An integrative taxonomic study of *Santolina* (Asteraceae) from southern France and north-eastern Spain reveals new endemic taxa

Antonio Giacò[1,*], Lucia Varaldo[2], Gabriele Casazza[2], Daniele De Luca[3], Paolo Caputo[3], Marco Sarigu[4], Gianluigi Bacchetta[4], Llorenç Sáez[5], and Lorenzo Peruzzi[1]

[1] PLANTSEED Lab, Department of Biology, University of Pisa, Via Derna 1, 56126 Pisa, Italy

[2] Department of DISTAV, University of Genoa, Corso Europa 26, 16132, Genoa, Italy

[3] Department of Biology, University of Naples "Federico II", via Cupa Nuova Cintia, 21, 80126, Naples, Italy

[4] Department of Life and Environmental Sciences, University of Cagliari, V.le S. Ignazio da Laconi 11-13, Cagliari, Italy

[5] Department BABVE (Systematics and Evolution of Vascular Plants, associate unit to CSIC), Faculty of Biosciences, Autonomous University of Barcelona, 08193, Bellaterra Barcelona, Spain

* Author for correspondence. E-mail: antonio.giaco@phd.unipi.it

Abstract: *Santolina* is a clear example of a genus lying in an alpha-taxonomic status, with species accepted only based on qualitative morphological descriptions. In particular, taxonomic issues still need to be resolved for *Santolina* populations from southern France and north-eastern Spain, and so, we carried out an integrative taxonomic study involving morphometrics, cypsela morphometrics, niche overlap, and phylogenetic analysis based on six plastid markers (*trnH-psbA*, *trnL-trnF*, *trnQ-rps16*, *rps15-ycf1*, *psbM-trnD*, and *trnS-trnG*). Our results revealed that the current taxonomic circumscription is not adequate. In particular, the *Santolina* populations at the foothills of eastern Pyrenees, previously included in the variability of *Santolina benthamiana*, have to be considered as a distinct species, namely *Santolina intricata*. In addition, despite their high phylogenetic relatedness, *S. benthamiana* s.str. and *Santolina ericoides* can be still considered as distinct species due to clear morphological and ecological differentiation. Finally, we demonstrated that three different subspecies can be recognized in *Santolina decumbens*, a species endemic to Provence. For one of these subspecies, due to its extremely restricted distribution range, conservation issues are pointed out.

Key words: Anthemideae, image analysis, Mediterranean Basin, molecular systematics, morphometrics, niche overlap

1. Introduction

In plant taxonomy, the study of morphology is the most traditional and basic approach to infer relationships among taxa (Radford, 1986). Species delimitation is thus based on the observation of qualitative or quantitative diagnostic characters showing scarce or no overlap among the studied taxonomic units (Wiens, 2007). However, morphology alone can be deceptive and can lead scholars to arrive at conclusions that do not reflect the actual relationships among taxa (Bateman, 2018; Gaudeul et al., 2018; Liu et al., 2022). Despite this, a great number of species are still accepted nowadays based only on qualitative morphological description (Rouhan & Gaudeul, 2014). Turrill (1938) defined this condition as "alpha taxonomy" and hypothesized that future scholars would have tried out new methods to infer taxonomic conclusions. Indeed, cytogenetics, molecular systematics, and also methods based on ecological niche analysis are now commonly used in systematics and taxonomy (Raxworthy et al., 2007; Rouhan & Gaudeul, 2014). In addition, morphology, the most traditional approach, could be studied in depth thanks to the implementation of advanced statistical tools and classification methods based on machine learning, the so-called morphometry (Rohlf & Marcus, 1993; Moffat et al., 2015). To overcome the limits of a taxonomic approach based on a single method, Dayrat (2005) defined the "integrative taxonomy" as "the science that aims to delimit the units of life′s diversity from multiple and complementary perspectives". Since then, an integrative approach has been successfully used in plenty of studies to untangle the complex relationships within several groups of plants (e.g., Herrando-Moraira et al., 2017; Andriamihaja et al., 2022; Liu et al., 2022, Tiburtini et al., 2022) and animals (e.g., Vacher et al., 2017; Venkatraman et al., 2019; Sakaguchi et al., 2020).

The *Santolina chamaecyparissus* L. species complex is an example of a group with an alpha-taxonomic status, for which an integrated systematic approach is currently being carried out (Giacò et al., 2022; De Giorgi et al., 2022). Except for the cytotaxonomic studies conducted by Marchi & D′Amato (1973), Marchi et al. (1979), and Valdés-Bermejo & Antúnez (1981), the taxonomic circumscription of species within this complex is mostly based on qualitative morphological observations (Arrigoni, 1982; Tison & de Foucault, 2014; Carbajal

et al., 2019). This complex includes 13 subshrub species from the western Mediterranean Basin. In a recent paper, the two traditionally recognized species occurring in Corsica and Sardinia, namely, *Santolina insularis* (Gennari ex Fiori) Arrigoni and *Santolina corsica* Jord. & Fourr., were investigated. Using different lines of evidence, the authors were able to conclude that, despite the different ploidy levels (tetraploid and hexaploid), these two putative species are actually included in the wide morphological variation of a single species (De Giorgi et al., 2022). In a revision of the nomenclature of the *S. chamaecyparissus* complex, Giacò et al. (2021) pointed out several taxonomic issues concerning *Santolina* populations from southern France and north-eastern Spain. Indeed, 15 names were validly published for that area. Of all these names, just three are currently accepted, whereas the others are considered as heterotypic synonyms. However, based on the study of type specimens, Giacò et al. (2021) concluded that the taxonomic position of some of these synonyms is unclear. The three accepted species are *Santolina decumbens* Mill., endemic to Provence (south-eastern France); *Santolina ericoides* Poir., widely distributed in Occitanie (southern France) and eastern Spain; and *Santolina benthamiana* Jord. & Fourr., endemic to the eastern Pyrenees (France and Spain) (Fig. 1). According to Tison et al. (2014) and Carbajal et al. (2019), the distribution ranges of the two latter species may overlap on both sides of the Spanish and French Pyrenees. Moreover, although *S. ericoides* and *S. benthamiana* are, respectively, recorded for lower and higher altitudes, their altitudinal ranges also overlap. In these putative contact areas, populations with a somewhat intermediate morphology are recorded (Carbajal et al., 2019). Giacò et al. (2021) also pointed out that four names, currently provisionally considered as synonyms of *S. benthamiana*, are based on type specimens that actually also show some morphological features that are otherwise typical of *S. ericoides*. A recent karyomorphological study (Giacò et al., 2022) demonstrated that all the *Santolina* populations native to southern France and north-eastern Spain are diploid, with 2n=18 chromosomes. Therefore, the current taxonomic delimitation of *Santolina* species from southern France and north-eastern Spain is based exclusively on qualitative morphological observations (Tison & de Foucault, 2014; Tison et al., 2014). The aim of the present study is to test the current three-species taxonomic hypothesis, using an integrated approach already successfully applied in *Santolina* (De Giorgi et al., 2022), involving morphometric analysis of adult plants and cypselae, niche similarity, and phylogenetic analysis (plastid markers *trnH-psbA*, *trnL-trnF*, *trnQ-rps16*, *rps15-ycf1*, *psbM-trnD*, and *trnS-trnG*).

2. Material and Methods

2.1 Collection of data

The populations analyzed in this study (Table 1) are the same as those sampled by Giacò et al. (2022) in southern France and north-eastern Spain: the populations from type localities of *Santolina benthamiana*, *Santolina decumbens*, and *Santolina ericoides* (Fig. 1); a population collected at the foothills of the French Pyrenees, from an area where populations morphologically intermediate between *S. benthamiana* and *S. ericoides* are reported (Giacò et al., 2021) (Ben-rou in Table 1); a population of *S. decumbens* located at the northernmost portion of its species range (Dec-sis in Table 1); an isolated population of *S. decumbens* located at the westernmost portion of its species range (Dec-lfo in Table 1); and two populations of *S. ericoides* from north-eastern Spain. For morphometrics, 20 flowering individuals were collected for each population. A subset of three individuals was used for the molecular phylogenetic analysis. A set of 100 cypselae for each population was used for the cypselae morphometric analysis. Finally, for the niche analysis, occurrence points were obtained from online databases (http://flore.silene.eu and http://www.anthos.es, both accessed on July 7, 2022), herbarium specimens georeferenced with sufficient precision, personal communications, and field investigations.

2.2 Molecular phylogenetic analysis

The same markers as those used by De Giorgi et al. (2022) were screened for variability (ITS region and the plastid markers *trnH-psbA*, *trnL-trnF*, *trnQ-rps16*, *rps15-ycf1*, *psbM-trnD*, and *trnS-trnG*). As outgroups, *Achillea millefolium* L. from NCBI and two populations of *Santolina villosa* Mill. (the same as those sampled in the karyo-morphological study by Giacò et al., 2022) were used. For A. millefolium, we extracted the above-mentioned markers from the plastid genome available on NCBI (accession no. MT500583), while for the two populations of *S. villosa*, they were de novo amplified and sequenced. DNA was extracted using the kit ExgeneTM Plant SV Mini (GeneAll Biotechnology, Seoul, South Korea). Amplification was carried out using a polymerase chain reaction (PCR) using the primers and amplification conditions listed in Table 1. The following reagents were added in 0.2 mL Eppendorf tubes: 12.5 µL of Kodaq 2× PCR MasterMix (Applied Biological Materials, Richmond, Canada), 1 µL of the forward primer at 10 µM, 1 µL of the reverse primer at

10 µM, and the DNA template (0.5–1.5 µL); then, tubes were filled with distilled water up to the volume of 25 µL. PCR products were subjected to quality checks and purification. The final product (5 µL) was then diluted with 8 µL of distilled water, and, finally, subjected to capillary electrophoresis using a 3130 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA). The electropherograms were edited with the software Sequence Scanner 2 (Applied Biosystems), while sequences were edited and aligned with the software Bioedit v. 7.2.5 (Hall, 1999) and Clustal W (Thompson et al., 1994). The sequences were submitted to DDJB (accession numbers are shown in Table S2). The nuclear markers were not further analyzed because of the total absence of variation among sequences. The nucleotide evolution models were computed separately for each of the six plastid markers using jModelTest ver. 2.1.10 (Darriba et al., 2012) with the corrected Akaike information criterion (Akaike, 1974). A Bayesian analysis was then carried out using the software MrBayes ver. 3.2.6 (Ronquist et al., 2012), with default values for 2 000 000 generations, sampling chains every 2000 generations. Convergence and effective sample sizes (ESSs) were investigated using Tracer v.1.7 (Rambaut et al., 2018). The first 10% of the samples were discarded as burn-in. The majority-rule consensus tree was observed in FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

2.3 Cypsela morphometric analysis

This approach is based on image analysis and its aim is to detect differences in the external size and shape of diaspores of different populations/groups of populations. It was successfully used in several studies to detect differences among taxonomically critical species groups, providing data useful for taxa delimitation (Sarigu et al., 2019; De Giorgi et al., 2022; Tiburtini et al., 2022). A set of 100 cypselae for each population was acquired with a flatbed scanner (Epson Perfection V550) with a digital resolution of 1200 dpi. Before image acquisitions, the scanner was calibrated as suggested by Shahin & Symons (2003). One hundred cypselae were then placed on the scanner. To avoid possible interference of environmental light, images were first acquired covering the scanner with a white box and then with a black box. These images were processed using the software package ImageJ v1.53a (available for free online: http://rsb.info.nih.gov/ij (accessed on June 27, 2022).

For each cypsela, to extract and analyze 20 colorimetric and 26 morphometric features and 78 variables of the Elliptic Fourier Descriptors (EFDs), the same procedure as that described in De Giorgi et al. (2022) and in

Tiburtini et al. (2022) was used. In this study, only the morphometric variables were used, since the colorimetric variables and elliptic Fourier descriptors were found to be scarcely discriminant among populations.

A Pearson correlation test was conducted on the 26 morphometric variables, and highly correlated variables ($r > 0.9$) were discarded. The final dataset was composed of 12 variables. After standardization of data, a Linear Discriminant Analysis (LDA) was conducted using SPSS 16.0 (Statistical Package for Social Science, IBM Corp., Armonk, NY, USA).

