UNIVERSITA' DEGLI STUDI DI PARMA

Dottorato di ricerca in ECOLOGIA

Ciclo XXIV

Integrating Taxonomy, Population Genetics And Ecology For The Conservation Of Endemic Plants In The Alps

Coordinatore:
Chiar.mo Prof. Paolo Menozzi

Tutor:
Chiar.mo Prof. Stefano Leonardi

Dottoranda: Margherita Lega

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Research issue

a) Conservation biology: a crisis discipline

Global biodiversity is highly endangered and has been continuously declining over the past four decades, as proved by the constant increase in the levels of biodiversity stress indicators (Butchart et al., 2010). According to a very conservative calculation, current extinction rate is estimated to be 1.000 - 10.000 times higher than during the 5 major mass extinction events of the past (http://cmsdata.iucn.org). Unlikely the past events, this sixth mass extinction is mainly due to anthropic activities through habitat fragmentation and destruction, biological invasions caused by intentional or unintentional transport of species outside their native distribution area, pollution and climate change.

Conservation biology was born in the context of this global emergency as a "crisis discipline", aimed at preserving biodiversity by taking rapid decisions based on available data. This problem-solving oriented discipline is the result of the gradual inclusion of many different fields which are fundamental to accelerate and increase the accuracy of conservation decision-making.

b) Species delimitation

Integration of systematics into conservation biology is important in slowing down the process of biodiversity erosion, through recognition and delimitation of the distinct lineages worthy of protection (Soltis & Gitzendanner, 1999). In this regard, the species represents the most prominent and readily recognizable form of biodiversity (Myers et al., 2000). The taxonomic significance that the variation between taxa might bear in terms of species delimitation is nonetheless a controversial issue, which has been debated by evolutionary biologists and taxonomists for decades (de Queiroz, 2011). Different views on the features that distinguish groups of organisms from others has led to different - and sometimes conflicting - species concepts, which, if taken alone, could possibly lead to development and implementation of very distinct conservation strategies.

Nevertheless, any attempt to abolish the use of species would come at the cost of sacrificing the usefulness of discrete, identifiable units of conservation which are also strategic to sensitize public awareness (Agapow et al., 2004). Thus, despite not straightforward, the species concept remains a fundamental tool in conservation biology.

However, the crucial point is not even in the plurality of species concepts, as biologists actually do share a common concept of species as evolutionary groups for their delimitation (Hey, 2006). There is general consensus among the biologists involved in the species concept controversy about the fact that species arise by evolution and that organisms within a species share more evolutionary history with each other than they do with organisms of other species. The real problem lays in the fact that several different criteria for identifying species were elevated to the level of concepts (Hey, 2006). This semantic shift was first introduced by Mayr (1942) who opened to a new usage of 'concept' based on the operational criteria useful for species delimitation, rather than the theoretical ideas on the causes of existence of species.

As a consequence of this confusion between species delimitation and species definition, a plurality of so-called species concepts was formulated from that moment on, and every criterion was treated as a necessary property of species. Thus, for example, presence of reproductive isolation between heterospecific organisms is a necessary requirement for the biological species concept (Wright, 1940; Mayr, 1942; Dobzhansky, 1950), while sharing of the same niche or adaptive zone is fundamental under the ecological species concept (Van Valen, 1976; Andersson, 1990). The reason why these alternative criteria come into conflict lays in the fact that while each of these is treated as necessary, not all of them are usually verified at the same time during the process of speciation (de Queiroz, 2007).

Only in recent years these alternative "species concepts" have been considered as complementary aspects of the process of differentiation that leads to the evolution of species (de Queiroz, 1998, 2007). Under such unified species concept (USC), species are seen as the result of the evolution of populations along independent lineages, where the newly developing species acquire the properties that characterize them at different times during the process of divergence. Any property that provides evidence of lineage separation is an important line of evidence for deciding where to put the boundaries among species, but none of these property is necessary. Indeed, the USC described by de Queiroz (1998, 2007) is not even to be considered a new species concept but simply the clear separation of the theoretical concept of species from the operational criteria used to delimit species. Only in this perspective the complementation of different species criteria becomes a fundamental tool to overcome the problem of species delimitation, for example when morphology taken alone is not sufficient to unambiguously distinguish the different lineages, and integration with genetic and

ecological data is necessary (French et al., 2008). If, on the contrary, pluralism is not seen as a pluralism of criteria for specie identification, it will remain problematic in the sense that one could always ask which species concepts are worthy to be taken into consideration and which are to be refused (Hey, 2006).

c) Population genetics

Together with species delimitation, population genetic studies represent another fundamental tool in conservation biology. The integration of genetics into conservation biology was promoted by the development of numerous molecular techniques in the last decades (DeSalle & Amato, 2004; Kramer & Havens, 2009), and more recently also by various high-throughput genomic technologies, which are providing an unprecedented precision in describing the processes of biodiversity erosion (Ekblom & Galindo, 2011). In fact, genetic diversity is necessary for evolution to occur, and it was recognized as one of the 3 forms of biodiversity by the World Conservation Union (IUCN) deserving conservation, together with species and ecosystems. A meta-analysis on 34 data sets proved the negative correlation between the loss of heterozygosity and the population fitness (Reed & Frankham, 2003), and inbreeding was observed to reduce reproduction and survival in essentially all well studied species (Frankham, Ballou & Briscoe, 2002). Furthermore, another comprehensive meta-analysis of 170 species and independent computer simulations demonstrated that most taxa are not driven to extinction before genetic factors affect them adversely (Spielman, Brook & Frankham, 2004). The key role played by genetic factors in species extinctions had further been verified by some isolated studies showing association between inbreeding and reduced genetic diversity with elevated extinction risk (Saccheri et al., 1998; Newman & Pilson, 1997). For these reasons, population genetic analyses constitute an effective instrument to evaluate the level of threat of endemic taxa. Moreover, the assessment of the conservation status of a species, or sub-specific taxon, becomes even more relevant if it has been realized considering genetic variation patterns in different populations of the same taxon (Smith & Waldren, 2010) and/or the sister taxa (Reisch, Kaiser, Horn & Poschlod, 2010; O'Reilly, Cowan & Hawkins, 2007; Chung et al., 2004). However, even in systematically very close species, ecological and biological factors together with ancestral population history can lead to completely different diversification processes (Kholodova, 2009). Hence, not all population genetic structures of sibling taxa could be easily comparable, and analyses of taxa sharing similar reproductive traits and ecological niches, in addition to a very recent

common evolutionary history, may be preferred to a "simple" sister taxon approach. Finally, population genetic analyses applied to conservation should be interpreted within the landscape where the endangered taxa under study evolved. Indeed, landscape variables can play a fundamental role in shaping the components of genetic diversity, for example through barriers that limit gene flow, like mountain ridges, waterways, roads, but also microhabitats with very specific ecological requirements (Manel et al., 2003; Storfer, 2007). Therefore, integration of genetic information with spatially explicit techniques (e.g. Guillot, Estoup, Mortier & Cosson, 2005) is important to detect barriers that are not normally identifiable using traditional population genetic tools.

d) Quantifying genetic differentiation: methodological issues

The formulation of effective conservation actions relies on correct estimations of genetic structure, which, in turn, depend on the particular statistics applied to a data set. If some statistics are used but do not fit the research questions asked, they can provide misleading results.

Among the currently available tools to describe genetic structuring, the traditional F_{ST} statistics (Malécot, 1948; Wright, 1951; Weir & Cockerham, 1984) and their analogues (e.g. $G_{ST,}$ for multiallelic loci; Nei, 1973a) are still widely applied and interpreted as true measures of genetic differentiation.

However, as Wright himself (1978) and Gregorius (1987) already pointed out, F_{ST}, G_{ST} and their relatives are only informative about fixation. As an example, F_{ST} may equal 1 even in the case of 10 demes, of which 9 are fixed for an identical allele (monomorphic), and only 1 is fixed for an alternative allele.

Later, the widespread use of microsatellite (SSR) data in population genetic studies led to the empirical discovery that these fixation indexes were often inadequate when applied to these highly polymorphic markers, as they tended to underestimate population differentiation (e.g. Balloux et al. 2000; Carreras-Carbonell et al., 2006). As a solution, Hedrick (2005) proposed a standardization of G_{ST}, and Meirmans (2006) developed an analogous method for the ψ_{ST} based on AMOVA. Jost (2008) also mathematically demonstrated that F_{ST}, G_{ST} and related measures are intrinsically limited on the maximum possible differentiation they can attain, as they strictly rely on the ratio of average within-subpopulation heterozigosity to total heterozigosity or, alternatively, on the additive partitioning of heterozygosity. As a consequence, they do not increase monotonically with

increasing diversity (Jost, 2008), and thus not only population differentiation could be underestimated, but many problems of interpretation could arise when a comparison is made among markers with different mutation rates or groups with different effective population sizes, or areas of a species distribution range with differing levels of diversity (Meirmans & Hedrick, 2011).

Jost (2008) additionally proposed a true measure of population differentiation, D, and its analogous D_{est}, an estimator of actual differentiation corrected for small sample size. The innovative aspect of this index based on the effective number of alleles lies in the total separation of whole genetic diversity into independent within- and between-deme components (Jost, 2008, 2009). Gerlach at al. (2010) confirmed through simulations Jost's mathematical demonstration (2008) that with increasing allele numbers, the range of values acquired by G_{ST} was drastically reduced, while Jost's D_{est} continued to vary from 0 (no differentiation among groups) to 1 (total differentiation among groups). The authors concluded that G_{ST} is an appropriate index of population differentiation only in the presence of 2 alleles, estimating fixation, while Jost's D_{est} measures true differentiation (Gerlach et al, 2010).

Conversely, some other authors criticized Jost's D_{est} pointing out that this measure is not influenced by local effective population size (N) in a finite island model (Li, 1976) and thus it is insensitive to the evolutionary processes controlling population differentiation. In fact, as a measure of differentiation, its equilibrium value depends strongly on migration rate (m), mutation rate (μ) and number of demes (n) (equations 15–17 in Jost, 2008). Since G_{ST} and its relatives depend only weakly on these parameters, while being influenced by N, they are preferable to D_{est} for estimating migration rate under a finite island model. Simulations by Ryman & Leimar (2009) and Leng & Zhang (2011) proved that when initial heterozigosity is low (e.g. in case of population bottlenecks) D_{est} increases much slower than G_{ST} across generations with fixed mutation rate and no migration, as its value is only determined by population divergence caused by mutations.

Ryman & Leimar (2009) and Whitlock (2011) further warned against the fact that D_{est} is highly dependent on μ. This property is considered inconvenient because it reflects the characteristics of a single locus and does not give information on the general processes acting on the populations under study. However, as Jost himself stated (2009), dependence of μ is not a defect, but an obvious characteristic of a measure of allelic differentiation that describes reality. In any case, none of the traditional statistics is completely independent of μ. For example, the low sensitivity of G_{ST} with regards to mutation model assumption is valid only with an island model of population structure assumed under equilibrium (Leng & Zhang, 2011), but populations in nature are very often violating these assumptions.

To summarize, it seems that the 2 classes of measures quantify different aspects of population structure: D_{est} is best suited for measuring the actual degree of differentiation (Jost, 2008; 2009) while F_{ST} and similar measures are useful when the focus is on the causes of this structure, such as the demographic processes of genetic drift and gene flow (Jost, 2009; Ryman & Leimar, 2009; Meirmans & Hedrick, 2011).

For this reason, the choice of the measure strictly depends on the particular research question asked, and their combined use is often recommended for a more complete examination of population structures (Meirmans & Hedrick, 2011). As an example, in an empirical AFLP study on 27 high-alpine plants (Meirmans, Goudet, & Gaggiotti, 2011) D_{est} showed no correlation with any of 6 key life-history traits for the species' dispersal (previously thought to indirectly affect its genetic structure), while F_{ST} was significantly correlated with the mode of seed dispersal. Consequently, the authors concluded that D_{est} was not a good statistic to study demographic processes. On the other hand, D_{est} was proved to be very useful to advise conservation genetics of rare species, particularly in studies based on highly polymorphic markers like SSR (e.g. Casado-Amezúa, 2012).

e) Refugial areas to preserve biodiversity

Southern Europe, and particularly the peninsulas of Italy, Iberia and the Balkans are well known hotspots of endemicity (Hewitt, 2011), thus representing very interesting areas to study the evolutionary processes that shaped the genetic structure of currently endangered taxa.
Furthermore, a perfect correspondence between areas of endemism and refugia was found both in the Mediterranean basin (Médail & Diadema, 2009) and at the periphery of the Alps (Schönswetter, Stehlik, Holderegger & Tribsch, 2005). Indeed, in an extensive analysis of the scientific literature from 1993 to 2007 treating plant intra-specific phylogeographical studies, Medail & Diadema (2009) detected 52 main refugia in the Mediterranean bioclimatic region, with several of them also found to be significantly associated to 10 regional hotspots of plant biodiversity. In this context, the three Mediterranean peninsulas were confirmed to have played a major role with approximately half of the total refugia recovered there, although other important refugia were recognized in the southern and northern parts of Mediterranean (i.e. North Africa, Turkey, etc.). Along with the less drastic climate conditions distinguishing southern Europe from northern continental territories, the topographic complexity entailing differences in altitude, temperature, humidity, substrate, sun and wind exposure, and the consequent richness in different habitats typical of the southern European

peninsulas are certainly the main factors explaining the high concentration of disjunct refugia recovered (Canestrelli, Sacco & Nascetti, 2012). Broadly speaking, refugia consist of climatically stable, warmer -or cooler- and wetter areas with regards to the neighbouring regions, whose relative ages are often uncertain. It is probable that what we call today "refugia" are in fact "cumulative" refugia (Médail & Diadema, 2009) which allowed the survival of living organisms during several, and not only the most recent, adverse long-term climatic events (Médail & Diadema, 2009). By consequence, refugial areas of the past, if not too fragmented by anthropic activities, could possibly continue to play a similar role with regards to future climate changes. In particular, inter-glacial refugia or glacial microrefugia whose microclimate is decoupled from the regional atmospheric conditions (Dobrowski, 2011) represent potential safe harbors against the current global warming, which is variously affecting biodiversity and eroding its evolutionary potential depending on species vulnerabilities (Dawson et al., 2011). The alpine vegetation as a whole, for example, is likely to be heavily affected by increasing temperatures. Competitive displacement of alpine species has already been observed at the lowest summits by trees, shrubs and clonal graminoids, and increase in species number in the summit areas was demonstrated, either on long term (1959-2005; Parolo & Rossi, 2008) and on short timescales (2001-2006; Erschbamer et al., 2009; Pauli *et al*., 2012;), with a species-specific response to the increasingly warmer climate associated to dispersal mode (Parolo & Rossi, 2008). Thus, the persistence of ancient refugial areas in future, coupled with the ability of species to survive in populations of low density and small effective size for long adverse periods of time, could potentially contribute in lowering the extinction risk provoked by global warming and worsened by the reduced migration capacity of many plant species (Pearson, 2006). Indeed, survival of populations in these refugia entails the preservation of the species evolutionary potential, through maintenance of some of the ancestral characters of their genetic diversity. Under this light, a correct identification, description and protection of refugia, including glacial, inter-glacial and current climate change refugia is of paramount importance for an effective conservation of biodiversity in the near future (Ashcroft, 2010). A growing body of studies is going towards this direction. For example, Leroy & Arpe (2007) reconstructed all potential long-term glacial refugia for cool and warm summer-green trees from Europe to the Caspian region by applying climate modelling, and proposed them to be integrated in the guidelines for setting new natural reserves. Klein et al. (2009) incorporated data from evolutionary refugia, there defined as areas where certain species manage to survive during temporary unsuitable long-term climatic conditions, for identifying spatial priorities in conservation planning across Australia. These evolutionary refugia included islands and mountain areas rich in relictual or threatened species, and species which were proved to have evolved new distinctive characteristics in response to past

climate changes.

