Mountains of Anatolia are one of the main Mediterranean biodiversity hotspots and their richness in endemic species amounts for 30% of the flora. Two main factors may account for this high diversity: the complex orography and its role as refugia during past glaciations. We have investigated seven narrow endemics of Centaurea subsect. Phalolepis from Anatolia by means of microsatellites and ecological niche modelling (ENM), in order to analyse genetic polymorphisms and getting insights into their speciation. Despite being narrow endemics, all the studied species show moderate to high SSR genetic diversity. Populations are genetically isolated, but exchange of genes probably occurred at glacial maxima (likely through the Anatolian mountain arches as suggested by the ENM). The lack of correlation between genetic clusters and (morpho) species is interpreted as a result of allopatric diversification on the basis of a shared gene pool. As suggested in a former study in Greece, post-glacial isolation in mountains would be the main driver of diversification in these plants; mountains of Anatolia would have acted as plant refugia, allowing the maintenance of high genetic diversity. Ancient gene flow between taxa that became sympatric during glaciations may also have contributed to the high levels of genetic diversity.
in the Flora of Turkey are narrow endemics\textsuperscript{11}. Two of them, *C. hieropolitana* Boiss. and *C. tosiensis* Freyn & Sint. ex Freyn, fall outside *Phalolepis* according to the results of the latest molecular survey of *Centaurea* and should be included in a new section, sect. *Hieropolitanae*\textsuperscript{9}. The species that remain in *Phalolepis* are *C. amaena* Boiss. & Balansa, *C. antalyensis* H. Duman & A. Duran, *C. aphrodisea* Boiss., *C. cadmea* Boiss., *C. dursunbeyensis* Uysal & Köse, *C. luschaniana* Heimerl ex Stapf, *C. lycaonica* Boiss. & Heldr., *C. lycia* Boiss., and *C. wagenitzii* Hub.-Mor.

On morphological grounds, Turkish species of *Phalolepis* share a very similar morphology, which was reviewed in depth recently\textsuperscript{14}. In sharp contrast with Greece where most of species were at some point subordinated to other widespread taxa\textsuperscript{9}, species of *Phalolepis* described from Turkey are fairly undisputed. They are usually mountain chasmophytes on limestone, with the exception of *C. wagenitzii*, which grows on serpentines at sea level.

We chose the group of *Phalolepis* species as a subject for phylogeographic studies for three reasons: firstly, the group is morphologically cohesive and its limits have been recently reviewed and redefined\textsuperscript{9}; secondly, there are no doubts on the taxonomy of the species, thanks to a recent morphological review\textsuperscript{14}; and thirdly, subsection *Phalolepis* has been subject of a recent study on speciation in Greece\textsuperscript{15}, and a similar study in Turkey would allow a comparison of results of high interest. The study in Greece concluded that it was a case of allopatric speciation by area fragmentation mediated by the special topography of the mountains of Greece, which favored isolation, and pointed out the character of refugium of the region during the Pleistocene glaciations\textsuperscript{15}.

As in the case of Greece, we selected microsatellites, a highly variable type of marker, for analyzing subsect. *Phalolepis* in Turkey. Microsatellites have the advantages of codominant inheritance, reproducibility and high levels of variability, and are extensively used for analysing genetic structure. Previous studies using microsatellites have resolved the population structure of other species of *Centaurea*\textsuperscript{16–20} and have unraveled the speciation process in *Phalolepis* in the refuge of continental Greece\textsuperscript{15}. Besides population analyses using microsatellites, ecological niche modelling (ENM) has also been employed to get insights into the paleodistribution of species. ENM represents an independent method of biogeographic inference on a given species, thus being an ideal complement to the classical genetic methods\textsuperscript{21}.

Our specific objectives are the following: (1) to investigate population structure by estimating genetic diversity within and between narrow endemics using microsatellites, and compare the results with those in Greece for verifying whether the evolution of the endemity in *Phalolepis* follow a similar pattern; (2) to analyse the correlation of morphological species and genetic clusters; (3) to explore the levels of ancient and recent gene flow and verify the extent of introgression; (4) to verify whether the range expansions/contractions of the studied *Centaurea* endemics is related to the special physiography of Anatolia using ecological niche modelling in *Centaurea cadmea*, the only species in this study that shows enough occurrences as to build a model.

**Results**

**Genetic diversity.** All surveyed loci were polymorphic for all the populations (except 21D9 for LY, LU, LYC1, AMA, and CA3). We detected a total of 97 alleles, ranging between four (locus 21D9) and 29 (locus 12B1); average was 13.9 alleles per locus. At population level, the mean number of alleles per locus (*A*) ranged between 4.286 in populations LYC1, LYC2, WA2, and CA2 and 7.286 in population LU (Table 1). We detected exclusive alleles for many populations, even though numbers were low (from one to five; Table 1) and frequencies were low too (mean = 0.063). Population LYC2 showed the lowest value of expected heterozygosity (*H*\textsubscript{e} = 0.491), whereas...
population WA1 harboured the highest value ($H_e = 0.669$; Table 1). The least variable species was *C. lycaonica* ($H_e = 0.498$), whereas the highest genetic diversity was shown by *C. antalyensis* ($H_e = 0.655$; Table 1).

Only one loci pair showed evidence of linkage disequilibrium (28A7 vs. 17E3). Many individual loci (42 out of 77 possible cases) exhibited significant deviations from H–W equilibrium expectations ($P < 0.05$); accordingly, $F_{ST}$ values at population and species level were positive and significantly different from zero (Table 1), which can be attributed to inbreeding and/or to the occurrence of null alleles. Given that the frequency of null alleles in our dataset is very low (range = 0.003 (21D9 locus) – 0.178 (12B1 locus); average = 0.095), we do not expect significant biases in our genetic analyses (it has been suggested that biases are negligible when null alleles are present at frequencies below 0.200\(^2\)). However, there are positive correlations between their frequency and $F_{ST}$ values within ($N = 77, R^2 = 0.760, P = 0.000$) and across populations ($N = 7, R^2 = 0.573, P = 0.049$), indicating that the significant deviations from H–W equilibrium can be partly attributed to the presence of null alleles.

All but one genetic divergence values between populations based on $F_{ST}$ were significant ($P < 0.05$). These values ranged from 0.027 (LU–LYC1) to 0.297 (LYC2–AMA), with a mean of 0.198 (Table 2). The $F_{ST}$ values between pairs of species were significant in all the cases, and somewhat smaller than the $F_{ST}$ values between populations (Table 3): they ranged from 0.057 (*C. lycaonica–C. luschaniana*) to 0.217 (*C. lycia–C. amaena*), with a mean of 0.135. Corrected $F_{ST}$ values for null alleles showed very similar to those found using the uncorrected dataset (the differences were only ca. 5–10%), both between populations and between species (Supplementary Tables S1 and S2); the effects of null alleles on the genetic structure of populations, thus, would be negligible.