2.4 Morphometric analysis

For each individual, 33 character (29 quantitative and 4 qualitative) were measured (Table 2). Measures were taken on dried material using a digital caliper or imageJ, depending on the character (Table 2). For the latter case, 1200 dpi scans were acquired using a flatbed scanner (Epson perfection 2480 photo). The degree of tomentosity of leaves and stems (fs_hair, ss_hair, fsl_hair, and ssl_hair in Table 2) was obtained after acquisition of HD images of a selected area of leaf or stem using a digital camera (Canon PowerShot S45) mounted on a WILD Heerbrugg M420 stereomicroscope. The degree of tomentosity was then calculated by dividing the area covered by tomentum by the total selected area. The character ss_hair (tomentosity of the sterile stem) was categorized as follows: 0%–5% (hairless or almost hairless), 6%–30% (slightly pubescent), 31%–60% (pubescent), 61%–90% (tomentose), and 91%–100% (densely tomentose). After measurements, the specimens were conserved at the Herbarium of the Botanic Museum of the Pisa University (PI; herbarium acronyms follow Thiers, 2022). Digital images of all the herbarium specimens used in this study can be found at JACQ Virtual Herbaria: https://www.jacq.org/ (the list of voucher specimens is reported in Table 1).

After standardization of data, a Principal Coordinate Analysis (PCoA) based on Gower distance was carried out to explore the overall morphological variability of populations. The Random Forest (RF) classification method was used to verify the correct classification of groups. In particular, we tested with RF the current taxonomic hypothesis and other alternative grouping hypotheses derived from the morphometric and phylogenetic results. RF was successfully used in other taxonomic studies that used a morphometric approach (Moffat et al., 2015; De Giorgi et al., 2022). The main difference and advantage with respect to other classification methods (e.g., LDA) is that qualitative variables can also be included. In addition, it is a nonparametric method, and the classification model is robust even when correlated variables are used (Strobl

et al., 2008; Touw et al., 2013). The RF analysis was conducted in R environment using the package "randomForest" (version 4.6-14, Liaw & Wiener, 2002). The morphometric data set was randomly and equally split into two subsets (training and testing). One hundred iterations of the function "tuneRF" were run to calculate the optimal number of characters to use at each node. The algorithm was calibrated to take into account the covariation among characters. Then, a forest of 800 decision trees was built and a confusion matrix based on the results of RF was obtained. This process was reiterated 100 times, randomly splitting the training and testing subsets each time. Finally, a confusion matrix reporting the mean values of classification was calculated.

Univariate analyses comparing populations were conducted on each quantitative character. Homoskedasticity was checked on each character using the Bartlett test. When $P > 0.05$, the analysis of variance (ANOVA) was conducted, followed by the Tukey–Kramer post-hoc test. Otherwise, Welch t tests were performed for each population pair, using the Hochberg correction to reduce the family-wise error rate. When a significant difference was detected ($P < 0.05$), the index of effect size Cohen′s d (Cohen, 1988) was calculated. For characters with unequal variance, an adjusted Cohen′s d index was calculated (Aoki, 2020). Cohen′s d index is a measure of the standardized difference between two means (Cohen, 1988). According to Sawilowsky (2009), significant differences with $d < 0.2$ should be considered negligible; instead, when $d > 0.8$, differences are considered as large and when $d > 1.2$, differences are considered very large. For the purposes of this study, to exclude characters with high overlap between population pairs, significant morphological differences were considered relevant only when $d > 1.2$, that is, when the two means are distant as 1.2 standard deviations.

2.5 Niche analysis

We quantified niche overlap among morphotypes using the PCA-based method developed by Broennimann et al. (2012). The climatic niche is defined in a multivariate space that is built from real occurrences of taxa. Nineteen bioclimatic variables for current (2000–2016) time slices at about 1 × 1 km spatial resolution were downloaded from the CHELSA v.1.2 data set (Karger et al., 2017, www.chelsa-climate.org, accessed on the 7 July 2022). We performed Principal Component Analysis (PCA) using the "ade4" package in R (Dray & Dufour, 2007) to reduce the transferability issue (Petitpierre et al., 2017). The first two axes of PCA were used

as new variables. To measure niche overlap, we used Schoener′s D index (Schoener, 1970), which ranges from 0 (no overlap) to 1 (full overlap). We used the niche similarity test to test whether the environmental niche occupied by a morphotype is more similar to the one occupied by another morphotype than would be expected at random (Warren et al., 2008). The niche similarity test compares the set of environmental conditions occupied by two taxa taking into account the background environmental conditions that are available in both distributional ranges. In brief, the observed niche overlap was compared with the overlap measured between the niche of one morphotype and the niche obtained by randomly sampling (100 times) occurrence points in the background area of the other morphotype. The randomized niche was obtained by randomly shifting the entire observed density of occurrences in one range (the background). We repeated the randomization procedure 100 times and we used a 5-km background area calculated around the points of occurrence. Significant results suggest that the ecological niches of morphotypes are either more or less similar than expected by chance. The analyses were conducted in R (R Core Team, 2022) using the "ecospat" package.

2.6 Distribution of circumscribed taxa

We studied the distribution of the taxa circumscribed at the end of our study by surveying the following online herbaria: AIX, ANG, AUR, BESA, CBPF, CLF, GAP, LIP, LY, LYJB, MARS, MHNM, MPU, NCY, P, SLA, VIL, and VTA (herbarium acronyms follow Thiers, 2022). The specimens were reidentified according to the new identification key, georeferenced, and used to build a distribution map.

3. RESULTS

3.1 Molecular phylogenetic analysis

The ITS region was excluded from further analysis since it did not provide any informative site. The marker *trnQ-rps16* shows the highest number of informative sites (14), followed by *rps15-ycf1* and *trnH-psbA* (9) (Table 3). Conversely, the marker *trnS-trnG* shows the lowest number of informative sites (2). The concatenated matrix was 3485 bp long and the total number of informative sites is 45 (1.2%). Runs of the Bayesian analysis were mutually converging (SD < 0.01) and ESS were all ≫200. The phylogenetic tree based on Bayesian inference is reported in Fig. 2. The studied taxa are collectively monophyletic with respect to the selected outgroups. Two well-supported clades (P = 1) can be observed: one is partly unresolved and consists

of the typical *S. benthamiana* and the three populations of *Santolina ericoides*; the second clade consists of *S. benthamiana* from Le Roumenga (Ben-rou), which is sister to the three *S. decumbens* populations.

3.2 Cypsela morphometric analysis

Of the 26 morphometric variables considered, 12 noncorrelated variables were retained (Table S3). The LDA applied to the current taxonomic hypothesis (Table 4) returned a percentage of correct classification of 85.3%. *Santolina decumbens* is well classified (97%), whereas *S. benthamiana* and *S. ericoides* show lower values. Then, we applied the LDA considering each population as a distinct group (Table 5). In this case, the value of correct classification was 53.1%. However, the typical *S. benthamiana* (Ben-LC) is well classified (95.9%), followed by *S. decumbens* from La Fare-les-Oliviers (Dec-lfo), with 86%. All the other populations show lower values and are largely misclassified.

3.3 Morphometric analysis

The first two axes of the PCoA account for 58.68% of the overall morphological variability (Fig. 3). The three populations of *S. ericoides* almost completely overlap. The three populations of *S. decumbens* do not overlap, although they are located very close in the multivariate space. The population of *S. benthamiana* from the type locality (Ben- LC) is distinct with respect to *S. ericoides* and *S. decumbens*, whereas the population of *S. benthamiana* from Le Roumenga (Ben-rou) overlaps in part with the typical *S. benthamiana*, with *S. decumbens* from Sisteron, and to a lesser extent with the typical *S. decumbens*. RF applied to the current taxonomic hypothesis (Table 6) shows a percentage of mean overall correct *a priori* classification of 95.2%. However, while *S. ericoides* and *S. decumbens* are highly correctly classified by the algorithm (100% and 97.1%, respectively), *S. benthamiana* shows lower values (88.3%) and is mostly misclassified with *S. decumbens*.

We tested an alternative grouping hypothesis that can be inferred from the phylogenetic results. If we consider *S. benthamiana* from the type locality (Ben-LC) and *S. ericoides* as belonging to the same group and *S. benthamiana* from Le Roumenga (Ben-rou) as a distinct group, the mean overall value of correct *a priori* classification decreases to 88.5% (Table 7). The group composed of *S. benthamiana* s.str. + *S. ericoides* is well

classified (99.3%), as well as *S. decumbens* (97.8%). However, *S. benthamiana* from Le Roumenga is misidentified with other groups in 31.7% of cases.

Finally, we also tested an alternative grouping hypothesis based on the results of the PCoA (Table 8). According to this hypothesis, each population is considered as a distinct group, with the exception of *S. ericoides*. The percentage of mean correct *a priori* classification is 91.4%. The typical *S. benthamiana* is well classified (95.6%) and only scarcely misclassified with *S. benthamiana* from Le Roumenga (2.1%) and with *S. ericoides* (2.4%). *Santolina benthamiana* from Le Roumenga (76.5% of mean correct classification) is misclassified with *S. decumbens* from Sisteron (14.1%), with the typical *S. decumbens* (5.9%), and with the typical *S. benthamiana* (3.5%). The three populations of *S. decumbens* are well classified, except for the population from Sisteron, which is misclassified with *S. benthamiana* from Le Roumenga (7.5%). On the contrary, *S. ericoides* is correctly classified with a value of 100%. In Table 9, the mean value ± standard deviation of each character is reported for each population, whereas in Table 10, the number of significant scarcely overlapping characters (Tukey–Kramer or Welch t test with a P value < 0.05 and Cohen's d > 1.2) is reported.

The lowest number of scarcely overlapping significant characters was recorded among the three populations of *S. ericoides*, which differ by two to four characters (the lowest Cohen′s d value is 1.26, while the highest value is 1.88). *Santolina benthamiana* from Le Roumenga and the typical *S. benthamiana* differ by five scarcely overlapping significant characters (Cohen′s d ranging from 1.31 to 3.74), whereas *S. benthamiana* from Le Roumenga and *S. decumbens* from Sisteron differ by 6 (Cohen′s d ranging from 1.24 to 1.73). The three populations of *S. decumbens* differ each other by 7–10 scarcely overlapping significant characters (Cohen′s d ranging from 1.25 to 4.01). The highest number of scarcely overlapping significant characters (17) was recorded between the typical *S. benthamiana* and the typical *S. decumbens* (Cohen′s d ranging from 1.3 to 8.45).

3.4 Niche Analysis

The occurrence points used for this analysis were categorized based on the morphotypes detected by morphometric results (Table S4). The niche overlap between each pair of morphotypes ranged from no (0.0) to moderate (0.38) overlap (Table 11). The highest value was recorded between the typical *S. benthamiana* and the Ben-rou morphotype. Moderate niche overlap was also detected between *S. ericoides* and *S. decumbens*

s.str., the Dec-Sis morphotype and the Dec-lfo morphotype, and the Ben-rou morphotype. The lowest overlap values were detected between the morphotypes of *S. decumbens* and those of *S. benthamiana*, and between Dec-lfo and Dec-sis morphotypes. The significant values in the niche similarity test suggested that the ecological niche of *S. ericoides* was generally more similar to that of Ben-rou, Dec-lfo, and *S. decumbens* s.str. morphotypes than expected by chance, given their environmental backgrounds. In the other cases, the lack of significance in the similarity test suggests differences in optimal niche positions without niche shift.

4. Discussion

As for the *Santolina* populations from Corsica and Sardinia studied by De Giorgi et al. (2022), the ITS region was uninformative. This result suggests a high overall phylogenetic relatedness among all species of the *S. chamaecyparissus* complex and a relatively recent diversification in the Mediterranean context, as already hypothesized by Oberprieler (2005) for this genus. Nevertheless, the six plastid markers allowed to infer the phylogenetic relationships among the diploid taxa native to southern France and north-eastern Spain that are monophyletic with respect to the polyploid *S. villosa*. This first result fully confirms the taxonomic distinction of *S. villosa* and *S. ericoides* that, before Carbajal et al. (2019) and Giacò et al. (2021), were for a long time included into the variability of a single species, namely, *S. villosa* (Greuter, 2008; Tison & de Foucault, 2014; Tison et al., 2014). Based on our results, three main issues need to be discussed more in detail: the taxonomic position of the populations at the foothills of the Pyrenees; the taxonomic distinction of *S. benthamiana* and *S. ericoides*; and the evaluation of the infraspecific variability found within *S. decumbens*.

