# UNIVERSITY OF SASSARI

DISSERTATION FOR THE DEGREE OF DOCTOR OF ENVIRONMENTAL BIOLOGY
PRESENTED AT SASSARI IN 2007

# "Population genetic analysis of the endemic Centaurea spp. in Sardinia


*Candidate:*

**Giulia Mameli**

*Tutor:*

**Prof. Rossella Filigheddu**

*Supervisor:*

**Prof. Marco Apollonio**

*Co-tutor*

**Prof. Giorgio Binelli**

**Dr. Nuria Garcia-Jacas**

# “Population genetic analysis of the endemic *Centaurea spp*. in Sardinia”

## Index

## *Introduction*

Plant population size varies in space and time both within and among species. This variability is the result of complex interactions among the life-history features of populations, local environmental conditions and the historical ecology of particular species (Barrett & Kohn, 1991). Genetic diversity reflects the differences among individuals for many characters and represents the variety of alleles and genotypes present in the group under study (population, species or group of species).

Mutation, migration, selection and change determine evolution in both small and large populations. These factors interact to produce the different levels of genetic diversity. Such patterns could have a profound influence on the genetic dynamics of threatened populations and suggest that theoretical models based entirely on random mating need to be revisited. The key question is whether these trends reflect what actually happens in nature (Amos & Balmford, 2001).

Small or declining populations of threatened and endangered species are more prone to extinction than large stable populations. The total genetic diversity of a species is a key factor in its persistence and conservation.

Maintenance of genetic diversity is a major objective in conservation programs, as genetic diversity represents evolutionary potential, because it is the raw material for adaptive evolutionary change. Loss of genetic diversity in small populations reduces the ability to evolve in response to ever-present environmental change. The importance of genetic diversity over the long term (maintenance of adaptive evolutionary potential) as well as the short term (maintenance of reproductive fitness) makes it a primary focus for conservation genetics. Conservation biologists need to understand how genetic diversity is maintained through natural processes if conservation programs are to be designed for its maintenance in managed populations of endangered species (Frankham *et al*., 2004).
The perspective importance of genetic problems in the conservation of endangered species has fluctuated considerably over the last two decades and remains the subject of debate (Amos & Balmford, 2001). The significance of genetic variation as one of several currencies for biodiversity evaluation is widely recognized (Humphries *et al*., 1995) and protection of genetic diversity is incorporated into many national and international conventions.

Insular plant populations are mainly prone to an extinction risk: 82% of the populations from 202 islands has a lower genetic diversity level versus populations coming from mainland (Frankham, 1998).

The Mediterranean is the largest inland sea in the world. The areas of the Mediterranean Basin have recognized as 'hotspots of biodiversity' for the immense wealth of the plant species (Myers *et al*., 2000). The Mediterranean region is an ideal place to study plant endemism. The basin's location at the intersection of two major landmasses, Eurasia and Africa, has contributed to its high diversity. Furthermore, many of endemic plant species in the basin are narrow endemic: they are confined to very small areas and thus very extremely vulnerable to habitat loss, overgrazing and urban expansion. Indeed, it is likely that more plant species have gone extinct here than in any other hotspot Endemic plants are mainly concentrated on islands, peninsulas, rocky cliffs and mountain peaks. Médail & Quézel (1997), proposed the delimitation of 10 biodiversity hotspots within the Mediterranean basin. Tyrrhenian Islands are one of these Mediterranean hotspots (Médail & Quézel, 1999). Within the Mediterranean basin the Sardinian-Corsican system shows one of the highest densities of endemic plant species, therefore it is so original in terms of vegetation cover, land use and landscape, that a biogeographic autonomy as a province can be easily justified (Arrigoni, 1983; Contandriopoulos, 1981).

Comparison and interpretation of the degree of endemism is particularly difficult owing to the wide disparities between the regions considered (territory origin, geographic situations, area and different altitudes) and the selection of endemism (Mèdail & Verlaque, 1997). In the Mediterranean context, the important south-east France and Corsica endemisms result from the very disturbed history (tectonic, geological and climatic) since the middle Tertiary. Due to the moderate direct impact of the Quaternary glaciations, especially the Würm, several zones have acted as refugia. As a consequence, some genera are limited to contemporary areas associated with ancient plates and some have greatly diversified within the limits of the zone. For example, the Iberian peninsula has 16 paleo-endemic genera and is also centre of diversification for many genera (e.g. *Genista*, *Thymus*, *Teucrium*, *Linaria*, *Narcissus*).

In spite of this only limited information is available on the genetic structure of endemic Mediterranean plant species (for review, see Thompson, 1999).

The occurrence of high numbers of endemic species, particularly on islands and in mountain ranges in the Mediterranean region, attests to the high levels of geographic differentiation that occur in its flora. Many species have a disjunct distribution such that

geographically isolated populations may also exhibit high levels of differentiation (Quilichini *et al.*, 2004).

The extent to which such differences among populations compares to differences among what are suggested to be different but closely related endemic taxa in the Mediterranean flora, is an issue which has recently attracted attention (Debussche & Thompson, 2002). This issue is particularly important in order to identify and delimit taxa which merit conservation status (Olfelt *et al.*, 2001). In fact, for only a few endemic and protected species do we have information concerning levels of population differentiation (Affre & Thompson, 1997).

The island of Sardinia has a consistent richness of endemic plants evolved as a result of its geological history (Thompson, 2005). Several species are intuitively known as palaeo-endemics (Arrigoni, 1976) because the island could have played a significant role during the last glacial maximum, and as schizo-endemics because a great number of endemic species could be evolved after the actual separation of Sardinia from the mainland and from Corsica, finished 20,000 years ago. On 347 endemic species 26.2% are in common to both islands whereas 45.8 % are exclusive to Sardinia (Bacchetta *et al.*, 2005). Among these, five species of the *Centaurea* genus are present: *C. horrida, C. filiformis, C. corensis, C. ferulacea* and *C. magistrorum*.

Following Heywood (1960) microendemic vicariants is a term used for those groups of endemic plants whose parentage is obvious and which are specially rather than genetically isolated. In these plants, morphological differentiation is usually weak and the groups are not widely separated geographically. Populations have been fragmented into discrete units (e.g., on separate mountain peaks or mountain ranges), and often the morphological differences between taxa, although small, are constant. As some cases of microendemic, presumably schizo-endemic, species are present within the *Centaurea* genus in western Mediterranean (Suárez-Santiago *et al.* 2007), we would here assess the genetic variability of Sardinian endemic *Centaurea* species and their taxonomical relationships.

## References:

**Affre L., Thompson J., 1997**. Variation in the population genetic structure of two *Cyclamen* species on the islnd of Corsica. *Heredity* 78: 205-214.

**Amos W. & Balmford A., 2001**When doers conservation genetics matter *Heredity*. **87**: 257-265..

**Arrigoni PV. 1976**. Le piante endemiche della Sardegna. Introduzione. *Bollettino della Società Sarda di Scienze Naturali* **16**: 259-264.

**Bacchetta G, Iiriti G, Pontecorvo C. 2005**. Contributo alla conoscenza della Flora vascolare endemica della Sardegna. *Informatore Botanico Italiano* **37(1, parte A):** 306-307.

**Barrett C.H.S. and Kohn J**., **1991**. Genetic and Evolutionary Consequences of Small Population Size in Plant: Implication for Conservation. Genetic and Conservation of Rare Plants. Edited by Donald A. I. Falk & Kent E. Holsinger. Oxford University Press.

**Debussche M., Thompson J. D., 2002.** Morphological differentiation among closely related species with disjunct distributions: a case study of Mediterranean *Cyclamen* L. subgenus *Psilanthum* (Primulaceae). *Botanical Journal of the Linnaean Society*. **132**: 133-144.

**Frankham R., 1998**. Inbreeding and extinction: island population. *Conservation Biology*. **12**: 665-675.

**Frankham R., Ballou J.D. and Broscoe D.A., 2004**. Introduction to conservation genetics.. Cambrige University Press.

**Heywood V. H. 1960**. Problems of geographical distribution and taxonomy in the Iberian Peninsula. Feddes Repert. Spec. Nov.Regni Veg. **63:** 160-168.

**Humphries C. J., Williams P. H., Wright Vane R: I:., 1995.** Measuring biodiversity value for conservation. ***Annuals review of ecology and systematic* 26:**93-111

**Médail F., Quézel P.,** 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. Annals of the Missouri Botanical Garden, 84: 112-127.

**Médail F, Quezel P. 1999**. Biodiversity hot-spots in the Mediterranean Basin: setting global 25 conservation priorities. *Conservation Biology* **13**: 1510-1513.

**Myers N., Mittermeier R.A., Mittermeier C.G., Da Fonseca G.A.B., Kents J.**, 2000. Biodiversity hotspots for conservation priorities. Nature, 403: 853-858.

**Olfelt JP, Furnier GR, Luby JJ. 2001**. What data determine whether a plant taxon is distinct enough to merit legal protection? A case study of *Sedum integrifolium* (Crassulaceae). *American Journal of Botany* **88**: 401–410.

**Quilichini A., Debussche M., Thompson J.D., 2004**. Geographic differentiation in the Mediterranean island endemic *Anchusa crispa*: implication for the conservation of a protected species. *Biological Conservation*. **118**: 651-660.

**Suárez-Santiago, V.N., Salinas, M.J., Garcia-Jacas, N., Soltis, P.S., Soltis, D.E., Blanca, G., 2007**. Reticulate evolution in the *Acrolophus* subgroup (*Centaurea* L., Compositae) from the western Mediterranean: origin and diversification of section *Willkommia* Blanca. *Molecular and Phylogenetic Evolution.* 43, 156–172.

**Thompson, J.D.,** 2005. *Plant Evolution in the Mediterranean*. Oxford: Oxford University Press.

# The Genetic Structure of the Remnant Populations of *Centaurea horrida* Badarò in Sardinia, a major island of the Mediterranean Sea

MAMELI G.[1], FILIGHEDDU R.[1] , BINELLI G.[2] and MELONI M.[2,3]

[1]*Dipartimento di Botanica ed Ecologia vegetale, Università di Sassari, via F. Muroni 25, I-07100 Sassari, Italy.* [2]*Dipartimento di Biotecnologie e Scienze molecolari, Università dell'Insubria, via J.H. Dunant 3, I-21100, Varese, Italy.* [3]*School of Biotechnology and Biomolecular Science, University of New South Wales, Sydney 2052, Australia.*

**ANNALS OF BOTANY** in press.

# The Genetic Structure of the Remnant Populations of *Centaurea horrida* Badarò in Sardinia and Associated Islands

GIULIA MAMELI [1], ROSSELLA FILIGHEDDU [1], GIORGIO BINELLI [2,*] and MARILENA MELONI [2,3]

[1]*Dipartimento di Botanica ed Ecologia Vegetale, Università degli Studi di Sassari, via Muroni 25, 07100 Sassari, Italy and* [2]*Dipartimento di Biotecnologie e Scienze Molecolari, Università degli Studi dell'Insubria, via J.H. Dunant 3, 21100 Varese, Italy and* [3] *School of Biotechnology and Biomolecular Science, University of New South Wales, Sydney 2052, Australia*

**Running Title:** Genetic structure of *Centaurea horrida*

***For correspondence:** e-mail giorgio.binelli@uninsubria.it

**Abstract**

- *Background and Aims* The Mediterranean region is of prime importance to biodiversity at a global level, mainly due to the abundance of endemic plant species. However, information about these species is still scarce, especially at the genetic level. In this paper we report the first assessment of the genetic structure of *Centaurea horrida* Badarò (Asteraceae), an endemic, sea cliff-dwelling plant from Sardinia.
- *Methods* The study was conducted on seven populations covering the entire natural range of the species, by means of SSR (microsatellite) markers.
- *Key Results* A considerable amount of genetic variation was found (average *He* = 0.603 - 0.854), together with a medium-high differentiation among populations, as estimated both by $F_{ST}$ (0.123) and $R_{ST}$ (0.158). Both Bayesian analysis and AMOVA were employed to detect genetic structuring in this species. The results suggest that the origins of the current populations of *C. horrida* lie in two gene pools.
- *Conclusions* Despite the restricted range, *C. horrida* displays high levels of genetic diversity, structured in such a way that three management units could be deemed viable for its conservation. The protected status of the species will probably suffice to prevent the impoverishment of its genetic resources.

**Key words**: Genetic diversity, *Centaurea horrida*, endangered species, narrow endemic, conservation, Mediterranean, Sardinia.

## INTRODUCTION

The Mediterranean Basin displays such an abundance of endemic species (about 13,000 species, corresponding to 4.3% of the plant species described worldwide) that it may be considered one of the biodiversity hotspots at a global level (Myers *et al*., 2000). At the local level, ten more biodiversity hotspots have been recognised (Médail and Quézel, 1999) in the Mediterranean region. In particular, the larger islands may have played a key role in the conservation of mid-tertiary floras (Greuter, 1995). Many species are characterised by disjoint distributions, thus leading to high levels of differentiation between geographically isolated populations (Quilichini *et al*., 2004). The extent of genetic differentiation among con-specific populations relative to the extent of differentiation present between closely related endemic taxa in the Mediterranean flora is an issue which has recently attracted attention (Debussche and Thompson, 2002). This issue is of particular importance for species delimitation in biodiversity inventories and in order to identify and delimit taxa for specific conservation measures (Olfelt *et al*., 2001).

Studies on the amount and distribution of genetic diversity of endemic and protected Mediterranean plant species are still scarce and very limited information is available on their genetic structure. Many species live in harsh environments, such as cliffs and steep slopes characterised by the presence of drought and wind, thus forming patches that are isolated by other environments which they cannot populate because of their limited dispersal ability. This is the case for *Centaurea horrida* and other plants of the Asteraceae family, which have recently been studied: the congener *Centaurea corymbosa* has a very low colonizing ability and survives in six small populations (Freville *et al*., 2001), while *Femeniasia balearica* (formerly *C. balearica*) now lives in a very restricted habitat (Vilatersana *et al*., 2007). Both have been analysed for their genetic composition by means of allozymes, microsatellites and AFLP and display quite high levels of genetic variation and genetic differentiation between

populations. In this paper we undertake the first study of the genetic structure of the remaining populations of the endangered species *Centaurea horrida* Badarò, by means of microsatellite markers. *Centaurea horrida* (Fig. 1) is a long-living spinous dwarf scrub that grows to a height of 70 cm (Valsecchi, 1977). Its distribution is limited to sea-cliffs in islands and peninsulas where it forms patches of isolated populations, both in primary and secondary dwarf communities (Desole, 1956; Valsecchi, 1977). *Centaurea horrida* is a diploid taxon with $2n = 18$ (Desole, 1954), that reproduces sexually, by way of cross-pollination carried out by insects. It flowers in late spring (April-May) and bears fruit in summer (July-August) (Pisanu, 2007). It is a protected species according to the Berne Convention (Appendix I), a priority species according to the EU Directive 43/92 "Habitat" (Annex II) and a vulnerable species according to the 1997 IUCN Red List of threatened plants. It is thus of importance to assess the amount of genetic variation available to the species and to suggest possible guidelines for conservation.

## MATERIALS AND METHODS

### Plant material

The distribution range of *Centaurea horrida* is highly fragmented and consists of only four coastal locations, from North-West to North-East Sardinia (Western Mediterranean), the characteristics of which are reported in Table 1; its geographical position is displayed in Figure 2. The study was conducted on two populations from the island of Asinara (FOR and STR), two from Stintino (FAL and DON), two from Alghero (LIO and BAR) and one from Tavolara (TAV), the latter consisting of the total of the plants living on Tavolara island.

Samples of fresh leaves were collected from a total of 385 individuals (Table 1) throughout the seven populations studied, and were stored at -80°C until DNA extraction. Total DNA was extracted by grinding the frozen leaves in a mortar in liquid $N_2$ and by using the DNeasy

Plant Mini Kit (Qiagen, Italy), according to the manufacturer's instructions. The average concentration of the extracted DNA was 20 ng/μL.

**Amplification conditions**

Simple Sequence Repeat (SSR) primers from *Centaurea corymbosa* (Freville *et al*., 2000) were tested for their ability to amplify single genomic regions in *Centaurea horrida*. Four out of seven were selected because they yielded an unambiguous amplification pattern. The SSRs chosen, their primer sequences and the fluorophore used are listed in Table 2.

Amplification reactions were modified with respect to Freville *et al*., 2000. They were performed in a total volume of 15 μL, containing HotMasterTaq (Eppendorf®) buffer 1X, 2.5 mM $MgCl_2$, 2 μM of each dNTP, 0.5 μM of each forward and reverse primer, 25 ng genomic DNA and one unit of *Taq* polymerase HotMasterTaq (Eppendorf®). Amplification was carried out in a PTC-100 thermal cycler (MJ Research, Watertown, MA, USA), under the following conditions: an initial cycle at 94°C for 2 min, followed by 30 amplification cycles, at 94°C for 1 min, annealing temperature (Ta; Table 2) for 30 s, 65°C for 1 min and a final step of extension at 65°C for 5 min.

The amplification products were run on a capillary MegaBACE® DNA sequencer (Amersham). The raw data were analysed using allied MegaBACE Fragment Profiler software, to score the single-plant genotypes.

**Data Analysis**

Allele frequencies and observed and expected heterozygosities were estimated at each locus for all populations. Fisher's exact test using the Markov Chain algorithm (Guo and Thompson, 1992) was used to assess deviations from the Hardy-Weinberg equilibrium for each population and each locus. Genotypic disequilibrium between pairs of loci was tested at the single population level by Fisher's exact test. Weir and Cockerham's (1984) estimators of *F*-statistics were used to analyse genetic diversity both within and between populations. In

particular, $F_{IS}$ was calculated in order to estimate which proportion of the total genetic variation was due to a departure from the Hardy-Weinberg equilibrium at the population level. $F_{ST}$ was calculated in order to estimate the proportion of the total genetic variation due to differentiation between populations. $F_{ST}$ was also used to estimate gene flow by calculating the number of migrants per generation ($N$m). The $F_{ST}$ analogue for microsatellites $R_{ST}$ (Slatkin, 1995) was also used, so as to include molecular information relating to the size of differences between the alleles in the differentiation estimates. The statistical methods implemented by BOTTLENECK (Piry *et al*., 1999) were used for detecting genetic bottlenecks in our populations either under the infinite allele model (IAM) or the stepwise mutation model (SMM). The Two-phased model of mutation (TPM) was also tested, because most microsatellite data better fit the TPM than the SMM or IAM. The TPM is intermediate to the SMM and IAM.

A Mantel (1967) test was applied to the matrices of pairwise $F_{ST}/(1- F_{ST})$ and log-transformed geographical distances between populations to assess isolation-by-distance, i.e. the presence of migration-drift equilibrium between populations.

Analysis of molecular variance (AMOVA) was performed to partition the total genetic variation among regions and between populations within regions (Excoffier *et al*. 1992). The test of significance for the AMOVA was carried out on 1000 permutations of the data.

The problem of inferring the number $K$ of clusters present in a data set has been addressed by Pritchard and colleagues (2000) by using the Bayesian paradigm and *ad hoc* software called STRUCTURE. They placed a prior distribution on $K$ and based inference for $K$ on the posterior distribution Pr $(X|K)$ = Pr $(K|X)$ Pr $(K)$, where $X$ is the multilocus genotype of individuals. More recently, it has been suggested that a better estimator of $K$ is the modal value of $\Delta K$ (Evanno *et al*., 2005), the second-order rate of change of the likelihood function with respect to $K$. The latter approach was used in our work to estimate $K$. The analysis was based on the

admixture model, correlated allele frequencies between populations, and was run with a length of burn-in period of $10^5$ and the same number of MCMC replications. Twenty runs were carried out for each *K* value from 1 to 10 (the number of real populations plus three) tested.

The software packages used to analyse the genetic data were GENEPOP (Raymond and Rousset, 1995), GENETIX (Belkhir *et al*., 1996), BOTTLENECK (Piry *et al*., 1999), GenAlEx v.6 (Peakall and Smouse, 1996-2001), RST CALC (Goodman, 1997) and STRUCTURE 2.1 (Pritchard *et al*., 2000).

## RESULTS

### Genetic variability

A total of 385 plants of *Centaurea horrida* were analysed using four microsatellite markers, identifying a total of 80 alleles. All the loci studied are highly polymorphic: the number of detected alleles per locus across all the populations ranged from 15 (locus *21D9*) to 25 (locus *13D10*). There were no indications for null alleles at any of the loci. No alleles were found fixed at any of the loci; neither was evidence found that a given population harboured specific alleles.

Genetic diversity (Table 3) was measured using Nei's heterozygosity (*He*) and ranged from 0.449 (locus *21D9*, TAV population) to 0.925 (locus *13D10*, DON population). The high estimates of genetic variability are confirmed by the average *He* values, ranging from 0.603 (LIO) to 0.854 (FAL and DON). These values are higher for the populations of the Stintino – Asinara region than for the two populations of the Alghero region and the isolated population of Tavolara.

The Hardy-Weinberg equilibrium was tested for all the loci and populations by testing the departure of $F_{IS}$ from zero under the null hypothesis. $F_{IS}$ values are significantly different from zero for all the loci except locus *28A7* for the STR, FAL and DON populations, locus

*12B1* for the FOR and LIO populations and locus *13D10* for the BAR and TAV populations. In the vast majority of cases, deviation from the Hardy-Weinberg equilibrium was associated with positive $F_{IS}$ values, while negative $F_{IS}$ values were mainly associated with the locus *28A7* (four populations).

**Genotypic disequilibrium**

The non-random association of the alleles at different loci, or linkage disequilibrium (LD), was investigated. A significant departure from equilibrium at the 5% level was found for almost all pairs of loci within population. Only five comparisons out of 42 were not significant, for the pairs of loci *21D9* - *13D10* (LIO), *21D9* - *28A7* (LIO and FOR) and *28A7* - *12B1* (DON and FOR).

**Genetic differentiation among populations**

The genetic divergence among populations was measured using both $F_{ST}$ and $R_{ST}$ (Table 4). Their significance was tested by a permutation procedure: all $F_{ST}$ and $R_{ST}$ values differed significantly from zero. The maximum $F_{ST}$ value was found between the LIO and TAV populations and the maximum $R_{ST}$ value between the BAR and TAV populations. It is to be noted that the pairwise $R_{ST}$ values are constantly higher than the respective $F_{ST}$ values, with the exception of the values relating to the LIO population.

The overall genetic differentiation between populations was significant. By means of $F_{ST}$ = 0.123 (confidence interval at 95% results in $0.072 \leq F_{ST} \leq 0.178$) we estimated that more than 12% of the genetic variance can be attributed to differentiation between populations. The same procedures for $R_{ST}$ yielded an estimated overall $R_{ST} = 0.158$, with a confidence interval at 95% of $0.137 \leq R_{ST} \leq 0.196$.

**Isolation by distance**

The presence of correlation between genetic differentiation (estimated as $F_{ST}/1 - F_{ST}$) and geographic distance (log km) between populations was demonstrated by a Mantel test ($p$ =

0.004, G = 2.41, Z = 10.6), indicating that the present distribution of genetic variation among the remnant populations of *Centaurea horrida* is, at least in part, the result of an equilibrium between drift and gene flow. Gene flow was estimated on the basis of either $F_{ST}$ or $R_{ST}$. The maximum value of *Nm* was 8.37 (populations FAL and DON), whereas the minimum value was 1.33, (populations LIO and TAV).

Under the assumption of drift-gene flow equilibrium, the distribution of the expected heterozygosities was compared to the Hardy-Weinberg heterozygosity for each locus and for all populations, to identify those populations which could have experienced a reduction of $N_e$ in recent times. Of the three statistical methods used by the BOTTLENECK software, sign test, Wilcoxon test and standardized differences test, the latter was not employed, because it requires at least 20 polymorphic loci to be reliable. Even so, the four polymorphic SSRs do not guarantee high statistical power. The presence of genetic bottlenecks was tested under the IAM, the SMM and the TPM models of evolution. In neither case we found evidence of a recent (within approx. the past $2N_e$ - $4N_e$ generations) bottleneck.

**Analysis of the population structure**

Since we were dealing with a rare and endangered species, it was of paramount importance to estimate *K*, the most probable number of 'genetic units' or 'gene pools' present in the data, in order to be able to suggest possible mechanisms that have shaped their genetic variability, and to reach conservation recommendations. This was done by applying the Bayesian clustering method as implemented by STRUCTURE (Pritchard *et al*., 2000). The estimate of *K* was based on $\Delta K$, the second-order rate of change of the likelihood function with respect to *K*, as suggested by Evanno *et al*. (2005). We found a sharp signal at $K = 2$ (Table 1SM) **[Supplementary Information]**, therefore suggesting that two homogeneous gene pools shaped the genetic structure of the populations analysed. To check the composition of each individual population and each plant with respect to the inferred populations, further analysis

was conducted based on $K = 2$. The results are shown for the populations in Figure 3. Analysis of the genetic components of the populations shows that the STR, FOR, FAL, DON and TAV populations derive the major component of their genetic composition from the first inferred population and the LIO and BAR populations from the second. Quantitative analysis of this process is shown also in Figure 1SM **[Supplementary Information]**, where the contribution of the two inferred gene pools is reported in graphical form for each of the plants analysed.

**AMOVA**

The total amount of genetic variation was also partitioned by AMOVA into components according to the geographic subdivision of the populations. First, based upon the analysis of the population structure, the hypothesis that the populations fall into two geographic regions was tested, separating the Alghero area from the rest of the range. The AMOVA results (Table 5*a*) show that the within population component accounts for 82% of the total variance and that both the differences between regions and the differences between populations within a region account for smaller, but significant, amounts of the total genetic variation. Second, we tested the hypothesis that all *three* geographic areas (Fig. 1) harbour significant amounts of variation. This partitioning of the data revealed that 10% of the genetic variance resided between regions and 7% between populations within regions (Table 5*b*).

**DISCUSSION**

**Genetic variability**

*C. horrida* is the only species belonging to the *Horridae* section of subgenus *Acrolophus* (Dostál, 1976), which previously included also *C. balearica* now re-classified as *Femeniasia balearica* Susanna. Today, this species is rare and survives only in a few scattered populations in Northern Sardinia, occupying less than 50 hectares (1/2 $km^2$) of Real Area Of Occupancy (RAOO) in four different areas. In this paper we analysed seven natural populations of *C.*

*horrida* covering the entire distribution range of the species using four microsatellite genetic markers. This represents the first attempt at assessing the amount and distribution of genetic variability of this species and therefore constitutes a first step towards the planning of sound conservation strategies.

The amount of genetic variability found was medium-high, as indicated by the values of *H*e, ranging from 0.603 (LIO) to 0.854 (FAL & DON). The north-western populations were those showing the highest levels of heterozygosity, while the lowest value was observed in the Alghero area. In the congener species *C. corymbosa*, estimates of *H*e by means of SSR markers in six natural populations yielded values in the range of 0.36–0.62 (Freville *et al*., 2001). It is to be noted that the four SSRs used in our work are the same used by Freville and colleagues, making these results directly comparable. In another rare species belonging to Asteraceae, *Femeniasia balearica*, with a lifestyle very similar to that of *C. horrida* and a comparably small habitat, quite high levels of genetic variation were found by means of AFLP (Vilatersana *et al*., 2007). Allozyme analysis of seven species of the *Centaurea* genus endemic to Sicily (Bancheva *et al*., 2006) revealed heterozygosity values ranging from *He* = 0.126 for *Centaurea cineraria* L. subsp. *Cineraria* to *He* = 0.276 in *Centaurea todari* Lacaita. All these species grow on limestone cliffs. In another endemic species, *Centaurea tenorei* Guss. ex Lacaita, in the Sorrentina peninsula, in Southern Italy, which has populations irregularly located in an area including coastal zones and internal ridges, the amount of genetic variability was assessed again by means of allozymes (Palermo *et al.*, 2002). The lowest *H*e value was observed in *C. tenorei* subsp. *tenorei* (0.08), while the highest was observed in *C. parlatoris* Heldr. (0.34). We note that estimates of genetic diversity obtained with AFLP, microsatellite, and allozyme markers are not directly comparable due to differences in mutation rates. Nevertheless, the data at hand suggest that high genetic diversity values may have played a role in allowing the survival of these species in a harsh

and (presumably) stressful highly-stressed environment. This is particularly true for *C. horrida*, which lives on shallow soil on rocky sea cliffs and is exposed to strong winds and high levels of salinity.