f) South-eastern Alps: hotspots of endemicity

Within the extremely genetically diverse peninsulas of southern Europe, south-eastern Alps are, as mentioned above, particularly rich in refugia, which actually represent important hotspots of endemicity (Tribsch, 2004; Schönswetter et al., 2005). Schmitt (2009), in his description of the 4 main recurring genetic lineages in alpine plants (south-western, western central, eastern central and eastern Alps), recognized to the eastern Alps lineage the highest levels of genetic diversity. He hypothesized in the eastern Alps better conditions for the persistence of species during the Pleistocene, and/or presence of much more spatially extended refugia compared to the western part. Tribsch (2004), studied 288 vascular plant taxa endemic to the eastern Alps and found a strong correlation between endemism and impact of Pleistocene glaciation, with snowline playing a major role than the elevation of the ice sheet. Thus, the high number of endemisms was associated to unglaciated areas, as already stated by Tribsch & Schönswetter (2003). The higher genetic diversity of endemic species in south-eastern Alps was underlined by Schmitt (2007) even with regards to the whole complex of Artic-Alpine species, that were widespread in the ice-free grasslands placed between the glaciers of northern Europe and those of the Alps. In these Artic-Alpine taxa the homogenizing effects of gene flow at large scales were preponderant over drift, thus limiting the overall process of differentiation. On the contrary, the other taxa not adapted to the dry conditions of the steppes remained localized in the wetter areas peripheral to the ice-shields of the diverse mountain chains, and particularly at the edges of the meridional and oriental Alps. The disjoint distribution of populations over an heterogeneous mountain landscape confined gene flow and therefore enhanced the process of genetic differentiation within the south-eastern alpine taxa both during the glacial and interglacial periods (Schmitt, 2007). South-eastern alpine endemics are thus of particular interest for the study of speciation and intra-specific differentiation processes.

g) Research statement

The present work focused on 2 stenoendemics of south-eastern Alps, potentially endangered because of their small population number and size: *Brassica repanda* (Willd.) DC. subsp. *baldensis* (Prosser & Bertolli) Prosser & Bertolli, comb. nov. and *Aquilegia thalictrifolia* Schott & Kotschy, growing within the area that carries the highest number of endemisms and local endemisms of oriental Alps, between Lago di Como and Lago di Garda (Tribsch, 2004). Additionally, 5 other endemic angiosperm taxa were included as a comparison: *Brassica repanda* subsp. *glabrescens* Poldini, a stenoendemic of south-eastern Prealps which are also characterized by a conspicuous number of endemics (Tribsch, 2004); *Aquilegia vestinae* Pfenninger & Moser, a recently described taxon overlapping with *A. thalictrifolia* distribution range; *Aquilegia julia* Nardi, confined in the eastern Alps of Slovenia; *Aquilegia bertolonii* Schott and *A. reuterii* Boss, located in the glacial refugia of Apuan and Maritime Alps, respectively (Médail & Diadema, 2009). Finally, *Aquilegia einseleana* Schultz, a widespread species encompassing populations with large geographic distances one to another, ranging from Rhaetic to Austrian and Slovenian Alps, was also studied.
An approach integrating taxonomy, population genetics and ecology was used to gain more insights into the evolutionary history and identify conservation priorities for the endemics. Importantly, comparisons were made throughout the study, among phylogenetic and/or population genetic data and information coming from traditional taxonomy and ecology.

More in detail, the present study aimed at:

- clarifying, with Internal Transcribed Spacers (ITS) and Amplified Fragment Length Polymorphisms (AFLP), whether *B. repanda* subsp. *baldensis* and *glabrescens* form evolutionary distinct lineages, both from one another, and from the remainder of the *B. repanda* complex, which is scattered throughout western Alps, the Pyrenees, the Iberian peninsula and the Atlas mountains;

- studying levels and spatial organization of genetic variation in *B. repanda* subsp. *baldensis* and *glabrescens,* and comparing the 2 endemics;

- studying, with SSR markers, levels and spatial organization of genetic variation in *A. thalictrifolia* populations, also by comparing the performances of alternative indexes of fixation (F_{ST}, G_{ST}) and genetic differentiation (D_{est});

- extending the study on *A. thalictrifolia* to *A. bertolonii*, *A. einseleana*, *A. vestinae, A. julia* and *A. reuterii* to explore the evolutionary processes causing the diversification processes within these European columbines;

- jointly considering the whole data sets to identify conservation priorities for the endemic angiosperms.

2. BRASSICA REPANDA

2.1. Introduction

a) Case study

Occurrence of closely related species, subspecies and populations of the same species with disjunct distributions is a frequent feature of the Mediterranean flora (Thompson, 1999). This fragmentation can be explained by the complex tectonic history of the Mediterranean basin in the Tertiary and/or possible population isolation during the subsequent glaciation events (reviewed in Thompson, 1999). Examining these disjunctions from a taxonomic perspective is of great interest both from an evolutionary and a conservation point of view, as they often include endemics threatened with extinction (Debussche & Thompson, 2002; Bellusci, Musacchio, Palermo & Pellegrino, 2010). In this context, a consensus on the taxonomic status of threatened entities allows the recognition of discrete units that are essential for the definition of extinction risks (e.g. IUCN Red List) and the development and implementation of appropriate conservation strategies (Agapow *et al.*, 2004). Current advances in the theory of species delimitation (de Queiroz, 1998, 2007, 2011; Schlick-Steiner *et al.*, 2010) and in the techniques used to characterize biodiversity (e.g. genetic markers) offer the opportunity to achieve accurate and uncontroversial taxonomic characterizations. In particular, the unified species concept (USC; de Queiroz, 1998) is intended to achieve a consensus definition of species reconciling the alternative concepts advocated by contemporary biologists to define the basic taxonomic unit. In his definition, de Queiroz argues that alternative species concepts are indeed based on a common element that is the primary property defining the species category: species are separately evolving metapopulation lineages (de Queiroz, 1998, 2007). What causes disagreement among concepts has to do with the criteria used to identify such lineages, as they refer to different secondary properties that may lead to incompatible species delimitations. Thus, for example, as already mentioned in Chapter 1, intrinsic reproductive isolation is necessary in the biological species concept (Mayr, 1942; Dobzhansky, 1970) while niche specialization in the ecological one (Van Valen, 1976; Andersson, 1990) and diagnosability in the phylogenetic one (Cracraft, 1983; Nixon & Wheeler, 1990). All such properties are indeed indicative of speciation, but their acquisition during the process of lineage divergence may occur at different times and in an

unpredictable order. Under the USC, the alternative species criteria are considered equally significant as each provides a line of evidence of lineage separation, but presence of multiple properties is associated with a higher degree of corroboration of species delimitation.

Independent evidence derived from multiple properties provides a powerful tool to investigate the taxonomy within traditionally complex groups, for which morphological similarity featured by strictly related taxa prevents direct observation of speciation (e.g. Ross, Gotzek, Ascunce & Shoemaker, 2010; Reeves & Richards, 2011; Barrett & Freudenstein, 2011). It is the case of critical species complexes that encompass a large number of morphologically similar taxa, where the use of genetic data may unveil cryptic divergence resulting from evolution of independent lineages. In this context, geographic isolation of disjunct taxa is a favourable condition for speciation following prolonged absence of gene flow (Coyne & Orr, 2004), and the application of the USC offers a robust and widely accepted conceptual framework to address their taxonomic recognition.

Brassica repanda (Willd.) DC. is a highly polymorphic species complex with scattered geographic distribution (Greuter, Burdet & Long, 1986; Heywood & Akeroyd, 1993; Gómez-Campo, 1999), including a large number of subspecific entities whose taxonomic classification largely relies on morphological characters (Gómez-Campo, 1993). In recent years various taxonomic adjustments have been adopted also in consideration of molecular phylogenetic results in the tribe Brassiceae (Gómez-Campo, 2003; Prosser & Bertolli, 2007b; Warwick & Sauder, 2005). The current classification of the group (Gomez-Campo, 1993) recognizes 19 subspecies, distributed in the southern part of the Iberian Peninsula and the Pyrenees (12 taxa), the Atlas Mountains (3 taxa), southern France and western Alps (2 taxa). Among these, all investigated European subspecies are diploid (2n=20), while the North African taxa are polyploid (Prosser & Bertolli, 2007). The easternmost distribution of the complex is represented by the disjunction of the alpine endemic *B. repanda* subsp. *glabrescens* (Poldini) Gómez-Campo (BRG; Picture 1C) and *B. repanda* subsp. *baldensis* (Prosser & Bertolli) Prosser & Bertolli (BRB; Picture 1A-B), whose geographic isolation and narrow distribution constitute a deep discontinuity in the European range of the complex. BRG had been originally classified as an independent species (Branca, Donnini, Dulloo & Kell, 2011; Poldini, 1973), and later downgraded to subspecific rank within *B. repanda* for insufficient morphological distinctiveness from the range of phenotypic variation that characterizes the species (Gómez-Campo & Martínez Laborde, 1998). It is a diploid taxon (2n=20; Jalas, Suominen & Lampinen, 1996) endemic to a very restricted area between the rivers Cellina and Meduna in the south-eastern Italian Alps, where approximately 1000 flowering individuals are known (Poldini, 1973). BRB was recently discovered on Mt Baldo, a mountain of the eastern Alps celebrated for its floristic richness. The distribution of the taxon is restricted to an area extending over about three

km, between 250 m and 820 m of altitude, on projecting rocks of dry, calcareous and sunny exposed ledges. Only 5 sites have been reported, for a total of approximately 1500 flowering individuals (Prosser & Bertolli, 2007a, 2007b; ploidy unknown). Interestingly, in their description of these species, the authors independently emphasize the similarity of either eastern endemic to different western representatives of the *B. repanda* complex (Poldini 1973; Prosser & Bertolli, 2007a), in contrast to what would be expected assuming a common origin of the eastern disjunction.

The extremely reduced population size and small geographic range of both BRG and BRB pose an important extinction threat to their populations. In fact, several studies have demonstrated that narrow endemics are more vulnerable to stochastic factors such as environmental change and/or habitat loss, which drastically reduce population size and elevate the extinction risk (Frankham, Ballou & Briscoe, 2002; Segarra-Moragues, Palop-Esteban, Gonzalez-Candelas & Catalán, 2005). Declining populations are normally characterized by loss of heterozygosity and reduced gene flow, which suggests increased inbreeding and reduction of population fitness (Frankham, Ballou & Briscoe, 2002; Reed & Frankham, 2003; Spielman, Brook & Frankham, 2004). Genetic data represent, therefore, a valuable and sensitive approach to both characterize the threat of extinction and evaluate conservation priorities.

b) Research aims

In this study, we aimed to determine whether the eastern endemics formed evolutionary distinct lineages, both from one another, and from the remainder of the *B. repanda* complex. This was achieved by applying the criteria of monophyly (Donoghue 1985; de Queiroz & Donoghue, 1988), diagnosability (Cracraft, 1983; Nixon & Wheeler, 1990) and genotypic clustering (Mallet, 1995). Compliance of the endemics with the above-mentioned criteria used to determine species delimitation was tested making use of AFLP data, and taxonomic recognition is discussed in light of the plurality of evidence provided by the analyses and of the morphology of the group. Further, the genetic data are used to investigate the genetic structure within BRG and BRB, in an attempt to quantify the relative levels of threat and to suggest guidelines for *in-situ* and *ex-situ* conservation.

2.2. Matherials and methods

a) Sampling

Sampling was carried out according to the current taxonomy of the group (Gómez-Campo, 1993; Fig. 1). Most of the diversity of the *B. repanda* complex is represented by the 12 subspecies occurring in the Iberian Peninsula and the Pyrenees, including a large proportion of endemics with local distribution [e.g. subsp. *almeriensis* Gómez-Campo, subsp. *dertosensis* Molero & Rovira, subsp. *galissieri* (Giraud.) Heywood]. Six representative subspecies of such western distribution were sampled among those with a larger range, together with the 2 subspecies occurring in southern France (i.e. subsp. *repanda* and subsp. *saxatilis*). Additionally, 2 samples from recently discovered populations in the western Alps that could not be classified in any of the currently recognized subspecies were included (here reported as *B. repanda* s.l.). Sampling of BRB covered the whole taxon distribution area on Monte Baldo, including 104 individuals from 5 sampling locations. Sampling of BRG included 38 individuals collected from the only 2 known populations described for this taxon. Detailed information on the samples is reported in Table 1 and in the Appendix.
Wild collected seeds of *B. repanda* subsp. *blancoana*, subsp. *cadevallii*, subsp. *confusa* (Emberger & Mayre) Heywood, subsp. *gypsicola* (Gómez-Campo), subsp. *latisiliqua* (Boiss. & Reuter) Heywood and subsp. *maritima*, were obtained from the Seed Bank of the Department of Vegetal Biology of the Polytechnic University of Madrid (Spain). The National Alpine Botanic Conservatory of Gap-Charance (France) provided seeds from an additional natural population of *B. repanda* subsp. *repanda* (Willd.) DC.

b) DNA extraction

Seeds were germinated on 2% agar with 250 ppm gibberellic acid and leaves from plants grown in the greenhouse were collected for DNA extraction. Total genomic DNA was isolated using either the CTAB extraction method (Doyle & Doyle, 1987) or the Qiagen DNeasy 96 Well Plate Kit.

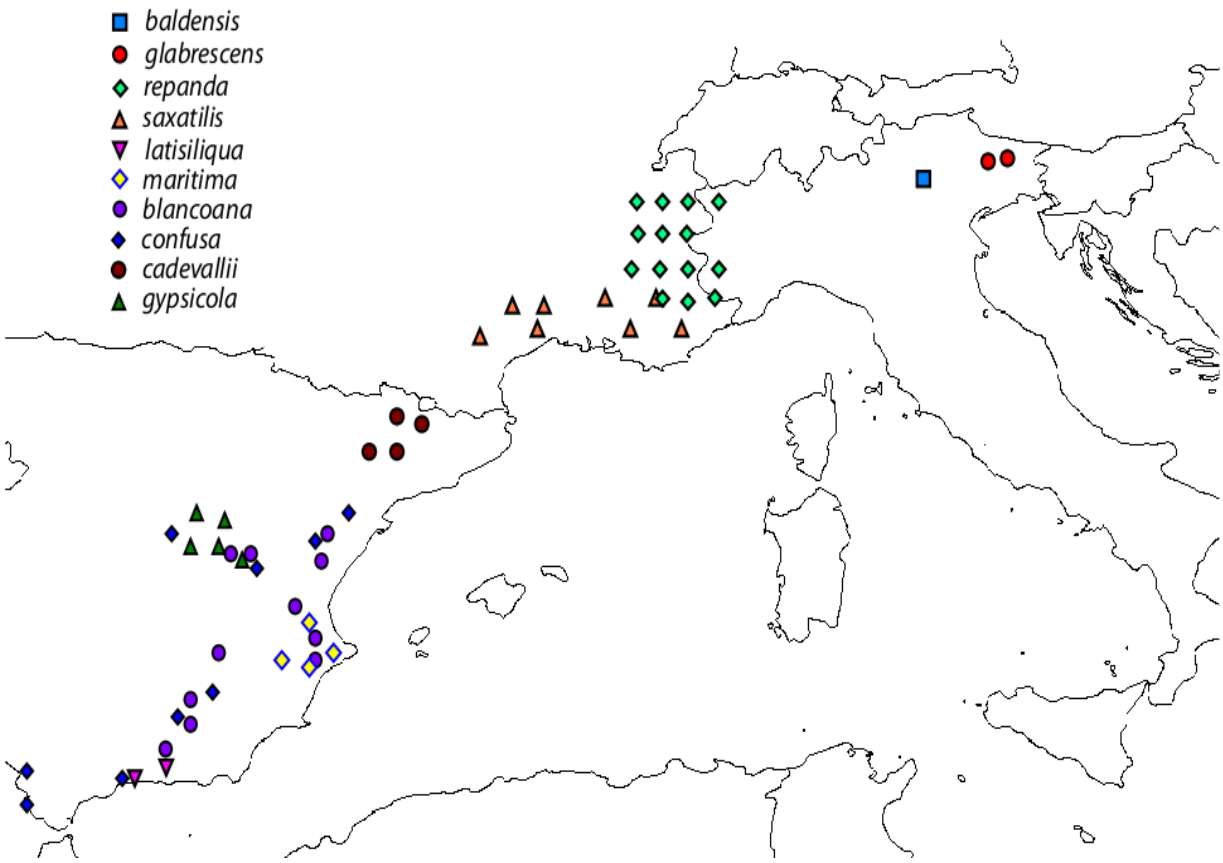


Figure 1. Main distribution areas of the *B. repanda* subspecies under study, from Atlas Flora Europaea (http://www.luomus.fi/english/botany/afe/index.htm) and Prosser and Bertolli (2007a).