According to Evanno’s approach, $K = 10$ was the most likely number of genetic clusters for STRUCTURE simulations, as this is the only clear peak. The ln Pr(X|K) also reached a plateau when $K = 10$ (Supplementary Fig. S1). Although several of the studied populations had their “own” clusters, other populations (LY, LU, LYC1, and WA1) showed a high degree of admixture (Fig. 2). Notably, not all the grouping schemes input in AMOVA were significant. The within-population component accounted for most of the total variation (84.42%), the among-population component contributed much less (15.53%), whereas the among-taxa component (i.e., when the populations were grouped into species) was negligible (0.05%) (Table 4). Even though the percentage of variance explained by the first two components was not very high (52.90%), the PCoA analysis at population level was in close agreement with both the STRUCTURE and (particularly) with AMOVA results: populations are relatively isolated among them, with no signs of taxonomically meaningful aggrupation (i.e. all populations belonging to a same species were invariably located in different quadrants; Fig. 3).

### Table 1. Main parameters of genetic diversity for each population computed after the seven polymorphic loci.

| Species/populations | Coordinates       | N   | A   | PA  | $P_{95}$ | $H_o$ | $H_e$ | $F_{IS}$ |
|---------------------|-------------------|-----|-----|-----|----------|-------|-------|---------|
| *C. lycaonica*      | N37° 45.054' E32° 04.529' | 33  | 5.571 | 3   | 0.857    | 0.269 | 0.409*|         |
| LY                  |                    | 37  | 7.286 | 5   | 1.000    | 0.346 | 0.582 | 0.409*  |
| *C. luschaniana*    | N36° 56.681' E30° 09.617' | 28  | 4.286 | 0   | 1.000    | 0.468 | 0.491 | 0.084  |
| LU                  |                    | 18  | 4.286 | 1   | 0.000    | 0.671 | 0.618 | 0.088  |
| Mean                |                    | 31  | 5.857 | 1   | 1.000    | 0.317 | 0.615 | 0.491*  |

**Main parameters of genetic diversity for each population computed after the seven polymorphic loci.** N, sample size; A, mean number of alleles per locus; PA, number of private alleles; $P_{95}$, percentage of polymorphic loci (95% criterion); $H_o$, observed heterozygosity; $H_e$, unbiased expected heterozygosity; $F_{IS}$, inbreeding coefficient; *P < 0.05.* Populations are identified by an alphanumeric code. See Supplementary Table S4 for more details on the sampled localities.
The first two assigned barriers in the Barrier analysis (Supplementary Fig. S2), mainly separated the populations located on western Taurus Mountains (that is, those around the Gulf of Antalya). With the third barrier ($B = 3$), separations in other regions (Pontic Mountains, central Taurus Mountains) started to be evident, although some separations (LU vs. LYC1, LY vs. CA2, and WA1 vs. WA2) did not appear until $B = 8$. The BayesAss analysis confirmed the absence of recent gene flow between populations: all but one pairwise $m$ values were at least one order of magnitude below 0.114 (this figure separates real migration from noise; Table 5). Regarding historical gene flow, as estimated with the software MIGRATE-N, $Nm$ values were relatively low (from 0.263 to 1.193, averaging 0.466; Table 6). Total immigration rates were highest for 1LYC, 1WA, and 2WA (all with $Nm > 5$), whereas the populations that showed total emigration rates $> 5$ were LU, 2WA, and 2CA (Table 6).

**Ecological niche modelling.** According to the MaxEnt jackknife tests of variable importance, the precipitation variables were more informative for the model than the temperature ones (the more informative variables were, by far, precipitation of the driest month and precipitation seasonality; Supplementary Fig. S3). The
present-day distributional predictions for *C. cadmea* were largely congruent with the known species occurrences, although other areas appear as suitable (shaded in Fig. 4), such as mountains of NW Anatolia and some scattered areas along the western and central Taurus Mountains (Fig. 4A). Projections of the species niche to the LGM climate produced considerably different maps of presence/absence. With the CCSM model almost all the Anatolian Peninsula appears as suitable, leaving only as unsuitable the eastern and southern coastal areas (Fig. 4B). The MIROC model showed that most of the mountainous areas of the Anatolian Peninsula were suitable for *C. cadmea* at the LGM with the exception of the plateau region of central Anatolia (Fig. 4C). Both LGM projections show a considerable increase of the potential area of *C. cadmea* compared to the present time, even for the likely more “realistic” MIROC model (with an increase of about four-fold of the suitable area, compared to an increase of ten-fold for the CCSM). The LGM projections should be, however, treated with extreme caution because the uncertainty of projecting the present-day ENM to the LGM should be added to that resulting from the use of a very small number of occurrences.
Table 6. Median historical gene flow (Nm) among the studied populations, estimated from seven microsatellite loci using the formula \(4\text{Nm} = \Theta M\) (with \(M\) values obtained with MIGRATE-N). As suggested by Beerli (2006), the median is used instead of the mean since the latter is heavily influenced by outliers. In parentheses, 95% confidence interval. \(\Theta\) is the mutation-scaled effective population size, as obtained from MIGRATE-N.

| From | LY | LU | 1LYC | 2LYC | AMA | IWA | 2WA | 1CA | 2CA | 3CA | AN | Total Nm (as immigration rates) |
|------|----|----|------|------|-----|-----|-----|-----|-----|-----|----|--------------------------------|
| To   | \(\Theta\) | \(\text{(0–2.000)}\) | \(\text{(0–0.855)}\) | \(\text{(0–0.900)}\) | \(\text{(0–0.630)}\) | \(\text{(0–1.260)}\) | \(\text{(0–0.810)}\) | \(\text{(0–0.945)}\) | \(\text{(0–0.945)}\) | \(\text{(0–0.855)}\) | \(\text{(0–0.630)}\) | \(\text{(0–1.260)}\) | \(\text{(0–0.945)}\) |
| LY   | \(0.900\) | \(0.383\) | \(0.428\) | \(0.473\) | \(0.293\) | \(0.518\) | \(0.383\) | \(0.473\) | \(0.383\) | \(0.383\) | \(0.383\) | \(0.473\) | \(0.383\) |
| LU   | \(0.900\) | \(0.383\) | \(0.428\) | \(0.473\) | \(0.293\) | \(0.518\) | \(0.383\) | \(0.473\) | \(0.383\) | \(0.383\) | \(0.383\) | \(0.473\) | \(0.383\) |
| 1LYC | \(0.700\) | \(0.578\) | \(0.648\) | \(0.683\) | \(0.368\) | \(0.575\) | \(0.438\) | \(0.403\) | \(0.543\) | \(0.718\) | \(0.702\) | \(0.702\) | \(0.702\) |
| 2LYC | \(0.700\) | \(0.368\) | \(0.473\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) |
| AMA  | \(0.700\) | \(0.263\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.298\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) | \(0.368\) |
| IWA  | \(0.900\) | \(0.608\) | \(0.923\) | \(0.653\) | \(0.518\) | \(0.608\) | \(1.193\) | \(0.518\) | \(0.518\) | \(0.518\) | \(0.518\) | \(0.518\) | \(0.518\) |
| 2WA  | \(0.700\) | \(0.333\) | \(0.963\) | \(0.823\) | \(0.858\) | \(0.543\) | \(1.068\) | \(0.928\) | \(1.068\) | \(0.928\) | \(0.928\) | \(0.928\) | \(0.928\) |
| 1CA  | \(0.900\) | \(0.473\) | \(0.473\) | \(0.518\) | \(0.383\) | \(0.383\) | \(0.383\) | \(0.608\) | \(0.608\) | \(0.608\) | \(0.608\) | \(0.608\) | \(0.608\) |
| 2CA  | \(0.700\) | \(0.333\) | \(0.473\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) | \(0.508\) |
| 3CA  | \(0.700\) | \(0.368\) | \(0.508\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) | \(0.333\) |
| AN   | \(0.900\) | \(0.338\) | \(0.338\) | \(0.293\) | \(0.338\) | \(0.338\) | \(0.338\) | \(0.338\) | \(0.338\) | \(0.338\) | \(0.338\) | \(0.338\) | \(0.338\) |