Linkage disequilibrium (LD) was pronounced in the populations studied, all loci being in LD, with a few exceptions. LD can arise as a consequence of a reduction in effective population size that enhances drift. We failed, however, to detect evidence of a relatively recent and severe genetic bottleneck, which could have been the result of habitat fragmentation. The results we obtained need to be confirmed on a larger data set, because a low number of genetic markers greatly reduces the power of the statistical tests used, under both IAM and SMM (Cornet and Luikart, 1996). It is recommended (Piry et al., 1999) that at least 10 polymorphic loci are analysed to achieve a statistical power higher than 0.8. Even under the TPM, arguably the more appropriate model of evolution for SSRs (Di Rienzo *et al*., 1994), our data failed to display any evidence of reduction in $N_e$.

We cannot rule out the possibility that LD has arisen as a consequence of physical linkage between the loci, since no genetic map is available. A third explanation is that LD has arisen as a result of positive selection acting on loci linked to the SSRs used (Kim and Stephan, 2000). However, the presence of LD is an indication that further investigation into the mating system of *C. horrida* is needed, in order to assess the relationship between $N$ and $N_e$ in this species. In fact a reduction in census size, similar to that probably undergone by *C. horrida*, may not also imply a genetic bottleneck, which would result from a reduction of $N_e$.

Despite the strong LD signal in our populations, the species does not display a reduction of genetic variability, as shown by the very high values of $H_e$ and by the absence of private alleles. This behaviour is peculiar, since other rare and endangered species of the Mediterranean basin, such as *F. balearica* (Vilatersana *et al*., 2007), are characterised by both a lower amount of genetic variability and by higher differentiation between populations. This

issue will probably be clarified by the use of a larger set of genetic markers on the population studied.

**Genetic structure**

When dealing with conservation issues, it is often necessary to detect *K*, the number of panmictic units or 'gene pools' in the data, in order to be able to suggest possible mechanisms that have shaped the genetic variability observed. The use of a Bayesian approach to the detection of *K* has become increasingly popular in the last decade (Bertorelle & Excoffier, 1998; Pritchard et al., 2000). In the present study it was possible to estimate $K = 2$ as the number of inferred populations from which the studied populations derive. The most precise interpretation of this value is that two homogeneous gene pools contributed to the seven populations sampled. The LIO and BAR populations may have originated from the same ancestral population (see below; analysis of genetic differentiation between populations).

**Genetic differentiation**

The genetic divergence between populations, as estimated by $F_{ST}$ and $R_{ST}$, was high ($F_{ST} = 0.123$ and $R_{ST} = 0.158$) even though lower than that observed in *Femeniasia balearica*, where the amount of genetic variation found between populations was 30% of the total genetic variation observed, based on an AMOVA analysis of AFLP genotypes and in *C. corymbosa*, where an overlapping set of microsatellite markers estimated $F_{ST} = 0.23$. The high levels of genetic differentiation observed are those expected for a species characterised by a scattered distribution pattern, which may well limit gene flow, thus determining the differentiation values observed in *C. horrida* populations. In a similar study conducted on the rare *Eryngium alpinum* (*Umbelliferare*), a species which bears evidence of comparable biological and ecological traits (seed set production and short distance dispersion), the differentiation observed was $F_{ST} = 0.23$ between 12 populations genotyped by seven SSRs (Gaudeul *et al*., 2004).

Genetic differentiation was evaluated also between pairs of populations and proved significant in all cases, based on a permutation test. The lowest differentiation was found for the population pair FAL - DON (0.046), which are located close to each other in the Stintino area. In general, the populations of the Asinara–Stintino groups display lower levels of differentiation. Their isolation is in fact recent: given the shallow nature of the sill between Stintino peninsula and Asinara island, which is only about 20 metres deep, it dates back only to the end of the Würmian, about 13 ka cal BP (Antonioli et *al*., 2004)

The highest $F_{ST}$ values were found for the populations LIO - TAV (0.24), which are at the extremes of the distribution on an East–West axis, but also for the populations FOR - LIO (0.23), which are separated by about 30 kilometres of coastline. While in the first case we can assume that geographic distance is responsible for the high differentiation, in the second case we must search an alternative explanation.

Most of the area between FOR and LIO is an unsuitable habitat for *C. horrida*, and has been so for the last 100,000 years (S Andreucci, University of Sassari, Italy, pers. comm.), as it hosts dense juniper woods and more competitive shrub communities. Taking into account both the very low dispersal ability and the habitat specificity of *C. horrida*, we could argue that genetic differentiation is more affected by biological barriers than by geographical distance.

We also estimated the genetic divergence between populations by $R_{ST}$, the $F_{ST}$ analogue based on the stepwise mutation model. The highest $R_{ST}$ value was again found between the Tavolara island population and that of the Alghero area; the lowest was observed between the pairs DON - LIO and DON - TAV. All the $R_{ST}$ values were significantly different from zero, and consistently higher than those for $F_{ST}$. An exception to this trend is presented by the LIO population; for five out of six pairwise population comparisons involving LIO, $F_{ST}$ was greater than $R_{ST}$. This can be interpreted as ongoing differentiation because of recent genetic

drift, due to the peculiar ability of $R_{ST}$ to detect differentiation events older than those revealed by $F_{ST}$. This hypothesis is at least in part corroborated by the presence, in the LIO population, of two out of five loci pairs showing linkage disequilibrium, a characteristic typical of small isolated populations.

Mantel's test, used to confirm the presence of isolation-by-distance (IBD) between the populations studied, was significant, thus IBD played a role in shaping the present distribution of genetic variability. This is in agreement with the separation of the populations studied in different geographical regions, as indicated also by the AMOVA results. The amount of gene flow, however, is quite low, estimated at about 1.7 migrants / generation. This is probably due to both restricted pollen dispersal and to the poor ability of *C. horrida* to disperse achenes (Pisanu *et al*., 2007).

**AMOVA**

The hierarchical partitioning of the total variation between the gene pools found by Structure was significant (8.4%; Table 5a). The populations of the Alghero region again appear to be quite well differentiated from the other populations of the habitat. However, AMOVA was significant also when the seven populations were grouped according to their geographic distribution (Fig. 2; Table 5b). This suggests that the three population groups should be considered as separate entities under the point of view of the conservation of genetic resources.

**Implications for conservation**

The current distribution area of *Centaurea horrida* consists of tracts of land that have neither been below sea level nor subjected to volcanic or sedimentary events since the Miocene (Carmignani *et al*., 2001). The divergence we observed between the populations studied is therefore to be ascribed to events linked to the life-cycle, the mating system and, in recent years, anthropogenic impacts on the species. The position of re-assessing what is meant by a

"population" is of the utmost importance, especially when dealing with conservation problems and in cases where the geographical proximity of individuals is not always indicative of their provenance from a single Mendelian unit. The combined results of Mantel's test, Bayesian analysis and AMOVA that were obtained suggest that three distinct conservation units exist, from the point of view of management. To successfully preserve the genetic diversity of the species, special regard should be given to *in situ* strategies, since the amount of genetic variation harboured in each population is still high and the number of individuals, with the exception of the Tavolara population, is not low. However, fragmentation of the populations should be avoided, to prevent problems due to loss of diversity. All the areas where *C. horrida* grows are included in the *Natura 2000* network, each at different levels of protection.

A more thorough characterisation of the ecological features of *Centaurea horrida* is under way, which should provide further useful insights for conservation. For example, a significant effect of the site on seed production and germination has been found (Pisanu, 2007), which could affect patterns of genetic diversity. Given the changes in climate that the Mediterranean area is likely to undergo in the future, the genetic composition of the populations of *C. horrida*, a plant adapted to harsh conditions, could also provide us with an interesting model to understand ecological and evolutionary responses to drought stress due to climate change.

## SUPPLEMENTARY INFORMATION

**Table 1SM.** Estimates of $K$, the number of inferred populations of origin, based upon the "$\Delta K$" method (see text) for *Centaurea horrida*. For each value of $K$, the value of $\Delta(K)$ based upon 20 replicates is reported. The number of sampling localities analysed was seven.

**Figure 1SM.** Quantitative analysis of the genetic structure in the seven populations of *Centaurea horrida* studied in this work. Each plant can derive its genotypic composition from

two different gene pools ("inferred populations of origin") according to a Bayesian analysis (see text). In the histogram, each bar represents a single plant and the different colours of the bar are proportional to the contribution of each inferred population of origin to the genotype of the plant. The plants are numbered progressively within each population and the populations are indicated by the bars drawn across the histogram.


## ACKNOWLEDGEMENTS

This study was supported by the Management Plan of I.C.S. Asinara National Park, by Management Plan for the Protected Marine Area Tavolara – Punta Coda Cavallo to R.F. and by grant FAR2006 to G.B., and constitutes part of the Ph.D. program of G.M.. We also wish to thank two anonymous reviewers for their useful comments on the first version of this manuscript.


## LITERATURE CITED


**Antonioli F, Lambeck K, Amorosi A, Belluomini G, Correggiari A, Devoti S, Demuro S, Monaco C, Marocco R, Pagliarulo R., Orrù P, Silenzi S. 2004.** Sea level at 8 and 22 ka cal BP along the Italian coastline. In: Antonioli F., Vai GB, eds. *Climex Maps Italy - explanatory notes*. Bologna: Museo Geologico Giovanni Capellini, 11-14.

**Bancheva S, Geraci A, Raimondo FM. 2006**. Genetic diversity in the *Centaurea cineraria* group (*Compositae*) in Sicily using isozymes. *Plant Biosystems* **140**: 10 – 16.

**Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 1996.** GENETIX 4.02, logiciel sous Windows$^{TM}$ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS, Université de Montpellier II, Montpellier (France).

**Bertorelle G, Excoffier L. 1998**. Inferring admixture proportions from molecular data. *Molecular Biology and Evolution* **15**: 1298-1311.

**Carmignani L**, **Oggiano G**, **Barca S**, **Conti P**, **Salvadori I**, **Eltrudis A**, **Funedda A**, **Pasci S 2001**. Note illustrative della Carta Geologica della Sardegna a scala 1:200.000. Roma: Istituto Poligrafico e Zecca dello Stato.

**Cornuet J-M, Luikart G. 1996**. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001-2014.

**Debussche M, Thompson JD. 2002**. Morphological differentiation among closely related species with disjunct distributions; a case study of Mediterranean *Cyclamen* L., subgenus *Psilanthum* (Primulaceae). *Botanical Journal of the Linnean Society* **132**: 133-144.

**Desole L. 1954**. Secondo contributo alla conoscenza dello sviluppo embriologico del genere *Centaurea* L. (*Asteraceae*). *Centaurea horrida* Bad.. *Nuovo Giornale Botanico Italiano nuova serie* **61**: 256-273.

**Desole L. 1956**. Nuove stazioni e distribuzione geografica della "*Centaurea horrida*" Bad. *Webbia* **12**: 251-324.

**Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. 1994.** Mutational processes of simple sequence repeat loci in human populations. *Proceedings of National Academy of Science* USA **91**: 3166-3170.

**Dostàl J. 1976.** *Centaurea horrida* Badarò. In Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA. *Flora Europea IV*. Cambridge: Cambridge University Press, 273.

**Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611-2620.

**Excoffier L, Smouse PE, Quattro JM. 1992.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.

**Frèville H, Imbert E, Justy F, Vitalis R, Olivieri I. 2000**. Isolation and characterization and microsatellites in the endemic species *Centaurea corymbosa* Pourret (Asteraceae) and other related species. *Molecular Ecology* **9**: 1671-1672.

**Frèville H, Justy F, Olivieri I. 2001.** Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). *Molecular Ecology* **10**: 879–889.

**Gaudeul M, Taberlet P, Till-Bottraud I. 2000.** Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers**.** *Molecular Ecology* **9**: 1625–1637.

**Goodman SJ. 1997**. RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology* **6:** 881-885.

**Greuter W. 1995**. Origin and peculiarities of Mediterranean island floras. *Ecologia Mediterranea* **21**: 1-10.

**Guo SW, Thompson EA**. **1992.** Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics* **48:** 361-372.

**Kim Y, Stephan W. 2000.** Joint effects of genetic hitchhiking and background selection on neutral variation. *Genetics* **155**: 1415-1427.

**Mantel N. 1967.** The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209-220.

**Médail F, Quezel P. 1999.** Biodiversity hot-spots in the Mediterranean Basin: setting global conservation priorities. *Conservation Biology* **13**: 1510-1513.

**Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kents J. 2000.** Biodiversity hotspots for conservation priorities. *Nature* **403**: 853-858.

**Olfelt JP, Furnier GR, Luby JJ. 2001.** What data determine whether a plant taxon is distinct enough to merit legal protection? A case study of *Sedum integrifolium* (Crassulaceae). *American Journal of Botany* **88**: 401–410.

**Palermo AM, Pellegrino G, Menale B, Musacchio A. 2002.** Allozymic variability in *Centaurea tenorei* Guss. ex Lacaita and in other species of *C. parlatoris* Heldr. group (*Asteraceae*). *Plant Biosystems* **136**: 331 – 338.

**Peakall R, Smouse PE. 2001.** GenAlEx V5: *Genetic Analysis in Excel. Population Genetic Software for teaching and research*. Australian National University, Canberra, Australia.

**Piry S, Luikart G, Cornuet J-M. 1999**. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *The Journal of Heredity* **90**: 502-503.

**Pisanu S. 2007**. Analisi della struttura e biologia di popolazione in *Centaurea horrida* Badarò (Asteraceae) come base per l'analisi di vitalità delle popolazioni (PVA). PhD Thesis, Università di Sassari, Italy.

**Pritchard JK, Stephens M, Donnelly O. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

**Quilichini A, Debussche M, Thompson JD. 2004.** Geographic differentiation in the Mediterranean island endemic *Anchusa crispa*: implications for the conservation of a protected species. *Biological conservation* **118**: 651-660.

**Raymond M, Rousset F. 1995.** GENEPOP (Ver. 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248-249.

**Slatkin M. 1995.** A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139:** 457-462.

**Valsecchi F**. **1977**. Le piante endemiche della Sardegna: *Centaurea horrida* Bad. *Bollettino Società Sarda Scienze Naturali* **16**: 299-303.

**Vilatersana R, Susanna A, Brochmann C. 2007.** Genetic variation in *Femeniasia balearica* (Compositae, Cardue), an endemic and endangered monotypic genus from the Balearic Islands (Spain). *Botanical Journal of the Linnean Society* **153**: 97-107.

**Weir BS, Cockerham CC. 1984.** Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

## CAPTIONS TO FIGURES

**Figure 1.** Specimens of *Centaurea horrida* from the Falcone (FAL) population in the Stintino area. Picture taken in late April.

**Figure 2.** Schematic map of Sardinia (Western Mediterranean Sea) showing the geographic localisation of the populations of *C. horrida* studied (see also Table 1).

**Figure 3.** Analysis of population structure according to a Bayesian clustering method. The populations studied derive their genetic structure from two inferred populations ("gene pools" 1 and 2) of origin. A pie diagram indicates the proportion of membership of each inferred population (black or white) in the real populations studied.

**Table 1.** Natural populations of *Centaurea horrida* Badarò used in this study and characteristics of the study sites.

| Location area | Coordinates | Status | Sample number | Population size (n°individuals) | Population name (code) | Surface area | Lithology |
|---|---|---|---|---|---|---|---|
| Asinara Isle | 40°59'N-41°07'N 8°12'E-8°19E' | National Park | 59 | >300 | Fornelli (FOR) | 30.83 ha | Schist and Granite |
| | | | 56 | >300 | Stretti (STR) | | |
| Stintino Peninsula | 40°50'N-40°58'N 8°10'E-8°15'E | Natura 2000 site | 60 | >300 | Capo Falcone (FAL) | 12.42 ha | Schist |
| | | | 59 | >300 | Coscia di Donna (DON) | | |
| Capo Caccia Peninsula | 40°33'N-40°37'N 8°08'E-8°10'E | Regional Park | 58 | >300 | Marina di Lioneddu (LIO) | 2.1 ha | Limestone |
| | | | 59 | >300 | Cala della Barca (BAR) | | |
| Tavolara Isle | 40°53'N-40°55'N 9°40'E-9°44'E | Marine Reserve | 33 | < 300 | Tavolara (TAV) | < 1 ha | Limestone and Granite |

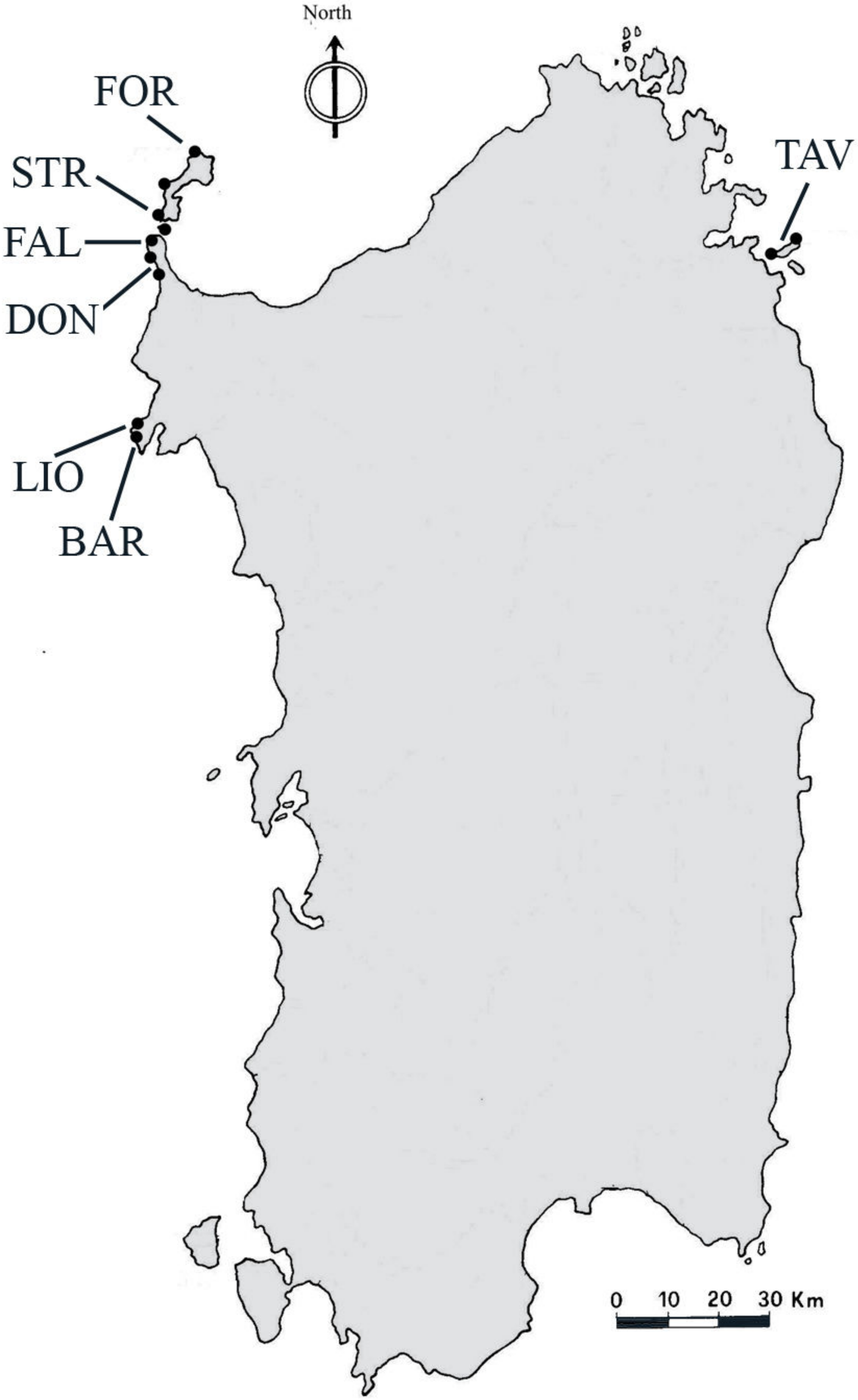

North
FOR
STR
FAL
DON
LIO
BAR
TAV
0 10 20 30 Km

**Table 2**. Features of the microsatellite markers used in this study[1].

| SSR locus | Repeat | Primer sequences (5'-3') | Ta (°C) | Fluorophore used | No. of detected alleles | Size of alleles (bp) |
|---|---|---|---|---|---|---|
| ***12B1*** | $(TA)_{27}$ $(GA)_{22}$ | F: CACACTCACGCTCAGCATTC<br>R:CATCGTTTCCAAACTTCCTC | 56 | HEX | 23 | 122-150 |
| ***13D10*** | $(AC)_7$ $ATAC(AT)_{10}$ | F:GGAGGCATGCGAACTAAAAG<br>R:CCGGTCTCATGAAAACAACT | 59 | FAM | 24 | 167-207 |
| ***21D9*** | $(CA)_{20}$ | F:CATATACACCCACGCACAGC<br>R:GGTGCAGCAAGGAGAGGAC | 60 | FAM | 15 | 101-125 |
| ***28A7*** | $(CA)_{16}$ | F:TTTCTATGCTGTTTGTTTTTGG<br>R:CCCATACGTCGTCTTCCC | 57 | HEX | 17 | 94-116 |

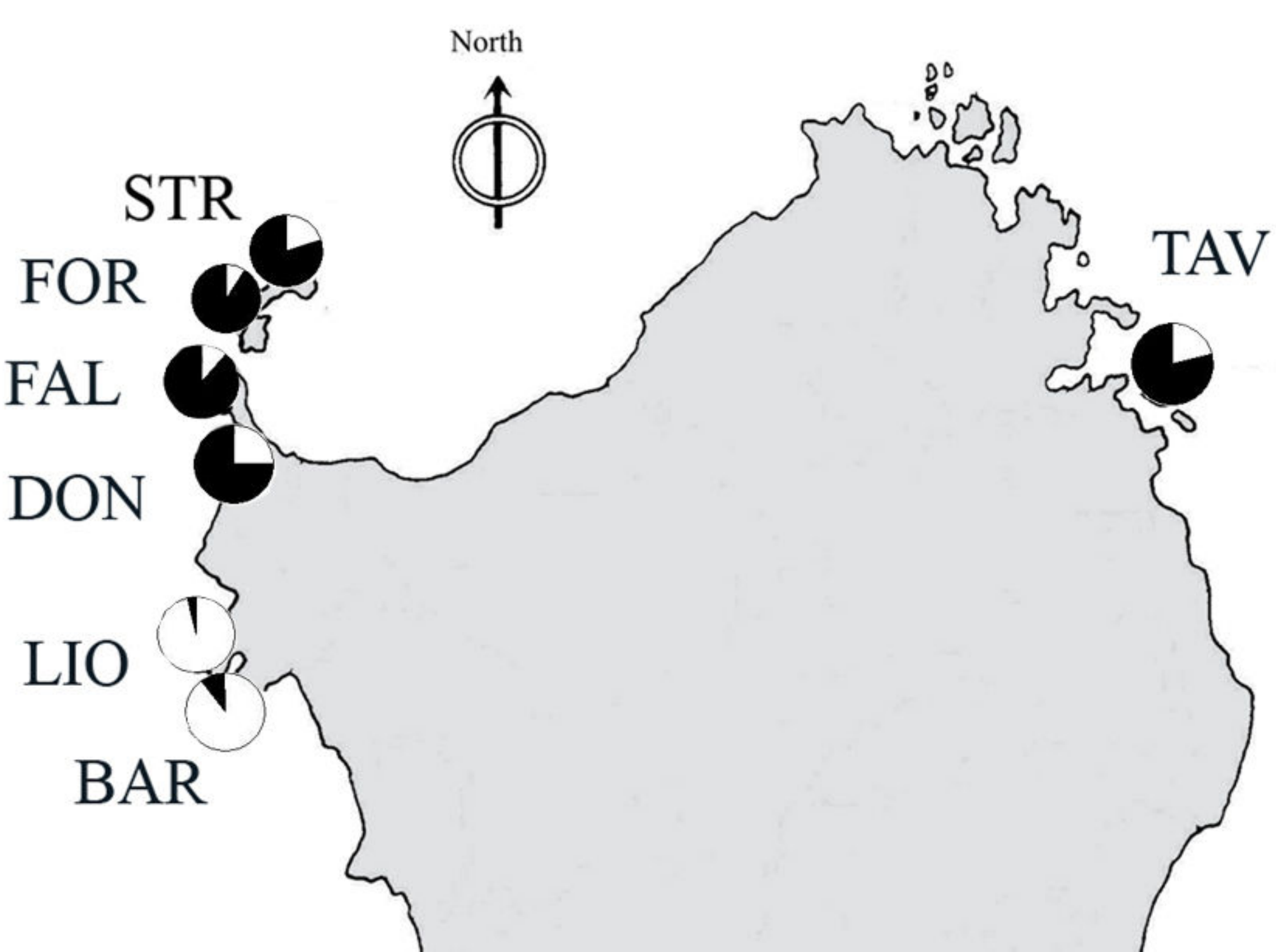
[1]**Frèville et al. 2000**. *Molecular Ecology* **9**: 1671-1672.

**Table 3.** Observed and expected heterozygosity measured at each locus for each population, and averages over loci and populations.

| | | STR | FOR | FAL | DON | LIO | BAR | TAV | *Average* |
|---|---|---|---|---|---|---|---|---|---|
| **21D9** | $H_o$ | 0.741 | 0.746 | 0.431 | 0.475 | 0.328 | 0.393 | 0.250 | |
| | $H_e$ | 0.856 | 0.866 | 0.842 | 0.798 | 0.287 | 0.610 | 0.449 | 0.673 |
| **13D10** | $H_o$ | 0.696 | 0.847 | 0.683 | 0.763 | 0.517 | 0.900 | 0.939 | |
| | $H_e$ | 0.843 | 0.909 | 0.862 | 0.925 | 0.734 | 0.879 | 0.869 | 0.860 |
| **28A7** | $H_o$ | 0.911 | 0.814 | 0.883 | 0.810 | 0.491 | 0.614 | 0.576 | |
| | $H_e$ | 0.789 | 0.760 | 0.827 | 0.796 | 0.524 | 0.697 | 0.674 | 0.724 |
| **12B1** | $H_o$ | 0.500 | 0.881 | 0.717 | 0.825 | 0.214 | 0.817 | 0.533 | |
| | $H_e$ | 0.837 | 0.873 | 0.887 | 0.899 | 0.867 | 0.908 | 0.759 | 0.861 |
| ***Average* $H_e$** | | 0.831 | 0.852 | 0.854 | 0.854 | 0.603 | 0.774 | 0.688 | |

**Table 4**. $F_{ST}$ (below diagonal) and $R_{ST}$ (above diagonal) values for each population pair.

| | **STR** | **FOR** | **FAL** | **DON** | **LIO** | **BAR** | **TAV** |
|---|---|---|---|---|---|---|---|
| **STR** | | 0.153 | 0.127 | 0.131 | 0.162 | 0.285 | 0.244 |
| **FOR** | 0.062 | | 0.136 | 0.052 | 0.125 | 0.327 | 0.141 |
| **FAL** | 0.084 | 0.071 | | 0.110 | 0.076 | 0.246 | 0.202 |
| **DON** | 0.075 | 0.072 | 0.046 | | 0.025 | 0.190 | 0.023 |
| **LIO** | 0.183 | 0.230 | 0.197 | 0.151 | | 0.111 | 0.082 |
| **BAR** | 0.108 | 0.112 | 0.107 | 0.089 | 0.082 | | 0.339 |
| **TAV** | 0.155 | 0.137 | 0.160 | 0.140 | 0.240 | 0.176 | |


North
STR
FOR
FAL
DON
LIO
BAR
TAV

**Tables 5a and 5b**. Analysis of Molecular Variance (AMOVA) based on four SSRs for the seven populations of *Centaurea horrida*. *P* values are estimated based on a permutation test (1000 randomizations).