Table 1. Population identifications, country, sampling locations and number of individuals sampled per population for AFLP and ITS data sets. In the current work we refer to the taxonomy prior to the work of Gómez-Campo (2003), and nomenclature follows Flora Europaea (Heywood & Akeroyd, 1993) for all subspecific entities except for *B. repanda* subsp. *gypsicola,* BRB and BRG, for which we refer to Gómez-Campo (1992), Prosser & Bertolli (2007b) and Gómez-Campo & Martínez Laborde (1998), respectively.

Taxon	Pop. ID	Country	Location	No. samples (AFLP)	No.samples (ITS)
BRG	A	Italy	S. Foca, S. Quirino, (*locus classicus*, Pordenone)	20	/
	B	Italy	Montereale, Valcellina, (Pordenone)	18	1
BRB	C	Italy	Brentino-north, Monte Baldo (Verona)	20	/
	D	Italy	Brentino-south, Monte Baldo (Verona)	23	/
	E	Italy	Preabocco-north, Monte Baldo (Verona)	22	/
	F	Italy	Preabocco-south, Monte Baldo (*locus classicus*, Verona)	20	1
	G	Italy	Monte Cimo, Monte Baldo (Verona)	19	/
B. repanda subsp. *blancoana*	H	Spain	Los Chorros, S. de Alcaraz (Albacete)	4	1
B. repanda subsp. *cadevallii*	I	Spain	Portell Dells Torradells (Lérida)	3	/
	J	Spain	Sopeira (Huesca)	2	2
B. repanda subsp. *confusa*	K	Spain	Sierra de Grazalema (Cadiz)	2	1
B. repanda subsp. *gypsicola*	L	Spain	Ribatejada-Arcos (Cuenca)	5	1
B. repanda subsp. *latisiliqua*	M	Spain	Trevenque, Sierra Nevada (Granada)	5	2
B. repanda subsp. *maritima*	O	Spain	Montgó (Alicante)	5	2
B. repanda subsp. *repanda*	P	France	Le Monêtier-les-Bains (Hautes Alpes)	1	/
	Q	Italy	Val Stura di Demonte (Cuneo)	1	1
	R	Italy	Monte Furgon, Val Thures, (Torino)	1	1
B. repanda s.l.	S	Italy	Foresto, Val Susa (Torino)	2	1
B. repanda s.l.	T	Italy	Macra, Val Maira (Cuneo)	1	1
B. repanda subsp. *saxatilis*	U	France	Pic de Mouches (*locus classicus*, Bouches-du-Rhône)	2	2
	V	France	Valbelle (Alpes-de-Haute-Provence)	1	/
B. repanda cf. subsp. *saxatilis*	W	France	La Breole (Alpes-de-Haute-Provence)	2	/

c) Sequencing and AFLP reaction

To ascertain the phylogenetic position of the taxa used in this study, we sequenced the Internal Transcribed Spacer (ITS) region of the 18S–5.8S–26S nuclear ribosomal cistron of at least one representative of each subspecific entity (17 individuals in total). Despite criticized for its potentially high levels of homoplasy (Alvarez & Wendel, 2003), the ITS region remains a widely used marker in plant molecular phylogenetics since 2 decades, mainly because of its biparental inheritance, universality, small size and high copy number that facilitate its amplification (Baldwin, 1995).

Primers and thermal cycling conditions used for PCR amplification are reported in Table 2. Direct sequencing was performed on a 3730xl DNA Analyzer sequencer (Applied Biosystems, Carlsbad, USA) using Big Dye Terminator v3.1 (Applied Biosystems, Carlsbad, USA). Contigs were assembled with the Staden Package (Staden, Beal & Bonfield, 1998) and nucleotide sequences were deposited at EMBL and GenBank databases (see Appendix for accession numbers).

Phylogenetic and population genetic analyses within *B. repanda* were conducted based on Amplified Fragment Length Polymorphisms (AFLPs).

AFLPs were chosen because they are highly reproducible, multilocus markers not requiring any previous knowledge of the genome of the studied taxa. Nonetheless, they can be highly homoplasious, due to their dominance, their codification as binary characters and to the fact that fragments of the same length are not always homologous. Despite these well known intrinsic limitations, the evolutionary information inferred from AFLP data represent a widely accepted and powerful tool for shallow phylogenetics and population genetics (Koopman, 2005; Meudt & Clarke, 2007; Simmons *et al*., 2007; Pereira-Garcia, Caballero & Quesada, 2010).

Data were produced according to the method described by Vos et al. (1995). Briefly, a total of 78 EcoRI/MseI primer combinations with three selective nucleotides per primer were tested on a set of 8 taxa. Nine primer combinations providing scorable AFLP patterns were selected and used to perform reactions on 179 individuals from 10 different taxa. Primers and thermal cycling conditions used for PCR amplification are reported in Table 2. Amplification products were loaded on a 3730xl DNA Analyser sequencer, using the 1200 GeneScan® LIZ as size standard (Applied Biosystems, Carlsbad, USA). Technical replicates using 10% of samples were included to calculate the error rate of genotyping (see Bonin et al., 2004). This was assessed to be 0.076. Fragments were detected using GENEMAPPER 4.0 (Applied Biosystems) applying an automated scoring procedure. Peak

width was set to 1.2 bp, allele calling was set to 0/1 for peaks below/above 100 Relative Fluorescent Units. Fragments were scored in a length range from 60 to approximately 300 base pairs. Scores for all primer combinations were compiled into the binary data matrices used for phylogenetic and population genetic analyses. The per-band scoring error rate was assessed by comparing manual and automatic scoring of one of the primer combinations. It was estimated to be 0.028.

Table 2. List of the sequences of ITS and EcorI/MseI primer pairs and thermal cycling conditions for PCR amplification (Sigma–Aldrich, St. Louis, USA) of ITS sequences and AFLPs respectively. "E" and "M" indicate EcoRI and MseI selective primers.

Molecular marker	Primer pairs	Sequences (5'-3')
ITS *		
	ITS-5F / ITS-2R	GGAAGTAAAAGTCGTAACAAGG / GATATGCTTAAACTCAGCGGG
AFLPs †		
	E-33 / M-32	GACTGCGTACCAATTCAAG / GATGAGTCCTGAGTAAAAC
	E-33 / M-37	GACTGCGTACCAATTCAAG / GATGAGTCCTGAGTAAACG
	E-35 / M-40	GACTGCGTACCAATTCACA / GATGAGTCCTGAGTAAAGC
	E-35 / M-42	GACTGCGTACCAATTCACA / GATGAGTCCTGAGTAAAGT
	E-36 / M-36	GACTGCGTACCAATTCACC / GATGAGTCCTGAGTAAACC
	E-36 / M-40	GACTGCGTACCAATTCACC / GATGAGTCCTGAGTAAAGC
	E-37 / M-40	GACTGCGTACCAATTCACG / GATGAGTCCTGAGTAAAGC
	E-37 / M-41	GACTGCGTACCAATTCACG / GATGAGTCCTGAGTAAAGG
	E-38 / M-40	GACTGCGTACCAATTCACT / GATGAGTCCTGAGTAAAGC

***** Thermal cycling conditions used for ITS amplification

Two min of initial denaturation at 94°C, followed by 35 cycles, each consisting of 94°C for 40 s, 50°C for 30 s and 72°C for 1 min. Final extension at 72°C for 2 min.

† Thermal cycling conditions used for AFLP primary amplification:

Twenty pre-amplification cycles of PCR, each consisting of 1 min of denaturation at 92°C, followed by 30 s of annealing at 60°C and finally 1 min of elongation at 72°C.

† Thermal cycling conditions used for AFLP secondary amplification:

One cycle consisting of 30 s at 94°C, followed by 30 s at 65°C and 1 min at 72°C. This was followed by 11 cycles with annealing temperature decreasing by 0.7°C at each cycle, and additional 24 cycles with 30 s of denaturation at 94°C, 30 s of annealing at 56°C and 1 min of elongation at 72°C.

2.3. Data analysis

a) Phylogenetic and genotypic data

ITS sequences

The 17 ITS sequences produced in this study and 108 sequences from the Brassiceae tribe phylogeny by Warwick & Sauder (2005; TreeBase Ref. M2227) were aligned using the MUSCLE algorithm (Edgar, 2004) implemented in the software Geneious v5.3 (Drummond et al., 2010). A small proportion of ambiguously aligned sites (positions 74-110) were excluded from the matrix used for further analyses (TreeBase Ref. S12132). Phylogenetic analyses were conducted in PAUP* 4.0b10 (Swofford, 2003) and search procedures followed Warwick & Sauder (2005). Briefly, most parsimonious trees were generated using the heuristic search algorithm with 1000 replicates of random addition sequence with equal weight, tree-bisection-reconnection (TBR) branch swapping and 200 trees retained per replicate. Support values for clades in the strict consensus tree were estimated by running 1000 bootstrap replicates with TBR branch swapping, random addition of taxa and saving multiple trees.

AFLP data

To further investigate intra-specific relationships among *B. repanda* subspecies and test the monophyly of genotypes of BRB and BRG, phylogenetic analyses were conducted on the AFLP data set.

A first phylogenetic reconstruction was obtained using a reduced AFLP matrix including a proportionate number of samples for each subspecific entity. The matrix was composed of 47 accessions, including 5 individuals for each of the representative subspecies, except for *B. repanda* subsp. *blancoana*, subsp. *repanda*, subsp. *repanda* s.l. and subsp. *confusa* for which respectively 4, 3, 3 and 2 individuals were available. Individuals were randomly chosen for BRB and BRG. To further test for monophyly of the eastern endemics, a subsequent analysis was performed on the complete matrix including all samples of this study. Phylogenetic analyses were conducted using Maximum Parsimony (MP) and Bayesian Inference (BI) methods. Parsimony trees were

constructed following the same procedure described above. Bayesian analyses were carried out implementing the restriction model in MrBayes (Huelsenbeck et al., 2001; Ronquist & Huelsenbeck, 2003) with the option "noabsencesites". For the 47-accession matrix, searches consisted of 2 runs of $2*10^5$ generations, each with 4 Metropolis-coupled Markov Chain Monte Carlo (MCMC) chains with incremental heating of 0.1 and sampling every 100 generations. Analyses of the total matrix were run for $2*10^7$ generations, incremental heating on 0.05 and sampling every 1000 generations.

Assessment of convergence relied on the standard deviation of split frequencies as well as on the effective sampling size criterion for each parameter as implemented in Tracer v1.4 (Drummond & Rambaut, 2007). Trees were summarized in a maximum clade credibility tree computed after 20% burn-in.

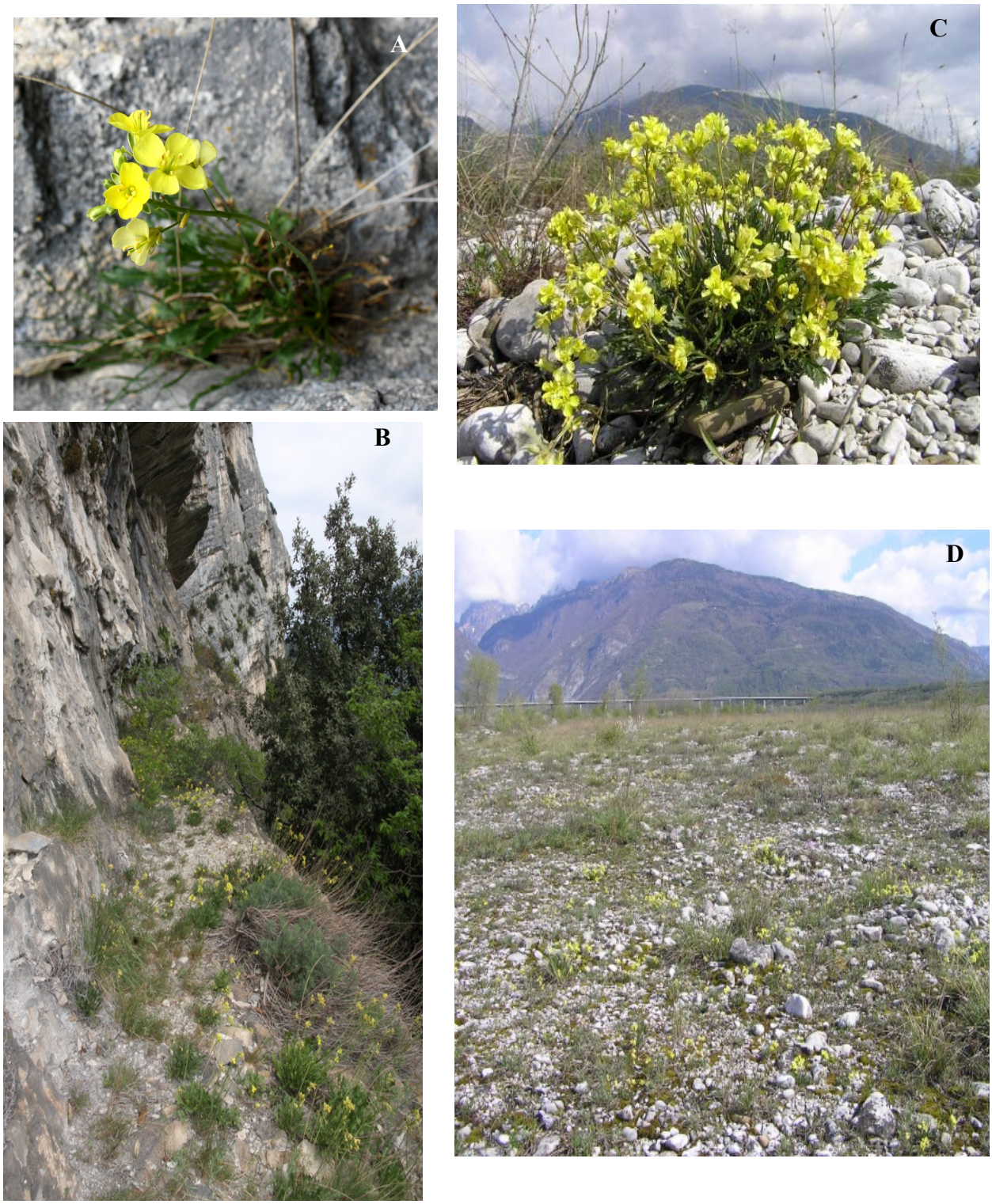
To test for diagnosability of the endemics according to the phylogenetic species concept (PSC; (Cracraft, 1983; Nixon & Wheeler, 1990), we used Population Aggregation Analysis (PAA; Davis & Nixon, 1992) based on AFLP data. Using this method, we aimed to assess whether BRB and BRG were diagnosable by a unique set of fixed characters both between each other, and with respect to the remainder of the *B. repanda* complex. As the great majority of missing data in the matrixes corresponded to peaks of ambiguous interpretation, the previously estimated scoring error rate (i.e. 0.028) was used as a threshold to exclude from the analyses those loci containing a number of missing data greater than this value. Moreover, samples of thousands of individuals would be required to identify with certainty diagnostic characters carrying null frequencies of polymorphisms. As a practical solution to this problem, Wiens & Servedio (2000) proposed to use a non-zero frequency cut-off for polymorphisms. Here, a similar approach was used to address the issue of finite sample size, and characters were considered as fixed within each group allowing a frequency cut-off of polymorphisms corresponding to the genotyping error rate (0.076). Thus, given that 104/38/37 samples were available for BRB/BRG/*B. repanda* complex, the minimum number of samples required to define a character as 'fixed' was adjusted to 97/36/35.

