red List of the
Pyrenean vascular flora. Percentages in each axis indicate the
amount of variance explained by each component.
the interplay between the environmental features of the study region and species traits, niche breadth or interspecific interactions (Wamelink et al. 2014, Zacaï et al. 2017). Despite this heterogeneity, the phylogenetic signal we observed in the flora of a temperate mountain system adds new evidence to the existence of this pattern, observed in other systems and organisms like tropical floras (Dexter and Chave 2016, Loza et al. 2017) and North American continental fishes (Giam and Olden 2018), and partly consistent with patterns observed in European birds (Cortgeave and Harvey 1992, Pearman et al. 2014), amphibians (Bonetti and Wiens 2014) and terrestrial vertebrates (Pie et al. 2021a, b). Altogether, this suggests that processes ruling over the phylogenetic patterns of rarity might be consistent throughout regions despite differences in evolutionary history and climate, and that there is a certain degree of congruence in these processes between groups of plants and animals (Liu et al. 2020).

Nevertheless, disentangling the evolutionary processes behind the phylogenetic patterns of different aspects of rarity can be a daunting task. An endemic species, for instance, can be limited to a particular area for different reasons such as recent speciation and not having had enough time to expand beyond its initial range, or because its formerly wider geographical range has been reduced and its conspecific species have gone extinct around the world (Stebbins and Major 1965, Kruckeberg and Rabinowitz 1985). The phylogenetic imprint left by either of these processes would be very different. Recently evolved species are expected to present significant phylogenetic clustering due to several closely related species coexisting in the region. Meanwhile, paleoendemics are presumed to show phylogenetic overdispersion, as these species would evolve evolutionarily isolated from the other extant species in the tree (Mishler et al. 2014). Our phylogenetic analysis, along with the high variation in endemism at the genus level, highlights the presence of clusters of endemic species throughout the phylogeny. Previous studies found an increase in diversification rates throughout the Pyrenees during the late Neogene due to the Alpine uplift (Boucher et al. 2016), Campanula (Roquet et al. 2022) and Saxifraga (Vargas et al. 2018) within the Pyrenees (Ninot et al. 2017). This increase in diversification rates likely led to allopatric speciation and the high proportion of neoendemic species found in genera Androsace (Boucher et al. 2016), Campanula (Roquet et al. 2022) and Saxifraga (Vargas et al. 2018) within the Pyrenees (Ninot et al. 2017).

Regional geographic range and habitat specialization follow a similar pattern of relatively weak phylogenetic signal accompanied by a few clusters of very closely related rare species. This pattern is further supported by the low importance of taxonomy regarding variation in rarity across the phylogenetic tree. Concerning geographic range, these sparse groups of phylogenetically close species with small geographic ranges could be the result of allopatric speciation coupled with limited abilities to disperse beyond their original distribution (Böhning-Gaese et al. 2006, Zacaï et al. 2017). These results are consistent with the geological history of the Pyrenees, where mountain uplifts and glacial cycles have expanded and contracted the geographic ranges of species, favoring allopatric speciation (Kadereit et al. 2004, Boucher et al. 2016, Wallis et al. 2016, Ninot et al. 2017). However, to fully understand the evolutionary mechanisms behind this pattern, it would be necessary to study not only the RGR of each species, but also the overlap between their spatial distributions. With respect to habitat specialization, the presence of separate clusters of species dispersed throughout the phylogeny highlighted by the analysis of local phylogenetic association suggests that the adaptations allowing species to inhabit very particular habitats have evolved several times and at different points in the evolutionary history of the region. The taxonomic analysis gives additional support to this hypothesis, as the limited variation explained by taxonomy is more or less evenly distributed among genera, families and orders, suggesting the diversification of habitat specialization at different evolutionary times. According to the IndVal measure of habitat specialization, around 75% of the specialist species with significant phylogenetic association (Supporting information) are present in aquatic habitats such as bogs, fens, mires and inland water bodies, indicating that the traits facilitating life in such particular environments have been conserved through evolutionary history after evolving at different times; this reflects the evolutionary pattern of adaptations to aquatic life observed in angiosperms by Philbrick and Les (1996). It is noteworthy that habitat specialists living in other habitats often considered stressful, such as scree, rocky cliffs or high altitude grasslands, do not show a clear phylogenetic pattern. This is in contrast with the patterns of phylogenetic closeness observed in plants living in similar habitats of high altitude summits of other temperate mountain regions like the Alps observed by Marx et al. (2017), suggesting differences in the evolutionary processes in the flora of each region.