5a (2 regions)

| Source of variation | d.f. | Percentage of variation | P - value |
|---|---|---|---|
| Among regions | 1 | 8.40 | 0.050 |
| Among populations | 5 | 9.33 | < 0.001 |
| Within regions | 763 | 82.27 | < 0.001 |

5b (3 regions)

| Source of variation | d.f. | Percentage of variation | P - value |
|---|---|---|---|
| Among regions | 2 | 10.01 | 0.009 |
| Among populations | 4 | 7.37 | < 0.001 |
| Within regions | 763 | 82.63 | < 0.001 |

**Table 1SM**. Estimate of *K*, the number of inferred populations of origin, based upon the "*ΔK*" method (see text) for *Centaurea horrida*. For each value of *K*, the value of Δ(K) based upon 20 replicates is reported. The number of real population analysed was seven.

| ***K*** | **1** | **2** | **3** | **4** | **5** | **6** |
|---|---|---|---|---|---|---|
| Δ(K) | - | 5.73 | 1.75 | 1.25 | 1.95 | 2.31 |

FOR
STR
STR
FAL
FAL
DON
DON
LIO
LIO
BAR
BAR
TAV
BAR

# Morphological and genetic traits in a natural homoploid hybrid between *Centaurea horrida* and *Centaurea filiformis* (Asteraceae).

PISANU S.[1], MAMELI G.[1], FARRIS E.[1], BINELLI G.[2] and FILIGHEDDU R.[1]

[1]*Dipartimento di Botanica ed Ecologia vegetale, Università di Sassari, via F. Muroni 25, I-07100 Sassari, Italy and* [2]*Dipartimento di Biotecnologie e Scienze molecolari, Università dell'Insubria, via J.H. Dunant 3, I-21100, Varese, Italy.*

Submitted to

**BOTANICAL JOURNAL OF THE LINNEAN SOCIETY**

**Morphological and genetic traits in a natural homoploid hybrid between *Centaurea horrida* and *Centaurea filiformis* (Asteraceae).**

PISANU S.[1], MAMELI G.[1], FARRIS E.[1], BINELLI G.[2] and FILIGHEDDU R.[1]

[1]*Dipartimento di Botanica ed Ecologia vegetale, Università di Sassari, via F. Muroni 25, I-07100 Sassari, Italy and* [2]*Dipartimento di Biotecnologie e Scienze molecolari, Università dell'Insubria, via J.H. Dunant 3, I-21100, Varese, Italy.*

**Running Title:** Morphological and genetic traits between *Centaurea horrida* and *Centaurea filiformis.*

**Corresponding Author:**
Rossella Filigheddu
Departimento di Botanica ed Ecologia vegetale,
Università di Sassari,
via F. Muroni 25,
I-07100 Sassari
Italy
Fax: +39-79-233600
Phone: +39-79-228640
E-mail: filighed@uniss.it

**Abstract**

Hybridization could have played a significant role in the evolution of several sections of the *Centaurea* genus, where several species could have an hybrid origin. Nevertheless to date no natural hybridization between Mediterranean endemic taxa of this genus has been documented. We have recently found in Tavolara island a patch of many fertile individuals showing intermediate morphological traits between *Centaurea horrida* and *Centaurea filiformis*, such as the morphology and size of capitula, appendages and leaves. The population of morphologically hybrid plants was found structured since individuals of different size classes were found at the study site. The hybrid population has a high level of seed production. Morphological variability among these individuals significantly differs from that of *C. horrida* and *C. filiformis*: the characters that mostly distinguished hybrid individuals from parent species are the length of leaves and the length and width of the heads.

ADDITIONAL KEYWORDS: endemic, population genetic structure, population structure, Sardinia.

## INTRODUCTION

Natural hybridization plays a fundamental role in the evolution of many plant taxa, sometimes resulting in the formation of entirely new species (Chapman & Burke, 2007). If hybrids are viable and fertile and if there are repeated opportunities for hybridization, extensive gene flow may results in the extinction of one of the hybridizing taxa via genetic assimilation (Genovart *et al.*, 2005) or even the merging of two taxa into a single evolutionary lineage. Persistent gene flow accompanied by reduced hybrid fitness can result in a stable hybrid zone, allowing for genetic exchange in certain genomic regions but preventing the merging of the taxa. Alternatively if the hybrids are fertile and viable, and at least partially reproductively isolated from their parents, the end result may be the production of a hybrid neospecies (Chapman & Burke, 2007). One possible path to reproductive isolation in hybrids is the segregation and recombination of chromosomal rearrangements or genetic incompatibilities that distinguish the parental taxa (homoploid hybrid species *sensu* Grant, 1981).

In the Mediterranean region many species of the *Centaurea* genus are currently present in rocky cliffs and crevices, steep slopes and coastal rocks (Hellwig, 2004), so that a considerable proportion of these taxa are endemic to one country or localized to a limited area, even a single mountain. Hybridization could have played a significant role in the evolution of several sections of this genus, where several species could have an hybrid origin (Garcia-Jacas, 1998). The presence of fertile hybrids is frequent in sect. *Acrocentron* and between sect. *Acrocentron* and *Chamaecianus* (Font *et al.*, 2002). Frequency of hybridization seems at present linked to human presence that can allow allopatric species to come in contact and introgress (Font *et al.*, 2002). Nevertheless to date no natural hybridization

between Mediterranean endemic taxa of the *Centurea* genus has been documented.

*Centaurea horrida* Badarò (Asteraceae) (Fig. 1) is a long-living spiny dwarf scrub, much-branched, that grows to heights of 70 cm, tomentous. Leaves are sessile, pinnatisect, tomentose, rigid and thorny, bearing a terminal segment with a single apical spine. Capitula are 5-6 mm in diameter, ovoid, oblong, cylindrical. Appendages are mucronate, shortly fimbriate at apex (Valsecchi, 1977). *C. horrida* reproduces sexually, by way of cross-pollination carried out by insects. It flowers in late spring (April-May) and bears fruit in summer (July-August), producing a seed that is 3.7 mm long, topped with a silky pappus that is 1.4 mm long. Its dispersal is of a mixed, ballistic/myrmicochorous type (Pisanu, unpubl. data).

*C. horrida* is a diploid species with 2n=18 (Desole, 1954), considered a paleoendemic *sensu* Contandriopoulos (Arrigoni, 1976) by Valsecchi (1977). Its distribution is limited to sea-cliffs in islands and peninsulas where it forms patches of isolated populations in dwarf communities. Its range extends in the Northern part of Sardinia (Fig. 2), with 5 locations..

*Centaurea filiformis* Viviani (Fig. 1) is a long living chamaephyte that grows to heights of 70 cm, woody below, corymbosely branched above. Leaves are glabrous, pinnatisect with linear or foliform, mucronulate laciniae. Capitula are ovoid, 1-2 cm in diameter. Appendages bear 6-10 fimbriae on each side. Seed has a pappus as long as achene (Arrigoni, 1972). *C. filiformis* is a diploid species (2n=18) (Arrigoni & Mori, 1971). It is a true chasmophytic plant, endemic of calcareous rocks in Eastern Sardinia (Fig. 2). This plant was recorded in 20 locations near each other, where several scattered individuals grow.

After Dostàl (1976) *C. horrida* belongs to subg. *Acrolophus* sect. *Horridae*, whereas *C. filiformis* to sect. *Maculosae*. These two species are then morphologically distinguishable endemic species. Despite of their systematic distance (Dostàl, 1976), phenotypic intermediates are present in the only location where the overlapping of the two species range occurs (Tavolara island, North-Eastern Sardinia), indicating a possible process of interspecific hybridization. Two morphological intermediate individuals were collected from Levier in 1885 at Tavolara island and named as *C. forsythiana* Levier (Arrigoni, 1972). Fiori (1903-1904: 332) traits these samples, from the nomenclatural point of view, as two different hybrids: *C. superfiliformis x horrida* Levier and *C. superhorrida x filiformis* (FI!). Another sample was then collected by Bocchieri in 1995 (CAG!).

We have recently found in Tavolara island (at the same locations of *specimina visa*) a patch of many fertile individuals showing intermediate morphological traits between *C. horrida* and *C. filiformis*, such as the morphology and size of capitula, appendages and leaves.

The aim of this work is therefore 1) to determine whether individuals observed and collected on the field, that appear to be morphologically intermediate between *C. horrida* and *C. filiformis,* are of hybrid origin; 2) quantify the population size, structure and seed production of intermediate forms; 3) verify whether hybrids are genetically distinguishable from the putative parents; 4) assess hybrids chromosomal number and 5) focus on the morphological characters of interest.

## STUDY AREA

The Tavolara island is 6 Km long, 1 Km large and extended on 600 hectares. The height is more of 565 m a.s.l. The island is constituted by a granitic base on which a mesozoic limestone rests, which is the prevailing geological substrate. The

bioclimate of the study site is of Mediterranean Pluviseasonal Oceanic type, with an Upper Thermomediterranean thermotype and a dry ombrotype. The flora is the richest among the circumsardinian islands, being composed by 463 entities that correspond to 19.2% of the Sardinian flora. Of these 34 (7.3%) are endemic. The endemic entities can be referred to a coastal component, in common with other coastal areas of Sardinia and to a limestone orophilous component, in common with the mesozoic limestone reliefs of central Sardinia. For this reason the island of Tavolara may be regarded as a plant biodiversity micro-hotspot. The biogeographical originality of its flora is stressed both by the presence of an exclusive species (*Asperula deficiens* Viv.) and the contact between the coastal and the mountain endemic contingents.

Unfortunately, the presence of military installations limits the opportunity to study and sample the plants present, but replaces the absence of special protection on the island.

Our samples come from the only limestone location, where also *C. horrida* is present. The nearest individual of *C. filiformis* grows about 500 m far, along a rocky wall (Fig. 2). The intermediate form is a perennial herb, woody at the medium height, that grows up to 70 cm, hardly tomentose. Leaves are sessile, pinnatifid and slashed. Flowers are white/rose wines. These individuals are very similar to *C. filiformis* regarding to habitus and leaves, that are divided in linear shape and are not spinous. Capitula instead seem much more similar to those of *C. horrida*, cylindrical and with the appendices briefly fimbriate at apex. Intermediate individuals are fertile: seeds easily germinate in lab and greenhouse. Pollen is vital: vitality was tested by using Alexander stain (1969) and controlled also on stamens from *C. horrida* and *C. filiformis* (pers. res.).

## METHODS

*Population structure and seed production*

In May 2007 all individuals (n=25) of putative hybrid origin were mapped. The major diameter was measured with a calibre and for each plant the number of branches was recorded. Population size was determined by counting all the mature individuals (adults) within the area. The structure of the population was estimated assigning each individual to one of three different stages: 1) *seedlings,* individuals developed to just beyond seed germination, with cotyledons, often also with one or two pairs of leaves and without stalks; 2) *saplings,* individuals non-reproducing in the year of study, with one or more stalks; 3) *adults,* all reproductive individuals. Population structure was expressed as the percentage of seedlings, saplings and adults present.

Seed production was estimated by counting capitula number on adults. In July 2007, since we found 8 adults severely damaged by browsing (feral goats), we collected 37 capitula from 11 adults, in order to estimate the seed production. The ratio ovary number / fertile seeds per capitulum was also verified on the field by using a stereoscope. In a way that was not damaging to the population, we left seeds from 29 capitula on the field and brought to the lab only 8 capitula, each from one individual.

*Genetic analysis*

In November 2006 green material was collected from 34 *C. horrida* adults, 15 *C. filiformis* adults and 21 intermediate individuals (19 adults and 2 saplings), at Tavolara island. Total genomic DNA was extracted by grinding the frozen leaves in a mortar in liquid $N_2$ and by using the DNeasy Plant Mini Kit (Qiagen, Italy),

according to the manufacturer's instructions. The average concentration of the extracted DNA was 20 ng/µL.

Due to the lack of information on the genome of the studied species, seven pairs of heterologous microsatellite primers, developed for the congener species *Centaurea corymbosa* Pourret (Fréville *et al.*, 2000), were firstly tested on *C. horrida*, and then on *C. filiformis* and intermediate individuals. Five of them (28A7, 13D10, 21D9, 12B1 and 13B7) have been insofar used to genotype our populations. SSRs reactions were performed in a total volume of 15µl, containing HotMasterTaq (Eppendorf®) buffer1X, 2.5mM $MgCl_2$, BSA (bovine serum albumin) 1.5µl, 2µM of each dNTPs, 10 µM of each forward and reverse primer, 25 ng genomic DNA and one unit of *Taq* polymerase (5U/µl) HotMasterTaq (Eppendorf®).

Polymerase Chain Reaction (PCR) amplifications were performed using a PTC-100 thermal cycler (MJ Research, Watertown, MA, USA), under the following conditions: an initial cycle at 94°C for 2 min, followed by 30 cycles of amplification, consisting of 94°C for 1 min, Ta for 30 s, 65C° for 1 min and a final step of extension at 65C° for 5 min. Microsatellites PCRs were processed using fluorescent-labelled primers, allowing PCR products to be simultaneously analyzed on a capillary MegaBACE® DNA sequencer (Amersham). The raw data were analysed by the allied MegaBACE Fragment Profiler software, to score the single-plant genotypes.

*Chromosomal number*

Root tip meristems were obtained from achenes collected on the field from adults of putative hybrid origin, by germinating them on wet filter paper in Petri dishes at room temperature. They were pretreated with 0.05% aqueous colchicines at

room temperature for 2h. The material was fixed in absolute ethanol and glacial acetic acid (3:1) for 24-48 h in the freezer and stored in 70% ethanol at -20°C. Samples were hydrolysed in 1N HCL for 12 min at 60°C and stained with Schiff's reagent (Feulgen and Rossenbeck, 1924) at room temperature for 30 min. They were mounted in a drop of acetocarmine following Ostergren and Haneen (1962). Preparations were made permanent by ethanol-dehydrating and mounted in Canada balsam. Observations were carried out in a Zeiss microscope and metaphase plates were photographed with a Pixelink Capture SE.

*Morphological analysis*

In a non destructive perspective we collected one capitulum and one leaf from 8 individuals (see above), to analyze size variability of: capitulum length (CL), capitulum width (CW), leaf length (LL), medium appendages length (ML) and width (MW), all important traits at the species level (Ertugrul *et al.*, 2004). The same analyses were carried out on samples from *C. horrida* and *C. filiformis* individuals, randomly chosen along the total range of these species and also in the sympatric populations of Tavolara island.

*Morphometric Data analysis*

Morphometric data (CL, CW, LL, ML, MW) were analysed by multivariate techniques using the PRIMER software package (Plymouth Marine Laboratory, UK: Clarke & Warwick, 1994). Data were not transformed. The Bray-Curtis similarity matrix was used to generate a cluster (Clarke, 1993). An analysis of similarity test (ANOSIM: Clarke, 1993) was performed to examine differences among populations. The similarity percentages procedure (SIMPER: Clarke,

1993) was employed to identify the major traits contributing to the differences among species.

*Genetic Data Analysis*

Allele frequencies and observed and expected heterozygosities were estimated at each locus for all populations, considering the intermediate form as a single population. Fisher's exact test using the Markov Chain algorithm (Guo & Thompson, 1992) was used to assess deviations from the Hardy-Weinberg equilibrium for each population and each locus. Weir and Cockerham's (1984) estimators of *F*-statistics were used to analyse genetic diversity both within and between populations. In particular, $F_{IS}$ was calculated in order to estimate what part of the total genetic variation was due to a departure from the Hardy-Weinberg equilibrium at the population level. $F_{ST}$ was calculated in order to estimate what part of the total genetic variation was due to differentiation between populations. $F_{ST}$ was also used to estimate gene flow by calculating the number of migrants per generation (Nm). The $F_{ST}$ analogue for microsatellites $R_{ST}$ (Slatkin, 1995) was also used, so as to include molecular information relating to the size of differences between the alleles in the differentiation estimates.

Nei's standard genetic distance (Nei, 1978) was calculated for pairwise comparisons of populations, under an infinite-allele-model. Principal coordinate analysis (PCoA) was performed using GenAlEx V6 (Peakall & Smouse, 2006), which provides a common pathway for the analysis of both binary and codominant data sets.

The software packages used to analyse the genetic data were GENETIX (Belkhir *et al.,* 1996), GenAlEx v.6 (Peakall & Smouse, 2006) and RSTCALC (Goodman, 1997).

## RESULTS

*Population structure and seed production*

In adult plants of putative hybrid origin the size ranges from 7 to 106 cm in diameter and from 10 to 45 cm in height.

The population of morphologically hybrid plants was found structured since individuals of different size classes were found at the study site. The 76% of the whole population (n=25) was constituted by adult plants (n=19) (Fig. 3). On average adults bore 32.8±7.73 capitula. In some cases seed production was ineffective: one adult had a capitulum without fertile seeds and another adult (bringing more capitula) had one capitulum without seeds. Capitula number was positively correlated to plant size ($n = 19$, $r = 0.89$, $p<0.05\%$ ) (Fig. 4) and to number of branches ($n = 19$, $r = 0.91$, $p<0.05\%$) (Fig. 5). On average were present 2.43 ± 0.35 intact seeds per capitulum (n=37), with a fecundity index of 0.22.

*Genetic analysis*

A total of 70 plants (*C. horrida, C. filiformis* and intermediate form) were analysed using five microsatellite markers, identifying a total of 86 alleles. The number of alleles per locus ranged from 11 (13B7) to 22 (13D10). No private alleles were detected at any locus for the intermediate form. In Table 1 the number of alleles shared by the hybrid with both *C. horrida* and *C. filiformis* is reported for the five SSR loci.

Genetic diversity at the loci studied was measured using Nei's heterozygosity (*He*) and the levels were medium-high; the highest value was found for *C. filiformis* (0.879, locus 13D10), the lowest value for *C. horrida* (0.449, locus 21D9) (Table 2).

The Hardy-Weinberg equilibrium was tested for all the loci and populations by testing the departure of $F_{IS}$ from zero under the null hypothesis. $F_{IS}$ values are significantly different from zero for all the loci except locus 28A7 for *C. filiformis* and *C. horrida*, locus 21D9 for the three species, locus 12B1 for the three species, locus 13D10 for *C. filiformis* and the intermediate form, finally locus 13B7 for the three species. In the vast majority of cases, deviation from the Hardy-Weinberg equilibrium was associated with positive $F_{IS}$ values, while negative $F_{IS}$ values were mainly associated with the intermediate form population. In particular, negative $F_{IS}$ values were found for loci 28A7 (intermediate form and *C. filiformis*), 13D10 (*C. horrida* and intermediate form) and 21D9 (intermediate form) (Table 2).

*Genetic differentiation*

The genetic divergence among species was measured using $F_{ST}$ and $R_{ST}$ (Table 3) and their significance tested by a permutation test based upon 1000 replicates. All $F_{ST}$ and $R_{ST}$ values differed significantly from zero. The overall $F_{ST}$ was 0.24 (confidence interval at the 95% level: $0.179 \leq F_{ST} \leq 0.299$), while the overall $R_{ST}$ was 0.286 (confidence interval at the 95% level: $0.235 \leq R_{ST} \leq 0.457$). As for pairwise comparisons between species, the maximum $F_{ST}$ value was found between *C. horrida* and *C. filiformis* (0.247) and the maximum $R_{ST}$ value between *C. horrida* and *C. filiformis* (0.315). It is to be noted that the highest values for both $R_{ST}$ and $F_{ST}$ are found between *C. horrida* and *C. filiformis*. The indirect estimate of gene flow (Nm) shows the highest value between *C. horrida* and intermediate form (5.140) and the lowest value between *C. horrida* and *C. filiformis* (0. 540).

Nei's genetic distances based upon the multilocus genotype of the individuals were also estimated. The lowest Nei's distance was found between the hybrid and *C. horrida* (0.687) and the highest between *C. horrida* and *C. filiformis* (3.470) (Table 4). The bi-dimensional scatter-plot of PCoA shows that the intermediate form population is in a central position between the two populations of putative parent species (Fig. 6). The multilocus genotype for the five SSRs used was also employed for an assignment test (Table 5). Only five plants out of 70 were misassigned with respect to their right species of provenance. In particular, two *C. horrida* and one *C. filiformis* plants were misassigned to the hybrid.

*Chromosomal number*

Tavolara Island, Sardinia, Italy. $2n = 18$ (Fig. 7). To date we have not metaphasic plates enough to describe the caryotype.

*Morphology*

Morphometric data used for multivariate analysis are shown in Table 6.

Furthermore we found the achenes of the hybrid individuals to be 2.75±0.05 mm long on average and pappus 1.75±0.07 mm (4.52±0.09 mm in total; n=90).

Multivariate analysis shows that three well distinct groups exist (Fig. 8). Simper demonstrated that the character that mainly contributes to the dissimilarity between the hybrid and *C. horrida* is LL (83.62%) followed by CL (8.54%), whereas between the hybrid and *C. filiformis* is LL (82.76%) followed by CW (9.26%).

## DISCUSSION

Since 1885 two morphological intermediate individuals between *C. horrida* and *C. filiformis* were known from Tavolara island. Their morphology however is different from our population recently discovered on the island. Even though preliminary, our results hint to the possibility that the "intermediate" form here shown is a real genetic homoploid hybrid between the two species *C. horrida* and *C. filiformis*. We can so consider Tavolara island as an original hybrid zone, where two endemic species, considered relictual and not allopatric, could give rise to repeated events of hybridization.

In this study we found that morphological variability among these individuals significantly differs from that of *C. horrida* and *C. filiformis*. The ratio CL/CW (2.5±0.50 mm) of capitula of hybrid population is very similar to that found in populations of *C. horrida* (2.14±0.50 mm), confirming the similarity of the capitula cylindrical shape. Also the form of appendages of the heads is more similar to that of *C. horrida*. The sizes of the heads, medium-sized appendages and leaves, are all intermediate between parent species. The characters that mostly distinguished hybrid individuals from parent species are the length of leaves and the length and width of the heads.

The hybrid population was found structured in different size classes and life stages, and this observation allows us to think that not only a F1 lineage is present and that an active recruitment is ongoing. The hybrid population has a high level of seed production, but not comparable with the higher values of *C. horrida* (pers. res.). Despite its apparent reproductive success, this natural hybrid population with intermediate morphology was only found on the limestone near a patch of *C. horrida* individuals. This may indicate that the rigid habitat requirements of *C. horrida* may also occur in hybrid plants, preventing their dispersion or that

hybridization has been so recent that the hybrids have not had yet the time to move.

Levels of genetic variation are moderately high in the intermediate form "hybrid", especially considering its endangered status and its narrow geographic range. Strong hints that these plants are real hybrids are the fact that all the alleles found are the same of the two "parental" species and that the Principal Coordinate Analysis puts the hybrid in an intermediate position between the two *Centaurea* species. However the relative contribution of *C. horrida* and *C. filiformis* is not the same in terms of the alleles present in the hybrid, in fact 21 *horrida* alleles can be found against only 11 alleles from *filiformis*. The possibility exists that the plants studied represent a second- or third-generation re-introgression of the original hybrid with *C. horrida*. To elucidate this aspect, we plan to analyse the haplotypes of the chloroplast DNA in the same plants to reveal both the origin of the female parent and possible phenomena of chloroplast capture through hybridization.

Intrinsic reproductive barriers among the species of *Centaurea* seem weak and genetic isolation is obtained mainly by geographical separation and ecological diversification, as shown by the fact that there are species of hybrid origin **(**Garcia-Jacas, 1992; Garcia-Jacas & Susanna, 1994; Garcia-Jacas, 1998). However to date no case of fertile homoploid natural hybrid population is reported within the *Centaurea* genus. Interestingly we can also exclude a hybridization process related to anthropogenic disturbance versus a more ancient hystorical process. The habitat of the hybrid population differs from that of *C. horrida*, especially in soil texture and plant community structure, and differs to an even higher degree from the habitat of *C. filiformis*, which is a complete chasmophyte, while the hybrids lives in the open. The importance of niche divergence is

corroborated by increased ecological tolerance in a number of putative homoploid hybrid species (Gross & Rieseberg, 2005). Perenniality also increases the likelihood of homoploid hybrid speciation (Chapman & Burke, 2007). But what could explain the maintenance of this hybrid zone and what ecological or geographical barrier has fallen?

The island of Sardinia has a consistent richness of endemic plants evolved as a result of its geological history (Thompson, 2005). Several species are intuitively known as palaeo-endemics (Arrigoni, 1976) because the island could have played a significant role during the last glacial maximum, and as schizo-endemics because a great number of endemic species could be evolved after the actual separation of Sardinia from the mainland and from Corsica, finished 20,000 years ago. On 347 endemic species 26.2% are in common to both islands whereas 45.8 % are exclusive to Sardinia (Bacchetta *et al*., 2005). Among these, five species of the *Centaurea* genus are present: *C. horrida*, *C. filiformis*, *C. corensis*, *C. ferulacea* and *C. magistrorum*. We here suggest that hybridization processes can still be present between two species of the *Centaurea* genus, which thus appear closely related, in contrast to Dostàl [1976] taxonomical point of view according to which *C. horrida* and *C. filiformis* belong to different sections. Several authors have argued that the rate of formation of fertile/viable hybrids between distantly related species should be lower than that between more closely related species (Schranz *et al*., 2005). Moreover the differentiation among *C. horrida* and *C. filiformis* could not be so old or we should hypothesize that the hybrid zone is also not so old. The highly complex geological and climatic history of Sardinia is likely to have provided ample opportunity for hybridization by breaking down ecological barriers and providing novel habitats for hybrids to establish.

The presence, if confirmed, of this hybrid population could bring to a reassessment of the systematic position of the parental species and of their role in the evolution of the Sardinian exclusive endemic contingent, with effects also on the development of genetic conservation strategies.

## REFERENCES

**Alexander MP. 1969.** Differential staining of aborted and non aborted pollen. *Stain Technology* **41:** 117-122.

**Arrigoni PV. 1972.** Sulla distribuzione e il rango sistematico di *Centaurea filiformis* Viviani e *Centaurea ferulacea* Martinelli. *Webbia* **27(1):** 279-287.

**Arrigoni PV. 1976.** Le piante endemiche della Sardegna. Introduzione. *Bollettino della Società Sarda di Scienze Naturali* **16:** 259-264.