4.1 The taxonomic position of the populations at the foothills of the Pyrenees

The morphology of the populations at the foothills of the Pyrenees was interpreted by Giacò et al. (2021) to be intermediate between *S. benthamiana* and *S. ericoides*. Based on the geographic position and qualitative morphological observations, these populations were hypothesized to possibly represent a hybrid zone between these two species (Carbajal et al., 2019; Giacò et al., 2021). However, our results do not support this hypothesis. According to the molecular phylogenetic analysis, the population from Le Roumenga (Ben-rou) falls in a different clade with respect to both *S. benthamiana* and *S. ericoides*. Instead, it is sister to *S. decumbens*. The relatedness between *S. benthamiana* from Le Roumenga (Ben-rou) and *S. decumbens* is also supported by the

moderate values of niche overlap detected (Table 11). On morphological grounds, the population from Le Roumenga shows a high variability (PCoA in Fig. 2). Indeed, RF suggests that individuals of this population can be confused with *S. decumbens* from Sisteron, with the typical *S. decumbens*, and with *S. benthamiana* s.str. (Table 10). However, thanks to the univariate analyses, we found that the population from Le Roumenga shows a peculiar combination of diagnostic characters. With respect to the typical *S. benthamiana* and *S. decumbens* from Sisteron, plants from Le Roumenga show, respectively, a higher degree of tomentosity and a different leaf morphology (Table 10). Based on these evidences, the populations at the foothills of the eastern Pyrenees can be considered as a distinct species, for which the name *S. intricata* Jord. & Fourr. is available. This name was published by Jordan & Fourreau (1869) based on material from Caudies (Pyrenées Orientales, France), just 25 km far from Le Roumenga. Carbajal et al. (2019) hypothesized that such kind of populations (here referred to *S. intricata*) may also occur in the Spanish side of Pyrenees. We were indeed able to detect a population from Spain whose morphology fits well with *S. intricata* (Fig. 4; Table S5), albeit more field and herbarium investigations are needed to obtain a more detailed picture of the distribution of this species in north-eastern Spain. Compared to other Mediterranean mountain chains, the Pyrenees host a lower percentage of endemic species, and the area of occurrence of *S. intricata* is one of the poorest (Gómez et al., 2017). According to the latter authors, besides biogeographic, climatic, and ecological reasons, this can be due to the exclusion of putative endemic species whose taxonomy was unresolved. Therefore, the recognition of *S. intricata*, whose "intricate" morphology has deceived previous taxonomists, is an important enrichment for the flora of the eastern Pyrenees.

4.2 The taxonomic distinction of S. benthamiana *and* S. ericoides

Our phylogenetic analysis did not resolve the clade that includes *S. benthamiana* and *S. ericoides*. Traditionally, these two species have been considered as distinct due to the different leaf morphologies (Tison & de Foucault, 2014; Tison et al., 2014; Carbajal et al., 2019). Our morphometric univariate analyses confirm previous observations (Table 10), and the RF never failed to correctly classify these two species. *Santolina benthamiana* and *S. ericoides* are also distinct in terms of cypsela morphometry. Indeed, the cypselae of *S. benthamiana* were correctly classified by the LDA with high percentages of correct classification, and only the topotypical population of *S. ericoides* was slightly misclassified (3.4%) with *S. benthamiana*. Finally, the relatedness

between these two species is also supported by climatic niche analysis that suggested that they have a moderate niche overlap (Table 11). Based on this evidence, we believe that *S. benthamiana* and *S. ericoides* could still be considered as distinct species. Regarding the high affinity detected in our molecular phylogenetic analysis, we can speculate that *S. benthamiana* is a species that just very recently differentiated from the widely distributed *S. ericoides*. The adaptation to a distinct environment at, on average, higher elevations (0–1800m a.s.l. for *S. ericoides* and (600) 1500–2000 for *S. benthamiana*) (Tison & de Foucault, 2014; Carbajal et al., 2019) may have contributed to the acquisition of a different morphology. A similar pattern was detected by Fréville et al. (1998) for *Centaurea stoebe* L. (= C. maculosa Lam. subsp. maculosa; Asteraceae), a widespread Mediterranean species, and *C. corymbosa* Pourr., a narrow endemic species occurring in Occitanie, southern France. According to these authors, the latter species may have differentiated from the former after a colonization event, and the new local environmental conditions may have contributed to the morphological differentiation. Another possible evolutionary interpretation may involve a chloroplast capture during a hybridization event and subsequent backcrosses. A phylogeny of the whole genus *Santolina* using NGS techniques is currently ongoing and will likely provide further data to better understand the relationships among these two related species.

4.3 Infraspecific variability within Santolina decumbens

The monophyly of *S. decumbens* is well supported by our phylogenetic analysis (Fig. 2). According to Tison & de Foucault (2014) and Tison et al. (2014), this species can be easily distinguished by other taxa by the high degree of tomentosity on stems and leaves. However, we found that the three studied populations, located at the extremes of the distribution of this species, significantly differ from each other by several morphological characters. Indeed, when considered as distinct groups, RF correctly classified these populations. Thanks to the univariate analyses, we found that each population can be distinguished by a combination of diagnostic characters. Indeed, the topotypical population, located at the southernmost portion of the species range, is the most tomentose, showing white stems and leaves; the fertile stems as well as the leaves are short. The population from Sisteron, located at the northernmost portion of the species range, shows less tomentose fertile stems and leaves, which are also longer. Finally, the population from La Fare-les-Oliviers, located at the westernmost portion of the species range, when compared to the topotypical population, shows less tomentose

fertile stems and leaves, shorter leaves, and cypselae that can be easily distinguished, too (Table 10). According to the niche analysis, these three morphotypes occur in different climatic conditions, suggesting that the observed morphological differences may be due to differences in environmental conditions. Based on our results, we believe that it is preferable to treat these groups as distinct subspecies (Hamilton & Reichard, 1992). For the "Sisteron morphotype", a name at the species level is available (*S. diversifolia* Jord. & Fourr.), that can be recombined at the subspecific level. Concerning the "La Fare-les-Oliviers morphotype", no previously published name is available, so it is described here as a subspecies new to science. Serious conservation issues must be pointed out for this latter taxon. Indeed, it occurs in a restricted area around Bouches-du-Rhône. This area is strongly anthropized and the natural environments are crucially affected by urbanization industrial activities, and agriculture (Tatoni et al., 2004). Approximately in the same area, *Teucrium pseudochamaepitys* L. (Lamiaceae) occurs. It is another southern French narrow endemic species that, according to Lhotte et al. (2014), is considerably threatened by urbanization and habitat fragmentation. Another case of a geographically close rare species is *Arenaria provincialis* Chater & G.Halliday (Caryophyllaceae), which is endemic to the hills around Marseille, and it is "near threatened" according to IUCN (Véla et al., 2008). Based on our field observations and the literature concerning endangered species occurring approximately in the same area, a conservation status assessment involving analyses of fitness and reproductive biology is urgently needed for this subspecies.

In conclusion, the integrative taxonomic approach carried out in this study allowed to reveal a biological diversity in *Santolina* diploids from southern France and north-eastern Spain that was not detected by previous qualitative taxonomic investigations. The *Santolina* populations at the foothills of Pyrenees that were previously interpreted as "intermediate" between *S. benthamiana* and *S. ericoides* actually belong to a distinct species (i.e. *S. intricata*). According to our herbarium research, this species occurs at both sides of French and Spanish Pyrenees (Fig. 4). For *S. decumbens*, a species endemic to Provence, three different isolated lineages (Fig. 4) worth of the subspecies rank were detected. Finally, for *S. benthamiana* and *S. ericoides*, despite a high phylogenetic relatedness, they can preserve the status of species thanks to clear morphological and ecological differentiation. The patterns detected in this study are in accordance with the rich literature available for recently diversified plant species in the Mediterranean Basin (De Castro et al., 2015; Bateman, 2018; Liu et al., 2022; Terlević et al., 2022). Indeed, the Mediterranean region is one of the most important hotspots for

biodiversity, and the factors that have driven plant speciation are many and are related to its complex climatic and geologic history (Greuter, 1991; Cañadas et al., 2014). In the case of *Santolina*, the current diversity can be explained by both biological and geological/climatic phenomena. In terms of biological factors, the role of polyploidy is one of the main factors that has driven the diversification of lineages currently occurring in Corsica, Sardinia, and the Iberian Peninsula (Rivero-Guerra, 2008a, 2008b; Giacò et al., 2021). However, this is not the case of the taxa studied here, since they are all diploid. Previous scholars (Arrigoni, 1979; Torricelli et al., 2000) suggested that diploid *Santolina* species may be interpreted as schizoendemics, that is, taxa originated from the fragmentation of a common, more widespread, ancestor (Siljak-Yakovlev & Peruzzi, 2012). A phylogenomic study involving NGS techniques is currently ongoing on the whole genus to understand the general patterns of diversification and to address biogeographical questions.

5. Taxonomic treatment

Santolina benthamiana Jord. & Fourr., Icon. Fl. Eur. 2: 10. 1869 ≡ *Santolina pectinata* Benth., Cat. Pl. Pyrénées: 117. 1826, nom. illeg., non Lag. 1816 ≡ *Santolina chamaecyparissus* subsp. pecten Rouy, Fl. France 8: 224. 1903 ≡ *Santolina chamaecyparissus* var. *pectinata* Fiori in Fiori & al., Fl. Italia 3(1): 270. 1903 – Lectotype designated by Giacò et al. (2021: 191): France, Occitanie, Pyrénées-Orientales, Prats-de-Mollo, inter Arles et Prats-de-Mollo, 11 Jun 1825, *G. Bentham 573* (K barcode K000929471!).

= *Santolina chamaecyparissus* subsp. *pecten* var. *ruscinonensis* Rouy, Fl. France 8: 225. 1903 – Lectotype designated by Giacò et al. (2021: 193): France, Occitanie, Pyrénées-Orientales, Prats-de-Mollo, Jun 1846*, A. Irat s.n.* (LY barcode LY0003145!; isolectotypes: LY barcode LY0341058!, P barcodes P00752629!, P00752630!, P00752631!, P00752632!, P00752633!, P00752636!, P00752673!, P00752674! & P03315514!).

= *Santolina chamaecyparissus* subsp. *pecten* var. *hispanica* Rouy, Fl. France 8: 225. 1903 – Lectotype by Giacò et al. (2021: 193): Spain, Aragon, Broto, Jul 1873, *A. Autheman s.n.* (LY barcode LY0715656!).

Distribution: Spanish and French central-eastern Pyrenees (Fig. 4; Table S5).

Santolina decumbens Mill., Gard. Dict., ed. 8: *Santolina* no. 3. 1768 – Neotype designated by Giacò et al. (2021: 194): France, Provence-Alpes-Côte d'Azur, Mont Caume, 23 Jun 1972, *L. Mercurin s.n.* (P barcode P00113904!).

subsp. ***decumbens***

= *Santolina incana* Lam., Fl. Franç. 2: 43. 1779 ≡ *S. chamaecyparissus* var. *incana* (Lam.) DC., Prodr. 6: 35. 1838 – Neotype designated by Giacò et al. (2021: 196): France, Occitanie, Aude, Narbonne [likely cultivated], s.d., *s.coll.* (P barcode P06898681!).

= *Santolina villosissima* Poir. in Lamarck, Encycl. 6: 505. 1805 ≡ *Santolina chamaecyparissus* var. *villosissima* (Poir.) DC., Prodr. 6: 35. 1838 – Neotype designated by Giacò et al. (2021: 200): France, Provence-Alpes-Côte d'Azur, Montagne Sainte-Victoire, 29 Jun 1906, *J.-J. Paulet s.n.* (LY barcode LY0340523!).