We observed a strong phylogenetic signal for LA, along with the highest proportion of variation explained by taxonomy of any rarity type. Although these results challenge our initial hypothesis, they follow the patterns of phylogenetic signal in LA found by other authors in birds (Cortgeave and Harvey 1992), tropical plants (Dexter and Chave 2016, Loza et al. 2017) and terrestrial vertebrates (Pie et al. 2021a). Most species in our data show low LA, which could favor the strong phylogenetic signal we observed. However, disentangling the mechanisms behind this pattern is quite difficult. First, these abundance distribution patterns have been recently questioned in the literature because they can arise from non-biological mechanisms, as seen in other study systems (Warren et al. 2011, Keil et al. 2018). Second, the determinants of species LA are very diverse, ranging from intraspecific variation in life history traits (Kolb et al. 2006) to local environmental conditions (Bertness and Ellison 1987), disturbance regimes (Guedo and Lamb 2013) or interspecific interactions (Levine and Rees 2002). However, the strong phylogenetic signal and the high proportion of variation explained by the family of each species suggests that certain traits determining the LA of species have been conserved at that taxonomic level. Several authors have highlighted the importance of traits such as plant growth form,
plant height or specific leaf area as determinants of local plant abundance (Murray et al. 2002, Cornwell and Ackerly 2010, Lauterbach et al. 2013). Studying the phylogenetic patterns of these traits, along with any other characteristics susceptible of influencing local abundance, would shed light on the evolutionary mechanisms that drive plant rarity.

**Implications for conservation of the Pyrenean flora**

The consequences derived from rare species being phylogenetically clustered in one way or another are straightforward from a conservation point of view: the higher extinction risk associated with these species implies a greater loss of phylogenetic diversity (PD) than expected under a random distribution of rarity (Heard and Mooers 2000, Thuiller et al. 2011). However, the magnitude of the loss depends on the degree of phylogenetic relatedness of those rare species, as well as the amount of phylogenetic diversity they represent. For instance, the loss of phylogenetically related habitat specialist species in the Pyrenees would have an important effect on the PD of the region, with a much higher loss than the amount expected by chance. This suggests that these species are not only phylogenetically close, but that they are also located in branches representing large amounts of PD. The opposite can be found in species with small RGRs, whose disappearance would imply a smaller loss compared to other species, most likely due to these kinds of rare species being located in shorter branches of the phylogenetic tree. The loss of PD permeates into other aspects of diversity because it may potentially act as a proxy for multiple species traits and functions whose loss could have an impact on ecosystems function, particularly in the case of specialists that tend to concentrate in one habitat (Srivastava et al. 2012, Winter et al. 2013).

The likelihood of extinction of a species depends on a combination of intrinsic factors like its rarity along with other external factors affecting these species and the scale at which they are assessed (Wilson et al. 2005, Veron et al. 2017). Regarding rarity, we found 28 species that we could consider very vulnerable to extinction, as they combine all four rarity types considered in this study. They belong to 17 different families, and approximately one third of them to the Asteraceae and Saxifragaceae families (Supporting information). This is congruent with our results from the phylogenetic PCA, which indicated a moderate phylogenetic signal for the combination of small RGR, habitat specialization and low LA. Most of our threatened plant species in the Pyrenees tend to be habitat specialists with small geographic ranges, which makes sense since one of the main criteria for threat assessment in the Red Lists is the geographic range of species. It is interesting to discover that threatened species do not necessarily have lower LA than other species. This indicates that, although these species might be threatened at a regional scale, they fare well within their local communities. Meanwhile, none of our rarest plant species were threatened, according to the Red List of the Pyrenean Flora. This may be caused by discrepancies between our criteria for rarity and those used for the Red List, because the rarest species are not necessarily the most threatened in the Pyrenees; or because our data set only included a fraction of the listed species. Regarding external factors to the vulnerability of diversity, a recent study by Miranda Cebrián et al. (2022) explored the distribution patterns of plant rarity in the Pyrenees and observed that rare species tend to concentrate in distinct habitats like wetlands, mires and fens, or rocky outcrops. Here we find two very different situations. On the one hand, 16 (57%) of the rarest species were found in rocky habitats like cliffs and scree. These are very stable habitats which are considered contemporary refugia resistant to the impacts of global change, acting as safe havens of biodiversity (García et al. 2020, Brighenti et al. 2021). On the other hand, wetlands and mires are highly vulnerable to global change according to the European Red List of Habitats (Janssen et al. 2016), which puts any rare species inhabiting those areas in a double jeopardy of intrinsic and extrinsic factors contributing to their extinction risk and the subsequent potential loss of PD in the region.