**Arrigoni PV, Mori B. 1971.** Numeri cromosomici per la Flora Italiana. n. 93. *Centaurea ferulacea* Martinelli. *Informatore Botanico Italiano* **3(3):** 226.

**Bacchetta G, Iiriti G, Pontecorvo C. 2005.** Contributo alla conoscenza della Flora vascolare endemica della Sardegna. *Informatore Botanico Italiano* **37(1, parte A):** 306-307.

**Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 1996.** GENETICS 4.02, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Population Interaction, CNRS, Université de Montpellier II, Montpellier (France).

**Chapman MA, Burke JM. 2007.** Genetic divergence and hybrid speciation. *Evolution* **61-7:** 1773-1780.

**Clarke KR. 1993.** Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* **18:** 117-143.

**Clarke KR, Warwick RM. 1994.** Change in marine communities: an approach to statistical analysis and interpretation. Natural Environment Research Council, Swindon.

**Corrias B, Diana S. 1988.** Isola di Tavolara. In: Camarda I, Cossu A eds. *Biotopi di Sardegna*. Sassari (Italy): Carlo Delfino Editore, 59-80.

**Desole L. 1954.** Secondo contributo alla conoscenza dello sviluppo embriologico del genere *Centaurea* L. (Asteraceae): *Centaurea horrida* Bad. *Nuovo Giornale Botanico Italiano* **vol LXI, n. 2-3:** 256-273.

**Dostàl J. 1976.** *Centaurea* L. In: Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA. *Flora Europea*. London New York Melbourne: Cambridge University Press, 254-301.

**Ertugrul K, Uysal T, Garcia-Jacas N, Susanna A, Garntje T. 2004.** The systematic position of *Centaurea ensiformis* and *Centaurea isaurica* from Turkey and the evolution of some characters in *Centaurea*. *Israel Journal of Plant Sciences* **52:** 257-263.

**Feulgen R, Rossenbeck H. 1924.** Mikroskopisch-chemischer Nachweis einer Nucleinsäure vom Typus der Thymonucleinsäure und die darauf beruhende selektive Färbung von Zellkernen in mikroskopischen Präparaten. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* **135:** 203–248.

**Fiori A, Paoletti G. 1903-1904.** Flora Analitica d'Italia **3:** 332. Edagricole.

**Font M, Garnatje T, Garcia-Jacas N, Susanna A. 2002.** Delineatio and phylogeny of *Centaurea* sect. *Acrocentron* based on DNA sequences: a restoration of the genus *Crocodylium* and indirect evidence of introgression. *Plant Systematics and Evolution* **234:** 15-26.

**Fréville H, Imbert E, Justy F, Vitalis R Olivieri I. 2000.** Isolation and characterization of microsatellites in the endemic species *Centaurea corymbosa* Pourret (Asteraceae) and other related species. *Molecular Ecology* **9:**1671-1672.

**Garcia-Jacas N, Susanna A. 1992.** Karyological notes on *Centaurea* sect. *Acrocentron*. *Plant Systematics and Evolution* **179:** 1-18.

**Garcia-Jacas N, Susanna A. 1994.** *Centaurea prolongi* and *Centaurea crocata* in Portugal: an old confusion. *Nordic Journal of Botany* - section of Holoarctic and general taxonomy **14 (1)**: 31-38.

**Garcia-Jacas N. 1998.** *Centaurea kunkelii* (Asteraceae, Cardueae), a new hybridogenic endecaploid species of sect. *Acrocentron* from Spain. *Annales Botanici Fennici* **35:** 159-167.

**Genovart M, Juste J, Oro D. 2005.** Two sibling species sympatrically breeding: new conservation concern for the critically endangered Balearic shearwater. *Conservation Genetic* **6:** 601-606.

**Goodman SJ. 1997.** RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology* **6:** 881-885.

**Grant V. 1981.** *Plant speciation*. New York: Columbia University Press.

**Gross BL, Rieseberg LH. 2005.** The ecological genetics of homoploid hybrid speciation. *Journal of Heredity* **96:** 241-252.

**Guo SW, Thompson EA. 1992.** Performing the exact text of Hardy-Weinberg proportions for multiple alleles. *Biometrics* **48:**361-372.

**Hellwig FH. 2004.** Centaureinae (Asteraceae) in the Mediterranean – history of ecogeographical radiation. *Plant Systematics and Evolution* **246:** 137-162.

**Kadereit JW, Uribe-Convers S, Westberg E, Comes HP. 2006.** Reciprocal hybridization at different times between *Senecio flavus* and *Senecio glaucus* gave rise to two polyploidy species in north Africa and south-west Asia. *New Phytologist* **169:** 431-441.

**Nei M. 1978.** Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

**Östergren G, Haneen KW. 1962.** A squash technique for chromosome morphological studies. *Hereditas* **48:** 332–341.

**Peakall R, Smouse PE. 2006.** GenAlEx V6: *Genetic Analysis in Excel. Population Genetic Software for teaching and research*. *Molecular Ecology Notes* **6:** 288-295.

**Repplinger M, Johannesen J, Seitz A, Comes HP. 2007.** Morphological and molecular evidence for hybrization and introgression in central European *Arctium* (Asteraceae). *Plant Systematics and Evolution* **268 (1-4):** 75-97.

**Schranz ME, Dobes C, Koch MA, Mitchell-Olds T. 2005.** Sexual reproduction, hybridization, apomixis and polyploidization in the genus *Boechera* (Brassicaceae). *American Journal of Botany* **92:** 1797-1810.

**Slatkin M. 1995.** A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139:** 457-462.

**Thompson JD. 2005.** *Plant Evolution in the Mediterranean*. Oxford: Oxford University Press.

**Valsecchi F. 1977.** Le piante endemiche della Sardegna: 9 - *Centaurea horrida* Bad. *Bollettino della Società Sarda di Scienze Naturali* **16:** 299-303.

**Weir BS, Cockerham CC. 1984.** Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

**CAPTIONS TO FIGURES**

Fig. 1. *Centaurea horrida* Badarò (1), *Centaurea filiformis* Viviani (2) and hybrid (3): (a) plant; (b) head.

Fig. 2. Distribution of *C. horrida* (black line), *C. filiformis* (grey line) and hybrid population (asterisk).

Fig. 3. Population structure of the hybrid form on Tavolara island (% of individuals, n=25).

Fig. 4. Correlation between capitula number and plant size of the hybrid plant.

Fig. 5. Correlation between capitula number and number of branches of the hybrid plants.

Fig. 6. Principal Coordinates Analysis (PCA) between *C. horrida* (white squares) *C. filiformis* (black rhomb) and hybrid (grey triangle) populations. Percentages of total variance explained by each axis are noted in brackets.

Fig. 7. Cromosomical number (2n=18) of the hybrid plant..

Fig. 8. CLUSTER showing existence of three well distinct groups (*C. horrida*: C. h.; *C. filiformis*: C. f.; hybrid: Hy.).

**CAPTIONS TO TABLE**

Table 1. Number of alleles found in the intermediate form at the 5 SSR loci studied and their provenance from the two putative parental species.

Table 2. Observed and expected heterozygosity measured at each locus for all species and

average He in each species.

Table 3. Genetic differentiation between population pairs as misured by $F_{ST}$ (belong diagonal) and $R_{ST}$ (above).

Table 4. Genetic differentiation between population pairs as misured by Nei (belong diagonal) and Nm (above).

Table 5. Assignment of individuals to populations and percentage of correct classification.

Table 6. Morphometric traits (Average±S.E., mm) used for multivariate analysis. Capitulum length (CL) and width (CW), medium appendages length (ML) and width (MW), leaf length (LL).

**FIGURES**:

(1a)  (1b) 

(2a)  (2b) 

(3a)  (3b) 

Fig. 1.

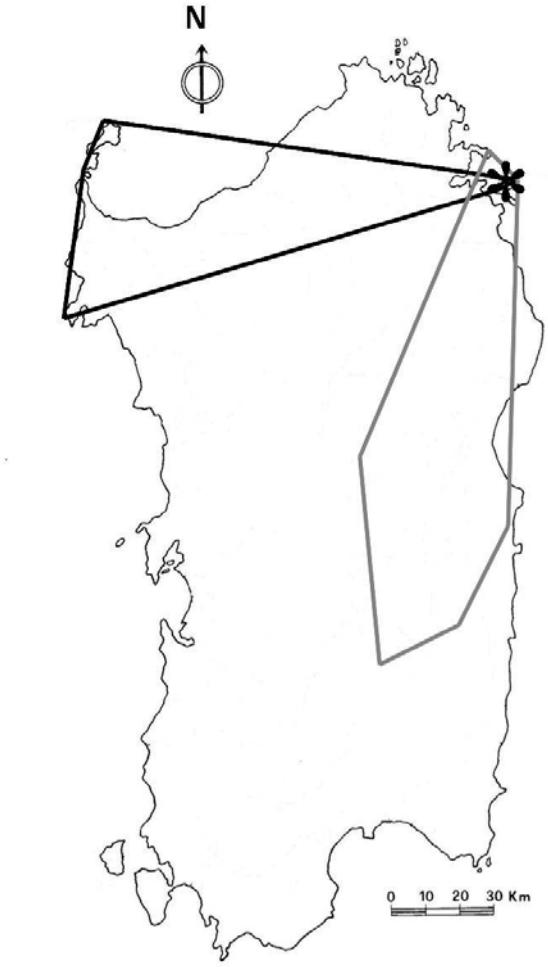



Fig. 2.

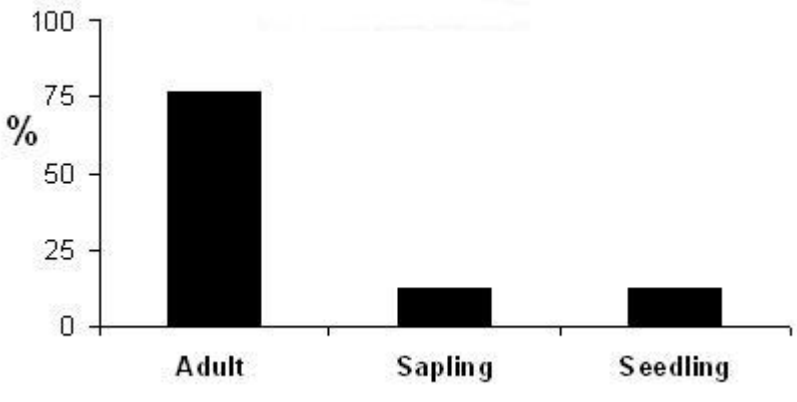



Fig. 3.

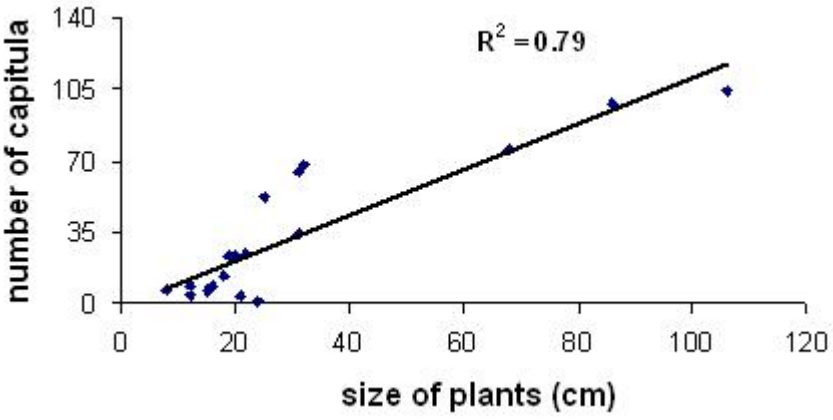



Fig. 4.

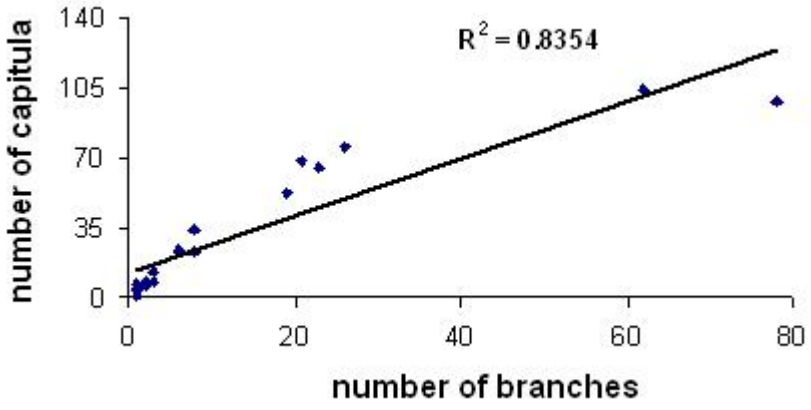



Fig. 5.

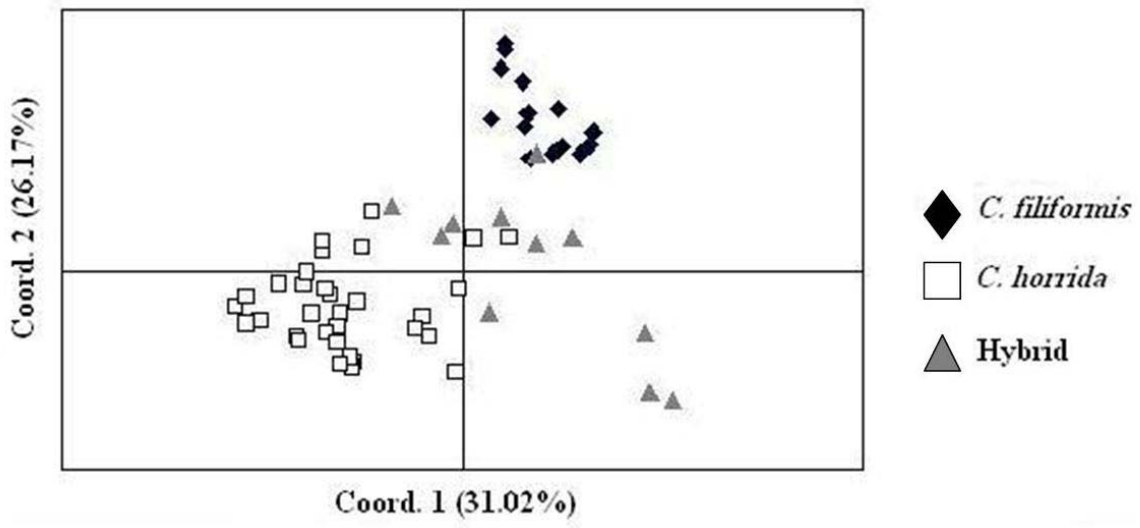



Fig. 6.

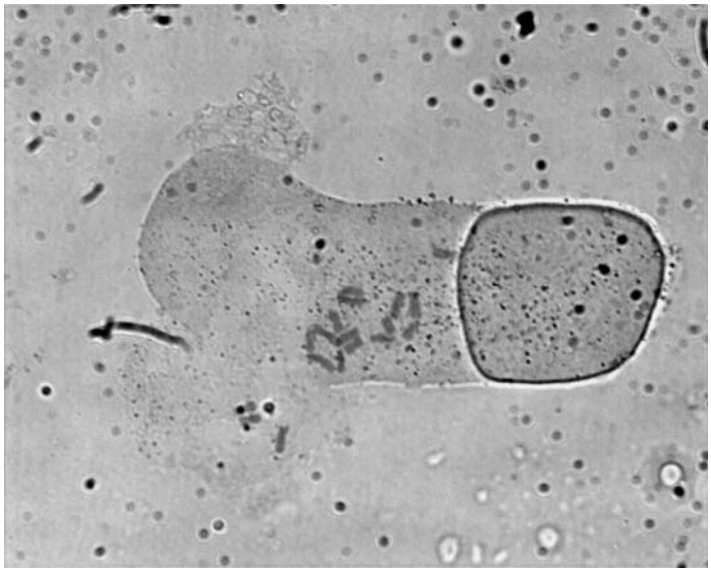


Fig. 7.

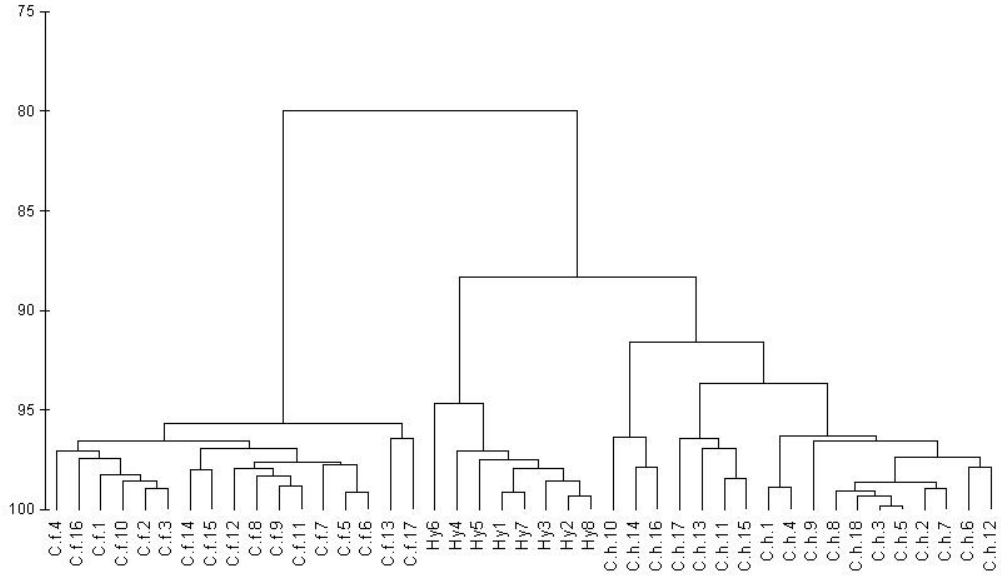

75
80
85
90
95
100
C.f.4
C.f.16
C.f.1
C.f.10
C.f.2
C.f.3
C.f.14
C.f.15
C.f.12
C.f.8
C.f.9
C.f.11
C.f.7
C.f.5
C.f.6
C.f.13
C.f.17
Hy6
Hy4
Hy5
Hy1
Hy7
Hy3
Hy2
Hy8
C.h.10
C.h.14
C.h.16
C.h.17
C.h.13
C.h.11
C.h.15
C.h.1
C.h.4
C.h.9
C.h.8
C.h.18
C.h.3
C.h.5
C.h.2
C.h.7
C.h.6
C.h.12

Fig. 8.

**TABLES**:

| | **28A7** | **13D10** | **12B1** | **21D9** | **13B7** |
|---|---|---|---|---|---|
| Total no. alleles | 9 | 6 | 6 | 6 | 5 |
| Shared (*C.h./C.f.*) | 5/4 | 5/1 | 4/2 | 4/2 | 3/2 |

Table 1.

| *Species* | ***Locus*** | ***Ho*** | ***He*** | **Fis** | **Signif** |
|---|---|---|---|---|---|
| ***C. filiformis*** | **28A7** | 0.800 | 0.782 | -0.023 | ** |
| | **21D9** | 0.692 | 0.784 | 0.117 | *** |
| | **12B1** | 0.714 | 0.834 | 0.144 | * |
| | **13D10** | 0.462 | 0.879 | 0.475 | ** |
| | **13B7** | 0.182 | 0.620 | 0.707 | *** |
| ***C. horrida*** | **28A7** | 0.576 | 0.674 | 0.146 | *** |
| | **21D9** | 0.250 | 0.449 | 0.443 | *** |
| | **12B1** | 0.533 | 0.759 | 0.298 | *** |
| | **13D10** | 0.939 | 0.869 | -0.081 | ns |
| | **13B7** | 0.345 | 0.640 | 0.461 | *** |
| **Hybrid** | **28A7** | 0.952 | 0.672 | -0.417 | ns |
| | **21D9** | 0.857 | 0.612 | -0.400 | * |
| | **12B1** | 0.333 | 0.505 | 0.339 | *** |
| | **13D10** | 0.857 | 0.757 | -0.132 | ** |
| | **13B7** | 0.190 | 0.522 | 0.635 | *** |

Table 2.

| | ***C. filiformis*** | ***C. horrida*** | **Hybrid** |
|---|---|---|---|
| ***C. filiformis*** | - | 0.315 | 0.216 |
| ***C. horrida*** | 0.247 | - | 0.046 |
| **Hybrid** | 0.227 | 0.201 | - |

Table 3. Pairwise $F_{ST}$ (above) and $R_{ST}$ (below) values between the three species analysed.

| Nei/Nm | *C. filiformis* | *C. horrida* | **Hybrid** |
|---|---|---|---|
| ***C. filiformis*** | 0.000 | 0.540 | 0.910 |
| ***C. horrida*** | 3.470 | 0.000 | 5.140 |
| **Hybrid** | 1.293 | 0.687 | 0.000 |

Tab 4.

| Population | *C. filiformis* | *C. horrida* | **Hybrid** | **Correctly assigned (%)** |
|---|---|---|---|---|
| ***C. filiformis*** | 14 | - | 1 | 93 |
| ***C. horrida*** | - | 32 | 2 | 94 |
| **Hybrid** | 1 | 1 | 19 | 90 |
| Misassigned | 1 | 1 | 3 | - |

Table 5.

| | **CL** | **CW** | **CL/CW** | **ML** | **MW** | **ML/MW** | **LL** |
|---|---|---|---|---|---|---|---|
| ***C. filiformis*** | 15.38±0.42 | 10.50±0.69 | 1.48±0.50 | 10.31±0.38 | 3.42±0.18 | 3.09±0.23 | 99.38±3.71 |
| ***C. horrida*** | 9.75±0.45 | 4.56±0.16 | 2.14±0.50 | 6.80±0.11 | 2.03±0.05 | 3.36±0.10 | 18.75±0.82 |
| **Hybrid** | 12.75±0.31 | 5.38±0.50 | 2.50±0.50 | 7.95±0.22 | 2.31±0.06 | 3.45±0.12 | 49.38±1.75 |

Table 6.

**Analyses of the Genetic Structure of the populations of two Sardinian endemics species *Centaurea filiformis* Viviani and *Centaurea ferulacea* Martelli.**

## INTRODUCTION:

***Centaurea filiformis*** Viv. is a true chasmophytic plant that grows to heights of 70 cm, woody below, corymbosely branched above. Leaves are glabrous pinnatisect with linear laciniae, mucronulate. Capitula are ovoid, 1-1.5 cm in diameter, with appendages acute with 6-10 fimbriae on each side. Achenes have a pappus as long as the achene (Arrigoni, 1972). *Centaurea filiformis* is a diploid species with 2n=18 (Arrigoni & Mori, 1971). It is endemic of calcareous rocks in Eastern Sardinia (Fig.1). This plant was recorded in 20 locations near each other, where several scattered individuals grow.

***Centaurea ferulacea*** Martelli,grows in a small area that lies south of the *C. filiformis*, as an appendix to the southern margin of the large calcareous formations of central Sardinia.

Both species are morphologically very similar, as shown from the iconography of Moris (1840-43) and Martelli (1896b). The two species differs almost exclusively the form of involucral bracts the capitula. From an ecological point of view, *Centaurea ferulacea* does not show different needs from those of *C. filiformis*; it is a calcicola limestone rock plant. The chromosomal number is identical in the two taxa (2n=18, Arrigoni and Mori, 1971), and both display two pairs of chromosomes with satellites.

Both *Centaurea* species are considered rocky endemics of the mesozoic limestones of middle-west Sardinia. Both entities are allopatric, but show, in the transition zone between the areas, some topodems morphologically *intergrading*, although constituted by homogenous individuals. Arrigoni (1972) considers that *C. filiformis* and *C. ferulacea* constitute an unique ologamodemus, and consequently that the

following taxonomic framing of the two entities can be justified: *Centaurea filiformis* Viv. ssp. *filiformis* and *C. filiformis* Viv. ssp. *ferulacea* (Martelli) Arrig. = (*Centaurea ferulacea*).

**Materials and Methods:**

**Plant material:**

We investigated the distribution area of the two species along the Northern West-coast of Sardinia during the autumn of 2006. We have not collected many individuals for each populations, because of the difficulty of access to several localities (Tab. 1). Fresh leaves were sampled non-destructively from a total of 46 individuals from four populations: 10 plants for *C.f1*, 11 for *C.f2* and 15 for *C.f3* populations of *Centaurea filiformis*, and 10 individuals for *C. fer*. population of *Centaurea ferulacea*. Leaves were stored at -80°C until DNA extraction. Genomic DNA was extracted of tissue of each plant by using the DNeasy Plant Mini Kit (Qiagen, Italy), leaf material (100mg) was ground to a fine powder in liquid $N_2$ in a mortar, according to the manufacturer's instructions. The average concentration of the extracted DNA was 20 ng/μL.

**Amplification conditions**

Simple Sequence Repeat (SSR) primers from *Centaurea corymbosa* (Freville *et al*., 2000) were tested for their ability to amplify single genomic regions in *Centaurea filiformis* Viviani and *Centaurea ferulacea* Martelli as already tested for *C. horrida* (Mameli et al.,2007). Five out of seven of them were selected because they yielded an unambiguous amplification pattern.

Amplification reactions were modified with respect to Freville *et al*., 2000. For genotyping of individuals, microsatellite amplifications were performed were performed in a total volume of 15 μL, containing HotMasterTaq (Eppendorf®) buffer 1X, 2.5 mM

$MgCl_2$, 2 µM of each dNTP, 0.5 µM of each forward and reverse primer, 25 ng genomic DNA and one unit of *Taq* polymerase HotMasterTaq (Eppendorf®). Amplification was carried out in a PTC-100 thermal cycler (MJ Research, Watertown, MA, USA), under the following conditions: an initial cycle at 94°C for 2 min, followed by 30 amplification cycles, at 94°C for 1 min, annealing temperature (Ta) for 30 s, 65°C for 1 min and a final step of extension at 65°C for 5 min.

The amplification products were run on a capillary MegaBACE® DNA sequencer (Amersham). The raw data were analysed using allied MegaBACE Fragment Profiler software, to score the single-plant genotypes.

**Statistical Data Analysis:**

The software packages used to analyse the genetic data were GENETIX (Belkhir *et al*., 1996), GenAlEx v.6 (Peakall & Smouse, 1996-2001), RST CALC (Goodman, 1997).

SSR loci were characterized for the number of alleles per locus and for the expected and observed heterozygosities under Hardy–Weinberg equilibrium for each locus and population (Nei,1978).

Therefore, SSRs polymorphism within samples was measured as allele frequencies and as observed and unbiased expected heterozygosity ($H_O$ and $H_E$) at each locus for all populations. Significance of deviation from HW equilibrium was estimated by means of a $\chi 2$ test, for each locus in each populations.

Weir & Cockerham's (1984) estimators of *F*-statistics were used to analyse genetic diversity both within and between populations. In particular, Wright's *F*-statistics $F_{IS}$ was calculated in order to estimate what part of the total genetic variation was due to a departure from the Hardy-Weinberg equilibrium at the population level.

$F_{ST}$ was calculated in order to estimate what part of the total genetic variation was due to differentiation between populations. $F_{ST}$ was also used to estimate gene flow by

calculating the number of migrants per generation ($Nm$). The $F_{ST}$ analogue for microsatellites $R_{ST}$ (Slatkin, 1995) was also used, so as to include in the differentiation estimates the molecular information relative to the size of differences between the alleles.