To test for the genotypic cluster criterion (Mallet, 1995), genetic structure in *B. repanda* was investigated with a PCO-MC analysis (Reeves & Richards, 2009). Assessment of the spatial distribution of the genetic variability through PCO was conducted on both the matrices described above. The clustering procedure followed the authors' recommendations (available at: http://lamar.colostate.edu/~reevesp/PCOMC/PCOMC.html): stability of clusters was first assessed with p-value cut-off set to 0.999 and a stability cut-off to 15. A second analysis was performed with

p-value cut-off set to 0.05 to test for statistical significance of the clusters discovered in the first analysis. Verification of the genotypic cluster criterion was also carried out with STRUCTURE software version 2.3.1 (Pritchard, Stephens & Donnelly, 2000; Falush, Stephens & Pritchard, 2003), using the approach of Reeves & Richards (2011) and Meudt, Lockhart & Bryant (2009). AFLPs were coded as dominant data following Evanno, Regnaut & Goudet (2005). Admixture and correlated frequencies models and non-informative priors were applied. The most likely number of populations (K) was estimated with 20 replicates each for K = 1 to K = 13 using 10^6 iterations in the burn-in period and 10^6 iterations in the data collection phase. To speed up the simulation analyses, we reduced the sample size of BRB to 38 individuals out of 104. These were randomly chosen within the three populations with greatest geographic distance in the range of the subspecies (i.e. C, F and G in Table 1). K was estimated using the Δ K statistic, based on the rate of change in the log probability of the data between successive clusters (Evanno et al., 2005). Finally, the fixation index between BRG + BRB taken together and the remainder of the complex was calculated in ARLEQUIN 3.5, and statistical robustness was evaluated through 1000 permutations.

b) Morphological, ecological and phenological data

To assess the variation in morphological, ecological and phenological characters between each other and with the remainder of the complex, we examined data obtained from inspection of herbarium specimens and from literature. Detailed information on the characters examined and literature sources is provided in Table 3.



Picture 1. **A**. *B. repanda baldensis* growing on the rock. **B.** Growing site of *B. repanda baldensis*. **C**. *B. repanda glabrescens*. **D.** Growing site of *B. repanda glabrescens*. Pictures **A** and **B** by Filippo Prosser and Alessio Bertolli; pictures **C** and **D** by Adriano Bruna.

Table 3. Morphological, ecological and phenological characters analysed in BRG, BRB and the remaining ssp. of *B. repanda* .

Morphology	***B. repanda glabrescens***	***B.repanda baldensis***	***B. repanda*** **(remaining ssp.)**
Floral stem length (cm)	10-25	30-60 (80)	8-60 (80)
Leaf morphology	pinnatifid	sinuate to pinnatifid (very rarely entire)	entire to pinnatifid or pinnatipartite
Leaf length (cm)	2-9	5-23	2-18
Leaf pubescence	glabrous except for the margins	ciliate at the margin (very rarely sparsely hairy)	glabrous to densely hispid
Number of flowers	2-12	15-30	2-35
Flower colour	pale yellow	bright yellow	between pale and bright yellow
Petal length (mm)	7-10	12-16	8-25 (30)
Siliqua length (mm)	25-60	30-70	20-90
Silique width (mm)	1.5-3 (max. width towards the end)	3-5 (6)	1-5
Silique position	horizontal to ascendent	horizontal to pendulous	pendulous to erecto-patent
Seed length (mm)	1.5-2	2.2	0.6-5.5
Seed disposition	uniseriate	uniseriate to biseriate	uniseriate to biseriate
Ecology			
Habitat	oligotrophic and calcareous river shore, sun exposure	arid rain shelters under projecting cliffs, on limestone, high temperature, sun exposure	arid rain shelters under projecting cliffs / rocks / clayey and thermophilous gullies / clayey and alpine gullies
Phenology			
Flowering time*	4-5	4-5-6	3-4-5-6-7
Fruiting time*	5-6-7	6-7	4-5-6-7-8

** Numbers refer to months of the year, from January (1) to December (12).*

c) Population genetic analyses

The number of genetically defined populations of BRG and BRB was investigated using 2 clustering methods based on Bayesian models. STRUCTURE software was first used, following the same procedure as explained above. As the method of Evanno et al. (2005) is never able to identify K = 1 as the most probable K, a combined data set of all individuals of BRG and BRB was first analysed. Secondly, a spatial analysis (Guillot et al., 2005) was carried out on BRB, for which more than 2 populations were available. The analysis was conducted using the extension for R (R Development Core Team, 2011) of the GENELAND program (version 3.2.4, Guillot, Mortier & Estoup, 2005; Guillot & Santos, 2010). We ran the MCMC 20 times with the uncorrelated frequencies model, spatial coordinates with an uncertainty of 30 m corresponding to the average size of sampling locations, 104 iterations, a burnin period of 2*104 iterations, allowing K to vary from 1 to 10. To further look into genetic relatedness between taxa and among individuals within taxa, PCO-MC was also applied following the same method described above. Diversity indices were further calculated for BRB and BRG populations with ARLEQUIN version 3.5 (Excoffier & Lischer, 2010) based on sampling locations. Gene diversity statistics, obtained with the method of Weir & Cockerham (1984), were PPL (proportion of polymorphic loci), H*exp* (expected heterozigosity under assumption of Hardy-Weinberg equilibrium [HWE]) within sampling locations and mean H*exp*. Furthermore, to assess the extent of partitioning of molecular variance among and within BRB and BRG pairwise distances were calculated according to Weir & Cockerham (1984), and hierarchical analyses of molecular variance (AMOVA, Excoffier, Smouse & Quattro, 1992) were carried out. AFLP data were subjected to a first AMOVA assuming between-taxa grouping and no within-taxa structure. A second AMOVA was then executed independently on BRB and BRG following sampling locations subdivision. Significance of group partitioning was tested against alternative random distribution of individuals among groups through 1000 random permutations. Finally, the fixation index among all subspecies of the *B. repanda* complex was measured with pairwise F_{ST} values in ARLEQUIN 3.5, and statistical robustness estimated with 1000 permutations.

2.4. Results

a) Phylogenetic and population genetic analyses

ITS sequences

Trees reconstructed using the ITS sequences of representatives of the *B. repanda* complex inserted within the large taxonomic sampling of the tribe Brassiceae by Warwick & Sauder (2005) were consistent in the resolution of supported groups with the results of the previous authors. *Brassica repanda* accessions formed a highly supported monophyletic group (MP=99%, BI=1) but no further resolution among the subspecific entities was recovered (Fig. A – Supplementary Material).

AFLP data

The primer combinations used in the data set including 179 individuals from 10 different taxa produced 743 scorable AFLP bands, of which 628 were polymorphic and included in the analysis. Trees constructed on the 47-accession AFLP data set (Fig. 2A and Fig. B – Supplementary Material) showed a number of strongly supported groups (MP>80%; BI=1) largely consistent with the taxonomic division of the *B. repanda* complex. However, some of the subspecies were not monophyletic (e.g. *B. repanda* subsp. *repanda*, subsp. *cadevallii*). Consistent results were recovered in the analyses conducted on the entire matrix (Fig. C – Supplementary Material). In both analyses, rooting of the tree in any of the western groups resulted in a monophyletic origin of the eastern disjunction (MP≥99%, BI≥0.98), i.e., a "clan" in the terminology proposed by Wilkinson et al. (2007), where BRB (MP≥80%, BI≥0.98) and BRG (MP≥85%, BI=1) are "adjacent groups".

PAA identified 6 AFLP loci as fixed characters that were constant within but variable among all the groups considered in the analysis (Table 4: E35/M42-135bp; E35/M42-173bp; E35/M42-219bp; E33/M37-105bp; E33/M37-221bp; E36/M36-122bp). In particular, BRB and BRG are individually diagnosable by 1 and 2 characters respectively (marked with asterisks in Table 4), whereas 3 characters were fixed among the representative samples of the remainder of the subspecies. Further, each endemic can be diagnosed by using several combinations of 1 or more characters that are fixed among their representatives. It should be noted that while these results support the compliance with

the diagnosability criterion required by the PSC for both BRB and BRG, they do not imply that the PSC applies to the group including the rest of the subspecies. In fact, PAA identifies fixed characters within and between defined groups (Davis & Nixon, 1992), and as such, it relies on the initial entities used for aggregation. In this study, as we aimed to test for the diagnosability of the eastern endemics, we used the remainder of the complex as a representative sampling of allele diversity outside these taxa. In this context, fixed characters among the rest of the subspecies are better interpreted as shared alleles between the endemics. Indeed, beyond the aims of the present study, it may be interesting to carry out a more exhaustive sampling including an higher number of populations and individuals of the remainder of the complex in order to check which of these characters is really fixed and which is variable among the subspecies.

Using the program STRUCTURE, the highest value for ΔK was obtained for K = 3, corresponding, respectively, to BRB, BRG and the remaining *B. repanda* subspecies. The estimated ln probability of the data for K set to 3 was similar for most replicate simulations (mean ± standard deviation [s.d.] = - 20123.3 ± 17.9). The percentages of individuals assigned to the clusters cited above with an a posteriori probability > 90% were 95%, 82% and 84% respectively. The graphical output of the STRUCTURE analysis is reported in Figure 2C.

Results of PCO analyses performed on both the reduced matrix with 5 individuals per taxon (Fig. D - Supplementary Material) and on the complete data set (Fig. 2B) consistently show the clear distinctiveness of BRB and BRG with respect to the remaining subspecies. In the larger data set, the first axis of principal coordinates analysis represents 12.9% of the total variance and clearly separates BRB from all other taxa, the second axis represents 7.4% of the total variance and separates BRG from the remaining *B. repanda* subspecies here represented, while the third axis explains the 5.9 % of the total variance. PCO-MC yields sound stability values (> 30) in support of this genetic structure, and groups are significantly distinct from each other (p-value < 0.05).

Lastly, the F_{ST} estimate of BRB and BRG taken together against the remainder of the complex was 0.26 (p value = 0.00).

b) Morphological, ecological and phenological data

The majority of morphological characters analysed in Table 3 were continuous and highly variable within groups (e.g. silique length and number of flowers). Moreover, their ranges of variability generally overlapped among groups (e.g. leaf, silique and seed length) or differed for a few millimetres (e.g. silique width and petal length). Discrete morphological characters (e.g. seed position, leaf morphology, leaf pubescence), as well as ecological and phenological ones (habitat, flowering and fruiting time), were also highly variable.

Notwithstanding, major morphological discontinuities were found between BRG and BRB in the number and colour of flowers, silique width, and in floral stem, petal and seed length. Habitat preferences are also known to be very different between the 2 taxa. On the other hand, comparison with the remainder of the complex revealed that none of these characters taken singularly could be interpreted as diagnostic for either endemic, as the range of variability was shared with that of other subspecies of the complex.

Table 4. Population profiles for diagnostic AFLP characters detected by PAA. Presence/absence of characters is indicated by 1/0; characters that are variable within a group (i.e. "traits" *sensu* Davis & Nixon,1992) are indicated by ±

Character †	BRB	BRG	*B. repanda* sp.
E-35 / M-42 - 78bp	1	0	±
E-35 / M-42 - 101bp	1	0	±
E-35 / M-42 - 135bp	1	1	0
E-35 / M-42 - 173bp	0	1 *	0
E-35 / M-42 - 181bp	±	1	0
E-35 / M-42 - 219bp	1 *	0	0
E-33 / M-32 - 81bp	±	1	0
E-33 / M-32 - 129bp	±	1	0
E-33 / M-32 - 194bp	1	0	±
E-33 / M-32 - 244bp	1	±	0
E-33 / M-37 - 105bp	0	1 *	0
E-33 / M-37 - 188bp	0	1	±
E-33 / M-37 - 221bp	1	1	0
E-33 / M-37 - 226bp	1	0	±
E-36 / M-36 - 122bp	0	0	1

† "E" and "M" indicate EcoRI and MseI selective primers. See Table 2 for the complete list of EcorI-MseI primer sequences.

* Indicates characters diagnostic for the endemics.