Distribution: southern France (Provence), from Marseille eastwards to Toulon and northwards up to Montagne de Sainte-Victoire (Fig. 4; Table S5).

subsp. ***diversifolia*** (Jord. & Fourr.) Giacò & Peruzzi, **comb. nov.**

Basionym: *S. diversifolia* Jord. & Fourr., Icon. Fl. Eur. 2: 9. 1869 ≡ *S. chamaecyparissus* var. *incana* subvar. *diversifolia* (Jord. & Fourr.) Rouy, Fl. France 8: 222. 1903 ≡ *Santolina chamaecyparissus* var. *diversifolia* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 42. 1970 – Lectotype designated by Giacò et al. 2021: 195): France, Provence-Alpes-Côte d'Azur, Montagne de Lure, 5 Jun 1864, *A. Jordan s.n.* (LY barcode LY0826364; isolectotypes: LY barcodes LY0826365, and LY0826366).

Distribution: southern France, Alpes-de-Haute-Provence (Fig. 4; Table S5).

subsp. ***tisoniana*** Giacò & Peruzzi, **subsp. nov.**

It is similar to *S. decumbens* s.str., but the fertile stems are not white-tomentose. With respect to *S. decumbens* subsp. *diversifolia*, all vegetative parts are distinctly shorter (fertile stems < 15 cm long, leaves of the sterile and fertile stems < 15 mm long). The leaf morphology reminds of *S. ericoides*, but the leaves of the sterile stems are densely tomentose, whereas in *S. ericoides* are almost glabrous. Named after Jean-Marc Tison, respected scholar of the French flora.

Holotype (here designated): France, Provence-Alpes-Côte d'Azur, La Fare-les-Oliviers, small spots around the aerodrome (WGS84: 43.539610 N, 5.172029 E), about 50 m, 28 June 2020, *A. Giacò & L. Peruzzi s.n.* (PI barcode PI043100; isotypes: PI barcodes PI043099, PI043101, PI043102, PI043103, PI043104, PI043105, PI043106).

Distribution: southern France (Provence), between Berre-L'Ėtang and Lançon-Provence (Fig. 4; Table S5).

Santolina ericoides Poir. in Lam., Encycl. 6: 504. 1805 ≡ *Santolina chamaecyparissus* var. *squarrosa* subvar. ericoides (Poir.) Rouy, Fl. France 8: 223. 1903 ≡ *Santolina chamaecyparissus* subsp. *ericoides* (Poir.) Sennen & Elias in Bol. Soc. Ibér. Ci. Nat. 28(1–2): 32. 1929 – Lectotype designated by Ferrer-Gallego et al. (2021: 235): [illustration] "ABROTONUM foemina, Dodonaei" in Daléchamps, Hist. Gen. Pl.: 938. 1586.

= *Santolina brevifolia* Jord. & Fourr., Icon. Fl. Eur. 2: 11. 1869 ≡ *Santolina chamaecyparissus* var. *incana* subvar. *brevifolia* (Jord. & Fourr.) Rouy, Fl. France 8: 222. 1903 ≡ *Santolina chamaecyparissus* var. *brevifolia* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 42. 1970 – Lectotype designated by Giacò et al. (2021: 191): [illustration] "*Santolina brevifolia* Jord. et Fourr." in Jordan & Fourreau, Icon. Fl. Eur. 2: t. 234. 1869.

= *Santolina chamaecyparissus* var. *mariolensis* O.Bolòs & Vigo in Collect. Bot. (Barcelona) 17(1): 90. 1987 – Holotype: Spain, Valencian Community, Serra de Mariola, 520 m, s.d., *A. Bolós & O. Bolós s.n.* (BC barcode BC-150851!).

= *Santolina chamaecyparissus* var. *virens* Willk., Prodr. Fl. Hispan. 2(1): 80. 1865 – Lectotype designated by Giacò et al. (2021: 194): Spain, Madrid, Alcalá de Henares, dans champs pirreux incultes, 27 Jun 1854, *J. Gay s.n.* (COI barcode COI00035925!).

= *Santolina glabrescens* Jord. & Fourr., Icon. Fl. Eur. 2: 11. 1869 ≡ *Santolina chamaecyparissus* var. *squarrosa* subvar. *laxa* Rouy, Fl. France 8: 223. 1903 ≡ *Santolina chamaecyparissus* var. *glabrescens* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 42. 1970 – Lectotype designated by Giacò et al. (2021: 195): [illustration] "*Santolina glabrescens* Jord. et Fourr." in Jordan & Fourreau, Icon. Fl. Eur. 2: t. 233. 1869.

= *Santolina homophylla* Jord. & Fourr., Icon. Fl. Eur. 2: 8. 1869 ≡ *Santolina chamaecyparissus* var. *incana* subvar. *homophylla* (Jord. & Fourr.) Rouy, Fl. France 8: 223. 1903 ≡ *Santolina chamaecyparissus* var. *homophylla* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 41. 1970 – Lectotype designated by Giacò et al. (2021: 196): [illustration] "*Santolina homophylla* Jord. et Fourr." in Jordan & Fourreau, Icon. Fl. Eur. 2: t. 223. 1869.

= *Santolina microcephala* Jord. & Fourr., Icon. Fl. Eur. 2: 11. 1869 – Lectotype designated by Giacò et al. (2021: 197): Spain, Castilla La Mancha, Hellin, Collines a Hellin, 24 May 1850, *J. Gay s.n.* (LY barcode LY0826375!).

Distribution: widespread from north-eastern Spain up to Nîmes in southern France (Fig. 4; Table S5).

***Santolina intricata* Jord. & Fourr.**, Icon. Fl. Eur. 2: 11. 1869 ≡ *Santolina chamaecyparissus* var. *intricata* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 42. 1970 – Lectotype designated by Giacò et al. (2021: 196): [illustration] "*Santolina intricata* Jord. et Fourr." in Jordan & Fourreau, Icon. Fl. Eur. 2: t. 232. 1869.

= *Santolina brevicaulis* Jord. & Fourr., Icon. Fl. Eur. 2: 10. 1869 ≡ *Santolina chamaecyparissus* var. *incana* subvar. *brevicaulis* (Jord. & Fourr.) Rouy, Fl. France 8: 222. 1903 ≡ *Santolina chamaecyparissus* var. *brevicaulis* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 41. 1970 – Lectotype designated by Giacò et al. (2021: 191): [illustration] "*Santolina brevicaulis* Jord. et Fourr." in Jordan & Fourreau, Icon. Fl. Eur. 2: t. 230. 1869.

= *Santolina rigidula* Jord. & Fourr., Icon. Fl. Eur. 2: 10. 1869 ≡ *Santolina chamaecyparissus* var. *rigidula* (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 42. 1970 – Lectotype designated by Giacò et al. (2021: 198): France, Occitanie, Pyrénées-Orientales, Perpignan, Estagel, Jun 1838, *s.coll.* (LY barcode LY0826378!; isolectotypes: LY barcodes LY0341579! & LY0341580!).

= *Santolina valida* Jord. & Fourr., Icon. Fl. Eur. 2: 9. 1869 ≡ *Santolina chamaecyparissus* var. *incana* subvar. *valida* (Jord. & Fourr.) Rouy, Fl. France 8: 223. 1903 ≡ *Santolina chamaecyparissus* var. valida (Jord. & Fourr.) Guinea in Anales Inst. Bot. Cavanilles 27: 42. 1970 – Lectotype designated by Giacò et al. (2021: 199): [illustration] "*Santolina valida* Jord. et Fourr." in Jordan & Fourreau, Icon. Fl. Eur. 2: t. 224. 1869.

These four names were published simultaneously at species level by Jordan & Fourreau (1869) for the distribution area typical of the "Le Roumenga" morphotype. We opt here for *Santolina intricata* as having priority over the competing *S. brevicaulis*, *S. rigidula*, and *S. valida* (Art. 11.5 of the ICN; Turland et al. 2019).

Distribution: southern France, Pyrénées-Orientales and Girona province in Spain (Fig. 4; Table S5).

Identification key for diploid *Santolina* native to France and north-eastern Spain

1. Leaves of the sterile stems, except for the ones at the tips, green and almost glabrous (tomentose at least on the leaf axis) ……... 2

1. Leaves of the sterile stems tomentose, grey or white ……..…… 3

2. Leaves of the sterile stems with 0.8–2.2 mm spaced out segments, (1.5)2–8 mm long …..... ***S. benthamiana***

2. Leaves of the sterile stems with 0.2–0.8 mm spaced out segments, usually shorter than 2 mm ..… ***S. ericoides***

3. Leaves of the sterile stems with 25–65, 0.8–2 mm spaced out, segments. Segments of the fertile stem leaves 1.5–3 mm long and 0.5–1.5 mm spaced-out ……..……...……. ***S. intricata***

3. Leaves of the sterile stems with 50–80, < 1 mm spaced out, appressed segments. Segments of the fertile stem leaves 0.5–2 mm long and 0.0–1.0 mm spaced-out ……...……. 4 (***S. decumbens***)

4. Fertile stems white, tomentose as (or almost as) the sterile stems ……. ***S. decumbens* subsp. *decumbens***

4. Fertile stems green, clearly less tomentose than the sterile stems ……. 5

5. Longest fertile stems 12–30 cm long. Longest leaves of the sterile stems 15–35 mm long ……..……..... ***S. decumbens* subsp. *diversifolia***

5. Longest fertile stems 5–15 cm long. Longest leaves of the sterile stems 8–15 mm long …… ***S. decumbens* subsp. *tisoniana***

Acknowledgements

We are grateful to J.-M. Tison, who kindly showed us precise localities for populations to be sampled in France. We also acknowledge R. Carballal for providing the precise locality of plants from Sisteron and for sampling *Santolina villosa*. This work was supported by the "Progetto di Ricerca di Rilevante Interesse Nazionale" (PRIN) "PLAN.T.S. 2.0—towards a renaissance of PLANt Taxonomy and Systematics" led by the University of Pisa under the grant number 2017JW4HZK (Principal Investigator: Lorenzo Peruzzi).

Conflicts of Interest

The authors declare no conflicts of interest.

References

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716–723.

Andriamihaja CF, Botomanga A, Misandeau C, Ramarosandratana AV, Grisoni M, Da Silva D, Pailler T, Jeannoda VH, Besse P. 2022. Integrative taxonomy and phylogeny of leafless *Vanilla* orchids from the South-

West Indian Ocean region reveal two new Malagasy species. *Journal of Systematics and Evolution*: https://doi.org/10.1111/jse.12858

Aoki S. 2020. Effect sizes of the differences between means without assuming variance equality and between a mean and a constant. *Heliyon 6*: e03306. https://doi.org/10.1016/j.heliyon.2020.e03306

Arrigoni PV. 1979. Le genre «*Santolina*» L. en Italie. *Webbia*: 34: 257–264.

Arrigoni PV. 1982. *Santolina* L. In: Pignatti S, ed. *Flora d'Italia*. 1st ed. Bologna: Edagricole, 64–65.

Bateman RM. 2018. Two bees or not two bees? An overview of *Ophrys* systematics. *Berichte aus den Arbeitskreisen Heimische Orchideen* 35: 5–46.

Bermejo EV, Antúnez C. 1981. Estudios caríológicos en especies españolas del género *Santolina* L. (Compositae). *Anales del Jardín Botánico de Madrid* 38: 127–144.

Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Thuiller W, Fortin MJ, Randin C, Zimmermann NE, Graham CH. 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global ecology and biogeography* 21: 481–497. https://doi.org/10.1111/j.1466-8238.2011.00698.x

Cañadas EM, Fenu G, Peñas J, Lorite J, Mattana E, Bacchetta G, 2014. Hotspots within hotspots: endemic plant richness, environmental drivers, and implications for conservation. *Biological Conservation* 170: 282–291. https://doi.org/10.1016/j.biocon.2013.12.007

Carbajal R, Ortiz S, Sáez L. 2019. *Santolina* L. In: Castroviejo SB, Benedí C, Buira A, Rico E, Crespo MB, Quintanar A, Aedo C, eds. *Flora Iberica*. Madrid: Real Jardín Botánico CSIC, 1938–1962.

Cohen J. 1988. *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, New Jersey: Erlbaum.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9: 772. https://doi.org/10.1038/nmeth.2109

Dayrat B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean society* 85: 407–417.