A Mantel test (1967) was applied to the matrices of pairwise $F_{ST}/(1- F_{ST})$ and log-transformed geographical distance between populations to assess isolation-by-distance, the model according to which genetic differentiation between populations is due to drift. Nei's standard genetic distance (Nei, 1978) was calculated for pairwise comparisons of populations, under an infinite-allele-model. UPGMA cluster analysis on pairwise Nei's (1978) unbiased genetic distances between populations was performed to construct an unrooted majority rule consensus tree with the programs NEIGHBOR and DRAWGRAM of the PHYLIP package (Felsenstein, 1995). The significance of the nodes was tested by bootstrapping with 1.000 replicates.

Analysis of molecular variance (AMOVA) was performed to partition the total genetic variation among regions and between populations within regions (Excoffier et al. 1992; Huff et al., 1993). The test of significance for the AMOVA was carried out on 1000 permutations of the data. Principal coordinate analysis (PCoA) was performed using GenAlex6 (Peakall & Smouse 2005).

**Results:**

**Genetic variability**

A total of 46 plants of *Centaurea filiformis* and *Centaurea ferulacea* were analysed using five microsatellite markers, identifying a total of 76 alleles. It was not possible to amplify the 12B1 locus for *C. ferulacea* thereby all estimates regarding this locus are missing. All the loci studied are medium polymorphic: the number of detected alleles per locus across all the populations ranges from 10 (locus *13B7*) to 20 (locus *13D10*).

Genetic diversity was measured using Nei's heterozygosity (*He*) and ranged from 0.377 (locus 13B7, *C.fer*. to 0.879 (locus *13D10*, *C.f*3). The medium-high estimates of genetic variability are confirmed by the average *He* values, ranging from 0. 360 of *C.fer* to 0.779 of *C.f*3 (Tab.2).

The Hardy-Weinberg equilibrium was tested for all the loci and populations by testing the departure of $F_{IS}$ from zero under the null hypothesis. $F_{IS}$ values are significantly different from zero for all the loci except locus *28A7* for the *C.f*1 and *C.f*2, locus *21D9* for *C.f*1, *C.f*2 and *C.fer*, and locus *12B1* and 13B7 for the *C.f*2, finally locus 13D10 for the *C.fer* populations. In the vast majority of cases, deviation from the Hardy-Weinberg equilibrium was associated with positive $F_{IS}$ values, while negative $F_{IS}$ values were mainly associated to *C.f*1 and *C.f*2. A monomorphic locus 1*2B1* was found for the *C.fer* populations.

**Genetic differentiation among populations**.

The genetic divergence among populations was measured using $F_{ST}$ and $R_{ST}$ (Table3). Their significance was tested by a permutation test based upon 1.000 replicates all $F_{ST}$ and $R_{ST}$ differed significantly from zero. The maximum $F_{ST}$ value was found between *C.f*2 and *C.f*3 (0.222), and the maximum $R_{ST}$ value between *C.f*2 and *C.f*3 (0.353).

Due to the absence of amplification at the *12B1* locus in *C. ferulacea* all estimates of genetic differentiation shown are based on four loci only.

It is to be noted that both $R_{ST}$ and $F_{ST}$ maximum values are found between *C.f*2 and *C.f*3. The minimum $F_{ST}$ value was found between *C.f*3 and *C.fer* (0.089), and the minimum $R_{ST}$ value between *C. f*1 and *C. fer* (0.240) (Tab 3).

The overall $F_{ST}$ was 0.24 (confidence interval at the 95% level: $0.179 \leq F_{ST} \leq 0.299$), while the overall $R_{ST}$ was 0.286 (confidence interval at the 95% level: $0.235 \leq R_{ST} \leq 0.457$).

The indirect estimate of gene flow (Nm) shows the highest value between *C.f*3 and *C. fer* (2.57) and the lowest value between *C. f*3 and *C. f*1 (0.088) (Tab.4).

A Mantel test was carried out, by correlating the amount of genetic differentiation between populations, as estimated by $F_{ST}$ /(1- $F_{ST}$), with the geographic distance between populations. The test was not significant ($r$ = -0.239; $p$ = 0.290), thus indicating that isolation – by – distance (IBD) was not a factor contributing to the differentiation among population.

The highest value of Nei's genetic distance was found between *C. f*1 and *C.f*3 (2.506) and the lowest between *C.f*3 and *C. fer* (0.431) (Table). An UPGMA tree based on Nei's genetic distance was built and is shown (Fig. 1).

The multilocus genotype for the four SSRs used was also employed for an assignment test (Tab.5). Only three plants out of 45 were misassigned with respect to their right species or populations of provenance. In particular, all misassignments involved *C. ferulacea*.

**AMOVA**

The total amount of genetic variation was also partitioned by AMOVA into components according to the geographic subdivision of the species. Based upon the analysis of the population structure, the hypothesis that the populations fall into the two species was tested, separating all three populations of *C. filiformis* from *C. ferulacea*. The AMOVA results (Table 6) show that the within population component accounts for 73% of the total variance and the remnant amounts of the total genetic variation was found for the difference between populations/region for 27%. The amount of the genetic variation among region is 0%, indicating that no differences regarding the distribution of the genetic variability exist between the two species.

**Principal coordinate analysis (PCoA):**

The bi-dimensional scatter-plot of PCoA display two distinct clusters (Fig.2), the first one formed by the individuals of the *C.f*1 and *C.f*2 populations, while the plants from *C.f*3 are grouped with those of *C. fer*., with a strong resemblance to the division already obtained by the phylogenetic analysis.

**Discussion**:

In this paper we analyse three natural populations of *C. filiformis* covering the entire habitat of the species and the only population of *C. ferulacea* known to date, using five microsatellite genetic markers. This represents the first attempt at assessing the amount and the distribution of genetic variability of these species and therefore constitutes a first step towards the planning of sound conservation strategies.

The amount of genetic variability found was medium-high, as indicated by the values of *H*e, with a peak of *H*e = 0.879 for locus *13D10* in the Tavolara populations of *C. filiformis*. In general, genetic variation was higher for *C. filiformis* than for *C. ferulacea*. Island plants generally have been found to have reduced levels of genetic variation. Frankham (1997) reviewed comparisons of closely related insular endemic and mainland plant taxa, and found that the insular endemic species is nearly always less heterozygous than its mainland congener. In the congener species *C. corymbosa*, estimates of *H*e by means of SSR markers in six natural populations yielded values in the range of 0.36–0.62 (Freville *et al*., 2001), while the congener and partially sympatric *C. horrida* display values of *H*e from 0.603 to 0.854 (Mameli *et al*., 2007). An exception to this trend is represented by endemic plants of the Canary Islands, which are more genetically variable (HT = 0.186 for 69 species in 18 genera) than species of other island archipelagos (HT =0.064) possibly due to the greater age of these islands

compared with their Pacific counterparts and to proximity to a continental source of migrants (Francisco-Ortega et al., 2000).

Allozyme analysis of seven species of the *Centaurea* genus endemic to Sicily (Bancheva *et al*., 2006) revealed heterozygosity values ranging from *He* = 0.126 for *Centaurea cineraria* L. subsp. *Cineraria* to *He* = 0.276 in *Centaurea todari* Lacaita. All these species grow on limestone cliffs. High genetic diversity values may thus have played a role in allowing the survival of these species in a harsh, highly-stressed environment. This is particularly true for *C. filiformis*, the populations of which live on shallow soil and/or on rocky sea cliffs, where are exposed to strong winds and high levels of salinity. *C. ferulacea* does grow on limestone cliffs both in the interior and in proximity of the sea, making it a true chasmophyte example.

**Genetic differentiation**

The genetic divergence between populations, as estimated by $F_{ST}$ and $R_{ST}$, is high ($F_{ST}$ = 0.24 and $R_{ST}$ = 0.29) comparable to that observed in *Femeniasia balearica*, where the amount of genetic variation found between populations was 30% of the total genetic variation observed, based on an AMOVA analysis of AFLP genotypes and in *C. corymbosa*, where an overlapping set of microsatellite markers estimated $F_{ST}$ = 0.23. The levels of genetic differentiation are however lower than those reported ($F_{ST}$ = 0.123) in *C. horrida*, by the same set of SSRs. The high levels of genetic differentiation observed are those expected for a species characterised by a scattered distribution pattern, which may well limit gene flow, thus determining the differentiation values observed in *C. filiformis* populations. In a similar study conducted on the rare *Eryngium alpinum* (*Umbellifer**ae***), that bears evidence of comparable biological and ecological traits (seed set production and short distance dispersion), the differentiation observed was $\boldsymbol{F_{ST}}$ = 0.23 between 12 populations genotyped by seven SSRs (Gaudeul *et al*.,

2004). Genetic differentiation was evaluated also between pairs of populations and proved significant in all cases, based on a permutation test. The lowest differentiation was found for the population pair *C.f*3 and *C.fer* (0.088), whilst the highest differentiation was found for the *C.f2* and *C.f3* pair (0.222). In this case we can assume that geographic distance is responsible for the high differentiation, because *C.f2* and *C.f3* are at the extremes of the distribution range of *C. filiformis*.

We also estimated the genetic divergence between populations by $R_{ST}$, the $F_{ST}$ analogue based on the stepwise mutation model. The results were identical: the highest $R_{ST}$ value was found, between the *C.f2* and the *C.f3* populations; the lowest was observed between the pair *C.f*3 and *C.fer*. All the $R_{ST}$ values are significantly different from zero, and consistently higher than those for $F_{ST}$. The peculiar ability of $R_{ST}$ to detect differentiation events older than those revealed by $F_{ST}$ indicates that the differentiation process has been uniform since a long period of time. Mantel's test, used to confirm the presence of isolation-by-distance (IBD) between the populations studied, was not significant, thus genetic drift has not recently played a role in shaping the present distribution of genetic variability, in agreement with the constant pattern indicated by both $F_{ST}$ and $R_{ST}$.

**AMOVA**

The hierarchical partitioning of the total $F_{ST}$ carried out by means of AMOVA was based on the difference between the species studied. The amount of variability resulting from this subdivision (0%) was not significant, while it was significant the amount of variability between the populations of the same species (23%). It appears like the pattern of distribution of the genetic variation is the same between the two species, indicating that both have undergone the same evolutive history.

**Implications for conservation**

The divergence we observed between the populations studied is to be ascribed to events linked to the life-cycle, the mating system and, in recent years, the anthropic impact on the species. The position of re-assessing what is meant by a "population" and a clear taxonomic indication of what a "species" is, is of the utmost importance, especially when dealing with conservation problems and in the case where the geographical proximity of individuals is not always indicative of their provenance from a single Mendelian unit. Our results seems to indicate that both *C. filiformis* and *C. ferulacea* have been differentiating in a similar way, to the point that at least one of the *C. filiformis* populations is less differentiated from the *C. ferulacea* one than from the co-specific populations.

To preserve successfully the genetic diversity of the species, special regard should be given to *in situ* strategies, since the amount of genetic variation harboured in each population is still high and the number of individuals, with the exception of the Tavolara population, is not low. However, fragmentation of the populations should be avoided, to prevent problems due to loss of diversity.

Finally, given the changes in climate that the Mediterranean area is likely to undergo in the future, the genetic composition of the populations of *C. filiformis* and *C. ferulacea*, plants adapted to harsh conditions could also provide us with an interesting model to understand mechanisms of drought tolerance and resistance.

**References:**

**Arrigoni PV. 1972. Sulla** distribuzione e il rango sistematico di Centaurea filiformis Viviani e Centaurea ferulacea Martinelli. *Webbia* **27**(1): 279-287.

**Arrigoni PV, Mori B. 1971.** Numeri cromosomici per la Flora Italiana. n. 93. Centaurea ferulacea Martinelli. *Informatore Botanico Italiano* **3**(3): 226.

**Bancheva S, Geraci A, Raimondo FM. 2006**. Genetic diversity in the *Centaurea cineraria* group (*Compositae*) in Sicily using isozymes. *Plant Biosystems* **140**: 10 – 16.

**Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 1996.** GENETICS 4.02, logiciel sous Windows$^{TM}$ pour la génétique des populations. Laboratoire Génome, Population Interaction, CNRS, Université de Montpellier II, Montpellier (France).

**Excoffier L, Smouse PE, Quattro JM. 1992.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.

**Felsenstein J. 2004**. PHYLIP (Phylogeny Inference Package) version 3.6. *Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.*

**Francisco-Ortega J, Santos-Guerra A, S-C. K, Crawford DJ**. 2000. Plant genetic diversity in the Canary Islands: a conservation perspective. American Journal of Botany 87: 909–919).

**Frankham R.** 1997. Do island populations have less genetic variation than mainland populations? Heredity 78: 311–327.

**Frèville H, Imbert E, Justy F, Vitalis R, Olivieri I. 2000**. Isolation and characterization and microsatellites in the endemic species *Centaurea corymbosa* Pourret (Asteraceae) and other related species. *Molecular Ecology* **9**: 1671-1672.

**Frèville H, Justy F, Olivieri I. 2001.** Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). *Molecular Ecology* **10**: 879–889.

**Gaudeul M, Taberlet P, Till-Bottraud I. 2000.** Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers**.** *Molecular Ecology* **9**: 1625–1637.

**Goodman SJ. 1997.** RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology* **6:** 881-885.

**Huff DR, Peakall R, Smouse PE** 1993. RAPD variation within and among natural populations of outcrossing buffalograss *Buchloe dactyloides* (Nutt) Engelm. *Theoretical and Applied Genetics*, 86, 927-934.

**Mameli G, Filigheddu R, Binelli G, Meloni M. In press.** The genetic structure of the remnant populations of Centaurea horrida Badarò in Sardinia, a major island of the Mediterranean Sea. *Annals of Botany (London).*

**Mantel N. 1967.** The detection of disease clustering and a generalized regression approach.

23 *Cancer Research* **27**: 209-220.

**Martelli U. 1896**. *Centaurea ferulacea* n. sp. sect. *Phalolepis*. Nuovo Giornale Botanico Italiano nov. ser., 3: 370-371.

**Moris JH. 1840-43.** Flora sardoa. 2: 455-456. Reg. Typ., Taurini.

**Nei M. 1978**. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583-590.

**Peakall R, Smouse PE. 2001.** GenAlEx V6: *Genetic Analysis in Excel. Population Genetic Software for teaching and research*. Australian National University, Canberra, Australia.

**Pritchard JK, Stephens M, Donnelly O. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

**Slatkin M. 1995.** A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139:** 457-462.

**Weir BS, Cockerham CC. 1984.** Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

Figures and Tables:

Tab.1 List of the populations studied for the two *Centaurea* species and their abbreviation.

| Species | Populations | Code |
|---|---|---|
| *Centaurea filiformis* | Oliena | *C.f1* |
| *Centaurea filiformis* | Cartoe | *C.f2* |
| *Centaurea filiformis* | Tavolara island | *C.f3* |
| *Centaurea ferulacea* | Baunei | *C.fer* |

Tab.2. Observed and expected heterozygosity measured at each locus for all species and average He in each species.

| ***Locus*** | | ***C. f1*** | ***C. f2*** | ***C. f3*** | ***C.fer.*** |
|---|---|---|---|---|---|
| **21D9** | **Ho** | **0.250** | **0.500** | **0.692** | **0.714** |
| **21D9** | **He** | **0.469** | **0.750** | **0.784** | **0.714** |
| **13D10** | **Ho** | **0.000** | **0.400** | **0.462** | **0.571** |
| **13D10** | **He** | **0.688** | **0.635** | **0.879** | **0.684** |
| **28A7** | **Ho** | **0.900** | **0.818** | **0.800** | **0.556** |
| **28A7** | **He** | **0.855** | **0.789** | **0.782** | **0.574** |
| **12B1** | **Ho** | **0.700** | **0.500** | **0.714** | **0.000** |
| **12B1** | **He** | **0.735** | **0.602** | **0.834** | **0.000** |
| **13B7** | **Ho** | **0.000** | **0.333** | **0.182** | **0.000** |
| **13B7** | **He** | **0.719** | **0.377** | **0.620** | **0.560** |
| **Average He** | **Ho** | **0.370** | **0.510** | **0,570** | **0,36** |
| **Average He** | **He** | **0.693** | **0.63** | **0,779** | **0.506** |

Tab.3. Genetic differentiation between population pairs as measured by $F_{ST}$ (below diagonal) and $R_{ST}$ (above)

| *Fst/Rst* | *C.f1* | *C.f2* | *C.f3* | *C.fer.* |
|---|---|---|---|---|
| *C.f1* | - | 0.277 | 0.260 | 0.240 |
| *C.f2* | 0.151 | - | 0.353 | 0.300 |
| *C.f3* | 0.194 | 0.222 | - | 0.245 |
| *C.fer.* | 0.149 | 0.193 | 0.089 | - |

Tab.4 Number of migrants per generation as estimated by means of $F_{ST}$ between the population studied.

| Nm | *C.f1* | *C.f2* | *C.f3* | *C.fer.* |
|---|---|---|---|---|
| *C.f1* | - | 1.41 | 1.04 | 1.42 |
| *C.f2* | | - | 0.88 | 1.05 |
| *C.f3* | | | - | 2.57 |
| *C.fer.* | | | | - |

Tab. 5 Assignment of individuals to populations and percentage of correct classification

| | Population | 1 | 2 | 3 | 4 | Correctly assigned (%) |
|---|---|---|---|---|---|---|
| 1 | *C.f1* | 10 | | | | 100 |
| 2 | *C.f2* | | 10 | | 1 | 91 |
| 3 | *C.f3* | | | 14 | | 100 |
| 4 | *C.fer.* | | 1 | 1 | 8 | 80 |
| | Misassigned | | 1 | 1 | 1 | |

Tab. 6 Analysis of Molecular Variance (AMOVA) based on four SSRs for 4 populations of *Centaurea* species . It is indicated the percentage of variation explained for the subdivision of the populations according to the hypothesis tested (see text). *P* values are estimated based on a permutation test (1000 randomizations).

| **Source of variation** | **d.f** | **% of variation** | **P-value** |
|---|---|---|---|
| **Among Regions** | 1 | 0% | 1.000 |
| **Among Pops/Regions** | 2 | 27% | 0.001 |
| **Within Pops** | 42 | 73% | 0.001 |

Fig. 1 UPGMA phylogenetic tree based on Nei's genetic distance for the species of the genus *Centaurea* studied. Number at the nodes indicate the bootstrap values (1000 replicates).

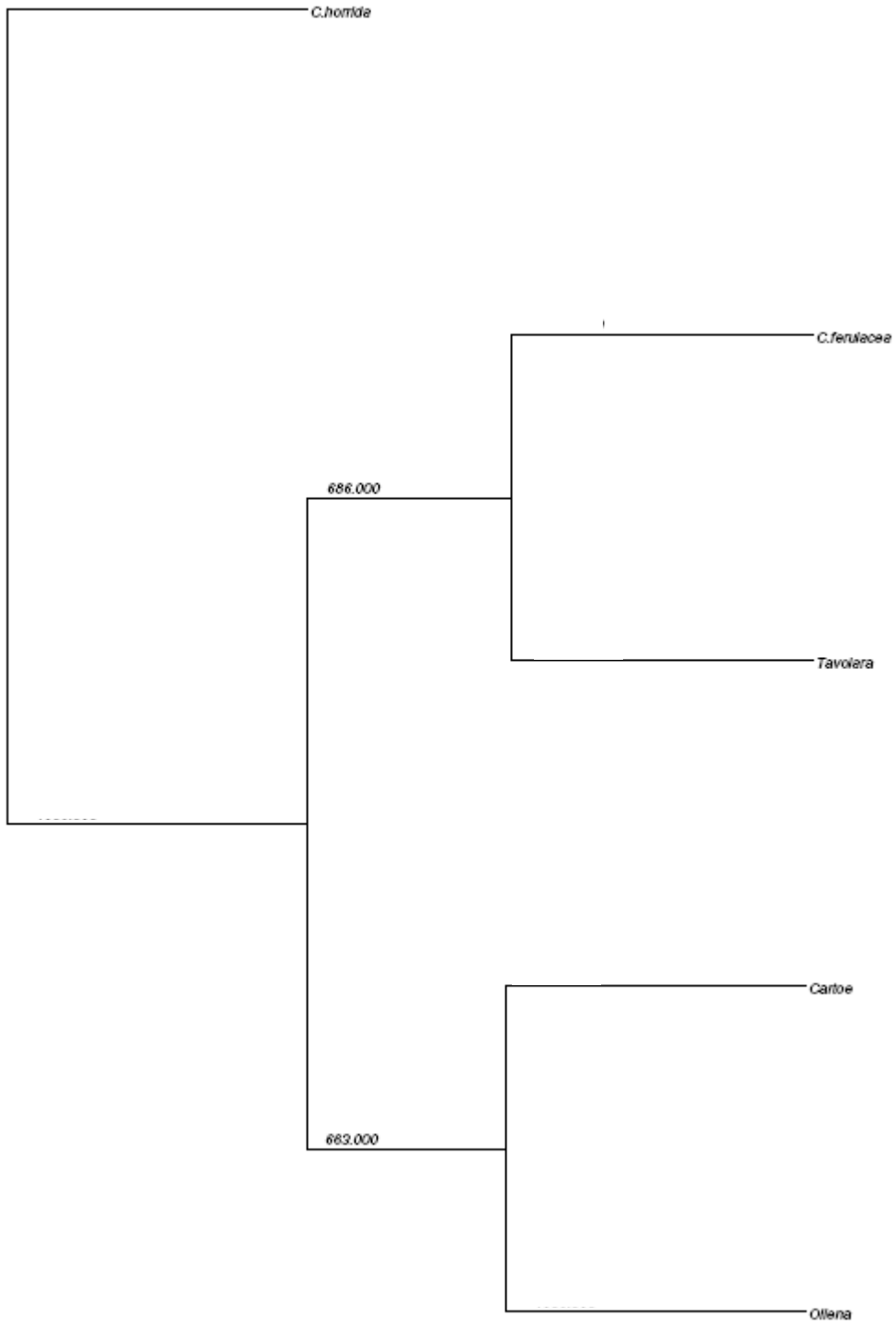
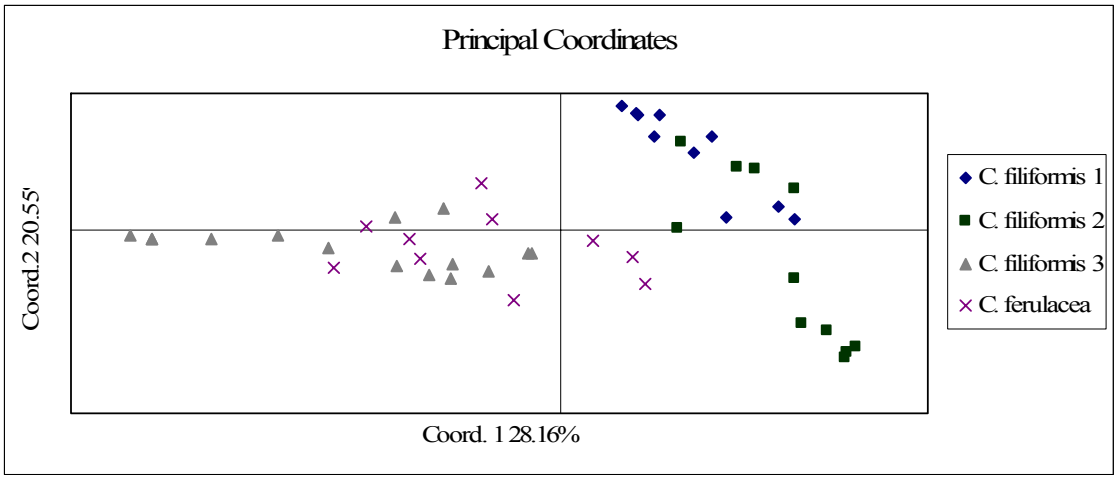


Fig. 2. Principal Coordinate Analysis (PCA) of the individuals of the three populations of *C. filiformis* and of the population of *C. ferulacea*, based upon their genotype at four SSR loci. Percentages of total variance explained by each axis are reported.



# **Phylogeny, systematics and hybridization in** *Centaurea horrida* **and** *Centaurea filiformis*: **evidence from nuclear-ribosomal DNA sequences**

Introduction:

The areas of the Mediterranean Basin are recognized as 'hotspots of biodiversity', for the immense wealth of the Mediterranean flora. Médail & Quézel 1999, proposed the delimitation of the 'hotspots'of the biodiversity within the Mediterranean region.

This setting represent one of the most geologically complex areas of the world and a example of a sea surrounded by different continents. The history of the Cyrno-Sardinian microplate is critical to our understanding of endemism in the western Mediterranean (Rosenbaum *et al*. 2002a). The basin's location at the intersection of two major landmasses, Eurasia and Africa, has contributed to its high diversity. The endemism is mainly concentrated on islands, peninsulas, rocky cliffs, and mountain peaks (www.biodiversityhotspots.org). In general, island populations have much higher risks of extinction than mainland populations and there is some evidence for higher extinction rates in island endemics than in nonendemics (Frankham, 1998).

The large number of endemic Mediterranean species has been interpreted to be the result of the diverse palaeogeographical history following the rotation of the corso-sardinian microplate which started at the oligo-miocene. The Mediterranean Region is an ideal place to study plants where you have a high plus a wealth of species and a higher rate of endemic entities that correspond to zones of high tectonic activity and/or microplate fragmentation and isolation (Cardona & Contandriopoulos, 1979). A common paleoendemism ties up the history of the corso-sardinian flora

(Contandriopoulos & Cardona, 1984; Contandriopoulos, 1981). During the Messinian salinity crisis, possibility for plant migration was greater as a result of land connections. Greuter (1979) calls it a key period for Mediterranean biogeography, being responsible for almost explosive speciation. This explains the high number of restricted endemic taxa at both ends of the Mediterranean (Hellwig, 2004).

Several species of the Mediterranean genus *Centaurea* (Compositae) segregated as local taxa; in fact, they are not separated by sex barriers, but only inability of cross-fertilization (geographical separation, fruits and heavy with pappus invalid or ineffective and inappropriate dissemination distance, or autogamy; Colas *et al.*, 2001, Pisanu S., in press). The genus *Centaurea* L., has traditionally been considered problematic. More recent molecular analyses of the genus and of subtribe Centaureinae, allowed definition of the natural limits of *Centaurea* (Susanna *et al*., 1995; Garcia-Jacas *et al*., 2000, 2001). Previous molecular phylogenies show the *Jacea* group to be a monophyletic clade divided into three major clades (Garcia-Jacas *et al*., 2000; 2006). The *Acrolophus* subgroup has traditionally been recognized to include species of three sections, i.e. *Acrolophus*, *Phalolepis,* and *Willkommia*. However, recently published studies (Garcia-Jacas *et al*., 2006; Suárez-Santiago *et al*., 2007) suggest the recognition of only two, *Willkommia* and *Acrolophus* (incl. *Phalolepis*). Their distribution area is restricted mainly to the two ends of the Mediterranean, with a group of species restricted to the western Mediterranean (species of the section *Willkommia* and several taxa of *Acrolophus–Phalolepis* complex), and other group mainly distributed in the eastern Mediterranean (species of *Acrolophus–Phalolepis* complex). The final analysis on molecular clock show places the divergence time of the *Jacea-Lepteranthus* and *Acrolophus* subgroup at the beginning of the Messinian (7.1 mya), The data confirm

the divergence of the *Acrolophus–Phalolepis* complex and *Willkommia* ribotypes at the end of the Messinian (Suárez-Santiago *et al*., 2007).