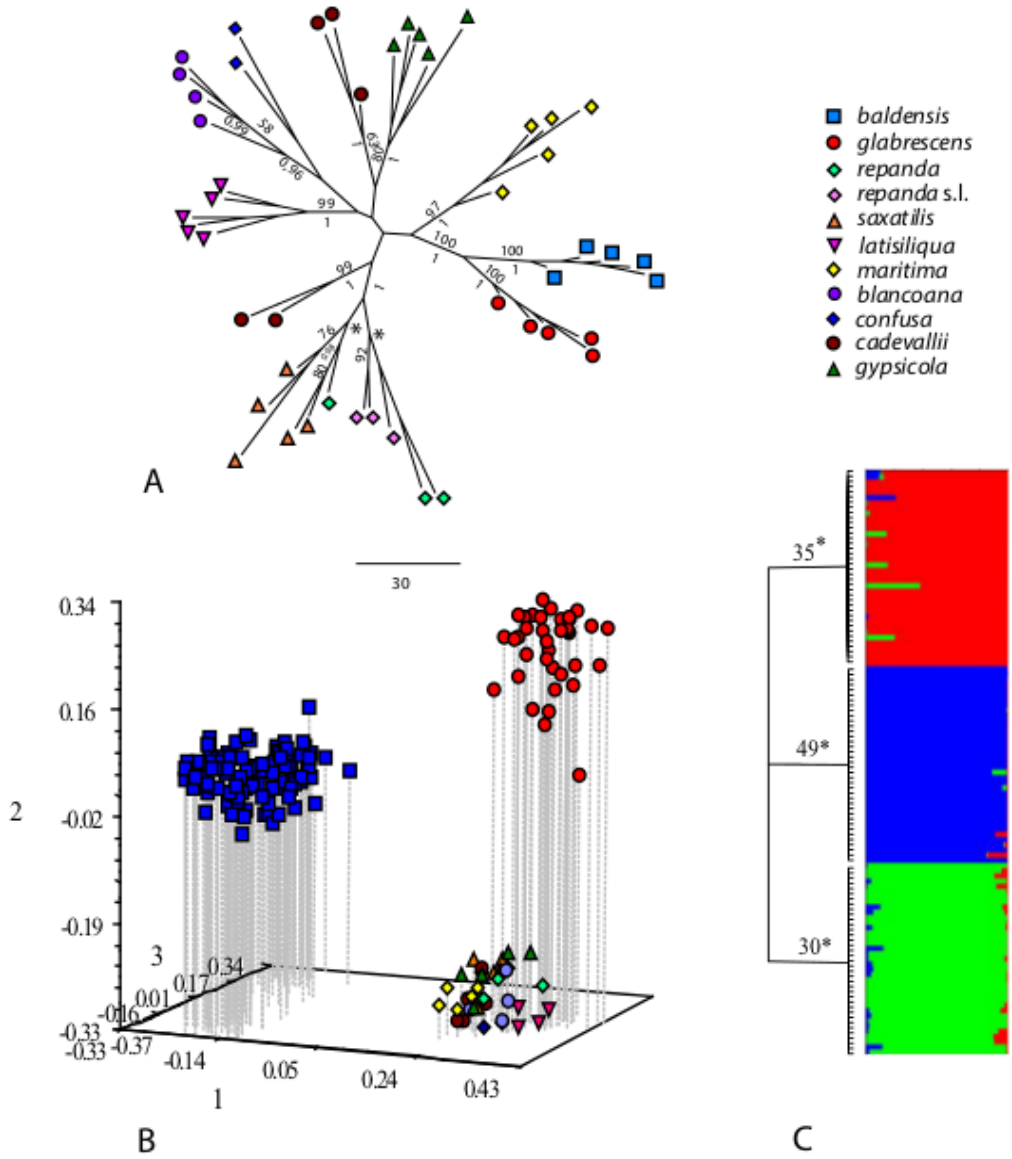


Figure 2. Phylogenetic and population genetic results from the analyses of AFLP data. **A.** Unrooted MP phylogram obtained for the 47-accession matrix. MP bootstrap support (>50%) and BI posterior probability (>0.95) values are shown above and below branches, respectively. Asterisks indicate nodes that collapse in the MP strict consensus tree. Scale bar indicates branch lengths. **B.** PCO plot obtained from the total matrix. Clear separation of BRB, BRG and the remaining *B. repanda* subspecies is evident along the first 2 PCO axes. **C.** Combined results from the STRUCTURE and PCO-MC analyses conducted on the total matrix. Tree at left is the hierarchical assignment resulting from PCO-MC analysis. Numbers at nodes are cluster stability values. Asterisks indicate significantly distinct clusters (p-value < 0.05). PCO-MC analyses were conducted on all individuals but only those included in the STRUCTURE simulations are shown here to permit direct comparison. Admixture within subspecies is shown in bar graph at right. The length of colored bars represents the fractional assignment of individuals to each of K = 3 genetic clusters inferred by STRUCTURE: red = Cluster 1 (BRG), blue = Cluster 2 (BRB), and green = Cluster 3 (the remaining *B. repanda* subspecies).

c) Comparative population genetic analyses

The combined data set of BRG and BRB used in the program STRUCTURE yielded K = 2 as the most probable a posteriori K, with both BRG and BRB forming a unique and consistent cluster in all 20 replicate runs.

The estimated ln probability of the data for K set to 2 was similar for most replicate simulations (mean ± 1 standard deviation [SD] = - 21146.3 ± 0.3). Thus, STRUCTURE showed no support for the existence of more than one population in either endemic. Similarly, GENELAND identified K = 1 as the modal number of populations in all 20 independent MCMC simulations carried out on the BRB data set. So, even with a spatial prior assigning the individuals to 5 distinct sampling locations, only one panmictic population was recognized.

PCO-MC output validated the results of the Bayesian analyses, producing the same 2 clusters (p-value < 0.05): individuals of BRG grouped together with a stability value of 90, separate from BRB individuals, which had a stability value of 52.

The diversity indices calculated on sampling locations are shown in Table 5. Genetic diversity in each sampling location of each taxon was generally low. Nevertheless, BRG showed a higher percentage of polymorphic loci (mean PPL = 38.6%) than BRB (mean PPL = 27.5%). Estimated heterozygosity was equal to 0.10 and 0.08 in BRG and BRB respectively, and the difference between means was not statistically significant (t-test, p-value = 0.76). Interestingly, considering the genetic diversity among BRB sampling locations, the individuals collected in Monte Cimo (population "G") showed both the highest values of PPL (31.1%) and Hexp (0.09).

The overall partitioning of molecular variance between taxa (Table 6) confirmed the distinctiveness of the 2 endemics of the eastern Alps (Φ = 0.43, p-value = 0.00). Further, the division of variance within BRB revealed a higher percentage of within (81.8%) than among (18.2%) sampling locations diversity. Nevertheless, the latter is highly statistically significant as well as the among population fixation index (Φ = 0.18, p-value = 0.00). BRG presents even higher levels of within (93.9%) than among (6.1%) groups diversity, with a lower, though statistically significant, fixation index (Φ = 0.06, p-value = 0.00) than BRB.

Pairwise F_{ST} values among all subspecies of the *B. repanda* complex are shown in Table 7. As a confirm of previous analyses, BRG and BRB carried the highest F_{ST} values, both between one

another and with regards to every subspecies of the *B. repanda* complex. The highest fixation index was found between BRG and *B. repanda* subsp. *confusa* (F_{ST} = 0.58, $p < 0.05$). On the other hand, the lowest levels of divergence across the whole data set were between *B. repanda* subsp. *repanda* and subsp. *saxatilis* (F_{ST} = 0.09, $p < 0.05$), as well as between the former and subsp. *cadevallii* (F_{ST} = 0.17, $p < 0.05$).

Table 5. Genetic diversity estimates within BRB and BRG. *PPL* indicates percentage of polymorphic loci; H*exp*: within location estimated heterozygosity under HWE; mean H*exp*: mean within taxon estimated heterozygosity under HWE.

Taxon	Sampling locality	*PPL*	Mean *PPL*	H*exp*	Mean H*exp*
BRG	A	39.1	38.6	0.10	0.10
	B	38.0		0.11	
BRB	C	24.7		0.07	
			27.5		0.08
	D	26.0		0.07	
	G	31.1		0.09	
	E	27.8		0.08	
	F	28.0		0.07	

Table 6. Partitioning of molecular variance among and within BRB and BRG. *df* indicates degrees of freedom, *SS* sum of squares.

Component of variation	*df*	*SS*	Variance component	% Variation	Fixation index Φ	Significance
All *taxa*						
Among *taxa*	1	1243.050	21.809	42.9	0.43	P= 0.000
Within *taxa*	140	4073.035	29.093	57.2		P= 0.000
BRB						P= 0.000
Among sampling localities	4	517.801	5.122	18.2	0.18	P= 0.000
Within sampling localities	99	2285.497	23.086	81.8		P= 0.000
BRG						P= 0.000
Among sampling localities	1	73.959	2.150	6.1	0.06	P= 0.000
Within sampling localities	36	1195.778	33.216	93.9		P= 0.000

Table 7. Pairwise F_{ST} values estimated for every pair of *B. repanda* subspecies (p-values < 0.05 over 1000 bootstrapped matrices, except for *con/bla* pairwise comparison marked with ^{ns}). *B. repanda* ssp. are identified with the first 3 letters of their subspecific epithet.

	bal	**gla**	**rep**	**sax**	**lat**	**bla**	**con**	**cad**	**gyp**	**mar**
bal	-									
gla	0.40	-								
rep	0.41	0.45	-							
sax	0.47	0.47	0.09	-						
lat	0.49	0.55	0.32	0.38	-					
bla	0.46	0.51	0.24	0.32	0.34	-				
con	0.55	0.58	0.26	0.33	0.36	0.24^{ns}	-			
cad	0.43	0.44	0.17	0.22	0.24	0.19	0.20	-		
gyp	0.44	0.49	0.28	0.35	0.33	0.29	0.31	0.19	-	
mar	0.49	0.50	0.33	0.42	0.40	0.40	0.41	0.26	0.30	-

2.5. Discussion

Taxonomic delimitation at the intra-generic or lower taxonomic levels is often difficult due to the low phenotypic divergence among taxa and/or the low information content of the genes normally used for phylogenetic reconstruction at higher phylogenetic distances. In this study, the use of the highly variable nuclear ITS region did not provide sufficient phylogenetic signal to reconstruct the evolutionary and genealogical relationships among representatives of the *B. repanda* complex. Recent studies (Meudt, Lockhart & Bryant, 2009; Koopman et al., 2008; Després, Gielly & Taberlet, 2003) reported the successful use of AFLP markers for taxonomic delimitation in case of closely related taxa, as they provide a reliable means to sample the genetic variation among homologous sites coupled with the advantage of a multilocus approach. In the case under study, AFLPs proved to be indeed highly informative and appropriate to identify genetic discontinuities of the taxa representing the easternmost disjunct distribution of the *B. repanda* complex.

a) Evolution and taxonomic status of *B. repanda baldensis* and *glabrescens*

Previous evidence on the phylogeny of the tribe Brassiceae showed a complex pattern of relationships within this tribe, where taxonomic classifications are challenged by the molecular evidence of highly polyphyletic genera (Warwick & Sauder, 2005). In this work, the inclusion of 10 subspecific entities of *B. repanda* within the ITS phylogeny of the tribe Brassiceae provides evidence of the monophyly of these taxa.

Within the *B. repanda* complex, the taxonomic recognition of BRG and BRB as species is investigated in light of the plurality of evidence gathered from the compliance with the criteria of monophyly (Donoghue 1985; de Queiroz & Donoghue, 1988), diagnosability (Cracraft, 1983; Nixon & Wheeler, 1990) and genotypic clustering (Mallet, 1995).

Support for monophyly comes from the phylogenetic analysis of AFLP variation, showing both BRG and BRB as strongly supported clades. Genetic diagnosability of the endemics is shown by PAA, which supports the compliance with the PSC for both BRG and BRB. In the PSC, the species is defined on the basis of fixed character differences, which are unique for the species, and thus diagnostic with regards to other populations or group of populations. Although fixed characters are

most easily described by private alleles, species delimitation is also inferred from a unique combination of characters that can be used to diagnose the groups (Davis & Nixon, 1992). In this study, BRG and BRB can be identified, respectively, by 2 and one private loci, as well as by the combination of several characters with fixed opposite state in at least 2 of the groups used in the analysis (see Table 4).

Distinctiveness of the endemics in compliance with the genotypic cluster criterion is provided by population genetics analyses. In fact, while PAA is a straightforward, qualitative approach to investigate how genotypic data support the distinctiveness of taxa, population genetics analyses offer a quantitative approach to test for genetic discontinuities. Results obtained by multivariate algorithms (PCO-MC) discriminate both the endemics with high stability and significance values, and consistent results are yielded by the Bayesian analyses with STRUCTURE. Here, K = 3 is found as the most probable number of clusters and there is little evidence of admixture among individuals of BRB and BRG.

The properties acquired by the endemics depict clear lineage divergence consequent to the independent evolutionary patterns of the 2 taxa. The possible evolutionary processes producing such patterns, indeed, constitute independent species criteria per se and might add further evidence to the distinctiveness of the endemics. Although we do not provide direct evidence of such processes, it is reasonable to assume that the geographic isolation of the endemics has played an important role in maintaining lineage separation. In fact, based on geographic distances, reproductive isolation of the endemics can be considered, at a minimum, extrinsic. This hypothesis is supported by the high value of the fixation index between BRB and BRG ($\Phi = 0.428$, p-value = 0.000), which are the 2 geographically closest groups in the eastern range of the species distribution, separated by approximately 150 km and with no intermediate populations reported so far. As *Brassica* species are entomogamous and mainly pollinated by hymenopters, lepidopters and dipters whose maximum flight distance usually does not exceed a few km (Hagen, Wikelski & Kissling, 2011; Van Rossum & Triest, 2010; Osborne et al., 2008), we consider gene flow between BRB and BRG very unlikely. Gene flow through seed dispersal is even less probable, because *Brassica* seeds are not equipped with wings, pappus or plumes for wind transport, and dispersal by water cannot occur since the hydrographic systems where the 2 endemics grow are isolated. The fact that we find indication of low level of admixture and shared fixed characters between the eastern endemics could be explained either by historical gene flow due to a past geographical continuum or shared ancestry followed by rare dispersal events causing the disjunction.

The evidence provided thus far shows that BRG and BRB have acquired multiple properties that satisfy the phylogenetic criteria of monophyly (Donoghue 1985; de Queiroz & Donoghue, 1988) and diagnosability (Cracraft, 1983; Nixon & Wheeler, 1990), and the genotypic cluster criterion (Mallet, 1995). Furthermore, it is plausible to hypothesize that reproductive isolation by virtue of geographic isolation may maintain divergence of the biology of the eastern endemics, although the intrinsic property of absence of interbreeding (i.e. the biological concept of species; Mayr, 1942; Dobzhansky, 1970) cannot be verified with the data in our hands. Based on these findings, we suggest that BRG and BRB should be regarded as distinct species from the rest of the complex, and we herein propose their taxonomic recombination (see below).

The findings of this work offer yet another instance within the Brassiceae where the phenotypic similarity featured by closely related taxa conceals the molecular divergence resulting from the ongoing process of independent evolution favoured by geographic separation. In fact, most of the properties that describe the morphology, ecology and phenology of the endemics appear homoplastic when compared across the taxonomic diversity of the *B. repanda* complex (see Table 3), where subspecific entities are better identified by a combination of features rather than a single, private character. However, the endemics are identifiable using a combination of characters that are diagnostic within the complex, as thoroughly described by the authors of their descriptions (Prosser & Bertolli, 2007a; Poldini, 1973). In fact, this is not surprising, as the systematic knowledge gathered in recent years by combining molecular and phenotypic analyses in many groups within the Brassicaceae suggests that morphological characters should not be used alone in establishing taxonomic boundaries, especially at the genus or lower level (Al-Shehbaz, Beilstein & Kellogg, 2006). Traits such as floral and fruit morphology, seed embryo type and seedling development have proven to be highly homoplasious throughout the whole family (Beilstein, Al-Shehbaz & Kellogg, 2006; Mummenhoff, Franzke & Koch, 1997a), and leaf morphology is known to vary both among and within species in Brassicaceae.

In our study, we took advantage of the geographic isolation of the endemics, which is reflected in the long genetic distance that separates them from the narrow cluster of the remainder of the taxa. The results of the PCO plot clearly show that the genetic variability that is used to infer the distinctiveness of the eastern endemics is considerably larger than that existing among several representatives of the taxonomy and distribution of the rest of the group, which includes the type of the complex, i.e. *B. repanda* subsp. *repanda*.

Despite ample evidence of lineage separation, it remains difficult to pinpoint the evolutionary

history of the eastern disjunction. Our phylogenetic results indicate that the origin of the endemics may be attributable to a single biogeographic event. Being our tree unrooted, however, the actual events that led to the disjunction remain unknown. A speculative hypothesis could be drawn from the PCO analysis performed on the reduced matrix, where more space is taken by the components of the allelic variation of the western alpine representatives of the *B. repanda* complex (Fig. D - Supplementary Material). Here, the subspecies of western Alps separate along the second axis of principal coordinates, owing to the acquisition of allelic variation that is not shared by either Spanish or eastern taxa. Although genetic distance is not always a good predictor of evolutionary relationships, the results of the PCO plot indicate that the eastern endemics are unlikely to have originated following a gradual eastwards colonization from the main centre of distribution of *B. repanda* (the Iberian peninsula), where the western alpine representatives would act as a "stepping -stone" (Kimura & Weiss, 1964) in such a process. The hypotheses of a long-distance colonization event from Spanish representatives or, rather, a fragmentation scenario where western and eastern alpine populations developed independent allelic variation appear more likely.