De Castro O, Véla E, Vendramin GG, Gargiulo R., Caputo P. 2015. Genetic structure in the *Genista ephedroides* complex (Fabaceae) and implications for its present distribution. *Botanical Journal of the Linnean Society* 177: 607–618. https://doi.org/10.1111/boj.12263

De Giorgi P, Giacò A, Astuti G, Minuto L, Varaldo L, De Luca D, De Rosa A, Bacchetta G, Sarigu M, Peruzzi L. 2022. An integrated taxonomic approach points towards a single-species hypothesis for *Santolina* (Asteraceae) in Corsica and Sardinia. *Biology* 11: 356. https://doi.org/10.3390/biology11030356

Dray S, Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software* 22: 1–20.

Ferrer-Gallego PP, Sáez L, Wajer J, Giacò A, Peruzzi L. 2021. Typification of the names *Santolina ericoides* and *S. villosa* (Asteraceae) revisited. *Phytotaxa* 509: 233–240. https://doi.org/10.11646/phytotaxa.509.2.6

Fréville H, Colas B, Ronfort J, Riba M, Olivieri I. 1998. Predicting endemism from population structure of a widespread species: case study in *Centaurea maculosa* Lam. (Asteraceae). *Conservation Biology* 12: 1269–1278.

Gaudeul M, Siljak-Yakovlev S, Jang TS, Rouhan G, 2018. Reconstructing species relationships within the recently diversified genus *Odontites* Ludw. (Orobanchaceae): evidence for extensive reticulate evolution. *International Journal of Plant Sciences* 179: 1–20.

Giacò A, Astuti G, Peruzzi L. 2021. Typification and nomenclature of the names in the *Santolina chamaecyparissus* species complex (Asteraceae). *Taxon* 70: 189–201. https://doi.org/10.1002/tax.12429

Giacò A, De Giorgi P, Astuti G, Varaldo L, Sáez L, Carballal R, Serrano M, Casazza G, Caputo P, Bacchetta G, Peruzzi L. 2022. Diploids and polyploids in the *Santolina chamaecyparissus* complex (Asteraceae) show different karyotype asymmetry. *Plant Biosystems*: https://doi.org/10.1080/11263504.2022.2029971

Gómez D, Ferrández JV, Tejero P, Font X. 2017. Spatial distribution and environmental analysis of the alpine flora in the Pyrenees. *Pirineos* 172: e027. http://dx.doi.org/10.3989/Pirineos.2017.172002

Greuter W. 1991. Botanical diversity, endemism, rarity, and extinction in the Mediterranean area: an analysis based on the published volumes of Med-Checklist. *Botanika Chronika* 10: 63–79.

Greuter W. 2008. *Santolina* L. In: Greuter W, Raab-Straube von E, eds. *Med-checklist. A Critical Inventory of Vascular Plants of the Circum-Mediterranean Countries. II. Dycotyledones (Compositae)*. Genève: Organization for the Phyto-Taxonomic Investigation of the Mediterranean Area (OPTIMA), 696–698.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Hamilton CW, Reichard SH. 1992. Current practice in the use of subspecies, variety, and forma in the classification of wild plants. *Taxon* 41: 485–498.

Herrando-Moraira S, Carnicero P, Blanco-Moreno JM, Sáez L, Véla E, Vilatersana R, Galbany-Casals M. 2017. Systematics and phylogeography of the Mediterranean *Helichrysum pendulum* complex (Compositae) inferred from nuclear and chloroplast DNA and morphometric analyses. *Taxon* 66: 909–933. https://doi.org/10.12705/664.7

Jordan A, Fourreau JP. 1869. Icones ad Floram Europaeam 2. Paris, France: F. Savy.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 4: 170122.

Liaw A, Wiener M. 2002. Classification and regression by randomForest. *R news* 2: 18–22.

Liu L, Astuti G, Coppi A, Peruzzi L. 2022. Different chromosome numbers but slight morphological differentiation and genetic admixture among populations of the *Pulmonaria hirta* complex (Boraginaceae). *Taxon*: https://doi.org/10.1002/tax.12721

Lhotte A, Affre L, Saatkamp A. 2014. Are there contrasted impacts of urbanization and land uses on population persistence? The case of *Teucrium pseudochamaepitys*, an endangered species in Southern France. *Flora* 209: 484–490. https://doi.org/10.1016/j.flora.2014.05.002

Marchi P, D'Amato G. 1973. Numeri cromosomici per la flora italiana: 145–150. *Informatore Botanico Italiano* 5: 93–100.

Marchi P, Capineri R, D'Amato G. 1979. Il cariotipo di *Santolina corsica* Jord. & Fourr. (Compositae) proveniente da Bastia (Corsica) ed altre osservazioni. *Annali di Botanica* 38: 1–13.

Moffat CE, Ensing DJ, Gaskin JF, De Clerck-Floate RA, Pither J. 2015. Morphology delimits more species than molecular genetic clusters of invasive *Pilosella*. *American Journal of Botany* 102: 1145–1159. https://doi.org/10.3732/ajb.1400466

Oberprieler C. 2005. Temporal and spatial diversification of circum-Mediterranean Compositae-Anthemideae. *Taxon* 54: 951–966. https://doi.org/10.2307/25065480

Petitpierre B, Broennimann O, Kueffer C, Daehler C, Guisan A. 2017. Selecting predictors to maximize the transferability of species distribution models: Lessons from cross-continental plant invasions. *Global Ecology and Biogeography* 26: 275–287.

R Core Team. 2022. R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Radford AE. 1986. *Fundamentals of plant systematics*. New York: Harper & Row.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biol*ogy 67: 901–904. https://doi.org/10.1093/sysbio/syy032

Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG. 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Systematic Biology* 56: 907–923. https://doi.org/10.1080/10635150701775111

Rivero-Guerra AO. 2008a. Cytogenetics, geographical distribution, and pollen fertility of diploid and tetraploid cytotypes of *Santolina pectinata* Lag. (Asteraceae: Anthemideae). *Botanical Journal of the Linnean Society* 156: 657–667. https://doi.org/10.1111/j.1095-8339.2007.00766.x

Rivero-Guerra AO. 2008b. Cytogenetics, biogeography and biology of *Santolina ageratifolia* Barnades ex Asso (Asteraceae: Anthemideae). *Botanical Journal of the Linnean Society* 157: 797–807. https://doi.org/10.1111/j.1095-8339.2008.00821.x

Rohlf FJ, Marcus LF. 1993. A revolution morphometrics. *Trends in Ecology and Evolution* 8: 129–132.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Rouhan G, Gaudeul M. 2014. Plant taxonomy: A historical perspective, current challenges, and perspectives. In: Besse P, ed. *Molecular Plant Taxonomy: methods and protocols*. New York, USA: Humana Press, 1–37.

Sakaguchi S, Yunus M, Sugi S, Sato H. 2020. Integrated taxonomic approaches to seven species of capillariid nematodes (Nematoda: Trichocephalida: Trichinelloidea) in poultry from Japan and Indonesia, with special reference to their 18S rDNA phylogenetic relationships. *Parasitology research* 119: 957–972. https://doi.org/10.1007/s00436-019-06544-y

Sawilowsky SS. 2009. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods* 8: 597–599.

Sarigu M, Porceddu M, Schmitt E, Camarda, I, Bacchetta G. 2019. Taxonomic discrimination of the *Paeonia mascula* group in the Tyrrhenian islands by seed image analysis. *Systematics and Biodiversity* 17: 801–810. https://doi.org/10.1080/14772000.2019.1685607

Schoener TW. 1970. Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology* 51: 408–418.

Shahin MA, Symons SJ. 2003. Color calibration of scanners for scanner-independent grain grading. *Cereal Chemistry* 80: 285–289. http://dx.doi.org/10.1094/CCHEM.2003.80.3.285

Siljak-Yakovlev S, Peruzzi L. 2012. Cytogenetic characterization of endemics: past and future. *Plant Biosystems* 146: 694–702. http://dx.doi.org/10.1080/11263504.2012.716796

Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9: 1–11. https://doi.org/10.1186/1471-2105-9-307

Tatoni T, Médail F, Roche P, Barbero M. 2004. The impact of changes in land use on ecological patterns in Provence (Mediterranean France). In: Mazzoleni S, Di Pasquale G, Mulligan M, Di Martino P, Rego F, eds. *Recent Dynamics of Mediterranean Vegetation and Landscape*. Chichester: John Wiley and Sons. 107–120.

Terlević A, Bogdanović S, Frajman B, Rešetnik I. 2022. Genome Size Variation in *Dianthus sylvestris* Wulfen sensu lato (Caryophyllaceae). *Plants* 11: 1481. https://doi.org/10.3390/plants11111481

Thiers B. 2022. *Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff*. Available online: http://sweetgum.nybg.org/science/ih/ (accessed on 21 July 2022).

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

Tiburtini M, Astuti G, Bartolucci F, Casazza G, Varaldo L, De Luca D, Bottigliero MV, Bacchetta G, Porceddu M, Domina G, Orsenigo S. 2022. Integrative taxonomy of *Armeria arenaria* (Plumbaginaceae), with a special focus on the putative subspecies endemic to the Apennines. *Biology* 11: 1060. https://doi.org/10.3390/biology11071060

Tison JM, de Foucault B. 2014. *Flora Gallica–Flore de France*. Mèze: Biotope Éditons.

Tison JM, Jauzein P, Michaud H, Michaud H. 2014. *Flore de la France méditerranéenne continentale*. Turriers: Naturalia publications.

Torricelli C, Garbari F, Bedini G. 2000. *Santolina ligustica* (Compositae), specie da proteggere della flora ligure. *Atti della Società Toscana di Scienze Naturali, Memorie, Serie B,* 106 (1999): 1–7.

Touw WG, Bayjanov JR, Overmars L, Backus L, Boekhorst J, Wels M, van Hijum SA. 2013. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle?. *Briefings in Bioinformatics* 14: 315–326. https://doi.org/10.1093/bib/bbs034

Turland NJ, Wiersema JH, Barrie FR, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Kusber W-H, Li D-Z, Marhold K, May TW, McNeill J, Monro AM, Prado J, Price MJ, Smith GF eds. 2019. *International Code of Nomenclature for Algae, Fungi, and Plants (Shenzhen Code) Adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*. Glashütten; Koeltz Botanical Books.

Turrill WB. 1938. The expansion of taxonomy with special reference to Spermatophyta. *Biological Reviews* 13: 342–373.

Vacher JP, Kok PJ, Rodrigues MT, Lima JD, Lorenzini A, Martinez Q, Fallet M, Courtois EA, Blanc M, Gaucher P, Dewynter M. 2017. Cryptic diversity in Amazonian frogs: Integrative taxonomy of the genus *Anomaloglossus* (Amphibia: Anura: Aromobatidae) reveals a unique case of diversification within the Guiana Shield. *Molecular phylogenetics and evolution* 112 : 158–173. https://doi.org/10.1016/j.ympev.2017.04.017

Véla E, Auda P, Léger JF, Gonçalves V, Baumel A. 2008. Exemple d'une nouvelle évaluation du statut de menace suivant les critères de l'UICN version 3.1.: le cas de l'endémique provençale *Arenaria provincialis* Chater & Halliday (Caryophyllaceae). *Acta Botanica Gallica* 155: 547–562.

Venkatraman MX, Deraad DA, Tsai WL, Zarza E, Zellmer AJ, Maley JM, Mccormack JE. 2019. Cloudy with a chance of speciation: integrative taxonomy reveals extraordinary divergence within a Mesoamerican cloud forest bird. *Biological Journal of the Linnean Society* 126: 1–15. https://doi.org/10.1093/biolinnean/bly156

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62: 2868–2883. https://doi.org/10.1111/j.1558-5646.2008.00482.x

Wiens JJ. 2007. Species delimitation: new approaches for discovering diversity. *Systematic Biology* 56: 875–878. https://doi.org/10.1080/10635150701748506

Table 1. Sampled *Santolina* populations from southern France and north-eastern Spain and related voucher information. The taxonomic assignment of populations follows Giacò et al. (2021); LC, type locality (*locus classicus*).