| Total Nm (as immigration rates) | \(4.040\) | \(5.590\) | \(4.840\) | \(4.840\) | \(4.570\) | \(4.170\) | \(4.830\) | \(5.085\) | \(4.760\) | \(5.010\) | \(4.540\) | \(3.875\) |

Discussion

Genetic diversity. Genetic theory predicts low levels of genetic diversity for narrow endemic species because of a combination of traits usually associated with rarity: low population sizes, isolation of populations, inbreeding, and ecological specialization\(^{23,24}\). In support of this view, several meta-analyses have demonstrated that there is an association between rarity and low genetic variability\(^{35-37}\); such association is even more compelling for those species that are limited to one or very few localities, with dozens or, at most, a few hundreds of individuals (often referred to as "extremely narrow endemics", ENEs\(^{38}\)). In the Mediterranean, two biological traits are often present in the endemic flora: production of few and small flowers, and low investment in pollen and seed production\(^{39}\). Both traits might also have a role in the association between rarity and low genetic diversity. Accordingly, examples of Mediterranean narrow endemics showing very low levels of genetic diversity are abundant, e.g. *Anchusa crispa*\(^{39}\), *Aquilegia barbicornis*\(^{40}\), *A. nuragica*\(^{41}\), *Coristospermum huteri*\(^{42}\), or *Zelkova sicula*\(^{43}\).

Contrary to the expected patterns for narrow endemics, Anatolian species of subsect. *Phalolepis* are not genetically depauperated at all (mean \(A = 5.643\); mean \(H_e = 0.580\); Table 1). Even the most restricted taxa such as *C. antalyensis*, *C. cadmea* subsp. *cadmea*, and *C. lycocina* (all are classified as CR, see below in Material and Methods) have moderate to high levels of genetic diversity, and *C. antalyensis*, only known from the type locality and having less than 250 individuals, harbors the highest levels of variability in terms of expected heterozygosity (\(H_e = 0.655\); Table 1) among all studied taxa. Genetic diversity values in Turkish species of subsect. *Phalolepis* are among the highest within the genus (mean \(A = 5.160\); range 3.100–7.286; mean \(H_e = 0.504\); range 0.225–1.260; Supplementary Table S3), and are comparable to the values obtained for a series of narrow endemics of the same subsection from the mountains of Greece using the same set of microsatellites (\(A = 5.363\); \(H_e = 0.587\)\(^{44}\)). Comparison with the "reference" values for plants (\(H_e = 0.420\) for endemic species, \(H_e = 0.620\) for widespread ones) provided in the meta-analysis of Nybom\(^{45}\) is another proof that our study species conserve considerable microsatellite genetic variability in spite of their rarity.

Our results, as well as the growing number of studies reporting narrow endemic species with unexpectedly high levels of genetic polymorphism\(^{33,34}\), suggest that the geographic range is not always a good predictor of genetic diversity in plant species. It is generally agreed that a complex network of factors shape genetic diversity in plant populations, which are usually classified into two categories\(^{35}\): (1) intrinsic biological properties of the species (i.e., life-history traits and ecological interactions) and (2) extrinsic dynamic processes which affect species (i.e., historical factors that may include occurrence of bottlenecks, divergence events, or Quaternary expansions/retreats). Species of *Centauraea* share most of their biological traits: they are usually protandrous and usually self-incompatible, and they have entomophilous pollination by a wide range of insects\(^{36-38}\). Thereafter, intrinsic
factors should be discarded and differences in polymorphism between the studied *Centaurea* species (Table S3) are more likely extrinsic and attributable to their evolutionary history.

The endurance of glacial-interglacial cycles in mountain glacial refugia in the Pindus Mountains of Greece was the main reason provided for the high levels of polymorphism detected in the Greek taxa of subsect. *Phalolepis*\(^{15}\).
Likewise, the relative environmental stability of another mountainous refugia in the Pre-Pyrenean Mountains has been suggested as an explanatory factor for the moderate microsatellite diversity in *C. emigrantis* and *C. tripontina*. The Pindus and the Pre-Pyrenean regions range among the “phylogeographic” glacial refugia recognized within the Mediterranean Basin. The southern Anatolian Peninsula and especially the western and central sections of Taurus Mountains is home of several refugia, and is also characterized by a high concentration of plant diversity and endemism.

Several factors have contributed to the refugial character of the Taurus: the rugged topography of these mountains with peaks surpassing 3000 m, its closeness to the sea, which would have constituted a continued source of moisture even at the most arid phases of the Pleistocene; and, in contrast to the Pyrenees or the Alps, the lack of major ice-sheets. Only the peaks over 1900 m were glaciated. This scenario is consistent with the most recent paleoecological reconstructions for the Taurus Mountains based on pollen sites or climatic simulations: the mountains would have been covered by temperate/boreal forests and parklands instead of more arid formations such as desert or steppe.

The results in the montane species of *Centaurea* studied by our team clearly support this scenario. *Centaurea* species, like other mountain endemics, would have found abundant favorable pockets throughout the unglaciated, relatively moist Mediterranean mountains, and survived the repeated glacial/interglacial cycles by altitudinal migrations without large geographical displacements. These movements would have favored secondary contacts at the glacial periods and populations would have admixed, largely blurring the genetic differentiation produced by periods of isolation at the interglacials. This process was favored by the much longer duration of glacial compared to interglacials and has also been proposed for other plants from the Mediterranean mountains. Genetic connectivity is also present in Anatolian species of subsect. *Phalolepis* is also present in Anatolian species of subsect. *Phalolepis* has been suggested as an explanatory factor for the moderate microsatellite diversity in *C. tripontina*. The results of the AMOVA analysis also indicate that the species cannot be defined as independent entities; the among-taxon component is notoriously low (0.05%; Table 4). This difficulty often arises when dealing with taxa that are scarcely divergent genetically, as in the present case (Table 3). The use of the ITS region and plastid markers in *Centaurea* has usually resulted in a general lack of resolution; the resulting phylogenies consist usually of large polytomies. Introggression and incomplete lineage sorting following recent speciation are usually invoked as the major causes of species-level polyphyly. Speciation of *Centaurea* subsect. *Phalolepis* in Anatolia is recent because an age of 1.55 (0.32–3.13) Ma was assigned to the Turkey/Cyprus haplotype. We will discuss in terms of current and former gene flow the reasons that make us think that cases of shared gene pool are due to introgression in a recently differentiated group.