Moreover, it has been suggested that BRB originated either following isolation on Mt. Baldo during the Pleistocenic glaciations or as a relic of a more widely distributed ancestor that colonized the slopes of the mountain during xerothermic post-glacial periods (Prosser & Bertolli, 2007a). On the other hand, BRG is part of the xerophylous vegetation of the ‘magredi’ formations, consisting of alluvial sediments of post-glacial origin (Poldini, 1973). Considering the shared ancestry recovered in this study, an origin of the endemics linked to the fragmentation of a continuum xerothermic distribution seems a plausible hypothesis, and recent methods using the coalescent approach may enable to test this scenario versus alternative evolutionary models.

b) Patterns of genetic divergence within the French-Iberian complex

It should be noted that the results of this study do not provide any conclusion on the taxonomy of the remainder of the *B. repanda* complex. Whether *B. repanda* satisfies any of the species criteria of the USC is not proved by any of the analyses of this study, which only addressed the issue of the distinctiveness of the eastern endemics, using the other subspecies of *B. repanda* as representatives of the western distribution. Presence of further distinct lineages among other subspecies is possible

(including those not sampled in this study, e.g., the 3 polyploid North African taxa [Prosser & Bertolli, 2007a]), but these should be investigated with a larger sampling of the taxonomic and geographic diversity of the western representatives suitable to this aim. Nonetheless, F_{ST} values calculated between every pair of taxa already provided some interesting hints concerning the levels of divergence among subspecies within the *B. repanda* complex. Indeed, while BRB and BRG were confirmed their distinct taxonomic position both between one another and with regards to the rest, the lowest fixation index across the whole data set was found between *B. repanda* subsp. *repanda* and *saxatilis*. The genetic similarity among these 2 taxa was also clearly indicated by the NJ tree, where a unique clan was detected that grouped the western alpine subspecies together, with one population of *B. repanda* subsp. *saxatilis* belonging to the subsp. *repanda* clan. This evidence is in accordance with the geographical proximity of the 2 subspecies, whose distribution ranges partially overlap. Similarly, the second lowest F_{ST} estimate was found between subsp. *repanda* and *cadevallii*, which is the immediately nearest taxon located in the Iberian peninsula. Although being the physically closest subspecies to the endemics of eastern Alps, subsp. *repanda* and *saxatilis* were not genetically more similar to them than to rest of the *B. repanda* complex, thus confirming morphological evidences (Prosser & Bertolli, 2007). *B. repanda* s.l. populations sampled in Val Maira and Val Susa were as well genetically differentiated from the eastern subspecies, despite sharing some morphological features with BRB. This genetic distinctiveness may be partly explained by the existence of an ecological discontinuity between the Provençal-western and the eastern Alps in the distribution of calcareous bedrock necessary for the growth of *B. repanda* (Schönswetter, 2005). Alvarez et al. (2009) demonstrated that soil substrate was an important determinant of spatial genetic structure in some alpine plants. Here, absence of suitable substrate possibly hampered the colonization of intermediate areas in central Alps, thus strengthening the process of genetic differentiation between the eastern disjunction and subsp. *repanda* / *saxatilis*, as well as the other *B. repanda* subspecies.

The pairwise F_{ST} values detected within the Iberian complex were generally lower than the pairwise comparisons with the eastern disjunction, suggesting that here the subspecies may be less differentiated from each other. Nevertheless, some important discontinuities were found, especially concerning the most geographically isolated endemics with very restricted distributions, that partly

reflected the separate clans of the Neighbor Joining phylogram, like subsp. *maritima* and subsp. *latisiliqua.*

However, as mentioned above, the sampling carried out within the present work is not sufficient to unravel accurately the taxonomic and evolutionary relationships among the French and Iberian *B. repanda* subspecies. A thorough sampling involving more individuals, population and subspecies across all the distribution area of *B. repanda* should be carried out for this purpose.

c) Comparative population genetics of *B. repanda baldensis* and *glabrescens*

Our comparative analyses did not show any significant difference in the levels and patterns of genetic variation in BRG and BRB. Bayesian analyses either with or without a spatial prior, as well as multivariate statistics, all recognize only one panmictic population both in BRB and BRG. Absence of genetic structure within the taxa is expected based on their narrow distribution and the close proximity of the sampling sites without apparent geographic barriers to pollination. For BRG, seed dispersal by water (Poldini, 1973) may further contribute to admixture. Existence of panmixis in both taxa suggests gene flow and recombination among the different sampled localities, however, the limited number of individuals censused in natural populations (see above) makes them possibly vulnerable to genetic drift and inbreeding (Hartl & Clark, 2007; Frankham, Ballou & Briscoe, 2002). The levels of polymorphic loci and estimated heterozygosity in both BRG and BRB (mean H*exp* = 0.09) are, indeed, moderately low if compared to those reported by other AFLP studies on herbaceous outcrossing and animal-pollinated species [e.g. Sánchez-Teyer et al. (2009), mean H*exp* = 0.26; Kreivi, Rautiainen, Aspi & Hyva (2005), mean Nei's H*exp* = 0.27; Chung, Gelembiuk & Givnish (2004), mean H*exp* = 0.13; Gaudeul, Taberlet & Till-Bottraud (2000), mean Nei's H*exp* = 0.20]. Despite the lack of field data concerning their mating system, AMOVA within taxa on the different sampling localities show for both BRB and BRG low levels of among-population variation which are typical of outcrossing species (Hamrick & Godt, 1996). Variance among groups is nonetheless significant, indicating that the amount of gene flow among locations is limited and has

possibly led to population differentiation.

d) Implications for conservation

There are indications of a progressive restriction of suitable habitat for both taxa, which could be responsible for restraint in gene flow mentioned above. Indeed, BRG is located in 2 European Sites of Community Importance (SCI) "Magredi del Cellina" and "Magredi di Pordenone" (European Environment Agency 2010), with a decreasing population trend and distribution range reported for the period 1972–2006 (Branca et al., 2011). The taxon is mainly threatened by habitat destruction and pollution caused by agricultural and zootechnical activities and by human-made morphological and hydrogeographical modifications of the rivers. Indeed, its "*locus classicus*" was devastated by setting up of vineyards. It was classified as "vulnerable" at a regional (Conti, Manzi & Pedrotti, 1997) and national scale (Conti, Manzi & Pedrotti, 1992) and recently listed as "vulnerable" also at the European (Bilz et al., 2011) and global level (Branca et al., 2011). As to the effective conservation of BRG, some measures are already in progress and need to be strengthen (http://www.magredinatura2000.it). These include the ongoing restoration of some recently degraded meadows which represent potential growing sites for the endemic and maintenance of the existing ones, as well as asbestos abatement from some contaminated sites.

BRB is located in the SCI “Mt. Baldo East”. The co-occurrence of some other xerothermic species with a disjunction on Mt Baldo (Prosser & Bertolli, 2007a) may indicate that the actual population of BRB is a relic of a more widespread taxon of pre-glacial origin which underwent distribution decline. The level of threat for this taxon is under investigation, as the increased afforestation of the exposed ledges constituting its habitat suggests that *in situ* conservation actions may be needed to decrease inter-specific competition and preserve population connectivity (Bertolli & Prosser, unpublished results). Effective conservation actions should thus include mowing of shrubs (namely young trees of *Quercus ilex*) and grasses in the neighbourhood of the calcareous projecting rocks were BRB is already settled or where it could potentially establish new populations. Moreover, the endemic is possibly threatened by rock climbers, whose rock routes partly overlap with the growing sites of populations (Prosser & Bertolli, 2007a). Though probably not alarming, the impact of rock

climbers may become important during the flowering and fruiting season, between April and June. Therefore, information panels in loco that warn against trampling of plants both on the rock routes and on the paths towards them could be useful to limit plants damage and assure seeds ripening.

In addition to pointing out the need for *in situ* conservation, genetic analyses provide useful information for guiding ex-situ germplasm collection. Branca et al. (2011) reported for BRG only one germplasm accession in European genebanks (EURISCO Catalogue 2010), and seeds of BRB have not yet been collected and stocked in the local Trentino Seedbank (http://www.mtsn.tn.it/seedbank/specie.html). Germplasm collection and duplicated *ex situ* storage are thus urgent for both taxa. The absence of population genetic structure or the low variance among sampling localities in both endemics indicates that sampling of as many allelic variants as possible in a cost and time effective manner can be achieved through seed collection from several individuals in a few sampling localities. Nevertheless, since AMOVA also detected a significant even if low percentage of variance among sampling localities, seeds collection from a few individuals in all remaining patches is suggested in any case. Patches with higher genetic variability should be further prioritized as they maximize genetic diversity sampling. This is the case of BRB individuals located on Mt Cimo (population "G"), which showed the highest percentage of polymorphic loci and expected heterozygosity. As this population represents a sink of biodiversity, it also deserves priority in terms of *in-situ* conservation. Moreover, seed bank persistence in the soil was widely demonstrated for conventionally bred and transgenic seeds of the congener *Brassica napus*, which are capable of entering in a "secondary dormancy" after being buried in the soil, and survive there for several years before germination (Munier, Brittan & Lanini, 2012; Gruber, Pekrun & Claupein, 2004). It is thus plausible to think that a similar seed longevity could characterize *B. repanda*. Honnay et al. (2008) in their meta-analysis on 42 published habitat fragmentation studies, found that the presence of a persistent seed bank in the ground lowered the rate of allele lost from fragmented plant populations, thus increasing effective population size and slowing down the influence of genetic drift. Specific tests to assess seed viability after long term burial and ageing should be also carried out for BRB and BRG. If seed longevity in the soil was demonstrated, even more stress is to be put on *in-situ* conservation of existing populations and their natural seed banks above the ground as durable sources of biodiversity, with the most genetically diverse populations

being, again, the first of the list.

e) Taxonomic recombination

Here below is the proposed taxonomic recombination based on the results of the current study. In order of appearance:

- New taxonomic classification proposed by the authors;
- Bas = basionym, i.e. the original, validly published name of a taxon, followed by its bibliographic reference;
- ≡ followed by the synonym, i.e. a scientific name referring to a taxon that now is classified with a different name, and its bibliographic reference.

Brassica *baldensis* (Prosser & Bertolli) Prosser & Bertolli stat. nov. Bas.: Guenthera repanda subs. *baldensis* Prosser & Bertolli, Willdenowia 37 (1): 192, 2007 ≡ Brassica repanda subsp *baldensis* (Prosser & Bertolli) Prosser & Bertolli, Ann. Mus. Civ. Rovereto 22: 295, 2007. Type: Prosser and Bertolli 21.6.2004 (ROV holotype; B isotypes).

Brassica *glabrescens* Poldini, Giorn. Bot. Ital., 107: 181-189, 1973 ≡ Brassica repanda subsp. *glabrescens* (Poldini) Gómez-Campo, Anales Jard. Bot. Madrid 56 (2): 379, 1998 ≡ Guenthera repanda subsp. *glabrescens* (Poldini) Gómez-Campo, Anales Jard. Bot. Madrid 60 (2): 306, 2003. Type: Poldini 31.4.1972 (TSB holotype).

2.6. Conclusions

Species delimitation is an important issue in terms of conservation priorities, especially for narrow endemics under threat of extinction. The Alpine endemics BRB and BRG belong to a highly polymorphic species complex, but their disjunct distribution suggests favourable conditions for independent evolution. In this work, we applied the unified species concept of species to test whether the endemics form distinct evolutionary lineages, both from one another, and from the remainder of the complex. Compliance with the criteria of monophyly, diagnosability and genotypic clustering was examined making use primarily of AFLP data. Both endemics were indicated as monophyletic by phylogenetic analyses, and diagnostic characters were found for both taxa. Population structure analyses showed clear genetic discontinuity for each of the endemics, with little admixture among the clusters. This evidence indicates that the endemics have acquired multiple properties that satisfy each of the species criteria considered. Hence, we suggest the taxonomic recognition of *B. baldensis* and *B. glabrescens* as separate species. Comparative population genetics analyses show the lack of marked genetic structuring within either taxon and low levels of heterozygosity, with important practical implications for *in situ* and *ex situ* conservation.

Sections of work described in this Chapter come from the following publication:

Lega M, Fior S, Prosser F, Bertolli A, Li M, Varotto C . 2012. Application of the unified species concept reveals distinct lineages for disjunct endemics of the *Brassica repanda* (Brassicaceae) complex. *Biological Journal of the Linnean Society*. 106: 482–497. DOI: 10.1111/j.1095-8312.2012.01887.x

The Biological Journal of The Linnean Society, The Linnean Society of London and Blackwell Publishing Ltd (‘Wiley-Blackwell’) are gratefully acknowledged for allowing the use of parts of the already published Article and Abstract.

2.7. Appendix

Voucher details including: species name, country, location, herbarium collection number, GenBank/EMBL accession number and collectors' names (in italics) for the taxa analysed in this study. All voucher specimens were deposited in the herbarium collection of Civic Museum of Rovereto (ROV) for future reference.

BRG - Italy, S. Quirino (*locus classicus*, PN), 42513, *Prosser*; Italy, Valcellina (PN), 53730, FR865928, *Prosser*; **BRB** - Italy, Brentino-north (VR), 47473, *Prosser & Bertolli*; Italy, Brentino-south (VR), 47475, *Prosser & Bertolli*; Italy Preabocco-north (VR), 46978, *Prosser & Bertolli*; Italy, Preabocco-south (*locus classicus*, VR), 46977, FR865935, *Prosser & Bertolli*; Italy, Monte Cimo (VR), 46981, *Prosser & Bertolli*; *B. repanda* **subsp. *blancoana*** - Spain, S. de Alcaraz (AB), 57529, grown from wild collected seeds, JQ042820, *Lega*; *B. repanda* **subsp. *cadevallii*** - Spain, Portell Dells Torradells (L), 57524, grown from wild collected seeds, *Lega*; Spain, Sopeira (HU), 55844, FR865929/FR865930, *Hilpold*; *B. repanda* **subsp. *confusa*** - Spain, Sierra de Grazalema (CA), 57526, grown from wild collected seeds, FR865931, *Lega*; *B. repanda* **subsp. *gypsicola*** - Spain, Ribatejada-Arcos (CU), 57527, grown from wild collected seeds, FR865932, *Lega*; *B. repanda* **subsp. *latisiliqua*** - Spain, Sierra Nevada (GR), 57528, grown from wild collected seeds, JQ042821/JQ042822, *Lega*; *B. repanda* **subsp. *maritima*** - Spain, Montgó (A), 57525, grown from wild collected seeds, FR865933/FR865934, *Lega*; *B. repanda* **subsp. *repanda*** - France, Le Monêtier-les-Bains (05), N.A.; Italy, Val Stura di Demonte (CN), 57124, JQ433546, *Prosser & Bertolli*; Italy, Val Thures (TO), S.N., JQ433547, *Selvaggi*; *B. repanda* **s.l.** - Italy, Val Susa (TO), 56200, JQ042824/JQ042825, *Prosser & Bertolli*; Italy, 54155, Val Maira (CN), JQ042823, *Prosser & Bertolli*; *B. repanda* **subsp. *saxatilis*** - France, Pic de Mouches (*locus classicus*, 13), S.N., JQ042826/JQ042827, *Garraud & Selvaggi*; France, Valbelle (04), 56503, *Bertolli, Garraud, Prosser & Selvaggi*; *B. repanda* **cf. subsp. *saxatilis*** - France, La Breole (04), 56558, *Bertolli, Garraud, Prosser & Selvaggi*.