Species	Population	Acronym	Vouchers
S. benthamiana	France, Occitanie, Prats-de-Mollo-la-Preste [WGS84: 42.407222 N, 2.523055 E]	Ben-LC	*A. Giacò* and *L. Peruzzi*, 29 June 2020, PI 043080–043098
S. benthamiana	France, Occitanie, Montalba-le-Château, Le Roumenga [WGS84: 42.699054 N, 2.552235 E]	Ben-rou	*A. Giacò* and *L. Peruzzi*, 28 June 2020, PI 043079, PI 057098–057114
S. decumbens	France, Provence-Alpes-Côte d'Azur, Mont Caume [WGS84: 43.184768 N, 5.908187 E]	Dec-LC	*A. Giacò* and *L. Peruzzi*, 27 June 2020, PI 043107–043118
S. decumbens	France, Provence-Alpes-Côte d'Azur, Sisteron [WGS84: 44.153341 N, 5.953744 E]	Dec-sis	*A. Giacò* and *L. Peruzzi*, 11 July 2021, PI 053348–053364
S. decumbens	France, Provence-Alpes-Côte d'Azur, La Fare-les-Oliviers [WGS84: 43.539610 N, 5.172029]	Dec-lfo	*A. Giacò* and *L. Peruzzi*, 28 June 2020, PI 043099–043106
S. ericoides	France, Occitanie, Béziers [WGS84: 43.28959 N 3.18539 E]	Eri-LC	*A. Giacò* and *L. Peruzzi*, 28 June 2020, PI 036086–036100
S. ericoides	Spain, Barcelona province, Sant Feliu de Codines [WGS84: 41.692294 N, 2.174761 E]	Eri-sfc	*L. Sáez*, 7 July 2020, PI 043077, PI 057135–057154
S. ericoides	Spain, Lleida province, Torà [WGS84: 41.814325 N, 1.404588 E]	Eri-tor	*L. Sáez*, 13 July 2020, PI 043076, PI 057115–057134

Table 2. Morphometric characters and their description for *Santolina* native to France and south-eastern Spain. CO, ordered factor; QC, quantitative continuous; QD, quantitative discrete; Y/N, bimodal.

Code	Description of the Character	Type	Tool
Vegetative parts			
fs_len	Length of the fertile stem (cm)	QC	Ruler
br_ratio	Ratio between the terminal ramification of the fertile stem and fs_len	QC	Ruler
dist_cap	Distance between the terminal leaf on the stem and the floral head (mm)	QC	Caliper
fs_n_br	Number of branches of the fertile stem	QD	
fs_n_nodes	Number of nodes of the fertile stem	QD	
ss_len	Length of the sterile stem (cm)	QC	Ruler
ss_n_nodes	Number of nodes of the sterile stem	QD	
ss_bicolor	Sterile stems densely tomentose on the upper portion, less tomentose at the bottom	Y/N	
ss_hair	Tomentosity of the sterile stem	CO	ImageJ
fs_hair	Degree of tomentosity of the fertile stem (%)	QC	ImageJ
fsl_n_seg	Number of segments on the fertile stem leaf (the longest)	QD	
ssl_n_seg	Number of segments on the sterile stem leaf (the longest)	QD	
ssl_len	Length of the sterile stem leaf (mm)	QC	ImageJ
ssl_pet_len	Length of the petiole of the sterile stem leaf (mm)	QC	ImageJ
ssl_seg_len	Length of the segment of the sterile stem leaf (mm)	QC	ImageJ
ssl_seg_dist	Distance between the segments of the sterile stem leaf (mm)	QC	ImageJ
fsl_len	Length of the fertile stem leaf (mm)	QC	ImageJ
fsl_pet_len	Length of the petiole of the fertile stem leaf (mm)	QC	ImageJ
fsl_seg_len	Length of the segment of the fertile stem leaf (mm)	QC	ImageJ
fsl_seg_dist	Distance between the segments of the fertile stem leaf (mm)	QC	ImageJ
ssl_hair	Degree of tomentosity of the sterile stem leaf segment (%)	QC	ImageJ
fsl_hair	Degree of tomentosity of the fertile stem leaf segment (%)	QC	ImageJ
Floral head			
cap_diam	Diameter of the floral head involucre (mm)	QC	Caliper
fl_len	Length of the floral tube (mm)	QC	ImageJ
fl_tooth_len	Length of the floral tooth (mm)	QC	ImageJ
sq_ext_len	Length of the external involucral bract (mm)	QC	ImageJ
sq_ext_wid	Width of the external involucral bract (mm)	QC	ImageJ
sq_int_len	Length of the internal involucral bract (mm)	QC	ImageJ
sq_int_wid	Width of the internal involucral bract (mm)	QC	ImageJ
sq_if_len	Length of the inter-floral bract (mm)	QC	ImageJ
sq_if_wid	Width of the inter-floral bract (mm)	QC	ImageJ
sq_if_n_hair	Tomentosity of the inter-floral bract (hairless/slightly pubescent/pubescent/hairy/densely hairy)	QD	ImageJ
sq_ext_hair	Tomentosity of the external involucral bract (hairless/only on the margin/everywhere)	CO	ImageJ
sq_int_hair	Tomentosity of the internal involucral bract (hairless/only on the margin/everywhere)	CO	ImageJ

Table 3. Length, consensus length, and number of informative sites of the six markers in the studied *Santolina* populations. Lengths refers to the ingroup.

Marker	Length (bp)	Consensus length (bp)	Informative Sites	Inferred DNA evolution model
psbM-trnD	714–715	721	7	GTR
rps15-ycf1	500–510	559	9	F81
trnF-trnL	381	386	4	HKY
trnH-psbA	420–429	463	9	HKY+I
trnQ-rps16	885–895	931	14	F81+G
trnS-trnG	420–425	425	2	F81
Concatenated matrix	3331–3336	3485	45	

Table 4. Cypsela morphometric analysis for Santolina native to France and south-eastern Spain. Results of the LDA applied to the current taxonomic hypothesis. Values are percentages.

	S. benthamiana	*S. decumbens*	*S. ericoides*
S. benthamiana	85.5	7.8	6.7
S. decumbens	1.3	97	1.7
S. ericoides	6.8	19.8	73.4

Table 5. Cypsela morphometric analysis for *Santolina* native to France and south-eastern Spain. Results of the LDA applied considering each population as a distinct group. Values are percentages. Ben-LC, *S. benthamiana* from type locality; Ben-rou, *S. benthamiana* from Le Roumenga; Dec-LC, *S. decumbens* from type locality; Dec-lfo, *S. decumbens* from La Fare-les-Oliviers; Dec-sis, *S. decumbens* from Sisteron; Eri-LC, *S. ericoides* from type locality; Eri-fsc, *S. ericoides* from Sant Feliu de Codines; Eri-tor, *S. ericoides* from Torà.

	Ben-LC	Ben-rou	Dec-LC	Dec-lfo	Dec-sis	Eri-LC	Eri-sfc	Eri-tor
Ben-LC	95.9	0	0.9	0	2.9	0.2	0	0.1
Ben-rou	0.3	35.8	12	5.7	8.3	13.3	8.3	16.3
Dec-LC	5.5	13.9	41.7	0.5	22.3	9.6	2	4.4
Dec-lfo	0	1.4	0	86	3.3	0.1	2.3	6.9
Dec-sis	0.2	13.1	23.2	4.4	35.7	13	6.8	3.8
Eri-LC	3.4	12.5	5.8	0.7	12.1	34.6	23	7.9
Eri-sfc	0	11.9	2.5	2.1	5.4	24.3	43.6	10.2
Eri-tor	0	6.5	6.8	12.1	8.9	5.1	8.8	51.7

Table 6. Morphometric analyses for *Santolina* native to France and south-eastern Spain. Results of the Random Forest applied to the current taxonomic hypothesis. Values are percentages.

	S. benthamiana	*S. decumbens*	*S. ericoides*
S. benthamiana	88.3	11.3	0.3
S. decumbens	2.9	97.1	0
S. ericoides	0	0	100

Table 7. Morphometric analysis for *Santolina* native to France and south-eastern Spain. Results of the Random Forest applied at the alternative grouping hypothesis inferred by the molecular phylogenetic analysis. Values are percentages. Ben-Eri, the typical population of *S. benthamiana* + the three populations of *S. ericoides*; Ben-rou, *S. benthamiana* from Le Roumenga; Dec, the three populations of *S. decumbens*.

	Ben-Eri	Ben-rou	Dec
Ben-Eri	99.3	0.7	0
Ben-rou	10.3	68.3	21.4
Dec	0.5	1.7	97.8

Table 8. Morphometric analysis for *Santolina* native to France and south-eastern Spain. Results of the Random Forest applied at the taxonomic hypothesis inferred by the results of the PCoA (Figure 3). Values are percentages. Ben-LC, *S. benthamiana* from type locality; Ben-rou, *S. benthamiana* from Le Roumenga; Dec-LC, *S. decumbens* from type locality; Dec-lfo, *S. decumbens* from La Fare-les-Oliviers; Dec-sis, *S. decumbens* from Sisteron; Eri, the three populations of *S. ericoides*.

	Ben-LC	Ben-rou	Dec-LC	Dec-lfo	Dec-sis	Eri
Ben-LC	95.6	2.1	0	0	0	2.4
Ben-rou	3.5	76.5	5.9	0	14.1	0
Dec-LC	0	0.4	96.6	1.6	1.4	0
Dec-lfo	0	0	0	99.2	0.8	0
Dec-sis	0	7.5	5.9	0.2	86.1	0.3
Eri	0	0	0	0	0	100

Table 9. Mean ± standard deviation values of the quantitative morphological characters for each *Santolina* population. Character codes follow Table 2. Ben-LC, *S. benthamiana* from type locality; Ben-rou, *S. benthamiana* from Le Roumenga; Dec-LC, *S. decumbens* from type locality; Dec-lfo, *S. decumbens* from La Fare-les-Oliviers; Dec-sis, *S. decumbens* from Sisteron; Eri-LC, *S. ericoides* from type locality; Eri-sfc, *S. ericoides* from Sant Feliu de Codines; Eri-tor, *S. ericoides* from Torà.