Recent gene flow is now almost absent and populations are genetically isolated, as deduced from the results of BayesAss (Table 5). The only case of current flow is the connection between one of the populations of *C. lycia* (LYC1) and *C. luschaniana*. In this case, Barrier results also indicate that there are no barriers between them, even with $B = 8$ (Supplementary Fig. S2). This may be due to lack of topographical barriers that do exist, however, with the other population of *C. lycia* (LYC2; Fig. 1), which grows on the other side of the mountains. The population LYC2 appears to be isolated and free from admixture in Structure (pale blue, Fig. 2) and the Barrier analysis effectively separates it from the rest of populations from $B = 5$ (Supplementary Fig. S2). Other instances of high admixture (C. lycoaica and population 1 of C. wagenitzii) are evident in Structure (Fig. 2), for which recent gene flow is not a valid explanation as reflected in BayesAss. For these two populations (especially for WA1) we can argue, instead, ancient gene flow as shown by the MIGRATE-N results (Table 6). Although the $Nm$ values suggested by this software are not high at all, at least there is some signal of exchange of genes; however, we cannot assign an exact temporal framework for these old contacts, as they may range from the Holocene (~450 ya) to the end of LGM (~18,500 ya) following our previous approximation. According to our reconstruction of the ecological niche at the LGM for one of the species studied here (Fig. 4; see below), gene flow would have been much more extensive during the glaciations than at present.

Lower but still noticeable levels of admixture are evident in almost all species studied (Fig. 2). For these cases, we can also argue ancient gene flow. In none of the cases of admixture, either recent or old, there is morphological evidence of introgression. That is, no intermediate characters have been observed between the species that have been connected. Ancient contacts between non sympatric species in the Mediterranean region have been explained in several cases by altitudinal migrations forced by the glaciations. In our case, the contacts may have been connected. Ancient contacts between non sympatric species in the Mediterranean region have been explained in several cases by altitudinal migrations forced by the glaciations. In our case, the contacts may have been connected. Ancient contacts between non sympatric species in the Mediterranean region have been explained in several cases by altitudinal migrations forced by the glaciations. In our case, the contacts may have been connected. Ancient contacts between non sympatric species in the Mediterranean region have been explained in several cases by altitudinal migrations forced by the glaciations. In our case, the contacts may have been connected.
LGM, these mountainous archs were probably covered by boreal/temperate forests and parkland instead of xeric grasslands or steppe, as revealed by climate simulations\(^4\)\(^5\). Fossil pollen records also points towards cold but relatively moist conditions (probably due to the orographic rain) for most of the mountainous southern regions of Anatolia\(^4\). A pollen record from Lake Iznik, in the mountains of NW Anatolia, is suggestive of steppe but with a still important fraction of arboreal vegetation (ca. 25%)\(^4\). There are no pollen sites for the Pontic Mountains, but there are currently some relict termophilous species in the region, such as Pterocarya fraxinifolia. Along the Black Sea coastline, Carya and Glyptostrobus persisted until the Holocene, and were probably extinct due to human disturbance\(^4\). Presence of termophilous species indicates that glacial refugia would have existed in the region. Contrarily, the Anatolian Plateau was dominated by Artemisia-steppe; at the only pollen site from the plateau, only 10% of the pollen is arboreal\(^4\). Presence of steppe suggests much drier conditions to those of the surrounding mountain archs. All the species of the study except C. wagenitzii share similar ecological requirements (see Introduction), and it is plausible that the same pathways favorable to C. cadmea have been exploited by other species of the study. The lack of barriers between the most isolated species, C. amaena, and the core of the SW species in the Taurus (Supplementary Fig. S2) reinforces the hypothesis of the existence of migration routes through the mountains of S Turkey. In addition to the ENM, migration routes though the northern mountain arches are also suggested by the lack of barriers between C. amaena and population 3 of C. cadmea (Supplementary Fig. S2) as well as by the relatively high \(N_m\) values between populations 2 and 3 of C. cadmea (located in the Pontic Mountains) and some of the studied populations of SW Taurus.

**Greece and Turkey models.** A comparison between the diversification patterns in Greece and Anatolia is straightforward. The similarities are relevant, with a main factor having shaped genetic diversity and diversification in Centaurea subsect. Phalolepis: the rugged and dissected topography. The Anatolian mountains favored the population isolation of a more widespread ancestor of the species of the group, triggering allopatric speciation. At the same time, the mountains would have constituted refugia during the glaciations, protecting the populations from genetic bottlenecks and promoting contacts between them as shown by the high genetic diversity found. Finally, the mountain arch that encircles the Anatolian Plateau, especially the Anatolian Diagonal at the east, offered a pathway during glacial for the migration of species.

The main difference between Greece and Turkey in the diversification of Centaurea subsect. Phalolepis is the time frame: speciation in Turkey is probably more recent. The main reasons for reaching this conclusion are: (a) Genetic clustering in Greek species using STRUCTURE show only limited level of admixture\(^15\), while in Turkey the levels of admixture are much higher (Fig. 2). (b) \(F_{ST}\) values between species are larger in Greece compared to Turkey (0.182 vs. 0.135). These results support that Greek species would have almost achieved complete isolation, while Anatolian ones would be still in progress. In sum, processes involving diversification in both hotspots are basically the same, with a slightly later temporal displacement in Turkey when compared to the Greek case.

**Evolutionary and conservation implications.** Mountains elsewhere are often regarded as suitable Quaternary refugia, where both persistence and differentiation of plant lineages would have occurred; thus, it is not surprising that mountains often harbor high taxonomic richness as well significant rates of endemism\(^5\). For the mountains of the Mediterranean basin, it was suggested a dual model to explain the patterns of endemism\(^3\): prevalence of paleo-endemics in its western section and dominance of neoendemics in its eastern half; according to these authors, the dominance of differentiation processes over conservative ones in the Eastern Mediterranean basin would have resulted from the relative geologic youth of this sector and, remarkably, the moderate role of Pleistocene glaciations\(^4\). The present study as well as our previous one\(^15\) in the subsect. Phalolepis of Centaurea (a very young group) are paradigmatic of the role of Mediterranean mountains as places where active speciation processes have taken, and are still taking place. The high levels of genetic diversity and the moderate genetic isolation between species support the role of mountains as glacial refugia and "species pumps". In conservation biogeography, one of the most debated questions at present is whether priority should be given to areas of active speciation ("species cradles") or, on the contrary, to areas that are a sink of ancestral taxa ("species museums")\(^15\). Although with the conservation of "species museums" we are avoiding the loss of relict taxa (that, in most cases, have unique, irreplaceable evolutionary history), conserving species pumps we are preserving those places that contribute most to the growth of the Tree of Life (i.e., we are guaranteeing the maximum levels of both present and future biodiversity). Conserving the eastern Mediterranean mountains in a more or less pristine state is, however, a challenging issue; overgrazing and erosion are rampant, and tourism impacts are quickly increasing\(^2\). Regrettably only 1.2% of Turkey's terrestrial area is strictly protected (much below the international standards), and high mountains, shrublands, and steppe ecosystems are underrepresented\(^10\). Therefore, although conserving large landscapes is perhaps the best tool in plant by preserving processes that create and maintain biodiversity\(^8\), it seems more realistic to ensure the conservation at population level of those species of conservation interest, such as narrow endemics. Unfortunately, most of the studied species do not enjoy at present of suitable conservation measures (even despite that some of them are classified as CR); for instance, less than 20% of all the populations of the studied species are located within a nature reserve and, to our knowledge, no specific conservation measures, either in-situ or ex-situ, exist.