---

3. AQUILEGIA THALICTRIFOLIA

3.1. Introduction

a) Case study

Aquilegia (Ranunculaceae) is a widespread genus composed of approximately 70 species distributed in the Northern hemisphere (Munz, 1946). Because of its rapid process of diversification, *Aquilegia* was defined as a "species flock", in virtue of the little phylogenetic signal that accumulated in DNA sequences during the process of divergence of major lineages (Hodges & Arnold, 1994a; 1995). At the species level, current genetic variation often proved insufficient to reflect taxonomic boundaries, even for entities otherwise diagnosed by discrete characters, including morphological features, ecological niches and reproductive modes (Hodges & Arnold, 1994b; Ro, Keener, & McPheron, 1997; Bastida, Alcántara, Rey, Vargas, & Herrera, 2010; Cooper, Whittal, Hodges & Nordborg, 2010; Fior et al., 2013). A benchmark work on the evolution of North American species (Whittall, Medina-Marino, Zimmer, & Hodges, 2006; Whittall & Hodges, 2007) confidently showed that speciation of North American columbines was strongly promoted by adaptation to different pollinators, as a result of co-evolution of colour and length of flower spurs. On the contrary, recent studies on Eurasian *Aquilegia* taxa suggested that the radiation of the Old World species was driven mostly by geographic isolation combined with habitat specialization (Bastida et al., 2010), with stronger selection acting on vegetative traits than on floral ones (Castellanos, Alcántara, Rey, & Bastida, 2011). In this context, selective forces such as edaphic factors, were suggested to spur species diversification in the European complex, in relation to the shift of some populations from forests and meadows to a more saxicolous habitat. Such transition was likely favoured during interglacial periods when mountainous regions were left free from the ice cover (Bastida et al., 2010). Thus, the saxicolous endemics presently distributed in Southern European Alps (*A. viscosa*, *A. thalictrifolia*, *A. pyrenaica*, *A. einseleana*, *A. bertolonii*) are to be considered stenoendemics of recent origin.

Recently, Fior et al. (2013) produced a chloroplast phylogeny of *Aquilegia* based on a ~24Kb matrix composed by the most rapidly evolving region of the plastome, and including multiple

accessions for some European taxa. Results from this work revealed that even the highly variable portion of the plastome could not provide sufficient information to resolve relationships among European taxa, nor to group multiple accessions in monophyletic groups. These results depict a complex scenario in which genetic patterns were likely shaped by repeated events of separation and introgression, and extensive work will be required to disentangle the evolutionary history of the taxa. In this context, the study of spatial genetic structuring of European alpine endemics at finer scales is set to shed new light upon the processes that regulate the distribution of genetic variation within infra-generic units well defined both taxonomically and as geographic distribution.

Gene flow and genetic drift represent the main micro-evolutionary processes configuring the arrangement of neutral genetic diversity at different hierarchical and spatial levels (Loveless & Hamrick, 1984; Hartl, 2000). If a regional equilibrium exists in the studied populations between the loss of alleles due to drift and their substitution by migration of seeds and pollen, an Isolation By Distance (IBD) pattern will be found, so that the differentiation process will get stronger as far as populations become more geographically distant one to each other. If the homogenizing effects of gene flow prevail, like in the case of large effective population sizes, absence of physical barriers to migration, outcrossing and long dispersal of diaspores, the differentiation process will slow down even at long distances. On the other hand, small effective population sizes, potential barriers in the landscape, self-pollination or reduced seed dispersal contribute to the maintenance of genetic discontinuities over long periods of time even at small spatial scales, thus enhancing the effect of random genetic drift (Hutchinson & Templeton, 1999; Hartl, 2000). The combined analysis of population genetic structure and components of the landscape where these populations are located helps to understand how these forces acted (Manel, 2003), and to unravel the relative roles of gene flow and drift (Hutchinson & Templeton, 1999).

In this sense, the heterogeneous alpine landscape represents a very interesting environment for the study of genetic differentiation processes, thanks to the presence of a variety of physical barriers (mountains, forests, rivers, etc.) which tend to limit gene flow. Additionally, its complex topography promotes differences in substrate, temperature, sun exposure and moisture, thus creating an extraordinary variety of micro-climates for local adaptation (e.g. Alvarez et al., 2009).

Only a few attempts have been made until now to study population genetic structure of *Aquilegia* in mountainous environment. Brunet, Larson-Rabin & Stewart (2012) recently focused on the alpine and sub-alpine *A. coerulea* in the American Rocky mountains and demonstrated that gene flow was

the main driver for the partitioning of genetic variability even in presence of some important barriers such as large deserts. Some localized differentiation among and within regions was found in certain cases, possibly caused by local fires and differences in flowering phenology. However, *A. coerulea* shows some notable differences compared to the endemics that occur in the Alps, which makes these results hardly applicable to understand differentiation in the Alpine system (Ozenda, 2009). In fact, this taxon is characterized by a wide distribution area, with pairwise distances between sampled populations varying from 2 to approximately 650 kilometers, a range that largely exceeds that of an Alpine endemic. Perhaps more importantly, it features a pollination syndrome that relies primarily on hawkmoths, as opposed to the possibly generalist syndrome of European taxa.

Insight on the process of diversification in Europe comes from the recent work by Garrido, Fenu, Mattana, & Bacchetta (2012), who investigated the spatial organization of genetic diversity in all known columbines populations distributed across the mountains of Sardinia. Here, populations are characterized by small distribution ranges and population sizes, and high habitat specificity. The authors proved genetic drift to prevail over gene flow in shaping genetic structure, putatively coupled with divergent selection causing local adaptation. So far, no studies addressing population structure of taxa occurring in the Alps have been reported.

A. thalictrifolia Schott & Kotschy non Rydberg is a strictly endemic alpine taxon with a very limited distribution area around the mountain chain of Tremalzo-Tombea, Judicarian mountains, between the provinces of Trento and Brescia, in the Italian south-eastern Alps. Only 22 populations were recovered, with approximately 4685 mature ramets on 1443 m^2 (Bonomi, Castellani & Longo, 2008). The species was first described by Schott and Kotschy (1853), who identified the bi-lateral viscosity of basal leaves as primary distinctive morphological character. On the contrary, the highly variable morphology and dimensions of leaves and flowers were judged diagnostic by several Italian botanists (Pampanini, 1909; Fiori, 1923-1929; Zenari, 1927; Luzzani, 1932), who variously defined the taxon as a variety of the morphologically similar *A. einseleana* Schultz. Later, Munz (1946) in his monography on the genus *Aquilegia*, and Akeroyd (1993) in Flora Europaea, recognized again *A. thalictrifolia* as a distinct species, mainly based on the glandularity and pubescence of leaves as opposed to the sub-glabrous and sparsely glandular leaves of *A. einseleana*. From that moment on, the rank of species has always been acknowledged to *A. thalictrifolia*.

Its distribution range is characterized by a complex topography, which includes orographic

discontinuities separating aggregates of populations in different valleys. Most importantly, *A. thalictrifolia* is distinguished by a very specific ecological niche, as it inhabits calcareous bedrock characterized by water springs at the base of mountain cliffs, or dripped with water that keeps a constant level of moisture. This ecological specialization determines the fragmented distribution of the species, and it limits the number and size of existing populations. In a conservation perspective, these scattered habitats appear to be particularly subject to increasing drought in virtue of climate change, characterized by raising temperatures and decreasing precipitations (Coumou & Rahmstorf, 2012). Indeed, 3 populations of *A. thalictrifolia* that had been signalled in the past were not recovered anymore, and 8 of the living populations grow on bedrock substrate where humidity has partly disappeared, with only a few apparently fertile ramets (Bonomi, Castellani & Longo, 2008).

On the other hand, it has been suggested that past refugial areas may acquire the same role also with regards to the current global warming (Dawson, Jackson, House, Prentice, & Mace, 2011), if the microclimate produced by fine-scale topographic complexity will be decoupled from regional climate (Dobrowski, 2011). In this case, the identification and protection of these microrefugia acquires a strategic importance to reduce the negative impacts of contemporary man-induced climate change (Keppel et al., 2012; Médail & Diadema, 2009), also for other endangered species presently associated to this ecological niche, like *Physoplexis comosa* (L.) Schur (Campanulaceae) and *Saxifraga aracnoidea* Sternb. (Saxifragaceae).

A. thalictrifolia has recently been classified as "critically endangered" at a global level (Bonomi, Castellani & Longo, 2008). Earlier, it had been defined "rare" in the “IUCN Red List of Threatened Plants” (Walter & Gillet, 1998), "vulnerable" in the regional red list of Italian plants (Conti et al., 1997), but at "lower risk" of extinction by the provincial red list of Trentino (Prosser, 2001), and it was not even described in the national red list of Italian plants published by Conti et al. (1992).

b) Research aims

The present work aims to study the genetic diversity of the alpine endemic *A. thalictrifolia*, in order to gain insight on the distribution of diversity patterns in relation to the heterogeneous alpine landscape, and pinpoint the role of moulding forces such as gene flow or genetic drift. Moreover, compared analyses relying on a traditional measure of fixation (G_{ST}; Nei, 1973) and an index of population differentiation (D_{est}; Jost, 2008) were used to generate hypotheses on the evolutionary history of *A. thalictrifolia*. Finally, quantification of genetic diversity and structuring served to discuss possible conservation strategies for the endemic.

Picture 1. *A. thalictrifolia*. **A.** Growing site on calcareous bedrock, AT12. **B.** Sirphid (Diptera) pollinating on one flower. **C.** Pubescent -glandular leaves. **D.** Growing site underneath watery calcareous bedrock, population AT14. **E.** Detail of flower with nectar spurs.
Pictures by Margherita Lega.

3.2. Materials and methods

a) Sampling

A total of 295 individuals were sampled from 11 locations representing the *A. thalictrifolia* distribution area (Bonomi, Castellani & Longo, 2008; Figure 1). Fresh and young leaves from a minimum of 16 to a maximum of 35 individuals per location were sampled in small patches throughout the area, choosing individuals 1-3 meters apart within the same patch. As *A. thalictrifolia* populations tend to be distributed linearly along the rocks, field sampling of some populations often followed a mono-dimensional scheme. Approximate population length and width, patch length, distance between patches and between individuals were recorded, together with information concerning the growing substrate of each single plant (rock, gravel, soil), humidity of the substrate, and any other detail concerning the health status of the plant (presence of pathogens, herbivores predation, etc.; data not shown). Detailed information about sampling locations, populations, voucher identifications and number of individuals sampled per location is given in Table 1.

Table 1. Sampling locations, identification for the populations and the valley where they are distributed, and number of individuals sampled per location for *A. thalictrifolia*. Voucher specimens were deposited in the herbarium collection of Museo Civico di Rovereto (ROV) and Museo Tridentino di Scienze Naturali (TR) for future reference.

Location	Population no.	No. samples	Valley ID
Rio Bragone, Val d'Ampola, (TN)	AT1	20	AM
Storo, Val Lorina, Val d'Ampola, (TN)	AT2	32	AM
Molina di Ledro, Val Pubregno (TN)	AT4a	27	PU
Molina di Ledro, Val Pubregno (TN)	AT4b	27	PU
Bocca di Valle, Valvestino (BS)	AT5	35	VE
Tiarno di Sopra, Val d'Ampola, (TN)	AT10a	16	AM
Turano / Magasa, Valvestino (BS)	AT11	23	VE
Loc. Pilaster, Valvestino (BS)	AT12	30	VE
Loc. Ponte Franato, Valvestino (BS)	AT13	28	VE
Costa Monte di Mezzo,Valle S. Michele (BS)	AT14	30	MI
Messane, Valvestino (BS)	AT23	27	VE

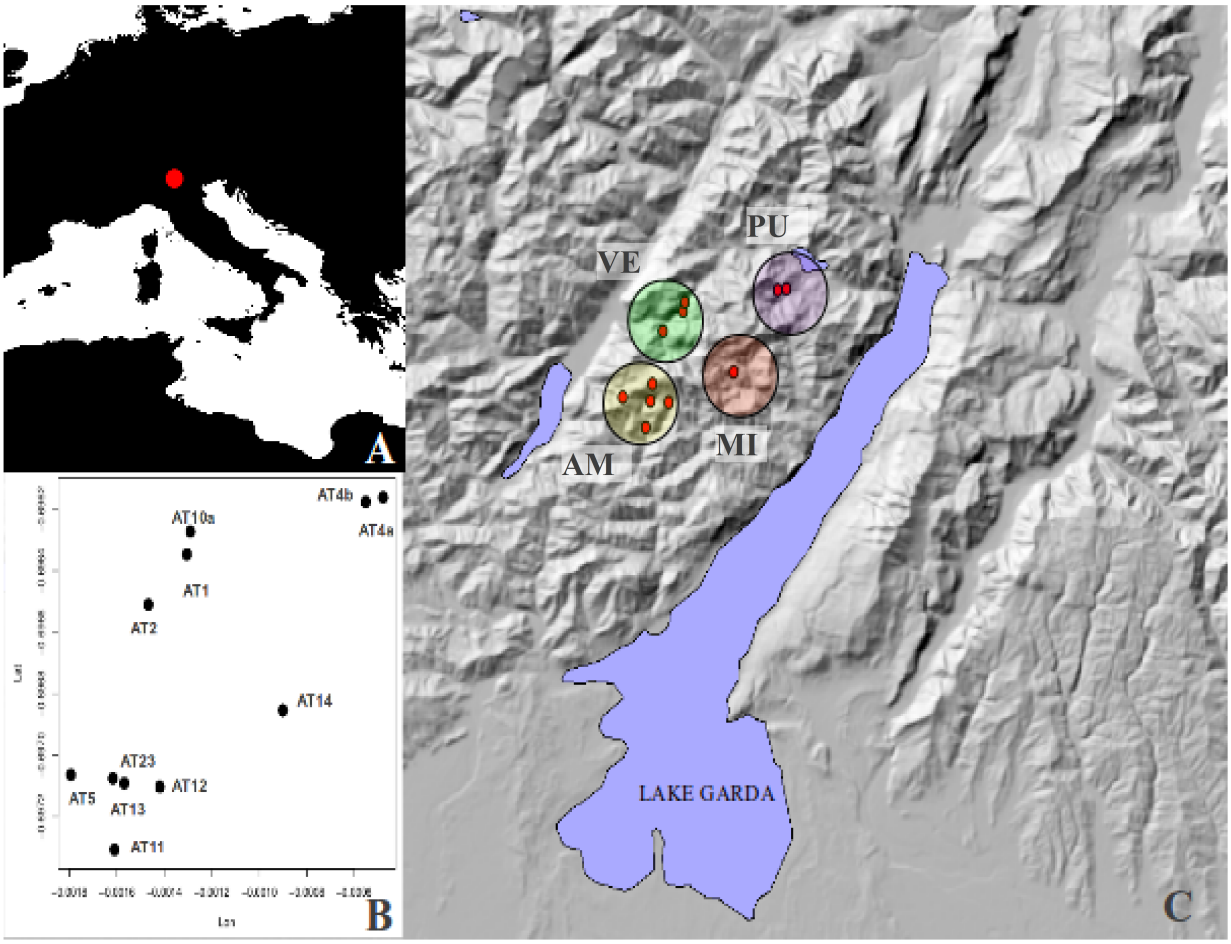


Figure 1. **A.** Overall position of *A. thalictrifolia* distribution range. **B.** Plot of the 11 sampled locations under study based on Lambert coordinates. **C.** Map of the sampled locations with coloured circles identifying the different valleys: yellow = AM, green = VE, red = MI, violet = PU.