Character	Ben-LC	Ben-rou	Dec-LC	Dec-lfo	Dec-sis	Eri-LC	Eri-sfc	Eri-tor
fs_len (cm)	22.9 ± 9.2	21.7 ± 6.3	10.8 ± 2.6	8.6 ± 2.9	18.1 ± 4.4	18.6 ± 6.7	18.6 ± 4.1	20.8 ± 5.1
dist_cap_lf (mm)	44.2 ± 26.7	35 ± 12.1	15.3 ± 7.4	10.6 ± 5.4	23.7 ± 11.8	20.4 ± 10.4	29.8 ± 15.7	33.4 ± 16.5
ss_len (cm)	11.8 ± 5.2	16.5 ± 6.4	6.3 ± 2.5	7.7 ± 2.6	10.1 ± 3.0	13.7 ± 6.8	15.3 ± 4.2	17.9 ± 7.5
cap_diam (mm)	6.6 ± 1.3	6.8 ± 1.1	6.9 ± 0.7	6.4 ± 1.0	7.0 ± 0.9	6.2 ± 1.3	7.2 ± 0.8	7.0 ± 0.9
sq_ext_len (mm)	2.8 ± 0.6	3.1 ± 0.5	2.8 ± 0.4	2.7 ± 0.4	3.2 ± 0.4	2.8 ± 0.3	2.8 ± 0.4	3.0 ± 0.4
sq_ext_wid (mm)	1.1 ± 0.2	1.1 ± 0.2	1.1 ± 0.2	1.1 ± 0.2	1.2 ± 0.2	1.2 ± 0.2	1.0 ± 0.2	1.0 ± 0.2
sq_int_len (mm)	2.8 ± 0.5	3.1 ± 0.4	3.1 ± 0.4	2.9 ± 0.3	3.4 ± 0.4	3.0 ± 0.4	3.2 ± 0.3	3.1 ± 0.4
sq_int_wid (mm)	1.2 ± 0.2	1.2 ± 0.2	1.2 ± 0.2	1.1 ± 0.2	1.3 ± 0.2	1.3 ± 0.2	1.3 ± 0.2	1.3 ± 0.1
sq_if_len (mm)	3.0 ± 0.3	3.0 ± 0.4	3.0 ± 0.3	2.8 ± 0.3	3.1 ± 0.2	2.9 ± 0.3	3.0 ± 0.2	2.7 ± 0.2
sq_if_wid (mm)	1.1 ± 0.2	1.0 ± 0.2	1.0 ± 0.2	1.0 ± 0.2	1.1 ± 0.2	1.1 ± 0.2	1.2 ± 0.2	1.1 ± 0.2
fl_len (mm)	3.4 ± 0.4	3.2 ± 0.4	3.0 ± 0.4	3.5 ± 0.5	3.3 ± 0.5	3.4 ± 0.6	3.5 ± 0.6	3.7 ± 0.4
fl_tooth_len (mm)	0.8 ± 0.1	0.7 ± 0.1	0.7 ± 0.1	0.7 ± 0.1	0.7 ± 0.1	0.5 ± 0.1	0.6 ± 0.1	0.6 ± 0.1
ssl_len (mm)	26.8 ± 9.9	21.3 ± 5.9	14.6 ± 3.8	11.4 ± 2.2	23.4 ± 4.1	11 ± 3.1	11.4 ± 2.1	10.6 ± 3.5
ssl_pet_len (mm)	5.7 ± 2.1	4.0 ± 2.3	2.9 ± 1.2	1.3 ± 0.7	3.7 ± 1.8	1.0 ± 0.8	1.6 ± 0.8	0.8 ± 0.4
ssl_seg_len (mm)	4.2 ± 1.5	2.9 ± 1.0	0.9 ± 0.2	0.8 ± 0.2	2.0 ± 0.4	0.9 ± 0.4	1.5 ± 0.4	1.4 ± 0.5
ssl_seg_dist (mm)	1.6 ± 0.7	1.1 ± 0.3	0.3 ± 0.3	0.3 ± 0.2	0.6 ± 0.3	0.5 ± 0.2	0.5 ± 0.3	0.6 ± 0.2
fsl_len (mm)	21.6 ± 8.9	18.9 ± 5.8	12.3 ± 3.2	8.0 ± 1.9	15.4 ± 3.1	10.7 ± 3.8	12.1 ± 2.2	10.3 ± 3.2
fsl_pet_len (mm)	6.4 ± 3.3	3.4 ± 2.4	2.1 ± 0.9	1.1 ± 0.6	2.3 ± 1.2	0.9 ± 0.6	1.8 ± 0.8	0.8 ± 0.5
fsl_seg_len (mm)	2.6 ± 0.9	2.3 ± 0.7	0.8 ± 0.3	0.6 ± 0.1	1.5 ± 0.4	0.9 ± 0.3	1.3 ± 0.4	1.4 ± 0.2
fsl_seg_dist (mm)	1.6 ± 0.9	1.0 ± 0.3	0.3 ± 0.3	0.3 ± 0.1	0.6 ± 0.2	0.6 ± 0.3	0.6 ± 0.2	0.6 ± 0.2
fs_n_nodes	21.0 ± 3.3	22.0 ± 4.3	19.8 ± 3.5	17.4 ± 4	19.4 ± 3.6	24.9 ± 6.6	26.2 ± 3.7	22.4 ± 4
ss_n_nodes	20.7 ± 4.7	23 ± 6.5	14.8 ± 3.4	20.0 ± 4.1	16.8 ± 2.7	24.6 ± 7.7	28.0 ± 5.9	26.8 ± 7.3
ssl_n_seg	29.7 ± 8.8	41.5 ± 8.9	60.5 ± 12.4	61.8 ± 13.7	60.7 ± 13.3	46.7 ± 10.6	35.4 ± 6.6	33.2 ± 4.3
fsl_n_seg	19.4 ± 5.4	31.8 ± 8	40.4 ± 9.4	38.8 ± 9.1	39 ± 9.7	40.5 ± 9.2	31.9 ± 8.3	31.6 ± 6.2
ssl_hair (%)	47.0 ± 22.0	76.1 ± 7.1	90.8 ± 7.5	61.4 ± 14.3	79.4 ± 9.9	23.8 ± 14.5	22.9 ± 14.0	13.6 ± 18.8
fsl_hair (%)	19.6 ± 18.5	67.3 ± 15.1	90.0 ± 6.2	29.1 ± 19.8	44.0 ± 18.8	15.2 ± 11.7	18.1 ± 10.9	11.3 ± 11.5
fs_hair (%)	15.2 ± 8.2	60.9 ± 14.8	90.3 ± 9.2	60.2 ± 14.0	62.9 ± 16.3	31.8 ± 12.2	31.8 ± 18.9	36.0 ± 21.2

Table 10. Pairwise comparisons of the morphological characters for which we found statistically significant differences between population pairs in *Santolina* native to France and south-eastern Spain. In the lower triangle of the table, the characters that are significantly different with Cohen's d > 1.2 are reported; in the upper triangle, the number of these characters is reported. In the following list, in the brackets, the number of times that a character occurs in the table is reported; a, fs_len (12); b, br_ratio (0); c, dist_cap_lf (10); d, ss_len (12); e, cap_diam (0); f, sq_ext_len (2); g, sq_ext_wid (1); h, sq_int_len (4); i, sq_int_wid (0); j, sq_if_len (3); k, sq_if_wid (0); l, fl_len (2); m, fl_tooth_len (8); n, ssl_len (15); o, ssl_pet_len (18); p, ssl_seg_len (19); q, ssl_seg_dist (12); r, fsl_len (16); s, fsl_pet_len (15); t, fsl_seg_len (22); u, fsl_seg_dist (15); v, fs_n_br (1); w, fs_n_nodes (5); x, ss_n_nodes (12); y, ssl_n_seg (17); z, fsl_n_seg (7); A, ssl_hair (24); B, fsl_hair (17); C, fs_hair (22). Populations acronyms are: Ben-LC, *S. benthamiana* from type locality; Ben-rou, *S. benthamiana* from Le Roumenga; Dec-LC, *S. decumbens* from type locality; Dec-lfo, *S. decumbens* from La Fare-les-Oliviers; Dec-sis, *S. decumbens* from Sisteron; Eri-LC, *S. ericoides* from type locality; Eri-sfc, *S. ericoides* from Sant Feliu de Codines; Eri-tor, *S. ericoides* from Torà.

	Ben-LC	Ben-rou	Dec-LC	Dec-lfo	Dec-sis	Eri-LC	Eri-sfc	Eri-tor
Ben-LC	-	5	17	13	11	13	13	13
Ben-rou	y,z,A,B,C	-	14	14	6	13	10	12
Dec-LC	a,c,d,n,o,p,q,r,s,t,u,x,y,z,A,B,C	a,c,d,n,p,q,r,t,u,x,y,A,B,C	-	7	8	9	11	13
Dec-lfo	a,c,n,o,p,q,r,s,t,u,y,z,C	a,c,d,n,o,p,q,r,s,t,u,y,A,B	o,r,s,x,A,B,C	-	10	8	13	9
Dec-sis	f,h,o,p,q,s,t,u,x,y,z,A,B,C	d,q,t,u,y,B	a,d,n,p,t,A,B,C	a,c,h,n,o,p,r,t,u,A	-	12	11	11
Eri-LC	c,m,n,o,p,q,r,s,t,u,y,z,A,C	c,m,n,o,p,q,r,s,t,A,B,C	a,d,m,o,s,x,A,B,C	a,m,t,u,w,y,A,C	f,m,n,o,p,r,s,t,x,A,B,C	-	4	2
Eri-sfc	h,m,n,o,p,q,r,s,t,u,w,x,z,A,C	n,o,p,q,r,t,u,A,B,C	a,d,o,p,t,w,x,y,A,B,C	a,c,d,h,p,r,t,u,w,x,y,A,C	d,g,n,o,r,w,x,y,A,B,C	m,p,t,y	-	2
Eri-tor	j,m,n,o,p,q,r,s,t,u,z,A,C	l,n,o,p,q,r,s,t,u,A,B,C	a,c,d,l,o,p,s,t,x,y,A,B,C	a,c,d,p,t,u,y,A,C	d,j,n,o,r,s,x,y,A,B,C	t,y	j,s	-

Table 11. Results of niche similarity tests for *Santolina* native to France and south-eastern Spain. Environmental spaces were obtained from the distribution of the morphotypes following the taxonomic hypothesis inferred by the results of the PCoA (Figure 3). Backgrounds were defined by applying 5 km buffer zones around the occurrence points. The asterisks indicate a significant difference: *P ≤ 0.05; ^{ns}P ≥ 0.05. Ben-LC, *S. benthamiana* s.str.; Ben-rou, "Le Roumenga" morphotype; Dec-LC, *S. decumbens* s.str.; Dec-lfo, "La Fare-les-Oliviers" morphotype; Dec-sis, "Sisteron" morphotype; Eri, *S. ericoides*.

Morphotype 1	Morphotype 2	D
Ben-LC	Ben-rou	0.379 $^{ns/ns}$
Ben-LC	Dec-LC	0.000 $^{ns/ns}$
Ben-LC	Dec-lfo	0.000 $^{ns/ns}$
Ben-LC	Dec-sis	0.000$^{ns/ns}$
Ben-LC	Eri	0.167 $^{ns/ns}$
Ben-rou	Dec-LC	0.048 $^{ns/ns}$
Ben-rou	Dec-lfo	0.184 $^{*/ns}$
Ben-rou	Dec-sis	0.178 $^{ns/ns}$
Ben-rou	Eri	0.212 $^{ns/*}$
Dec-LC	Dec-lfo	0.017 $^{ns/ns}$
Dec-LC	Dec-sis	0.013 $^{ns/ns}$
Dec-LC	Eri	0.204 $^{ns/*}$
Dec-lfo	Dec-sis	0.000 $^{ns/ns}$
Dec-lfo	Eri	0.023 $^{ns/*}$
Dec-sis	Eri	0.190 $^{ns/*}$

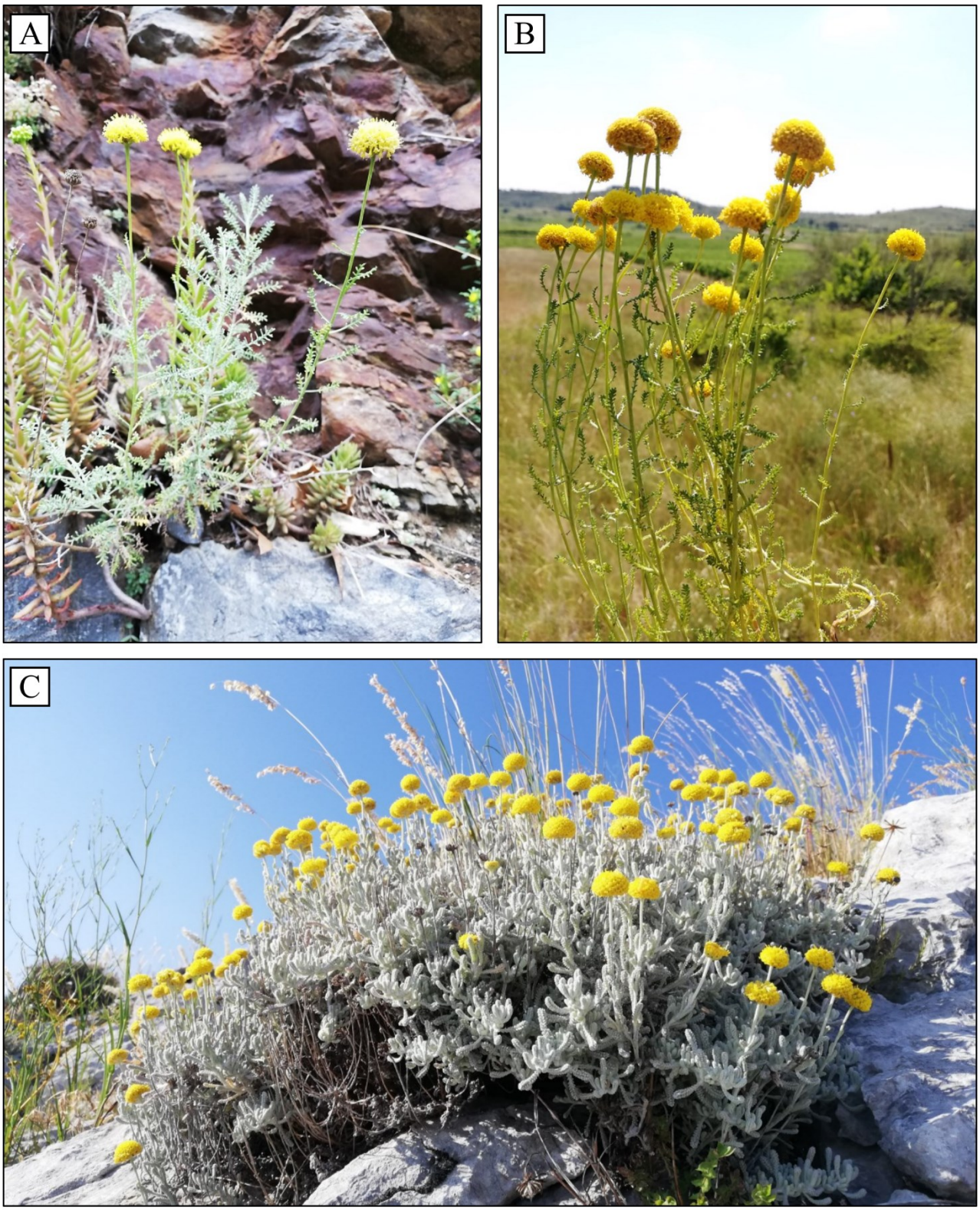

Figure 1. The three *Santolina* species currently recorded for southern France and north-eastern Spain from their type populations. *Santolina benthamiana* from Prats-de-Mollo-la Preste, June 29, 2020 (**A**); *S. ericoides* from Béziers, June 28, 2020 (**B**); *S. decumbens* from Monte Caume, June 27, 2020 (**C**). All photos by L. Peruzzi.