**Methods**

**Plant material.** Most species of subsect. Phalolepis classically recognized\(^11\) are included in the study, with the addition of a couple of taxa described afterwards. We excluded C. hieropolitana and C. tossiensis (see Introduction). Tetraploid C. aphrodisea and C. dursunbeyensis were also excluded due to the difficulties in interpreting microsatellite results in polyploids\(^5\). Thereafter, our survey includes seven species: C. amaena (one...
population), *C. antalyensis* (one population), *C. cadmea* (three populations), *C. luschaniana* (one population), *C. lycaonica* (one population), *C. pontica* (two populations), and *C. wagenitzii* (two populations).

*Centaurea amniens* grows in rocky places in the Kayseri province and it is restricted to a small area between Yalihan and Erzincay Mountains. The only two localities together cover an area of around 0.55 km², and the number of individuals was calculated to be 5672. Its threat category was suggested to be EN (Endangered)

However, the species displays fairly local distribution in an isolated area open to urbanisation as well as limited number of mature individuals. The species should be categorised as CR (Critically Endangered) according to 2001 IUCN criteria

Critical Endangered (CR)56. There are four–five very close populations that are threatened due to extreme live

and species (with and without the ENA correction for null alleles68).

method60. For verifying cross-amplification of the seven Turkey-endemic species, we carried out a preliminary

Genetic analysis. Software GenAlEx v. 6.163 and GENETIX v. 4.0564 were used for estimating genetic diversity parameters at both species and population levels: (a) mean number of alleles per locus (*A*); (b) number of private alleles (*PA*); (c) percentage of polymorphic loci (*Pp*); (d) observed heterozygosity (*H*); (e) unbiased expected heterozygosity (*H*) by the method of Weir & Cockerham65. GenePop v. 4.0.1066 was used for calculating possible deviations from Hardy-Weinberg (H–W) equilibrium and for checking genotypic linkage disequilibrium between pairs of loci at population level and across all populations; for both calcula

the spatial genetic structure was assessed through four different methods. First, Structure v. 2.3.469, a widely-employed clustering software that is based on a Bayesian algorithm, was used. On the basis of preliminary runs, *K* was run from 1 to 12 (20 iterations per *K*) assuming an admixture model with correlated allele frequencies, and with a priori grouping of individuals into populations (but not into species). The length of burn-in period and the MCMC replications were set to 10⁵ and 10⁶, respectively. The most likely value of *K* was determined both by choosing the smallest *K* after the log probability of data *ln Pr(X|K)* values reached a plateau

and by the Δ*K* statistic of Evanno et al.70 with the aid of Structure Harvester71. Second, a molecular variance analysis (AMOVA) was run with the aid of GenAlEx v. 6.1, establishing two hierarchical levels: (i) among taxonomic groups (species), (ii) among populations, and within populations. Third, a Principal Coordinates Analysis (PCoA) at population level was carried out with the same software (GenAlEx). Fourth, putative genetic barriers between populations were detected with the software Barrier v. 2.272; significance of barriers was tested

PA alleles (*A*); (b) percentage of polymorphic loci (*Pp*); (c) observed heterozygosity (*H*); (d) unbiased expected heterozygosity (*H*) by the method of Weir & Cockerham65. GenePop v. 4.0.1066 was used for calculating possible deviations from Hardy-Weinberg (H–W) equilibrium and for checking genotypic linkage disequilibrium between pairs of loci at population level and across all populations; for both calcula

the spatial genetic structure was assessed through four different methods. First, Structure v. 2.3.469, a widely-employed clustering software that is based on a Bayesian algorithm, was used. On the basis of preliminary runs, *K* was run from 1 to 12 (20 iterations per *K*) assuming an admixture model with correlated allele frequencies, and with a priori grouping of individuals into populations (but not into species). The length of burn-in period and the MCMC replications were set to 10⁵ and 10⁶, respectively. The most likely value of *K* was determined both by choosing the smallest *K* after the log probability of data *ln Pr(X|K)* values reached a plateau

and by the Δ*K* statistic of Evanno et al.70 with the aid of Structure Harvester71. Second, a molecular variance analysis (AMOVA) was run with the aid of GenAlEx v. 6.1, establishing two hierarchical levels: (i) among taxonomic groups (species), (ii) among populations, and within populations. Third, a Principal Coordinates Analysis (PCoA) at population level was carried out with the same software (GenAlEx). Fourth, putative genetic barriers between populations were detected with the software Barrier v. 2.272; significance of barriers was tested
by bootstrapping 1000 $D_p^{71}$ matrices that were previously obtained with Microsatellite Analyzer (MSA) v. 4.05 software$^{24}$. Gene flow was estimated with two time-frameworks. First, we estimated recent (i.e. within the recent 2–3 generations) migration rates between individual populations with the software BayesAss v. 1.3$^{72}$. As program settings, the default values were used (MCMC iterations, 3 $\times$ 10$^5$ length of the burn-in, 999,999; sampling frequency, 2000; delta value, 0.15). Second, historical migration-scaled migration rates ($M = m/\mu$, where $m$ is migration rate and $\mu$ is mutation rate per generation) were estimated using MIGRATE-N v. 3.6.4$^{76}$. Ten replicates were run under a Brownian motion model, assuming constant mutation rate for all loci. With a Bayesian approach, a long chain with 20,000 genealogies to sample was run, with a sampling increment of 100 (thus, totalling 2,000,000 genealogies for each replicate); the burn-in was set to 20,000. A static heating scheme was chosen (temperatures were specified to 1.00, 1.50, 3 and 1 $\times$ 10$^6$), with uniform prior distribution both for $\Theta$ and $M$ (min: 0; max: 100; delta: 10). The effective number of migrants per generation ($N_m$ among populations was estimated using the formula $4N_m = \Theta M^{0.25}$). We could not estimate migration rate ($m$) values, provided that those obtained from MIGRATE-N are mutation-scaled ($M$), and mutation rates ($\mu$) for microsatellites for the genus Centaurea are not available. Total immigration and emigration rates for each population were obtained by summing values of $N_m$. Analyses were carried out at the CIPRES bioinformatic facility$^{72}$.