In the same work he suggested an evolutionary scenario for the *Acrolophus* subgroup in the western Mediterranean involving recurrent hybridizations of parapatric ("microallopatric") lineages within the geographical range of a primary radiation, where the isolation-contact periods may have occurred repeatedly during the Pleistocene glacial/interglacial cycles, (Suárez-Santiago *et al*. 2007)

The genus *Centaurea*, in Sardinia, presents five interesting endemic species: *C. horrida* Badarò, *C. filiformis* Viv., *C. ferulacea* Martelli, *C. corensis* recently described by Valsecchi & Filigheddu (1991) and *C. magistrorum* Arrigoni & Camarda (2003). For this study were analyzed the first three, for which we have started a study to identify the genetic structure of populations.

Species subject of this analysis are part of subgen. *Acrolophus* and respectively are included in the following section; *C. horrida* in sect. *Horrida*, *Centaurea filiformis* in sect. *Maculosae* and *C. ferulacea* in sect. *Phalolepis*. The hybrid has not been included in any section.

Several species of the genus *Centaurea*, such as *C. horrida*, *C. filiformis* and *C. ferulacea*, are currently in open steppe-like landescapes, including rock cliffs and crevices, steep slopes and coastal rocks and as for many other chasmophytic species of subgenus *Acrolophus* these habitats are a Mediterranean biogeographical refuge (Hellwig, 2004).

All the species of this large group *Acrolophus*, tend to have the same characteristic habit; however, it is extremely difficult to separate the constituent species and intermediate forms (often considered to be hybrids) that are frequent, (Dostál, 1976).

***Centaurea horrida*** Badarò Gior. Fis. (Brugnat.) ser. 2, 7: 363 (1824).

This species is a long-living spiny dwarf shrub, much-branched, that grows to heights of 70 cm. Leaves are pinnatisect, tomentose, terminal segment with a single apical spine. Capitula are 3-4 mm in diameter, ovoid cylindrical, with mucronate appendages, shortly fimbriate at apex. *C. horrida* reproduces sexually, by way of cross-pollination carried out by insects. It flowers in late spring (April-May) and bears fruit in summer (July-August), producing a seed that is 3.7 mm long, topped with a silky pappus that is 1.4 mm long. Its dispersal is of a mixed, ballistic/myrmicochorous type (Pisanu in press). *Centaurea horrida* is also a species characterized by heavy achenes, fitted with elaiosoma and reduced pappus (Pisanu, unpublished data), all characters that agree with the so-called myrmekochory syndrome (authors' personal observation) and Wagenitz & Hellwig (1996). This is a diploid species with 2n=18 (Desole, 1954).

These characters, together with reproductive biology, do not favor a long-distance dispersal and thus determine a very restricted distribution, as happens to many entities of subtribe of Centaureinae in Mediterranean (Hellwig, 2004).

*Centaurea horrida* is a narrow endemic, sensu Contrandiopoulos (1981), exclusive of northern Sardinia (Valsecchi, 1977; Desole, 1956). It is a perennial polycarpous spiny dwarf included in *Horridae* section of *Acrolophus* subgenus (Dostál, 1976). In this sect. it was included another species, *C. balearica*; now classified in a new distinct genus *Femeniasia balearica (*J. J. Rodr.) Susanna. For this motive the species of *C. horrida* is isolated sistematically any from other.

It is a protected species according to the Bern Convention (Appendix I) and a priority species according to the EU Directive 43/92 "Habitat" (Annex II). It's a vulnerable species according to the 1997 IUCN Red List of threatened plants.

This species is located in highly fragmented habitats, ranging from North-Weast to North-East Sardinian sea-cliffs (Desole 1956; Pisanu S., in press). Its range includes two parasarde islands (islets of Asinara and Tavolara).

Particularly *C. horrida* is fragmented in 5 subpopulations, defined as geographically distinct groups into the population, according to the new IUCN guidelines (Standards and Petitions Working Group, 2006).

***Centaurea filiformis*** Viv., Fl. Cors. App.: 6 (1825).

This species is a true chasmophytic plant that grows to heights of 70 cm, woody below, corymbosely branched above. Leaves are glabrous pinnatisect with linear laciniae, mucronulate. Capitula are ovoid, 1-1.5 cm in diameter, with appendages acute with 6-10 fimbriae on each side. Achenes have a pappus as long as the achene (Arrigoni, 1972). *Centaurea filiformis* is a diploid species with 2n=18 (Arrigoni & Mori, 1971). It is endemic of calcareous rocks in Eastern Sardinia (Fig.1). This plant was recorded in 20 locations near each other, where several scattered individuals grow. According to Dostál (1976) *C. horrida* belongs to subg. *Acrolophus* sect. *Horridae*, whereas *C. filiformis* to sect. *Maculosa*. The sections are very difficult to establish because they present many hybrid species, after which occurs introgression many cases (Ochsmann, 2000).

These two species are then morphologically distinguishable endemic species. Despite of their systematic distance (Dostál, 1976), phenotypic intermediates are present, indicating a possible process of interspecific hybridization (Mameli *et al*., in press)

Two morphological intermediate individuals were collected by Levier in 1885 at Tavolara island (north-eastern Sardinia) and named as *C. forsythiana* Levier. Fiori (1903-1904) traits these samples, from the nomenclatural point of view, as two

different hybrids: *C. superfiliformis x horrida* Levier (FI!) and *C. superhorrida x filiformis* (FI?). Another sample was then collected by Bocchieri in 1995 (CAG!)

***Centaurea ferulacea*** Martelli, Nuovo Gior. Bot. Ital. nov. ser., 3: 370 (1896).

*Centaurea ferulacea* grown in a small area that lies south of the *C. filiformis*, as an appendix to the southern margin of the large calcareous formations of central Sardinia.

Both species are morphologically very similar, as evident from the iconography of Moris (1840-43) and Martelli (1896b). The two species differs almost exclusively the form of involucral bracts the capitula, which are combed-ciliate in *C.filiformis*, scabrid and brown ferrugineous, and are scariose and irregularly fimbriate lacerate in *C. ferulacea*. This type of differentiation of involucral bracts occurs frequently in *Centaurea*.

From an ecological point of view, *Centaurea ferulacea*, does not show different needs from those of *C. filiformis*; it is a calcicola limestone rock plant. The chromosomal number is identical in the two taxa (2n=18, Arrigoni and Mori, 1971), both with two pairs of chromosomes with satellites.

Both *Centaurea* are considered rocky endemics of the mesozoic limestones of middle-west Sardinia. Both entities are allopatric, but show, in the transition zone between the areas, some topodems morphologically *intergrading*, although constituted by homogenous individuals. Arrigoni (1972) considers that *C. filiformis* and *C. ferulacea* constitute an unique ologamodemus, and consequently that the following taxonomic framing of the two entities can be justified: *Centaurea filiformis* Viv. ssp. *filiformis* and *C. filiformis* Viv. ssp. *ferulacea* (Martelli) Arrig. = (*Centaurea ferulacea*).

The distinctions between sections are based mainly on the characteristic of an appendage of involucral bracts.

In the molecular level there are different phylogenetic studies including species of the subgroup *Acrolophus* (Susanna *et al*, 1995; Garcia- *et al*., 2000; 2006; Suárez-Santiago *et al*. 2007) using ribosomal nuclear DNA. Ribosomal DNA (rDNA) present three ribosomal gene subunits that are very conservative throughout organisms, and are useful in phylogenetic analyses at broad levels. The internal transcribed region (ITS) is more divergent in their nucleotide sequences; Baldwin (1992) used sequences of ITS region to study evolution in the Asteraceae.

Comparison of the ITS region has clarified phylogenetic relationship among many putative closely related species in diverse lineage of Asteraceae (Baldwin, 1992; Susanna *et al.,* 1999; Vilatersana *et al*., 2000) particularly in *Centaurea* genus (Susanna *et al*.,1995; Garcìa-Jacas *et al*., 2000; 2001; 2002). The aims of these study are clear up the phylogenetic position of species within *Centaurea* and investigate their possible hybrids which will help us to solve the complex systematic problems, that this group of plants possess.

***Material and Methods***:

**Plant material**:

Samples used for this analysis were collected respectively: for *Centaurea horrida* two populations from the island of Asinara, two from peninsula of Stintino, two from Alghero and one from Tavolara; for *Centaurea filiformis* one population for Oliena, one for Cartoe and one for Tavolara isle populations, for *Centaurea ferulacea* the only population for Baunei and finally all the individuals of the hybrid. The populations we have collected cover the entire previously known distribution area. As a reference, we have included a representation of the *Acrolophus-Phalolepis-Willkommia* complex.

The sequences of these species were taken from previous studies (Garcia-Jacas *et al.*, 2006), with the exception of *C. aeolica* that was downloaded from GenBank. The outgroup species were chosen among *Centaurea* section *Jacea*, which is sister to the *Acrolophus-Phalolepis-Willkommia* complex (Garcia-Jacas *et al*., 2006). Voucher and GenBank accession numbers are given in (Tab. 1).

*DNA extraction, amplification and sequencing*:

For each sample of field-collected leaf tissue (kept on ice or frozen in liquid nitrogen and subsequently stored at 80°C), total genomic DNA was extracted and purified, approximately (100mg), by grinding the frozen leaves in a mortar in liquid $N_2$ and by using the DNeasy Plant Mini Kit (Qiagen, Italy), according to the manufacturer's instructions. The average concentration of the extracted DNA was 20 ng/µL.

*nrDNA ITS region strategies*:

Double-stranded DNA of the ITS region was amplified using the 17SE, forward, and the 26SE, reverse, primers (Sun *et al*. 1994). The primer sequences are the following:

17SE F: ACGAATCGGTGAAGTGTTCGTCATGGTC;

26SE R: TAGAATTCCCCGGTTCGCTCGCCGTTAC.

Amplification reactions were modified in a total volume of 25 µL, containing 10mM 10X PCR buffer, 25 mM $MgCl_2$ solution, 20mM of each dNTP, 25pmol/ µL of each forward and reverse primer, 25 ng genomic DNA, one unit of Ampli*Taq*® polymerase (Applied Biosystems Foster City, CA) and DMSO [Dimethyl sulfoxide] (Sigma-Aldrich, Schnellidorf, Germany). Amplifications were carried out in a PTC-200 thermal cycler (MJ Research, Watertown, MA, USA).

The profile used for amplification included a warm start at 94°C for 2 min, followed by 35 cycles of 94°C denaturing for 1 min 30s, 57°C annealing for 2 min and 72°C

extension for 3 min, with an additional extension step of 15 min at 72°C (Galbany-Casals *et al*. 2004).

Double-stranded PCR products were purified with QIAquick® Purification Kit (Qiagen Inc., Valencia, CA, USA) and sequenced with the primers 17SE as forward primer and 26SE as reverse.

The sequences obtained in the first instance were unclear, and there afther it was necessary to clone the regions ITS 17SE/26SE. The PCR products of all species were cloned using TOPO TA Cloning kit (Invitrogen, Carlsbad, CA) following the manufacturer's instruction, except that only half reactions were used. When possible, 8 positive colonies from each reaction were screened with direct PCR using T7 and M13R universal primers under the following conditions: 10 min at 94°C, followed by 30 cycles of 94°C for 30 s, 55°C for 1 min and 72°C for 2 min, ending with 10 min at 72°C (Vilatersana *et al.*, 2007).

Direct sequencing of the amplified DNA segments was performed with a "Big Dye® Terminator v3.1 kit" (Applied Biosystems, Foster City, CA, USA), following the protocol recommended by the manufacturer. Nucleotide sequencing was carried out at the "Serveis Científico-Tècnics" of the University of Barcelona on an ABI PRISM 3700 DNA analyzer (Applied Biosystems, Foster City, CA, USA). Sequences were edited using BioEdit version 5.0.6 (Tom Hall, Noth Caroline State University, Department of Microbiology).

**Phylogenetic analysis**:

Sequences were aligned visually by sequential pairwise comparison (Swofford & Olsen, 1990). The data matrices are available on request from the author.

Phylogenetic analyses were performed using two optimality criteria: Maximum parsimony (MP) and Bayesian inference optimality criteria (BI).

Parsimony analyses of the ITS dataset involved heuristic searches conducted with PAUP version 4.0b10 (Swofford, 2002) using tree-bisection-reconnection (TBR); branch swapping with character states specified as unordered and unweighted. The indels were treated as missing data. All most-parsimonious trees (MPTs) were saved. To locate other potential islands of most-parsimonious trees (Maddison, 1991), we performed 100 replications with random taxon addition, also with TBR branch swapping. Consistency index (CI) and retention index (RI) are always given excluding uninformative characters. Bootstrap analyses (BS) (Felsenstein, 1985) were performed with 100 replications, with simple taxon addition and TBR branch swapping,

Bayesian inference (BI) estimation was calculated using MrBayes 3.1.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). The best-available model of molecular evolution required for Bayesian estimations of phylogeny was selected using hierarchical Likelihood Ratio Tests (hLRT) and Akaike Information Criteria (AIC) (Akaike, 1973) as implemented in the software MrModeltest 2.2 (Nylander, 2004), which considers only nucleotide substitution models that are currently implemented in PAUP and MrBayes 3.1.2. Four Markov chains were run simultaneously for 1,000,000 generations and sampled every 100 generations. Data from the first 1000 generations were discarded as the "burn-in" period, after confirming that likelihood values had stabilized prior to the $1000^{th}$ generation. The 50% majority rule consensus phylogeny and posterior probability of nodes (PP) were calculated from the remaining sample.

***Results:***

The results of this study are very preliminary and we will not discuss them in depth.

Numerical results of the analyses are shown in (Tab. 2). Both parsimony and Bayesian inference analyses showed highly congruent topologies for each dataset. Therefore, for

each dataset, we shall comment both the Bayesian inference an Parsimony strict consensus tree (Figs. 1 and 2).

*Phylogenetic analysis*:

The trees show that the complex *Acrolophus-Phalolepis-Willkommia* and the sections *Maculosae* and *Horridae* form a monophyletic group (BS = 95%; PP = 1.00).

There are two separate clades. The first one is mostly formed by sect. *Willkommia* (BS = 95%; PP = 1.00), and the second includes the remaining of *Acrolophus-Phalolepis*, *Maculosae* and *Horridae* (BS = 95%; PP = 1.00).

These results confirm that the sect. *Maculosae* is not independent from the other sections (Garcia-Jacas *et al*., 2006) and that *Centaurea filiformis* which is included in the sect. *Maculosae* should be placed in the *Acrolophus-Phalolepis* complex.

*Centaurea horrida*, now considered the only member of sect. *Horridae*, is also part of the *Acrolophus-Phalolepis* complex. Thereafter it makes no sense to keep this separate section.

Finally, in this clade sections *Acrolophus* and *Phalolepis* appear intermixed, which confirms the difficulties of morphological differentiation of this group of taxa (Wagenitz, 1989).

*Hybridization*

In the two represented trees, the purported hybrid *Centaurea horrida* × *Centaurea filiformis* (Figs. 1 and 2) is included in the clade of sect. *Acrolophus-Phalolepis* (BS = 95%; PP = 1.00) and is placed in the subclade that included the parental species. Support for the branches within this clade is lower, which is also an indicator of introgression. Sequences of the ITS region also show many informative nucleic substitutions indicating hybridization between the suggested parental species (Fig. 3).

The pattern of rybotypes found in one of the parental species, *C. filiformis*, is extremely complicated and constitutes a proof of ancient hybridization events. One population presents three different ribotypes, which are in turn different from a fourth ribotype that is present in the other populations (Figs. 1 and 2). One population of *Centaurea filiformis* (FI4, Figs. 1 and 2) also appears very closely connected to *C. ferulacea*.

The objectives for the future are to deepen the analysis through addition of other ITS sequences and use of other nuclear and organellar markers such as the ETS region and non-coding plastid regions like the *rps*16 and the *trn*T.

***References:***


**Arrigoni PV. 1972.** Sulla distribuzione e il rango sistematico di *Centaurea filiformis* Viviani e *Centaurea ferulacea* Martinelli. *Webbia* 27: 279-287.

**Arrigoni PV, Camarda I. 2003.** *Centaurea magistrorum* species nova (Asteraceae) di Sardegna. *Parlatorea* 6: 79-82.

**Arrigoni PV, Mori B. 1971.** Numeri cromosomici per la Flora Italiana. n. 93. *Centaurea ferulacea* Martinelli. *Informatore Botanico Italiano* 3: 226.

**Akaike H. 1973.** Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. *Second International Symposium on Information Theory*. Budapest: Akademial Kiado, 267-281.

**Colas B, Olivieri I, Riba M, 2001**. Spatio-temporal variation of reproductive success and conservation of the narrow-endemic *Centaurea corymbosa* (Asteraceae). *Biological Conservation* 99: 375–386.

**Cardona MA, Contandriopoulos J. 1979.** *Endemism and evolution in the island of the Western Mediterranean*. In *Plants and islands*, D. Bramwell (ed.). London, Academic Press, 133-169.

**Contandriopoulos J, Cardona MA. 1984.** Caractère original de la flore endémique des Baléares. *Botanica Helvetica* 94: 101-132.

**Contandriopoulos J. 1981**. Endemisme et origine de la flore de la Corse: mise au point des connaisances actuelles. *Bollettino Società Sarda Scienze Naturali* 20: 187-230.

**Desole L. 1954.** Secondo contributo alla conoscenza dello sviluppo embriologico del genere *Centaurea* L. (Asteraceae): *Centaurea horrida* Bad. *Nuovo Giornale Botanico Italiano* 61**:** 256-273.

**Desole L**. **1956**. Nuove stazioni e distribuzione geografica della "*Centaurea horrida*" Bad. *Webbia* 12: 251-324.

**Dostál J. 1976.** *Centaurea* L. In: Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA. *Flora Europaea*. London-New York-Melbourne: Cambridge University Press, 254-301.

**Felsenstein J. 1985.** Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39**:** 783–791.

**Fiori A, Paoletti G. 1903-1904.** Flora Analitica d'Italia 3. Edagricole

**Frankham R. 1998.** Inbreeding and extinction: island population. *Conservation Biology* 12**:** 665-675.

**Galbany-Casals M, Garcia-Jacas N, Susanna S, Sáez L, Benedí C. 2004.** Phylogenetic relationships in the Mediterranean *Helichrysum* (Asteraceae, Gnaphalieae) based on nuclear rDNA ITS sequence data. *Australian Systematic Botany* 17: 241-253.

**Garcia-Jacas N, Susanna A, Mozaffarian V, Ilarslan R. 2000**. The natural delimitation of *Centaurea* (Asteraceae: Cardueae): ITS sequence analysis of the *Centaurea jacea* group. *Plant Systematics and Evolution* 223: 185–199.

**Garcia-Jacas N, Susanna A, Garnatje T, Vilatersana R, 2001**. Generic delimitation and phylogeny of the subtribe Centaureinae (Asteraceae): a combined nuclear and chloroplast DNA analysis *Annals of Botany (London)* 87: 503–515.

**Garcia-Jacas N, Garnatje T, Susanna A,Vilatersana R. 2002.** Tribal and subtribal delimitation and phylogeny of the Cardueae (Asteraceae): a combined nuclear and chloroplast DNA analysis. *Molecular Phylogenetics and Evolution* 22: 51-64.

**Garcia-Jacas N, Uysal T, Romashchenko KY, Suárez-Santiago VN, Ertugrul K, Susanna A. 2006**. *Centaurea* revisited: a molecular survey of the *Centaurea jacea* group. *Annals of Botany (London)* 98: 741–753.

**Greuter W. 1979**. The origin and evolution of island floras as exemplified by Aegean archipelago. In *Plants and islands*, D. Bramwell (ed.). London, Academic Press, 87-106.

**Hellwig FH. 2004.** Centaureinae (Asteraceae) in the Mediterranean – history of ecogeographical radiation. *Plant Systematics and Evolution* 246: 137-162.

**Huelsenbeck JP, Ronquist F. 2001**. MRBAYES: Bayesian inference of phylogenenetic trees. *Bioinformatics* 17: 754–755.

**Maddison, D.R, 1991.** The discovery and importance of multiple islands of most-parsimonious trees. *Systematic. Zoology*. 40, 315–328.

**Mameli G, Filigheddu R, Binelli G, Meloni M. In press.** The genetic structure of the remnant populations of *Centaurea horrida* Badarò in Sardinia, a major island of the Mediterranean Sea. *Annals of Botany (London).*

**Martelli U. 1896**. *Centaurea ferulacea* n. sp. sect. *Phalolepis*. Nuovo Giornale Botanico Italiano nov. ser., 3: 370-371.

**Médail F, Quézel P. 1999.** Biodiversity hot-spots in the Mediterranean Basin: setting global conservation priorities. *Conservation Biology* 13: 1510-1513.

**Moris JH. 1840-43.** Flora sardoa. 2: 455-456. Reg. Typ., Taurini.

**Ochsmann J. 2000.** Morphologisque und molekularsystematische Untersuchungen an der *Centaurea stoebe* L.- Gruppe (Asteraceae-Cardueae) in Europa. *Dissertationes Botanicae* 324.

**Nylander JA. 2004.** Mr Modeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

**Ronquist F, Huelsenbeck JP. 2003**. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19**:**1572-1574.

**Rosenbaum G, Lister GS, Duboz C. 2002a.** Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. *Journal of the Virtual Explorer* 8: 107-130.

**Standards and Petitions Working Group. 2006.** Guidelines for Using the IUCN Red List Categories and Criteria: version 6.1. Prepared by the Standards and Petitions Working Group for the IUCN SSC Biodiversity Assessments Sub-Committee in July 2006. Downloadable from webfiles/doc/SSC/RedList/RedListGuidelines.pdf.

**Susanna A, Garcia-Jacas N, Soltis DE, Soltis PS. 1995.** Phylogenetic relationship in the tribe Cardueae (Asteraceae) based on ITS sequences. *American Journal of Botany* 82: 1056-1068.

**Susanna A, Garnatje T, Garcia-Jacas N. 1999.** Molecular phylogeny of *Cheirolophus* (Asteraceae) based on ITS sequences ribosomal DNA. *Plant Systematics and Evolution* 214**:** 147-160.

**Suárez-Santiago VN, Salinas MJ, Garcia-Jacas N, Soltis, PS, Soltis DE, Blanca G. 2007**. Reticulate evolution in the *Acrolophus* subgroup (*Centaurea* L., Compositae) from the western Mediterranean: origin and diversification of section *Willkommia* Blanca. *Molecular Phylogenetics and Evolution* 43: 156-172.

**Sun Y, Skinner DZ, Liang GH, Hulbert SH. 1994.** Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics* 89: 26-32.

**Swofford DL. 2002.** PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10. Sinauer Associates, Sunderland, MA.

**Swofford DL, Olsen GJ. 1990.** Phylogeny reconstruction. In: Hillis D, Moritz C, eds. *Molecular systematics*. Sunderland, MA: Sinauer Associates, 411-501.

**Pisanu S, Filigheddu R, Farris E. In press.** The conservation status of an endemic species of northern Sardinia: *Centaurea horrida* Badarò (Asteraceae). *Plant Biosystems.*

**Vilatersana R, Susanna A, Garcia-Jacas N, Garnatje T. 2000.** Generic delimitation and phylogeny of the *Carduncellus-Carthamus* complex (Asteraceae) based on ITS sequences. *Plant Systematics and Evolution* 221: 89-105.

**Vilatersana R, Brysting AK, Brochmann C. 2007.** Molecular evidence for hybrid origins of the invasive polyploids *Carthamus creticus* and *C. turkestanicus* (Cardueae, Asteraceae). *Molecular Phylogenetics and Evolution* 44: 610-621.

**Valsecchi F, Filigheddu R. 1991.** *Centaurea corensis* Valsecchi & Filigheddu, sp. nova (Compositae) in Sardegna. *Webbia* 45: 235-239.

**Viviani D. 1825.** *Appendix ad florae Corsicae prodromum*. Genoa.

**Wagenitz G. 1989**. Nahe Verwandtschaft zwischen Arten der *Centaurea*-Sektionen *Acrolophus* und *Phalolepis*. *Flora* 182: 341-351

**CAPTIONS TO FIGURES:**

Fig. 1. Majority-rule consensus tree based on Monte Carlo markov chains. Numbers above branches are Bayesian posterior probability.

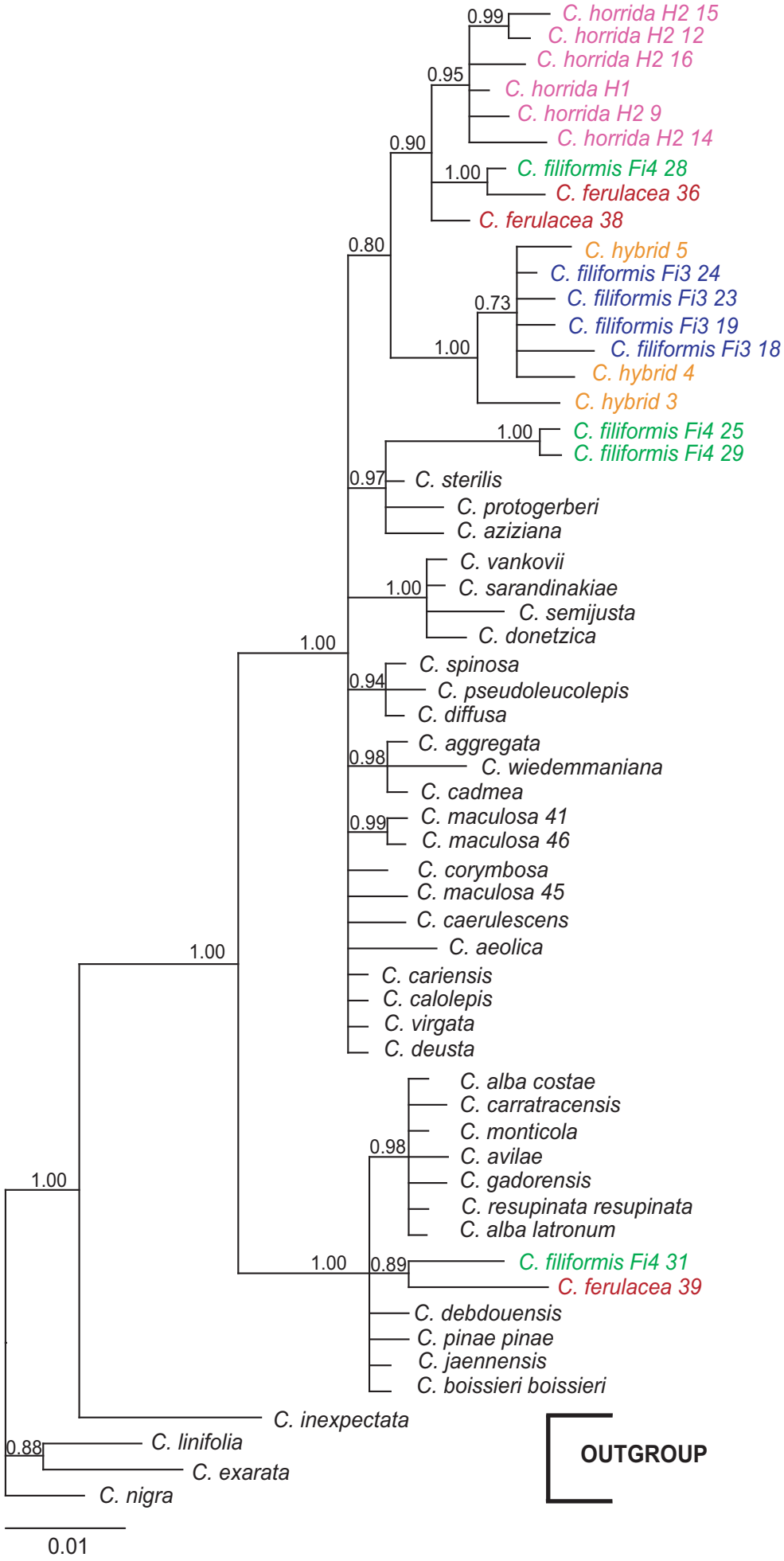
Fig. 2. Strict consensus tree of the 16 most parsimonius trees generated by the ITS matrix. Numbers above branches are bootstrap values.