One positive aspect to the phylogenetic clustering of rarity is that it might be informative for conservation efforts: knowing that rare species tend to be phylogenetically close can help to identify potentially vulnerable groups of species for which we do not have enough information for proper vulnerability assessment (Winter et al. 2013). In this study, we used a huge number of plant inventories, information that might not be available in other regions, and found a good match between threatened species according to the regional Red List and our classification of rare species in terms of reduced RGR. If our results are confirmed for other areas, then conservation planning in vast territories could benefit from the premise that species related to rare taxa are likely to be rare too, and also vulnerable.

**Caveats and limitations**

There are two main limiting factors in our study. The first comprises possible biases caused by uneven sampling effort of our data through time and space. Using data gathered by different sources and at different time periods can lead to differences in how LA has been assessed, or to changes through time in the same location. However, the nature of the Braun–Blanquet scale used in the inventories buffers against large differences between observers: the smallest scores are very close together and mistakes in using them would still return small abundances, whereas the bigger scores are less prone to be wrongly applied as they encompass wider ranges of cover. The second factor is the resolution of our phylogeny. Any study involving phylogenetic analyses is conditioned by the scope of species included and the taxonomic resolution at which the phylogeny is resolved. Fully resolved phylogenies that include both living and extinct species from the area of interest would be more informative and lead to more precise results. However, the phylogeny that we used contains all plant genera in the Pyrenees, which ensures a good representation of evolutionary patterns up to that level. In addition, this phylogeny
performed well when compared with other commonly used plant phylogenies (Miranda Cebrián et al. 2022), with the advantage of including more information specific to the Pyrenean flora.

**Conclusion**

Our results support the pattern of phylogenetic relatedness between rare species found in other regions and groups of organisms. By exploring the strength and importance of such phylogenetic signal for biodiversity conservation, we found that the loss of habitat specialists and locally scarce species would lead to significant losses in phylogenetic diversity, with important consequences for other aspects of plant diversity such as functional diversity and, in turn, ecosystem function. This highlights the importance of a much-needed integrated insight into the evolutionary relationships of species, their function and role in the ecosystems they inhabit. Exploring the evolutionary patterns of rare species can help us to identify the most vulnerable branches of the Tree of Life and guide the management of vulnerable species before it is too late. This kind of knowledge can be very helpful for the conservation of the biodiversity of any territory, as it allows the preparation of plans to face the effects of global change.

**Acknowledgements** – The authors would like to thank Daniel Gómez and the FLORAPYR project for their help with the vegetation data and taxonomic nomenclature.

**Funding** – This study is a product of the VULBIMON project (CGL2017-90040-R). HMC received a grant from the Spanish Ministry of Science (grant no. PRE2018-083868).

**Author contributions**

Héctor Miranda Cebrián: Conceptualization (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Visualization (lead); Writing – original draft (lead). María Begoña García: Data curation (equal); Funding acquisition (lead); Project administration (lead); Supervision (lead); Writing – original draft (supporting). Cristina Roquet: Formal analysis (supporting); Writing – review and editing (equal). Manuel Pizarro Gavilán: Data curation (equal). Xavier Font: Data curation (equal).

**Supporting information**

The Supporting information associated with this article is available with the online version.

**References**


IUCN 2012. IUCN Red List categories and criteria: ver. 3.1, 2nd edn. – IUCN.