Ecological niche modelling. Ecological niche modelling (ENM) was performed to evaluate the potential distribution of the Turkey species of Centaurea under present climatic conditions and to project it to the Last Glacial Maximum (LGM, ca. 21,000 yr BP). The ENM was only performed with two of the study species (C. cadmea and C. luschaniana) because for the other taxa the number of wild occurrences was not enough to get reliable models ($<5$)$^{78}$. We employed the maximum entropy algorithm, as implemented in MaxEnt v. 3.3$^{79}$. The current distribution information for both Centaurea species was obtained from presence records included in the Global Biodiversity Information Facility (www.gbif.org), from literature$^{12,14}$, and from the sampling sites of this study. In total, after removing duplicate records within each pixel (2.5 arc-min, ca. 5 km), we obtained 8 and 5 presence records for C. cadmea and C. luschaniana, respectively. A set of 19 bioclimatic variables at 2.5 arc-min resolution covering the distribution range (and neighboring areas) for both species under current conditions (1950–2000) were downloaded from the WorldClim website (www.worldclim.org). After a correlation analysis in a random sample of 10,000 points within the study area, we selected a smaller set of nine (relatively) uncorrelated ($r < 0.9$) variables: isothermality (bio3), temperature seasonality (bio4), mean temperature of the wettest quarter (bio8), mean temperature of the warmest quarter (bio10), mean temperature of the coldest quarter (bio11), annual precipitation (bio12), precipitation of the driest month (bio14), precipitation seasonality (bio15), and precipitation of the coldest quarter (bio19). The selection of variables of pairs or groups of highly correlated ones was done on the basis of their relative contribution to the model (percent contribution, jackknife tests of variable importance). The distribution model under current conditions was projected to the LGM using paleoclimatic layers simulated by both the Community Climate System Model (CCSM)$^{80}$ and the Model for Interdisciplinary Research on Climate (MIROC)$^{81}$. Replicate runs (10) of MaxEnt (using the “bootstrap” method) were performed to ensure reliable results.

Model testing when the number of occurrences is small ($<25$) can be problematic, because the “training” and “test” datasets would be too low$^{78}$. To overcome this, we used a methodology based on a jackknife (or “leave-one-out”) procedure$^{78}$ to test the model. With this procedure the model is built (or “trained”) using $n – 1$ occurrences, and tested using the discarded locality (that is, by evaluating the ability to predict the single locality excluded from the training dataset). Thus, eight and five predictions were obtained for C. cadmea and C. luschaniana, respectively. We used the “lowest presence threshold” (LPT, also commonly referred as “minimum training presence” in the MaxEnt terminology) as the cut-off value to decide whether the discarded locality is “suitable” or “unsuitable”. The LPT is considered as more conservative than other approaches (such as the 10% fixed threshold) since identifies the minimum predicted area possible whilst maintaining zero omission error in the training data set$^{78}$. Performance of models for both C. cadmea and C. luschaniana was evaluated through success rate ($q$, which is the proportion of right predictions) and statistical significance (a $P$-value computed across the set of jackknife predictions), which was done using the software provided by Pearson et al.$^{78}$. The jackknife analysis indicated that the model was successful at predicting known occurrences as suitable areas for C. cadmea (success rate $= 75\%$, $P < 0.001$) but not for C. luschaniana (success rate $= 40\%$, $P = 0.041$). Thus, definitive ENM models (that is, using all occurrence points) for the present and LGM were only built for C. cadmea. Before it, however, we carried out tuning experiments varying the level of regularization, in order to improve the performance of the model$^{25}$. After discarding regularization multipliers $> 2$ (as they diminished AUC values considerably according to exploratory runs), leave-one-out models were built with up to eight regularization multipliers (0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, and 2.00). Based on the optimal combination of AUC values and success rates ($q$)$^{25}$, the best model was that using a regularization multiplier of 0.5 (Supplementary Fig. S4).

To convert continuous suitability values obtained for each pixel to presence/absence (that is, a binary map), we chose the minimum training presence value as the threshold. Finally, a jackknife analysis was used to evaluate the relative importance of the nine bioclimatic variables employed. All ENM predictions were visualized in ArcGIS v. 9.3 (ESRI, Redlands, CA, USA), with the aid of Hawth’s Analysis Tools$^{84}$. References

1. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. Biodiversity hotspots for conservation priorities. Nature 403, 853–858 (2000).
2. Thompson, J. D., Laverge, S., Affre, L., Gaudeul, M. & Debussche, M. Ecological differentiation of Mediterranean endemic plants. Taxon 54, 967–976 (2005).
3. Davis, P. H. Distribution patterns in Anatolia with particular references to endemism in Plant life of south-west Asia (eds. Davis, P. H., Harper, P. C. & Hedge, I. C.) 15–27 (Botany Society of Edinburgh, 1971).
4. Médail, F. & Quézel, P. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. Ann. Mo. Bot. Gard. 84, 112–127 (1997).

5. Takhjtaian, A. Floristic regions of the world (University of California Press, 1986).

6. Médail, F. & Diadema, K. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. J. Biogeogr. 36, 1333–1345 (2009).

7. Susanna, A. & Garcia-Jacas, N. Tribe Cardueae in The families and genera of vascular plants (eds. Kadereit, J. W. & Jeffrey, C.) 123–147 (Springer, 2007).

8. García-Jacas, N. et al. Centaurea revisited: A molecular survey of the Jacae group. Ann. Bot. 98, 741–753 (2006).

9. Hilpold, A. et al. Phylogeny of the Centaurea group (Centaurea, Compositae) – geography is a better predictor than morphology. Mol. Phylogenet. Evol. 77, 195–215 (2014).

10. Ture, C. & Böckl, H. Distribution patterns of threatened endemic plants in Turkey: A quantitative approach for conservation. J. Nat. Conserv. 18, 296–303 (2010).

11. Wagenitz, G. Centaurea L. in Flora of Turkey and the East Aegean Islands, Vol. 5 (ed. Davis, P. H.) 465–485 (Edinburgh University Press, 1975).

12. Yusal, T. Centaurea L. in Türkiye bitkileri listesi (Damarruh bitkileri) [A check list of the flora of Turkey (Vascular plants)] (eds. Güner, A., Aslan, S., Ekim, T., Vural, M. & Babac, M. T.) 127–140 (Nezahat Gökyiğit Botanik Bahçesi ve Flora Araştırmaları Derneği Yayınları, 2012) [in Turkish].

13. García-Jacas, N. et al. The polyphylid series of Centaurea toletana: glacial migrations and introgression revealed by mtDNA and cpDNA sequence analyses. Mol. Phylogenet. Evol. 52, 377–394 (2009).

14. Köse, Y. B., Alan, S. & Yücel, E. Comparative investigation of the morphological characteristics of species belonging to the Centaurea L. section Phalépis (Cass.) DC. Biol. Divers. Conserv. 3, 10–22 (2010).