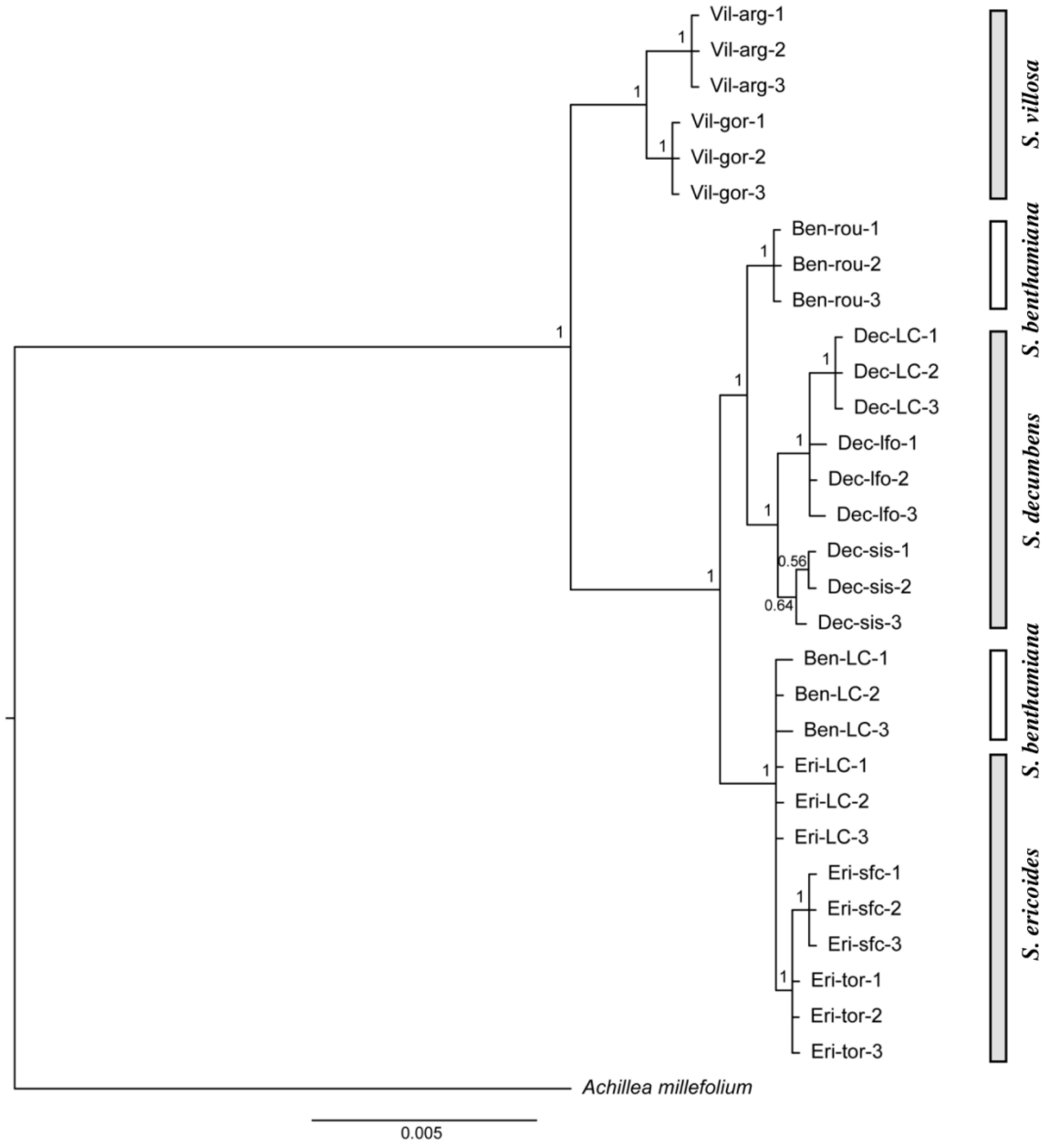

Figure 2. Bayesian consensus tree of plastid dataset for *Santolina* native to France and north-eastern Spain. LC, type locality.

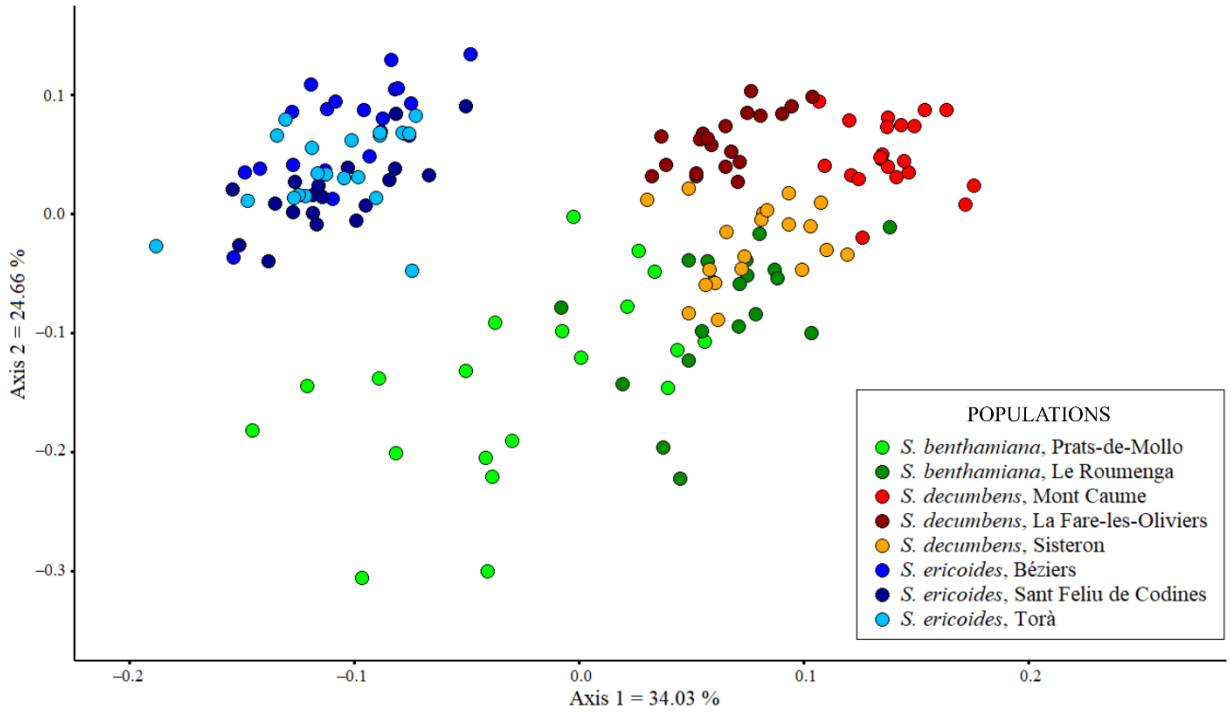

Figure 3. Morphometric analysis for *Santolina* native to France and south-eastern Spain. Scatter plot of the first two axes of the PCoA based on Gower distance. The taxonomic assignment of populations follows Giacò et al. (2021).

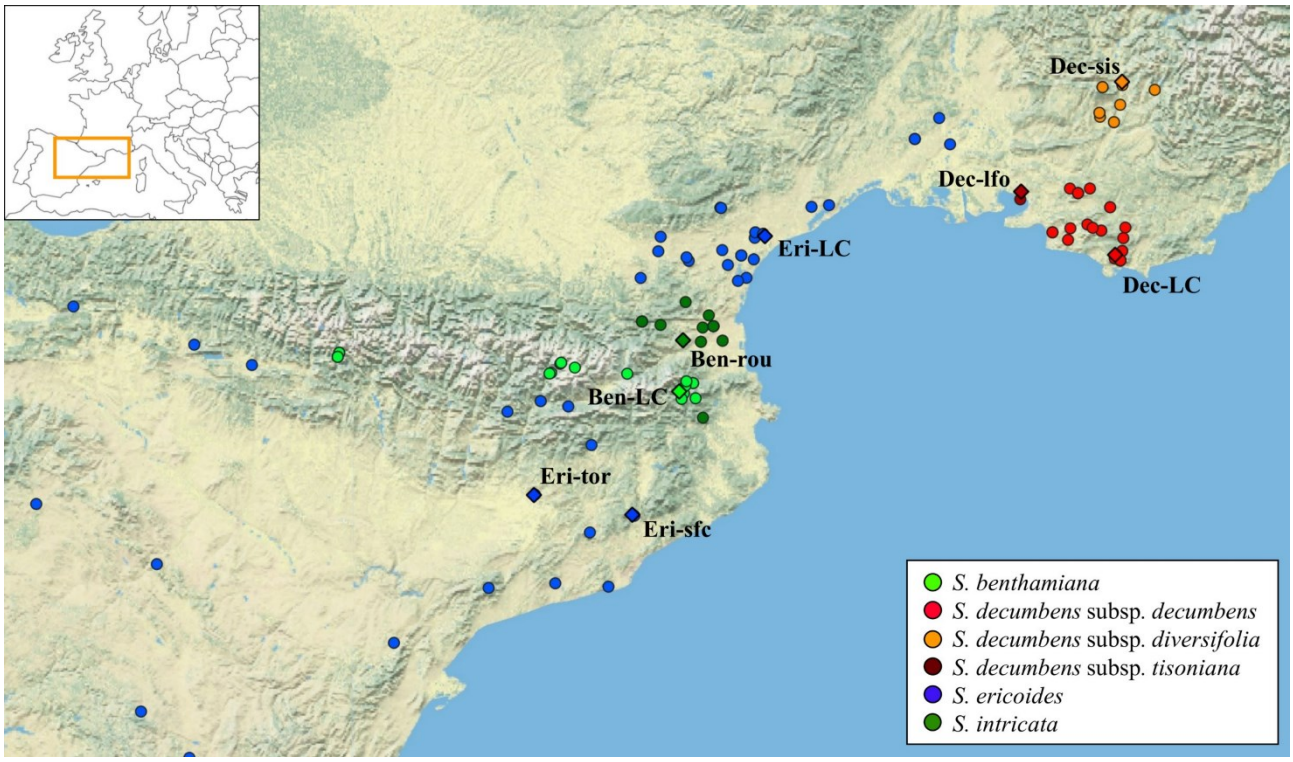

Figure 4. Distribution of the studied *Santolina* taxa native to France and south-eastern Spain, based on 163 herbarium specimens seen (Table S5). Diamonds correspond to the sampled populations: Ben-LC = Prats-de-Mollo-la-Preste, Ben-rou = Le Roumenga, Dec-LC = Mont Caume, Dec-lfo = La Fare-les-Oliviers, Dec-sis = Sisteron, Eri-LC = Béziers, Eri-sfc = Sant Feliu de Codines, Eri-tor = Torà. The taxonomic assignment of populations follows the final treatment presented in this study.