Fig. 3 Sequences of *C. horrida*, *C. filiformis*, and *C. hybrid* showing the nucleotide site variations.

**CAPTIONS TO TABLE:**

Table 1. Origin of the materials, herbaria where the vouchers are deposited and GenBank accession numbers

Table 2. Numeric results of the phylogenetic analyses


0.99
C. horrida H2 15
C. horrida H2 12
C. horrida H2 16
0.95
C. horrida H1
C. horrida H2 9
C. horrida H2 14
0.90
1.00
C. filiformis Fi4 28
C. ferulacea 36
C. ferulacea 38
0.80
C. hybrid 5
C. filiformis Fi3 24
0.73
C. filiformis Fi3 23
C. filiformis Fi3 19
1.00
C. filiformis Fi3 18
C. hybrid 4
C. hybrid 3
1.00
C. filiformis Fi4 25
C. filiformis Fi4 29
0.97
C. sterilis
C. protogerberi
C. aziziana
C. vankovii
1.00
C. sarandinakiae
C. semijusta
C. donetzica
1.00
C. spinosa
0.94
C. pseudoleucolepis
C. diffusa
C. aggregata
0.98
C. wiedemmaniana
C. cadmea
C. maculosa 41
0.99
C. maculosa 46
C. corymbosa
C. maculosa 45
C. caerulescens
1.00
C. aeolica
C. cariensis
C. calolepis
C. virgata
C. deusta
C. alba costae
C. carratracensis
C. monticola
0.98
C. avilae
C. gadorensis
C. resupinata resupinata
C. alba latronum
1.00
1.00
0.89
C. filiformis Fi4 31
C. ferulacea 39
C. debdouensis
C. pinae pinae
C. jaennensis
C. boissieri boissieri
C. inexpectata
0.88
C. linifolia
C. exarata
C. nigra
OUTGROUP
0.01

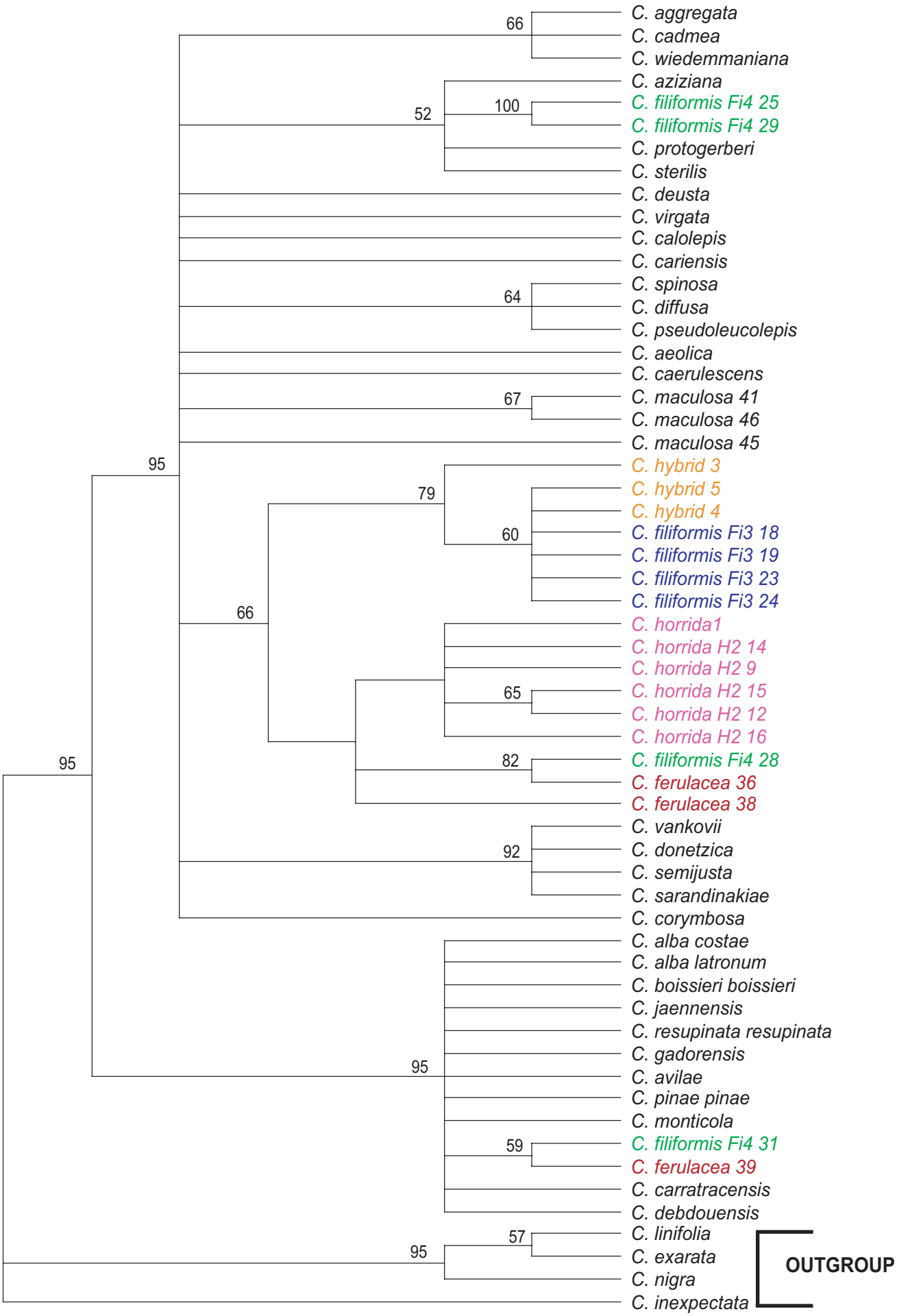
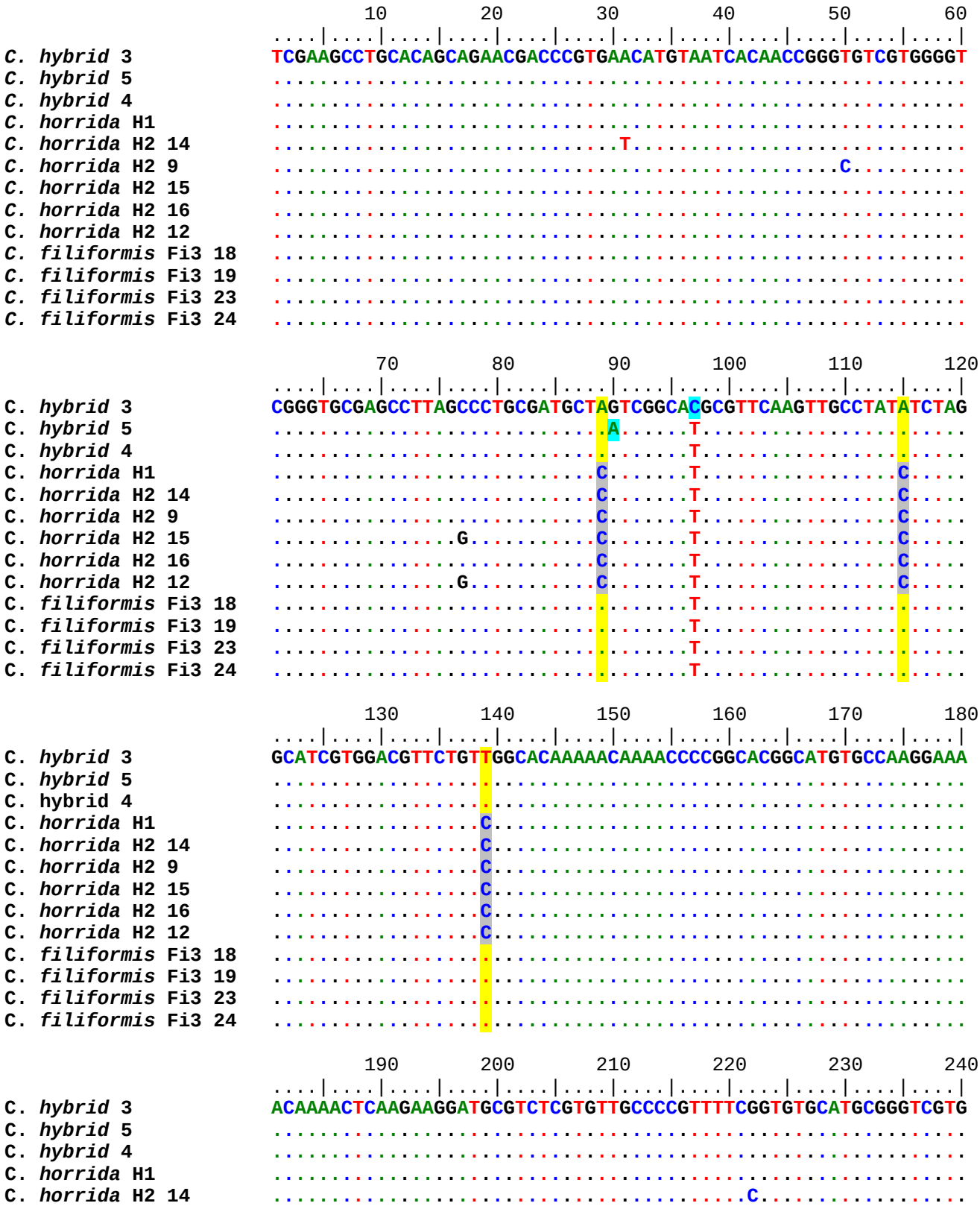
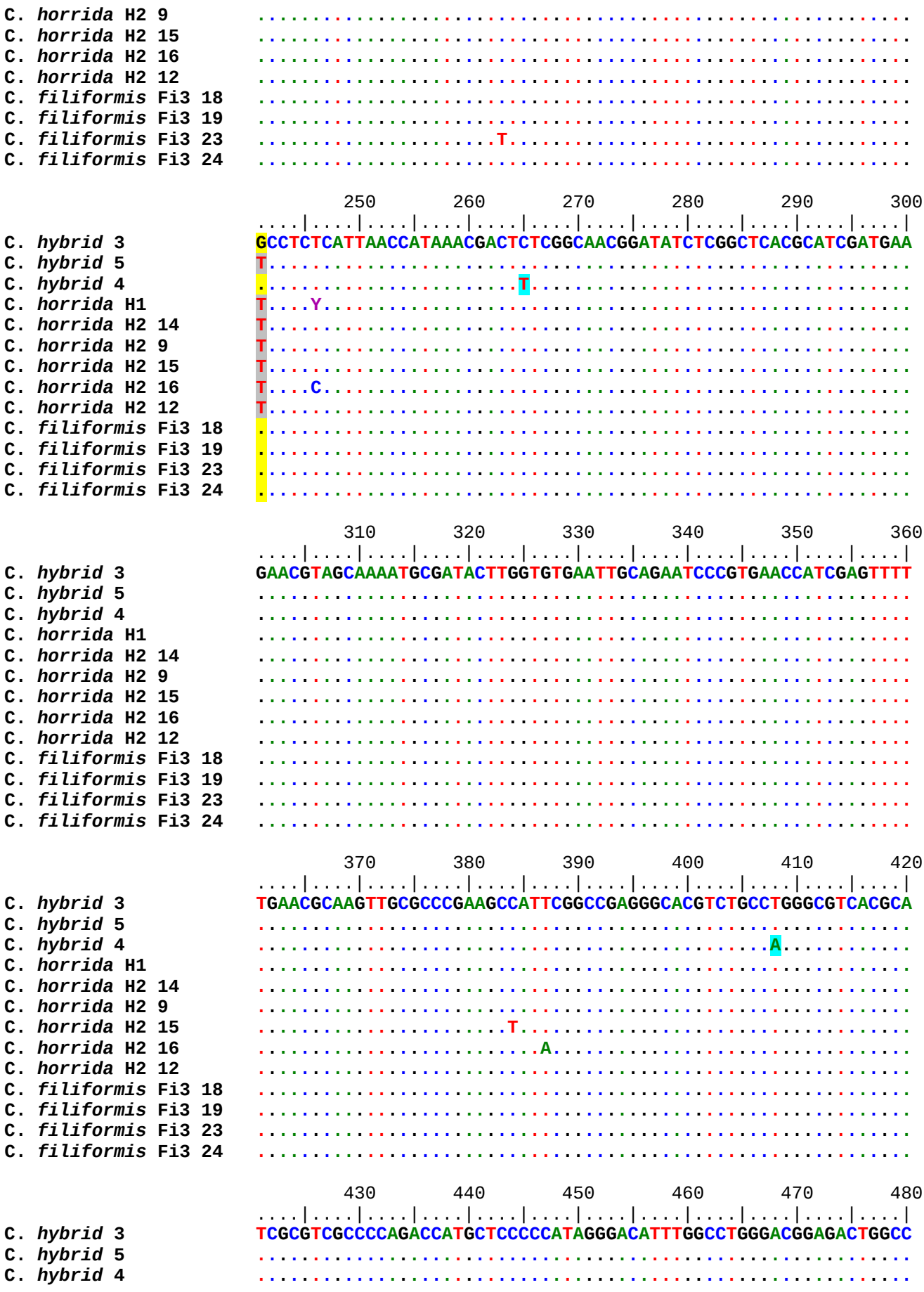

C. aggregata
C. cadmea
C. wiedemmaniana
C. aziziana
C. filiformis Fi4 25
C. filiformis Fi4 29
C. protogerberi
C. sterilis
C. deusta
C. virgata
C. calolepis
C. cariensis
C. spinosa
C. diffusa
C. pseudoleucolepis
C. aeolica
C. caerulescens
C. maculosa 41
C. maculosa 46
C. maculosa 45
C. hybrid 3
C. hybrid 5
C. hybrid 4
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
C. horrida1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 12
C. horrida H2 16
C. filiformis Fi4 28
C. ferulacea 36
C. ferulacea 38
C. vankovii
C. donetzica
C. semijusta
C. sarandinakiae
C. corymbosa
C. alba costae
C. alba latronum
C. boissieri boissieri
C. jaennensis
C. resupinata resupinata
C. gadorensis
C. avilae
C. pinae pinae
C. monticola
C. filiformis Fi4 31
C. ferulacea 39
C. carratracensis
C. debdouensis
C. linifolia
C. exarata
C. nigra
C. inexpectata
OUTGROUP

**Fig. 3** Sequences of *C. horrida*, *C. filiformis*, and *C. hybrid* showing the nucleotide site variations. Yellow nucleotides shared by *C. hybrid* and *C. filiformis*; Grey nucleotides shared by *C. hybrid* and *C.* horrida; Light blu autoapomorphic nucleotides of the hybrid.




C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
250 260 270 280 290 300
C. hybrid 3 GCCTCTCATTAACCATAAACGACTCTCGGCAACGGATATCTCGGCTCACGCATCGATGAA
C. hybrid 5
C. hybrid 4
C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
310 320 330 340 350 360
C. hybrid 3 GAACGTAGCAAAATGCGATACTTGGTGTGAATTGCAGAATCCCGTGAACCATCGAGTTTT
C. hybrid 5
C. hybrid 4
C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
370 380 390 400 410 420
C. hybrid 3 TGAACGCAAGTTGCGCCCGAAGCCATTCGGCCGAGGGCACGTCTGCCTGGGCGTCACGCA
C. hybrid 5
C. hybrid 4
C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
430 440 450 460 470 480
C. hybrid 3 TCGCGTCGCCCCAGACCATGCTCCCCCATAGGGACATTTGGCCTGGGACGGAGACTGGCC
C. hybrid 5
C. hybrid 4

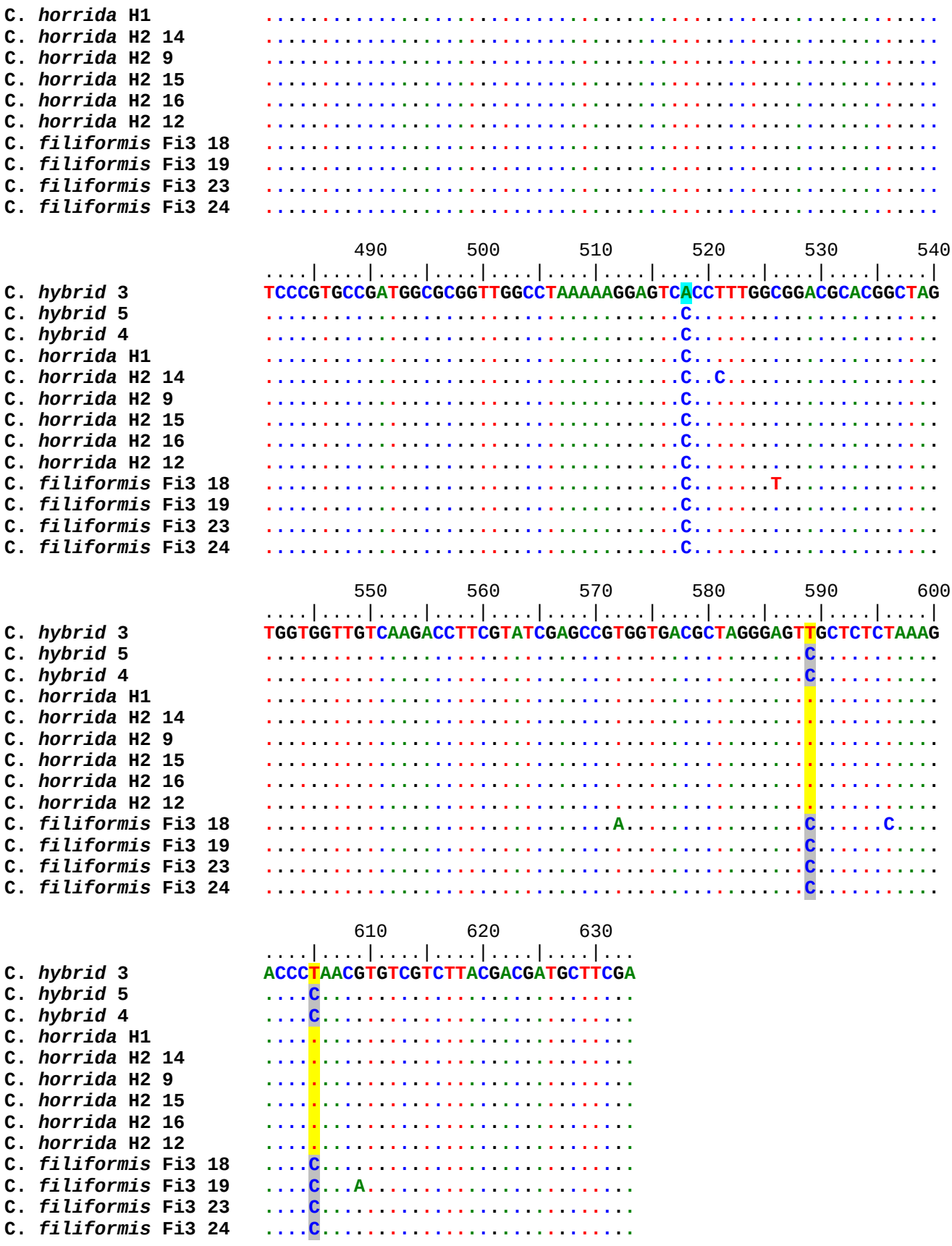

C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
490 500 510 520 530 540
C. hybrid 3 TCCCGTGCCGATGGCGCGGTTGGCCTAAAAAGGAGTCACCTTTGGCGGACGCACGGCTAG
C. hybrid 5
C. hybrid 4
C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
550 560 570 580 590 600
C. hybrid 3 TGGTGGTTGTCAAGACCTTCGTATCGAGCCGTGGTGACGCTAGGGAGTTGCTCTCTAAAG
C. hybrid 5
C. hybrid 4
C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24
610 620 630
C. hybrid 3 ACCCTAACGTGTCGTCTTACGACGATGCTTCGA
C. hybrid 5
C. hybrid 4
C. horrida H1
C. horrida H2 14
C. horrida H2 9
C. horrida H2 15
C. horrida H2 16
C. horrida H2 12
C. filiformis Fi3 18
C. filiformis Fi3 19
C. filiformis Fi3 23
C. filiformis Fi3 24

Table 1. Origin of the materials, herbaria where the vouchers are deposited and GenBank accession numbers

| SPECIES | RANGE | VOUCHER | ITS ACCESSION |
|---|---|---|---|
| *Centaurea aggregata* Fisch. & C. A. Mey. ex DC. | Caucasus, Iran, Turkey (weed) | Turkey, Adana: Ala Dağ above Dağdibi, 2000 m, *Ertuğrul, Garcia-Jacas, Susanna 2305 & Uysal*, 3.8.2002 (BC). | DQ319077 |
| *Centaurea alba* L. subsp. *costae* (Willk.) Dostál | Iberian Peninsula endemic | Spain, Huesca: Peña de Oroel, *Fernández-Galiano & Rivas Goday 23733*, 15.7.1947 (GDA). | AM114325 |
| *Centaurea alba* L. subsp. l*atronum* (Pau) Dostál | Iberian Peninsula endemic | Spain, Ávila: La Adrada, *Sánchez-Mata & Cantó 24946*, 27.7.1982 (GDAC). | AM114326 |
| *Centaurea aeolica* Guss. ex DC. | Italian Peninsula | - | AM117057 |
| *Centaurea avilae* Pau | Iberian Peninsula endemic | Spain, Ávila: Sierra de Gredos, *Blanca 6087*, 30.7.1979 (GDAC). | AM114309 |
| *Centaurea aziziana* Rech. f | Turkey endemic | Iran, Azarbayjan-e-Sharghi : between Tatar and Golfa, 85 km from Golfa, *Garcia-Jacas, Mozaffarian, Susanna 1680 & Vallès*, 7.8.1996 (BC). | DQ319089 |
| *Centaurea boissieri* DC. *subsp. boissieri* | Iberian Peninsula endemic | Spain, Granada: Sierra de Cázulas, *Blanca 6597*, 8.6.1979 (GDAC). | AM114278 |
| *Centaurea cadmea* Boiss. | Turkey endemic | Turkey, Burdur: 4 km from Burdur on the road to Sparta, mountains above Burdur, 1200 m, *Ertuğrul, Garcia-Jacas, Susanna 2249 & Uysal*, 28.7.2002 (BC). | DQ319094 |
| *Centaurea caerulescens* Willd. | Francia | Francia, Hérault: Cirque de Labeil, sobre la gruta, prados **V-1046** *Centaurea* cf *coerulescens Noemí Montes-Moreno & Roser Vilatersana* 20-07-07 | - |
| *Centaurea calolepis* Boiss. | Turkey endemic | Turkey, Burdur-Muğla: Dirimli mountain pass, 1600 m, *Ertuğrul, Garcia-Jacas, Susanna 2254 & Uysal*, 29.7.2002 (BC). | DQ319095 |
| *Centaurea cariensis* Boiss. | Turkey endemic | Turkey, Antalya: 40 km from Elmalı on the road to Korkuteli, N slopes of the Karamanbeli mountain pass, 1400 m, *Ertuğrul, Garcia-Jacas, Susanna 2258B & Uysal*, 30.7.2002 (BC). | DQ319097 |
| *Centaurea carratracensis* Lange | Iberian Peninsula endemic | Spain, Málaga: Carratraca, Sierra de Aguas, *Blanca 42802*, 4.7.1998 (GDAC). | AM114302 |
| *Centaurea corymbosa* Pourr. | S France endemic | France, Narbonne: La Clappe, *M. Riba*, 1995 (BC). | DQ319103 |
| *Centaurea debdouensis* Breitw. & Podlech | Morocco endemic | Morocco, Debdou: Gaada de Debdou, *Pasquier & Ch. Rungs*, 18.6.1954 (MPU). | AM114317 |
| *Centaurea deusta* Ten. | Italy endemic | Italy, Calabria: Crotone, Torrente Matassa near Caccuri, 360 m, *Vogt 15531*, Berlin Botanical Garden, Index Seminum 1997. | DQ319107 |
| *Centaurea diffusa* Lam. | Widespread (weed) | Armenia, Talin : between vil. Pokr Arthik and Bagravan, *Fajvush, Gabrielyan, Garcia-Jacas, Guara, Hovhannisyan, Susanna 1589, Tamanyan & Vallès*, 26.8.1995 (BC). | DQ319108 |
| *Centaurea donetzica* Klokov | Ukraine endemic | Ukraine, Donetzkaya: Krasny Liman, *Romashchenko*, 12.8.2002 (BC). | DQ319110 |
| *Centaurea exarata* Boiss. ex Coss. | Iberian Peninsula endemic | Spain, Huelva: road A-983, Almonte to Matalascañas km 25, *Roché & Susanna 1909*, 9.7.1999 (BC). | DQ319113 |
| *Centaurea gadorensis* Blanca | Iberian Peninsula endemic | Spain, Almería: Sierra de Gádor, Pico La Estrella, 1730 m, *Martínez Lirola & Salinas 44171*, 29.7.1996 (GDAC). | AM114298 |
| *Centaurea inexpectata* Wagenitz | Turkey endemic | Turkey, Antalya: Gevne valley, high of village Küçüklü, 1750 m, *Uysal 598*, 30.6.2004 (KNYA). | DQ319122 |
| *Centaurea jaennensis* Degen & Debeaux | Iberian Peninsula endemic | Spain, Jaén: Pozo Alcón, La Bolera dam, *Blanca & Varo 6724*, 19.6.1978 (GDAC). | AM114287 |
| *Centaurea linifolia* L. | Eurosiberian | Garcia-Jacas *et al.* (2000). | DQ319129 |

| | | | |
|---|---|---|---|
| *Centaurea maculosa* Lam. | | Italy: Aosta, *Roché 117*, 25.8.99 (BC). | - |
| *Centaurea monticola* Boiss. ex DC. | Iberian Peninsula endemic | Spain, Granada: Pantano del Cubillas, *Blanca 6750*, 6.6.1977 (GDAC). | AM114313 |
| *Centaurea nigra* L. | Eurosiberian | Garcia-Jacas *et al.* (2000). | DQ319138 |
| *Centaurea pinae* Pau var. *pinae* | Iberian Peninsula endemic | Spain, Teruel: Puerto Ragudo, 900 m, *Blanca, Socorro & Valle 6768*, 15.7.1978 (GDAC). | AM114310 |
| *Centaurea proto-gerberi* Klokov | Ukraine endemic | Ukraine, Luganskaya: Stanichno-Lugansk, *Romashchenko*, 5.9.2002 (BC). | DQ319149 |
| *Centaurea pseudoleucolepis* Kleopow | Ukraine endemic | Ukraine, Donetzkaya: Kamennye Mogily national reservation, *Romashchenko*, 1.8.2002 (BC). | DQ319150 |
| *Centaurea resupinata* Coss. subsp. *resupinata* | Iberian Peninsula endemic | Spain, Albacete: between Elche de la Sierra and Hellín, Cenajo dam, *Blanca & Varo 6714*, 6.7.1977 (GDAC). | AM114288 |
| *Centaurea sarandinakiae* N. B. Illar | Ukraine endemic | Ukraine, Crimea: Planerskoe, Kara-Dag mountain, *Romashchenko*, 16.8.2002 (BC). | DQ319160 |
| *Centaurea semijusta* Juz. | Ukraine endemic | Ukraine, Crimea: Simferopol, Chatyr-Dag mountain, *Romashchenko*, 1.9.2002 (BC). | DQ319162 |
| *Centaurea spinosa* L. | Aegean | Greece, Thrakia: Nomos Evrou, Samothraki, 2 m, *Raus/Sch* 18942, Berlin Botanical Garden, Index Seminum 1997. | DQ319165 |
| *Centaurea sterilis* Stev. | Ukraine endemic | Ukraine, Crimea: Planerskoe, Kara-Dag mountain, *Romashchenko*, 16.8.2002 (BC). | DQ319167 |
| *Centaurea vankovii* Klokov | Ukraine endemic | Ukraine, Crimea: Alupka, Ai-Petri mountain, *Romashchenko*, 30.8.2002 (BC). | DQ319173 |
| *Centaurea virgata* Lam. | Turkey endemic | Turkey, Muğla: Köyceğiz district, Sandras Dag range 13 km from Ağla, 1700 m, *Ertuğrul, Garcia-Jacas, Susanna 2252 & Uysal*, 29.7.2002 (BC). | DQ319174 |
| *Centaurea wiedemanniana* Fisch. & C. A. Mey. | Turkey endemic | Turkey, Bilecik: Selimiye, between Osmaneli and Bilecik, 100 m, *Davis & Coode*, 1.7.1962 (E). | DQ319175 |
| *Centaurea filiformis* Viv. | Sardinian endemic | Italy, Sardinia: Dorgali, Cartoe, *Filigheddu R.*, 30.4.2007 | - |
| *Centaurea horrida* Badarò | Sardinian endemic | Italy, Sardinia: Asinara island Piano degli Stretti, *Pisanu S.* 8.5.2007 | - |
| *Centaurea filiformis* × *Centaurea horrida* (Hybrid) | Sardinian endemic | Italy, Sardinia: Tavolara island, *Mameli &Pisanu,* 21.5.2007 | - |
| *Centaurea ferulacea* Martelli | Sardinian endemic | Italy, Sardinia: Baunei, *Mameli &Pisanu,* 12.10.2007 | - |

Table 2. Numeric results of the phylogenetic analyses.

| **Data set** | **ITS** |
|---|---|
| Taxa | 40 |
| Number of sequences | 58 |
| Total characters | 639 |
| Informative characters | 46 |
| Number MPTs | 16 |
| Number of steps | 66 |
| Consistency index (CI) | 0.6944 |
| Retention index (RI) | 0.9211 |
| Range of divergence, ingroup (%) | 0-0.4783 |
| Range of divergence, ingroup-outgroup (%) | 0-0.5652 |
| Model | GTR |

# APPENDIX

# HYBRIDIZATION BETWEEN *Centaurea horrida* AND *Centaurea filiformis* (ASTERACEAE) AS REVEALED BY SSR (Simple Sequence Repeat) AND ISSR (INTER-SIMPLE SEQUENCE REPEATs) markers.