15. López-Vinyallonga, S., López-Pujol, J., Constantinidis, T., Susanna, A. & Garcia-Jacas, N. Mountains and refuges: genetic structure and evolutionary history in closely related, endemic Centaurea in continental Greece. Mol. Phylogenet. Evol. 92, 243–254 (2015).

16. Fréville, H., Justy, F. & Olivieri, I. Comparative allozyme and microsatellite population structure in a narrow endemic plant species, Centaurea corymbosa Pourret (Asteraceae). Mol. Ecol. 10, 879–889 (2001).

17. Mameli, G., Filigheddu, R., Binelli, G. & Meloni, M. The genetic structure of the remnant populations of Centaurea horrida in Sardinia and associated islands. Ann. Bot. 101, 633–640 (2008).

18. Guarino, C., Salerno, G. & Cipriani, G. Effects of fragmentation phenomena on the genetic structure and gene flow in Centaurea cineraria (Asteraceae) in the Mediterranean Basin. Plant Biosyst. 147, 996–1005 (2003).

19. Marrs, R., Sforza, R. & Hufbauer, R. A. Evidence for multiple introductions of Centaurea stoebe micranthus (spotted knapweed, Asteraceae) to North America. Mol. Ecol. 17, 4197–4208 (2008).

20. Marrs, R., Sforza, R. & Hufbauer, R. A. When invasion increases population genetic structure: a study with Centaurea diffusa. Biol. Invasions 10, 561–572 (2008).

21. Huang, P. & Schaal, B. A. Association between the geographic distribution during the last glacial maximum of Asian wild rice, Oryza rupestris (Poaceae), and its current genetic variation. Am. J. Bot. 99, 1866–1874 (2012).

22. Dakin, E. E. & Avise, J. C. Microsatellite null alleles in parentage analysis. Heredity 93, 504–509 (2004).

23. Babbel, G. R. & Selander, R. K. Genetic variability in edaphically restricted and widespread plant species. Evolution 28, 619–630 (1974).

24. Frankham, R. Genetics and extinction. Biol. Conserv. 126, 131–140 (2005).

25. Hamrick, J. L. & Godt, M. J. W. Allozyme diversity in plant species in Plant population genetics, breeding and genetic resources (eds. Brown, A. H. D., Clegg, M. T., Kahler, A. L. & Weir, B. S.) 43–63 (Sinauer Associates, 1990).

26. Gitzendanner, M. A. & Soltis, P. S. Patterns of genetic variation in rare and widespread plant congeners. Am. J. Bot. 87, 783–792 (2000).

27. Nybom, H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol. Ecol. 13, 1143–1155 (2004).

28. López-Pujol, J., Martinell, M. C., Massó, S., Blanche, G. & Sáez, L. The `paradigm of extremes': Extremely low genetic diversity in an extremely narrow endemic species, Cirsium perpusillum (Umbelliferae). Plant Syst. Evol. 299, 439–446 (2013).

29. Lavergne, S., Thompson, J. D., Garnier, E. & Debussche, M. The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. Oikos 107, 505–518 (2004).

30. Quillichini, A., Debussche, M. & Thompson, J. D. Geographic differentiation of morphological traits and isozymes in the Mediterranean island endemic Anchusa crispa: implications for the conservation of a protected species. Biol. Conserv. 118, 651–660 (2004).

31. Garrido, J. L., Fenu, G., Mattana, E. & Bacchetta, G. Spatial genetic structure of Aquilegia taxa endemic to the island of Sardinia. Ann. Bot. 109, 953–964 (2012).

32. Fineschi, S., Cozzolino, S., Migliaccio, M. & Vendramin, G. G. Genetic variation of relic tree species: the case of Mediterranean Želkova abelicea (Lam.) Boissier and Z. scutellata Di Pasquale, Garfi and Quézel (Ulmaceae). For. Ecol. Manag. 273, 277–284 (2004).

33. Fernández-Mazuecos, M., Jiménez-Mejias, P., Rotllan-Puig, X. & Vargas, P. Narrow endemics to Mediterranean islands: moderate genetic diversity but narrow climatic niche of the ancient, critically endangered Naufraqui (Apiaceae). Perspect. Plant Ecol. Evol. Syst. 16, 190–202 (2014).

34. Jiménez-Mejías, P., Fernández-Mazuecos, M., Amat, M. & Vargas, P. Narrow endemics in European mountains: high genetic diversity within the monospecific genus Pseudomisopates (Plantaginaceae) despite isolation since the late Pleistocene. J. Biogeogr. 42, 1455–1468 (2015).

35. Gray, A. Genetic diversity and its conservation in natural populations of plants. Biodivers. Lett. 3, 71–80 (1996).

36. Marsden-Jones, E. M. & Turrill, W. B. British Knapeweeds: a study in systemic taxonomy (The Ray Society, 1954).

37. Jeffrey, C. Evolution of Composite flowers in Systematics, evolution, and biogeography of Compositae (eds. Funk, V. A., Susanna, A., Stuessy, T. F. & Bayer, R. J.) 293–313 (IAPT, 2009).

38. Pisano, S., Filigheddu, R. & Farris, E. The conservation status of an endemic species of northern Sardinia: Centaurea horrida Badarò (Asteraceae). Plant Biosyst. 143, 275–283 (2009).

39. López-Alvarado, J. Centaurea L. section Phalépis Pers.: Phylogeny and biogeography (PhD Thesis, Università degli Studi di Sassari, 2012).

40. Sarkaryan M. A., Çiner, A. & Zreda, M. Quaternary glaciations of Turkey in Quaternary glaciations—Extent and chronology: A closer look (eds. Ehlers, J., Gibbard, P. L. & Hughes, P. D.) 393–403 (Elsevier, 2011).

41. Prentice, I. C., Harrison, S. P. & Bartlein, P. J. Global vegetation and terrestrial carbon cycle changes after the last ice age. New Phytol. 189, 988–998 (2011).

42. Kageyama, M. et al. Mid-Holocene and last glacial maximum climate simulations with the IPSL model: part II: model-data comparisons. Clim. Dyn. 40, 2469–2493 (2015).

43. Şenol, Ç. & Doğan, U. Vegetation and climate of Anatolia and adjacent regions during the Last Glacial period. Quatern. Int. 302, 110–122 (2013).

44. Nieto Feliner, G. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. Perspect. Plant Ecol. Evol. Syst. 16, 265–278 (2014).
45. Duminnil, J. & di Micheile, M. Plant species delimitation: a comparison of 838 morphological and molecular markers. Plant Biosyst. 143, 528–542 (2009).
46. Funk, D. J. & Omland, K. E. Species-level paralogy and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423 (2003).
47. Gutiérrez Larena, B., Fuertes Aguilar, J. & Nieto Feliner, G. Glacial-induced altitudinal migrations in Armeria (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing. Mol. Ecol. 11, 1965–1974 (2002).
48. Miebach, A., Niestrath, P., Roesser, P. & Litt, T. Impacts of climate and humans on the vegetation in NW Turkey: palynological insights from Lake İznil since the Last Glacial. Clim. Past Discuss. 11, 5157–5201 (2015).
49. Billekin, D. et al. Anatolia: A long-time plant refuge area documented by pollen records over the last 23 million years. Rev. Palaeobot. Palynol. 215, 1–22 (2015).
50. Körner, C. & Ohshima, M. Mountain systems in Ecosystems and human well-being, current state and trends. Findings of the Condition and Trends Working Group of the Millennium Ecosystem Assessment, Vol. 1 (eds. Hassan, R., Scholes, R. & Ash, N.) 681–716 (Island Press, 2005).
51. Verlaque, R., Médail, F., Quézel, P. & Babinot, J. F. Endémisme végétal et paléogéographie dans le bassin Méditerranéen. Geobios 21, 159–166 (1997).
52. Nace, G. M., Gattlemam, J. L. & Purvis A. Preserving the tree of life. Science 300, 1707–1709 (2003).
53. Sekeçerioğlu, Ç. H. et al. Turkey's globally important biodiversity in crisis. Biol. Conserv. 144, 2752–2769 (2011).
54. Davis, E. B. et al. The California Hotspots Project: Identifying regions of rapid diversification of mammals. Mol. Ecol. 17, 120–138 (2008).
55. Andreakis, N., Koostira, W. H. C. F. & Procaccini, G. High genetic diversity and connectivity in the polypliod invasive seaweed Asparagopsis taxiformis (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. Mol. Ecol. 18, 212–226 (2009).
56. Ekim, T. et al. Türkçe bitkileri kırımızı kitabı (egrelti ve tohumlu bitikler) [Red data book of Turkish plants (Pteridophyta and Spermatophyta)] (Turkish Association for the Conservation of Nature-Van University, 2000) [in Turkish].
57. IUCN. IUCN Red List Categories and Criteria: Version 3.1. (IUCN Species Survival Commission, 2001).
58. Duran, A. & Duman, H. Two new species of Centaurea (Asteraceae) from Turkey. Ann. Bot. Fennici 39, 43–48 (2002).
59. Uysal, T., Özel, E., Bozkurt, M. & Ertugrul, K. Genetic diversity in threatened populations of the endemic species Centaurea lycaonica Boiss. & Heldr. (Asteraceae). Res. J. Biol. 2, 110–116 (2012).
60. Doyle, J. J. & Dickson, E. E. Preservation of plant samples for DNA restriction endonuclease analysis. Taxon 36, 715–718 (1987).
61. Fréville, H., Imbert, E., Justy, F., Vitalis, R. & Olivieri, I. Isolation and characterization of microsatellites in the endemic species Centaurea corymbosa Pourret (Asteraceae) and other related species. Mol. Ecol. 9, 1671–1672 (2000).
62. Marrs, R. A., Hufbauer, R. A., Bogdanowicz, S. M. & Sforza, R. Nine polymorphic microsatellite markers in Centaurea stoebe L. (subspecies C. s. stoebe and C. s. micranthos (S. G. Gmelin ex Gugler) Hayek) and C. diffusa Lam. (Asteraceae). Mol. Ecol. Notes 6, 897–899 (2006).
63. Peakall, R. & Smouse, P. E. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295 (2006).
64. Belkhir, K., Bours, P., Chikhi, L., Raufaste, N. & Bonhomme, F. NEWMIC 4. 05, Logiciel sous Windows™ pour la génétique des populations (Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, 1996–2004).
65. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370 (1984).
66. Raymond, M. & Rousset, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenism. J. Hered. 86, 248–249 (1995).
67. Raymond, M & Rousset, F. An exact test for population differentiation. Evolution 49, 1283–1286 (1995).
68. Chapuis, M.-P. & Estoup, A. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24, 621–631 (2007).
69. Pritchard, J. K., Wen, X. & Falush, D. Documentation for Structure software: version 2. 3 (Department of Human Genetics, University of Chicago). Available at: http://pritch.bsd.uchicago.edu/structure_software/release_versions/v2.3.4/structure_doc.pdf (Accessed 24th August 2015), (2010).
70. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters individuals using the software Structure: a simulation study. Mol. Ecol. 14, 2611–2620 (2005).
71. Neal, D. A. & vonHoldt, B. M. Structure Harvester: A website and program for visualizing Structure output and implementing the Evanno method. Conserv. Genet. Resources 6, 359–361 (2012).
72. Menghi, X., Guérard, E. & Meyer, E. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by Monmonner's algorithm. Hum. Biol. 76, 173–190 (2004).
73. Nei, M., Tajima, F. & Tateno, Y. Accuracy of estimated phylogenetic trees from molecular data. J. Mol. Evol. 19, 153–170 (1983).
74. Dieringer, D. & Schlötterer, C. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes 3, 167–169 (2003).
75. Wilson, G. A. & Rannala, B. Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163, 1177–1191 (2003).
76. Breehl, P. & Felsenstein, J. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by a coalescent approach. Proc. Natl. Acad. Sci. USA 98, 4563–4568 (2001).
77. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees in parallel. Proc. of the Gateway Computing Environments Workshop (GCE). 14 Nov. 2010, 1–8 (2010).
78. Pearson, R. G., Saxworthy, C. J., Nakamura, M. & Peterson, A. T. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeogr. 34, 102–117 (2007).
79. Phillips, S. I., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259 (2006).
80. Collins, W. D. et al. The community climate system model version 3 (CCSM3). J. Clim. 19, 2122–2143 (2006).
81. Hasumi, H. & Emori, S. K-1 Coupled GCM (MIROC) Description. K-1 Technical Report 1 (Center for Climate System Research, University of Tokyo, 2004).
82. Radosavljevic, A. & Anderson, R. P. Making better Maxent models of species distributions: complexity, overfitting and evaluation. J. Biogeogr. 41, 629–643 (2014).
83. Zhao, J.-L., Gugger, P. F., Xia, Y.-M. & Li, Q.-J. Ecological divergence of two closely related Rosoccea species associated with late Quaternary climate change. J. Biogeogr. 43, 1990–2001 (2016).
84. Beyer, H. L. Hawth's analysis tools for ArcGIS, Available at: http://www.spatial ecology.com/htools (Accessed: 5th August 2015) (2004).

Acknowledgements

This work was supported by the Spanish Ministerio de Ciencia e Innovación [project CGL2010/18631] and Generalitat de Catalunya [Ajuts a Grups de Recerca Consolidats 2014-SGR514-GREB].
Author Contributions
A.S., K.E., T.U. and O.T. collected samples; S.L.-V., A.G. and N.G.-J. conducted the experiments; J.L.-P., S.L.-V. and N.G.-J. analyzed the data; A.S. and N.G.-J. designed the study; J.L.-P., A.S. and N.G.-J. led the writing of the manuscript; K.E., T.U. contributed to the writing. All authors read and reviewed the final manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: López-Pujol, J. et al. Speciation and genetic diversity in Centaurea subsect. Phalolepis in Anatolia. Sci. Rep. 6, 37818; doi: 10.1038/srep37818 (2016).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2016