Mameli, G.[1]; Farris, E.[1]; Filigheddu, R.[1]; Binelli, G.[2]

[1]*Dipartimento di Botanica ed Ecologia vegetale, University of Sassari, Via Muroni 25, 07100 Sassari (Italy)*

[2]*Dipartimento di Biotecnologie e Scienze Molecolari, University of Insubria, Via J. H. Dunant 3, 21100 Varese (Italy)*

Author for correspondence: filighed@uniss.it

**Individual cases of natural hybridization are analyzed because this process is considered to be evolutionary important in its own right. It is important to examine the evolutionary consequences of recombination between divergent genomes.**
***Centaurea horrida* Badarò (Fig. 1) and *Centaurea filiformis* Viviani (Fig. 2) (Asteraceae) are morphologically distinguishable endemic species, whose habitat is restricted to Northern Sardinia (Fig. 3). On the Tavolara Island, where a partial overlap occurs (Fig. 4), many individuals showing morphological features common to both species have been found (Fig. 5) and have been studied either for morphological and genetic traits.**

**Fig. 1 – *Centaurea horrida* Badarò**

**Fig. 2 – *Centaurea filiformis* Viviani**

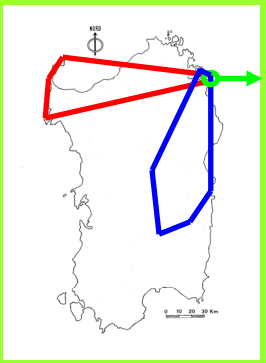


**Fig. 3 – Distribution range of *C. horrida* (red), *C. filiformis* (blue) and overlapping area (green).**

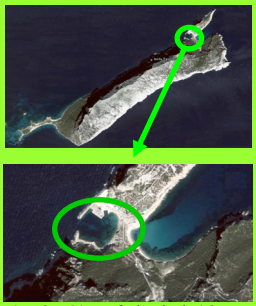


**Fig. 4 – Site of the hybrid population at Cala del Faro on Tavolara isle (40°91'N/09°72'E).**

**Fig. 5 – Hybrid individual under study.**

**The morphological analysis was carried out by examining the variability of capitula (Fig. 6) and ...ves (Fi...**

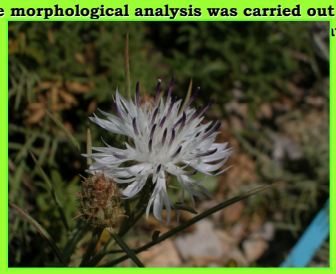


**Fig. 6 – Head of a hybrid individual.**

**Fig. 7 – Leaves of a hybrid individual.**

## Methods:

**Field sampling: samples were collected in November 2006.**
**We extracted the genomic DNA from 30 samples of *C. horrida*, seven of *C. filiformis* and 13 of the intermediate form.**

- **SSR (Simple Sequence Repeat or microsatellite) genetic analysis: due to the lack of information on the genome of the studied species, seven pairs of heterologous microsatellite primers, developed for the congener species *Centaurea corymbosa* (Fréville et al., 2000), were firstly tested on *C. horrida*, and then on *C. filiformis* and hybrid samples. Three of them (28A7, 13D10 and 12B1) have been insofar used to genotype our populations. The amplification products were analysed by a capillary MegaBACE® DNA sequencer. Simple population genetics parameters have been estimated.**

- **Primer names and sequences used in the ISSR (Inter-Simple Sequence Repeats) analysis, number of polymorphic bands *per* primer and range of molecular weight in base pairs (*bp*) amplified by PCR-ISSR. Tm, melting temperature; Ta, annealing temperature.**

| Primer | Sequence (5'-3') | Tm (°C) | Ta (° C) | No. of bands | Size range of bands (bp) |
|---|---|---|---|---|---|
| CH843 | (CT)8-RA | 48 | 55 | 7 | 480-1450 |
| CH844B | (CT)8-RC | 52 | 55 | 20 | 500-1200 |
| OMAR | (GAG)4-RC | 52 | 55 | 8 | 380-1250 |
| DAT | (GA)7-RG | 52 | 55 | 7 | 450-1500 |
| MAO | (CTC)4-RC | 48 | 51 | 8 | 380-1400 |
| UBC809 | $(AG)_8G$ | 52 | 55 | 13 | 380-1000 |
| UBC811 | $(GA)_8C$ | 52 | 55 | 13 | 400-1250 |
| UBC827 | $(AC)_8G$ | 52 | 55 | 14 | 500-1500 |

## Results :

**SSR: Genetic variability.**
**The number of alleles per locus ranged from 8 (28A7) to 16 (13D10).**
**At the 28A7 locus the hybrid samples showed 4 alleles, all shared with *C. horrida* but only one with *C. filiformis*.**
**At the 13D10 locus the hybrid samples showed 7 alleles, among which 2 privates, 4 in common to *C. horrida* and only one to *C. filiformis*.**
**Finally, at the 12B1 locus, the hybrid samples had 6 alleles, among which 1 private, 5 shared with *C. horrida* and none in common with *C. filiformis*.**
**The levels of observed and expected heterozygosity were medium-high; the highest value was found for *C. horrida* (0.862), the lowest value for *C. filiformis* (0.460).**

**Genetic differentiation: The overall genetic divergence between populations was estimated by $F_{ST}$ = 0.22. The lowest pairwise $F_{ST}$ value was found between hybrid and *C. horrida* (0.116), the highest between hybrid and *C. filiformis* (0.204).**

| Pairwise Population *Fst* values | | | |
|---|---|---|---|
| *Hybrid* | *C. filiformis* | *C. horrida* | |
| 0.000 | | | *Hybrid* |
| 0.204 | 0.000 | | *C. filiformis* |
| 0.116 | 0.202 | 0.000 | *C. horrida* |

**ISSR: genetic analysis.**
**Eight out of nine ISSR primers tested gave positive results, in terms of repeatability of amplification and band resolution. We found a number of polymorphic bands from 7 to 20, in the range of 380-1500 *bp*. The number of private bands found in *C. horrida* and *C. filiformis* was 16 and 9, respectively. The morphologically hybrid plants displayed bands from both putative parents.**

**Summary of ISSR products *per* sampling station. NI, number of individuals analysed; TB, number of total bands; UB, number of unique bands**

| | NI | TB | UB |
|---|---|---|---|
| *Hybrid* | 13 | 8 | 7 |
| *C.horrida* | 30 | 19 | 16 |
| *C.filiformis* | 7 | 15 | 9 |

## Future developments:

**Even though preliminary, these results hints to the possibility that the "hybrid" form is a real genetic hybrid between the two species.**
**Our results support the utility of genetic markers for addressing questions of population genetics and taxonomic differentiation also in these endangered, endemic plant species.**
**Ecological, cytogenetic and botanical studies are under way, together with a more detailed genetic analysis, to unerstand yhe nature of this hybrid form, potentially of evolutionary importance**

References:
Fréville, H., E. Imbert, F. Justy, R. Vitalis and I. Olivieri. (2000) Isolation and characterization and microsatellites in the endemic species *Centaurea corymbosa* Pourret (Asteraceae) and other related species. Mol.Ecol. 9: 1671-1672.

1st Congress of the Italian Society for Evolutionary Biology

# Genetic analysis of the populations of the endangered *Centaurea horrida* Badarò

Marilena Meloni[1], Giorgio Binelli[1], Giulia Mameli[2], Rossella Filigheddu[2]

[1]Dipartimento di Biotecnologie e Scienze Molecolari - University of Insubria, via J.H. Dunant 3, 21100 Varese; [2]Dipartimento di Botanica ed Ecologia vegetale – University of Sassari Via Muroni, 25 – 07100 Sassari; ITALY

marilena.meloni@uninsubria.it; filighed@uniss.it

***Centaurea horrida* Badarò** (Asteraceae) is a narrow endemic species located only in Northern Sardinia (Italy), where it occurs in four areas: Asinara, Stintino and Alghero, in the north-western Sardinia; Island of Tavolara, in the north-eastern Sardinia (Fig. 1).

It is protected under the Habitat 92/43 CEE Directive and is in the IUCN Red List.

It is a perennial, pulvinate, spiny species (Fig. 2), whose habitat is restricted to rocky cliffs where it is challenged by harsh environmental conditions, especially related to drought. *Centaurea horrida* promotes itself as a tool to understand the dynamics of genetic variation following the reduction of the habitat. The aim of this study is to estimate the amount and the distribution of the genetic variability of populations of *C. horrida*. Demographic and ecological analyses, on the other hand, will complement the genetic ones to reconstruct the genetic history of the species and to plan appropriate conservation strategies. Results presented here are relative to the north-western populations.

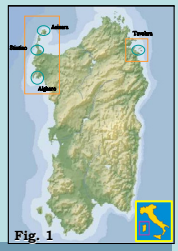


Fig. 1


Fig. 2

## Methods

SSRs (Simple Sequence Repeats) were used to assess the genetic structure of the populations. Seven pairs of microsatellite primers developed for the congener species *Centaurea corymbosa* (Fréville et al., 2000) were tested; four of them yielded simple amplification patterns and were used for genotyping.

Two populations (a cliff-dwelling one and a plain-dwelling one) were analysed for each of the three north-western areas. Green material was collected from about 60 plants per population, for a total of 352 samples. A preliminary analysis was conducted in order to verify whether close plants were clones originated by vegetative reproduction or different individuals. Since it was not possible to observe roots of the plants without damaging the individual itself, we genotyped them. The vegetative reproduction spans a 5m diameter at its maximum, thus we sampled accordingly.

Data were analysed to assess the amount of genetic variability and the degree of differentiation between the investigated populations. A Bayesian analysis was also conducted to analyse quantitatively the hybridisation process.

## Genetic diversity

The genetic diversity of this species was still high, despite the restricted range and the small number of plants/population. A total of 77 alleles were found for the four loci analysed. The number of alleles/pop ranged from 4 to 18, no fixed alleles were observed (data not shown).

Heterozygosity values were also high and are reported below for each population; the highest value was found for the Donna population (Stintino area, 0.93), the lowest value for the Lioneddu population (Alghero area, 0.6).

| | Asinara | | Stintino | | Alghero | |
|---|---|---|---|---|---|---|
| | Stretti | Fornelli | Falcone | Donna | Lioneddu | Barca |
| $H_o$ | 0.71 | 0.82 | 0.68 | 0.72 | 0.38 | 0.69 |
| $H_e$ | 0.83 | 0.85 | 0.86 | 0.85 | 0.60 | 0.77 |

## Genetic differentiation between populations

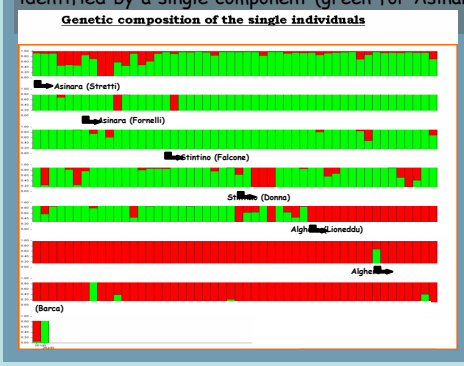
The populations appeared quite differentiated as indicated by an overall $F_{ST}$ = 0.11 and $R_{ST}$ = 0.15; both values were statistically significant. The lowest pairwise $F_{ST}$ values were found between the populations of the Stintino and Asinara areas, the Alghero populations showing the highest values. Nei's genetic distances confirmed the same pattern (data not shown).

The estimates of $R_{ST}$, which detects older differentiation events, suggested that the populations in the Alghero area have been differentiating since a longer time from those of the northernmost areas.

| $R_{ST}$ \ $F_{ST}$ | | Asinara | | Stintino | | Alghero | |
|---|---|---|---|---|---|---|---|
| | | Stretti | Fornelli | Falcone | Donna | Lioneddu | Barca |
| Asinara | Stretti | | 0.067 | 0.091 | 0.081 | 0.224 | 0.120 |
| | Fornelli | 0.150 | | 0.076 | 0.078 | 0.298 | 0.126 |
| Stintino | Falcone | 0.128 | 0.137 | | 0.049 | 0.245 | 0.120 |
| | Donna | 0.129 | 0.050 | 0.112 | | 0.178 | 0.098 |
| Alghero | Lioneddu | 0.155 | 0.119 | 0.076 | 0.024 | | 0.089 |
| | Barca | 0.283 | 0.320 | 0.251 | 0.186 | 0.108 | |

## Population structure analysis

Bayesian analysis, conducted by means of STRUCTURE (Pritchard et al., 2000), allowed us to *i*) estimate the number *K* of inferred populations from which the populations studied here could have arisen (according to the modification of the procedure proposed by Evanno and coll., 2005) and to *ii*) evaluate the coefficient of membership for each individual to the genetic clusters assumed. The six populations seem to derive their genetic structure from two different gene pools. The same can be seen at the level of the genetic composition of each plant, which appears fairly well identified by a single component (green for Asinara and Stintino and red

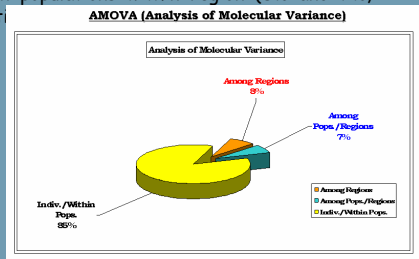



Proportion of membership of each pre-defined population in each of the 2 clusters

| Pops. investigated | Inferred populations 1 | 2 |
|---|---|---|
| Stretti | 0.160 | 0.840 |
| Fornelli | 0.076 | 0.924 |
| Falcone | 0.091 | 0.909 |
| Donna | 0.212 | 0.788 |
| Lioneddu | 0.970 | 0.030 |
| Barca | 0.901 | 0.099 |

## Analysis of Molecular Variance

Based on the information obtained from the Bayesian analysis, the total amount of genetic variation was partitioned by AMOVA into components according to the subdivisions between the northern areas and the southern one and between populations within regions. A significant amount of variation (15% of the total) was due to differences between regions, and between populations within region (8% and 7%, respect


| Source | df | Est. Variance | Φstatistics | p |
|---|---|---|---|---|
| | | | | |
| Among Regions | 1 | 0.143 | 0.076 | 0.010 |
| Among populations within regions | 4 | 0.129 | 0.075 | 0.010 |
| Within population | 698 | 1.601 | 0.145 | 0.010 |

## Some important conclusions can be drawn from our results:

- ✓The vegetative propagation of this plant stop at a few meters from the mother plant;
- ✓No significant differences can be found between the cliff-dwelling and a the plain-dwelling populations of the same area;
- ✓Despite its status as an endangered species, *Centaurea horrida* is not characterised by a low genetic variability, a fact which bodes well for its conservation;
- ✓The actual populations of western Sardinia are probably derived from two heterogeneous gene pools;
- ✓The high genetic differentiation observed requires, for conservation purposes, that the Alghero populations are considered as separate entities from the northernmost ones;
- ✓This study will be completed by the analysis of the population of Tavolara (Eastern Sardinia), which will bring us to the genetic definition of the whole range of this species

**References:**

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14, 2611-2620.

Frèville, H., E. Imbert, F. Justy, R. Vitalis and I. Olivieri. (2000) Isolation and characterization and microsatellites in the endemic species *Centaurea corymbosa* Pourret (Asteraceae) and other related species. Mol.Ecol. 9: 1671-1672.

Pritchard, J.K., Stephens, M. & Donnelly, O. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959.

1st European Congress of Conservation Biology
(22-26 August 2006 – Eger, Hungary)

# GENETIC ANALYSIS OF THE POPULATIONS OF *Centaurea horrida* Badarò (ASTERACEAE).

Giulia Mameli[1], Marilena Meloni[2], Giorgio Binelli[2], Rossella Filigheddu[1]
[1]Dipartimento di Botanica ed Ecologia vegetale – University of Sassari
Via Muroni, 25 – 07100 Sassari; [2]Dipartimento di Biotecnologie e Scienze Molecolari - University of Insubria, via J.H. Dunant 3, 21100 Varese, ITALY;
e-mail: magiul@uniss.it,filighed@uniss.it

**Area and study species:**
*Centaurea horrida* Badarò (Asteraceae) is a narrow endemic species located only in Northern Sardinia (Italy), where it occurs in five areas (Fig. 1): Island of Asinara and Piana Isle, Peninsulae of Stintino and Capo Caccia located in the North-west Sardinia; Island of Tavolara, sited in the East Sardinia.
It is a perennial, pulvinate, spiny species (Figs. 2, 3) whose ecological range is restricted to rocky cliffs, characterized by harsh conditions (especially drought) (Figs. 4, 5).
*C. horrida* is a priority species according to the EU Directive 92/43 Habitat, a protected species according to the Bern Convention, and it is listed as vulnerable (V) in the IUCN Red List (1997).
We have undertaken a study that aims to estimate the amount and the distribution of the genetic variability of populations of *C. horrida*, in order to plan appropriate conservation strategies.

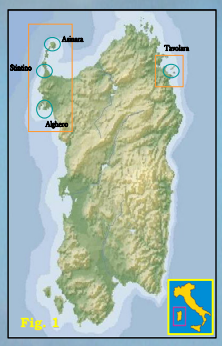


Fig. 1


Fig. 2


Fig. 3


Fig. 4

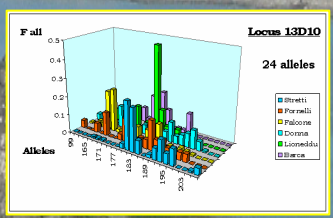

Fig. 5

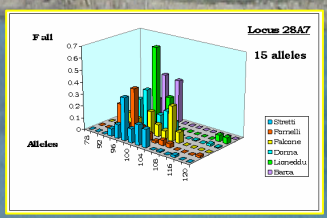
**Methods:**
**Field sampling**: samples were collected between March and April 2005, and between October and November 2005. Three representative areas were investigated: two populations were analysed per area, and 60 individuals were sampled per population (where possible), for a total of 352 individuals.
**Genetic analysis**: the genetic analysis was performed by SSR (Simple Sequence Repeat). Due to the lack of information on the genome of the studied species, seven pairs of heterologous microsatellite primers developed for the congener species *Centaurea corymbosa* (Fréville et al., 2000) were tested; four of them yielded simple amplification patterns and were used to genotype our populations. Amplification conditions for each primer are reported in Tab.1. The amplification products were analysed by a capillary MegaBACE® DNA sequencer. The amount of genetic variability and the degree of differentiation between the investigated populations were then estimated.

Tab.1. Amplification conditions

| 12B1 | | 13D10 | | 21D9 | | 28A7 | |
|---|---|---|---|---|---|---|---|
| 94°C | 2 min | 94°C | 2 min | 94°C | 2 min | 94°C | 2 min |
| 94°C | 1 min | 94°C | 1 min | 94°C | 1 min | 94°C | 1 min |
| Ta 56°C | 30 sec | Ta 59°C | 30 sec | Ta 60°C | 30 sec | Ta 57°C | 30 sec |
| 65°C | 1 min | 65°C | 1 min | 65°C | 1 min | 65°C | 1 min |
| | 30 cycles | | 30 cycles | | 30 cycles | | 30 cycles |
| 65°C | 5 min | 65°C | 5 min | 65°C | 5 min | 65°C | 5 min |
| 10°C | ∞ | 10°C | ∞ | 10°C | ∞ | 10°C | ∞ |

**Results:**




**Allelic frequencies**. The above panels show the allelic distribution of two representative locus. The number of alleles per locus ranged from 15 to 24.

***Ho* e *He***

| | | Asinara | | Stintino | | Alghero | |
|---|---|---|---|---|---|---|---|
| | | Stretti | Fornelli | Falcone | Donna | Lioneddu | Barca |
| 21D9 | *Ho* | 0.74 | 0.75 | 0.43 | 0.48 | 0.33 | 0.39 |
| | *He* | 0.86 | 0.87 | 0.85 | 0.8 | 0.28 | 0.61 |
| 13D10 | *Ho* | 0.69 | 0.85 | 0.68 | 0.76 | 0.51 | 0.9 |
| | *He* | 0.84 | 0.91 | 0.86 | 0.93 | 0.73 | 0.88 |
| 28A7 | *Ho* | 0.91 | 0.81 | 0.88 | 0.81 | 0.5 | 0.61 |
| | *He* | 0.78 | 0.76 | 0.83 | 0.79 | 0.52 | 0.69 |
| 12B1 | *Ho* | 0.5 | 0.88 | 0.72 | 0.82 | 0.21 | 0.82 |
| | *He* | 0.84 | 0.87 | 0.88 | 0.9 | 0.87 | 0.91 |
| Total | *Ho* | 0.71 | 0.82 | 0.68 | 0.72 | 0.38 | 0.69 |
| | He | 0.83 | 0.85 | 0.86 | 0.85 | 0.6 | 0.77 |

$He = 1 - \Sigma x_i^2$ $x_i$ = freq. of the $i^{th}$ allele

**Genetic variability**. The levels of observed and expected heterozygosity were high; the highest value was found for the Donna population (Stintino, 0.93), the lowest value for the Lioneddu population (Alghero, 0.6).

**Genetic differentiation between population pairs as measured by *Fst* (below diagonal) and *Nm* (above)**

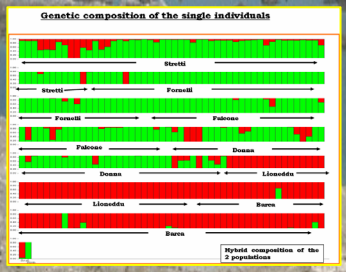
| *Fst* / *Nm* | Stretti | Fornelli | Falcone | Donna | Lioneddu | Barca |
|---|---|---|---|---|---|---|
| Stretti | - | 3.76 | 2.74 | 3.08 | 1.12 | 2.07 |
| Fornelli | 0.067 | - | 3.3 | 3.2 | 0.84 | 1.98 |
| Falcone | 0.091 | 0.076 | - | 5.14 | 1.02 | 2.08 |
| Donna | 0.081 | 0.078 | 0.049 | - | 1.4 | 2.55 |
| Lioneddu | 0.224 | 0.298 | 0.245 | 0.178 | - | 2.81 |
| Barca | 0.120 | 0.126 | 0.120 | 0.098 | 0.089 | - |

$Nm = (1-Fst)/4Fst$

**Proportion of membership of each pre-defined population in each of the 2 clusters**

| Pops. investigated | Inferred populations 1 | 2 |
|---|---|---|
| Stretti | 0.160 | 0.840 |
| Fornelli | 0.076 | 0.924 |
| Falcone | 0.091 | 0.909 |
| Donna | 0.212 | 0.788 |
| Lioneddu | 0.970 | 0.030 |
| Barca | 0.901 | 0.099 |

**Genetic composition of the single individuals**

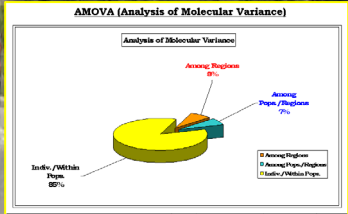



**Analysis of the population structure.** Bayesian analysis, conducted by means of STRUCTURE (Pritchard et al., 2000), allowed us to *i)* estimate the number *K* of inferred populations from which the populations studied here could have arisen and to *ii)* evaluate the coefficient of membership for each individual to the genetic clusters assumed. The six populations derive their genetic structure from two different gene pools and almost all the individuals were correctly assigned to the populations surveyed in this work.

**Genetic differentiation among populations:**
The overall genetic divergence between populations was estimated by $F_{ST}$ = 0.11. The lowest pairwise $F_{ST}$ value was found between Stintino and Asinara, the highest between Alghero and Stintino. Nei's genetic distances (not shown) follow the same pattern.

**Population partition for AMOVA**


**AMOVA (Analysis of Molecular Variance)**




| Source | df | Est.Variance | Φstatistics | *p* |
|---|---|---|---|---|
| Among Regions | 1 | 0.143 | 0.076 | 0.010 |
| Among populations within regions | 4 | 0.129 | 0.075 | 0.010 |
| Within population | 698 | 1.601 | 0.145 | 0.010 |

**Conclusions:**
Populations of *C.horrida* maintain a high level of genetic variability within populations, as shown by the values of *He*. A significant amount of genetic differentiation was detected among the populations analysed.
According to Bayesian analysis, the six populations might derive from two homogeneous gene pools: the first originated the Asinara and Stintino populations, the second originated only the Alghero population.
This finding is very important for the management of the species, because these populations are the result of differential selection pressures and environmental heterogeneity in space and time. Thus, this work will represent the starting point to understand what factors are determining the genetic structure in the current natural populations.

The total amount of genetic variation was partitioned by AMOVA into components according to the subdivisions between regions and between populations within regions. A significant amount of variation (15% of the total) was due to differences between regions, and between populations within region (8% and 7%, respectively).