Maps kindly provided by USGS (2004), Shuttle Radar Topography Mission, Global Land Cover Facility, University of Maryland, College Park, Maryland, USA; http://thematicmapping.org/downloads/world_borders.php, provided by Bjorn Sandvik; VMAP0 data, NGA,USA, http://geoengine.nga.mil/geospatial/SW_TOOLS/NIMAMUSE/webinter/rast_roam.html.

b) DNA extraction

Extractions of total genomic DNA were carried out using the Qiagen Dneasy 96 Well Plate Kit, as this proved more efficient compared to the CTAB extraction method (Doyle & Doyle, 1987), which produced lower DNA yield possibly because of viscous polysaccharides of glandular hairs on leaves. Extracted DNA was quantified on 1.0 % agarose gel stained with ethidium bromide and then diluted to approximately 5 ng/µl for PCR amplifications.

c) SSR genotyping

Sixteen primers for Simple Sequence Repeats (SSRs or microsatellites) developed from an F2 hybrid between *A. formosa* and *A. pubescens* by Yang, Counterman, Eckert & Hodges (2005) were tested on a sub-sample of 6 individuals collected in 6 different populations of *A. thalictrifolia*, *A. einseleana* and *A. bertolonii* (see Chapter 4 for details about *A. einseleana* and *A. bertolonii*).
Fourteen primers successfully amplified the whole sub-sample. Further sequencing of 4 individuals from 4 different populations of *A. thalictrifolia* showed polymorphism among and within populations in 9 loci (*7–27.2*, *200.2–4*, *25.6–16*, *50–7*, *25.3–33*, *11–3*, *50–21*, *10–15*, *1–40*), which were further used for SSR genotyping (Table 2). Primers were purchased from Sigma–Aldrich (St. Louis, USA) and from Applied Biosystems (Carlsbad, USA).
Each locus was amplified independently in a final volume of 10 µL containing 25 µM of each dNTP, 1X PCR buffer with $MgCl_2$ included, 0.02 µM of forward and reverse fluorescent-labelled primers, 0.5 units of Taq DNA polymerase and approximately 5-10 ng template DNA. The PCR amplification was conducted using a Mastercycler Gradient thermal cycler (Eppendorf, Hamburg, Germany) and the program followed Yang et al. (2005): initial denaturing step at 94 ° C for 2 min, followed by 30 cycles at 94 ° C for 45 s, annealing for 30 s, extension at 72 ° C for 30 s, and a final extension at 72 ° C for 10 min. Annealing temperature for each single locus again followed Yang et al. (2005), except for AY566439 and AY566440, for which it was optimized to 52°C to obtain a stronger amplification.
Ninety-six-well plates were used for amplification and 2 wells per plate were devoted to negative and migration controls in order to test for possible contaminations and peak shifts. While the first is

attributed to sample handling during laboratory procedures, the latter may arise because of the possible discrepancy between the actual and inferred allele size due to electrophoresis variability across time. Thus, the use of common reference samples among different runs can help to detect and correct this type of genotyping error (Morin, Manaster, Mesnick, & Holland, 2009). Approximately 20 amplification products from every plate were quantified on 1.5 % agarose gel stained with ethidium bromide and diluted accordingly. Post-PCR SSR duplex were produced by pooling together PCR products from pairs of loci marked with different fluorochromes, and loaded on a 3730xl DNA Analyzer sequencer using the 1200 GeneScan® LIZ as the internal size standard (Applied Biosystems, Carlsbad, USA).

Fragment lengths were scored in GENEMAPPER 4.0 software (Applied Biosystems) and manually assigned with a customized binning of 0.8-1.0 bp on each allele. Ambiguous electropherograms and peaks displaying less than 300-400 Relative Fluorescent Units (RFU) were considered as missing data, in order to decrease the occurrence of genotyping errors due to stuttering and large allele drop-out (Dewoody, Nason & Hipkins, 2006). Genotype tables produced for every duplex were concatenated among loci and finally converted into a single matrix.

Table 2. List of the sequences of primer pairs from Yang et al. (2005) used for SSRs genotyping.

Locus	Sequences (5'-3')	Genebank Accession no.
7–27.2	CCTCTCTTGTGTGTTTTCACTCTT / TAACTTTTGGTGGGGGTGCT†	AY566427
200.2–4	GCGAAATACAAATCTGGTTGAGA / CCTTCTTCCTTGACCACAAATCT†	AY566429
25.6–16	CGGAGATTTGAGAGAAGTAGATA / CGATTACGACAACTCAATTTCACA†	AY566430
50–7	CAATTCTTTGCGATTTCATCA / CCGCCAAAACAACTTTCACT†	AY566431
25.3–33	GAGGAGAAGAAAGCCATTGAAGAA / CACTTGAGCACCTTGATCCAGATA†	AY566433
11–3	GAAGAAATTGCAGAATCCATGA† / CGGCTTTGTCTTTTAGTTTCG	AY566437
50–21	TACAGGTTGGAGTGGTTGGA† / GGGTTTCTTCTTTTATCGTTGA	AY566438
10–15	GAATCGTCACTTCATTTTGCTG / TCGCCATTGTTGAAACTTGA†	AY566439
1–40	CAAAAACCCTTCTCCAAATCC / CAAACCCTACAATCAATCTCCA†	AY566440

† *indicates the fluorescently labelled primer pair for each locus*

3.3. Data analysis

a) Descriptive statistics

Linkage disequilibrium (LD) for each pair of loci in every sampling location and across locations within the species was checked using the log likelihood ratio statistic in GENEPOP 4.1.4 (Rousset, 2008). The Markov chain method was applied with 500 batches and 10^4 iterations per batch. Deviations from HWE were also verified in GENEPOP for each sampling location and across locations of the same species, using the Probability-test (Haldane, 1954; Weir, 1996; Guo and Thompson, 1992) and the score test (U test; Raymond & Russet, 1995), where the latter allowed to test both for heterozygote deficiency and heterozygote excess. The Markov chain settings were the same as above for loci with more than 5 alleles, while the complete enumeration method was applied for loci with up to 4 alleles per locus. When applicable, we controlled for multiple comparisons by calculating the P values adjusted for FDR (False Discovery Rate; Benjamini & Hochberg, 1995) with the function fdrtool for R 2.15.1 (R Core Team, 2012). The library of the function is available online at: http://strimmerlab.org/software/fdrtool/index.html. When FDR was not applicable, we used Bonferroni correction. Finally, locus-by-population frequencies of null alleles (i.e. null homozygotes caused by technical failures independent of genotype) were again estimated with GENEPOP choosing the default estimation method of maximum likelihood based on the EM algorithm (Dempster, Laird & Rubin, 1977).

Standard diversity indexes like observed heterozygosity, expected heterozygosity and total heterozygosity (H*obs,* H*exp*, H*tot*) and number of alleles averaged across loci (Nei, 1987) were calculated within every sampled location with ARLEQUIN software version 3.5.3.1 (Excoffier & Lischer, 2010). To determine whether there was a significant difference of H*exp* among loci and sampling locations, a 2-way ANOVA was performed. Moreover, as small samples usually contain less alleles than large ones (Kalinowski, 2004), an unbiased measure of allelic richness and private allelic richness corrected for differences in sample size was estimated for each sampling location with the statistical technique of rarefaction implemented in HpRare (Kalinowski, 2005).

b) Population structure

During scoring of microsatellite profiles repeat mutations were noticed to differ frequently from the stepwise pattern in practically all examined loci. In the previous work by Yang et al. (2005) dinucleotidic repeats were reported for all loci except *11-3* and *1-40* for which both dinucleotidic and trinucleotidic repeats were observed. We found this non-stepwise pattern of mutation to be more common in our case. Therefore, we applied the infinite allele model (IAM; Kimura and Crow, 1964), which states that every mutation event creates a new allele whose size is independent from the progenitor allele, as the basis for further analyses.

As a preliminary analysis of population structure, a 2-level AMOVA (among and within sampling locations; Excoffier et al., 1992) was performed in ARLEQUIN 3.5 in accordance with the mutation model assumed (IAM). The input consisted in a unique group containing the 11 sampling locations defined as different samples. Significance of group partitioning was tested against alternative random distribution of individuals among groups through 10000 random permutations. Differentiation between pairs of populations was also assessed in ARLEQUIN 3.5 calculating a global estimate across loci of F_{ST} (Weir & Cockerham, 1984), and their statistical significance was assessed with 1000 random permutations. A global estimate of Nei's G_{ST} (Nei, 1973a), the F_{ST} analogue for loci with multiple alleles, was also calculated for comparison. However, as already mentioned in Chapter 1, it was recently demonstrated that F_{ST}, G_{ST} and other related measures based on heterozygosity and entropy tend to decrease with increasing polymorphism, even if sub-populations are completely differentiated (Jost, 2008). As a consequence of this negative dependence on diversity of F_{ST} and its relatives, population genetic structure in presence of highly polymorphic microsatellites can be underestimated. Thus, we also calculated the D_{est} (Jost, 2008) estimate of genetic differentiation based on the effective number of alleles, that accounts for the bias. G_{ST} and D_{est} pairwise comparisons between populations were also estimated. All G_{ST} and D_{est} values were calculated in the R-package DEMEtics (Gerlach et al., 2010) and their statistical robustness evaluated with 1000 bootstrap replicates with a Bonferroni correction for multiple testing.

Correlations between G_{ST} and D_{est} measures were then verified through a Mantel test with 10000 permutations (Mantel, 1967) implemented in the R Package Ade4 (Dray and Dufour, 2007).

As the 2 classes of measures quantify different aspects of population structure, we performed correlation analyses to test which model best approximated the relationship between them. For this

purpose, the weighted least-squares estimates of the parameters of a linear ($y = a*x$) and an alternative nonlinear ($y = a*x / (1+b*x)$) model were determined with the function *nls* in R (R Core Team, 2012). A log-likelihood value for the 2 models was calculated using Akaike's Information Criterion (AIC). Moreover, as the 2 models were hierarchical, ANOVA was also applied as a further confirmation of the chosen model.

The null hypothesis of regional migration-drift equilibrium assuming a stepping-stone model of population structure (Kimura & Weiss, 1964) was tested following Hutchinson & Templeton method (1999). The interest of their approach lies in the possibility of evaluating the relative historical roles played by gene flow and genetic drift in shaping the structure of the populations under study. If the null hypothesis is accepted, one should expect that both pairwise genetic distances and the variance in those pairwise genetic distances (i.e. the level of dispersion of pairwise population comparisons when plotting genetic against geographic distances) increase monotonically with geographic distances (Hutchinson & Templeton, 1999). If instead the alternative hypothesis is accepted, no equilibrium between gene flow and drift exists at the regional level and other scenarios are possible. For example, drift could be much more important than gene flow (no increase of genetic distances with geographic distances, high and constant scattering of pairwise comparisons over geographic distances; possible presence of unnoticed discontinuities independent from physical distance) or viceversa (no increase of genetic distances with geographic distances, little and constant scattering of pairwise points over geographic distances), with possible intermediate situations among these extremes. To verify the null hypothesis of regional migration-drift equilibrium, a Mantel test (Mantel, 1967) was thus applied between genetic distances (both linearised G_{ST} and D_{est}) and geographic distances. Moreover, the absolute values of the residuals obtained from a standard linear regression between genetic and geographic distances were again correlated with geographic distances through Mantel tests. The analyses were carried out both on a global scale and within the single clusters identified by bayesian simulations in STRUCTURE (see below). Significance of correlations was estimated with 10000 random permutations. Geographic distances were calculated as straight-line distances in kilometers between sampling locations with the R Package Sp (Pebesma & Bivand, 2005; Bivand, Pebesma & Gomez-Rubio, 2008). Mantel tests were executed with the R Package Ade4 (Dray and Dufour, 2007).

A graphical representation of the population structure described by G_{ST} and D_{est} and a visualization of any possible qualitative differences between the alternative measures was achieved through a classical multidimensional scaling (CMDS) plot with the first 2 dimensions produced with

cmdscale function in R (R Core Team, 2012).

A non-spatial Bayesian assignment method was adopted to unravel the number of genetic clusters and their level of admixture using STRUCTURE v2.3.3 (Pritchard et al., 2000). The model assumed admixture, correlated frequencies and no prior population information. The following parameter settings were applied: 20 independent replicates each for K = 1 to K = 22 (i.e. the double number of sampling locations), a burnin period of 10^5 iterations, 10^5 subsequent MCMC repetitions. Simulations were performed at the freely available Bioportal server (www.bioportal.uio.no).

The most likely number of populations (K) was estimated with the ΔK statistic of Evanno et al. (2005) using STRUCTURE HARVESTER software (Earl & vonHoldt, 2012). Multimodality in individual memberships coefficients and label switching across different runs were accounted for using the permutation procedure in CLUMPP (Jakobsson & Rosenberg 2007). The resulting matrix of Q-values was graphically displayed through DISTRUCT (Rosenberg, 2004).

Finally, a spatial analysis was carried out on the geo-referenced genetic data using the extension for R (R Core Team, 2012) of the GENELAND program (version 4.0.2, Guillot, Mortier & Estoup, 2005; Guillot & Santos, 2010). We first launched 10 exploratory runs varying the burnin period and the number of iterations to check the convergence of the chains by the end of the MCMC runs. We then performed 20 independent MCMC runs with a burnin period of $2*10^4$ iterations, 10^4 subsequent iterations, uncorrelated frequencies model, spatial coordinates with uncertainty of approximately 50 meters, 400 pixels along the X axis and 250 along the Y axis (so as to attain a similar resolution on both axis and to have every population in a different pixel), allowing K to vary from 1 to 22.

3.4. Results

a) Descriptive statistics

The complete data-set contained 8.17 % of missing data, with loci *200.2-4*, *25.3–33* and *7–27.2* displaying alone 64.5 % of the whole missing information. None of the negative controls showed evidence of contamination. The 3 migration controls confirmed that no migration shifts happened among different runs of capillary electrophoresis for each SSR duplex. A slight shift (0.5-1 bp) was occasionally detected for some controls in loci *25.6–16* and *11-3* but it was due to presence of off-scale peaks produced by too much concentrated PCR products and readily corrected.

There was evidence of LD among 4 different pairs of loci in 3 *A. thalictrifolia* populations after FDR correction (AT4a, AT10a, AT13; $q \leq 0.01$). As no significant LD was found in the remaining data-set and no LD was evident across all sampled locations on the same pairs of loci, all markers were retained in the data matrix. Three of 99 probability-tests showed significant departures from HWE proportions following Bonferroni correction: *50-21* in AT2, *7–27.2* in AT12 and 50-*7* in AT10a. Four of 99 tests for heterozygote deficit were statistically significant: *50–21* in AT2 and AT13, *50–7* in AT12, *11-3* in AT4a. Tests for heterozygote excess showed no evidence of departure from HWE expectations. Proportions of null alleles ($0.16 < p < 0.26$) were estimated to be moderate according to Howes et al. (2006) in 4 loci (*50–21*, *50–7*, *25.3–33*, *11–3*) and 3 locations (AT4a, AT13, AT14), while the remaining data-set revealed rare null allele frequencies for all loci within all locations. Nevertheless, departure from HWE in the data set affected only a few populations and the mean population frequency of null alleles per locus was rare. Hence, these factors were considered unlikely to heavily bias the analyses results (Dakin & Avise, 2004) and the entire set of loci was retained.

All 9 microsatellites were moderately polymorphic across all populations (Table 3), with 8 alleles per locus on average, ranging from a minimum of 1 (AT14, *10–15*) to a maximum of 22 alleles per locus (AT5, *7–27.2*). Overall expected heterozygosity across loci and sampling locations was high (Mean H*exp* = 0.68 ± 0.19 s.d.; see Table 3). Pairwise comparisons of H*exp* among locations were significant only for AT4a/AT23 (Pairwise T-test, p-value = 0.051 after Bonferroni correction).

Pairwise comparisons among loci were significant only for *25.6–16* / *7–27.2* (Pairwise T-test, p-value = 1.9e-05 after Bonferroni correction). H*obs* (mean H*obs* = 0.63 ± 0.19 s.d.; see Table 3) were not significantly different from H*exp* (T-test, p-value = 0.1724), as already indicated by the HWE tests (see above).

Table 3. Genetic diversity estimates for each sampling location across 9 microsatellites in *A. thalictrifolia*. Mean observed heterozygosity (H*obs*); Nei's (1978) unbiased expected heterozygosity (H*exp*); mean H*exp* and mean H*tot* across locations; mean number of alleles per locus (ALoc); allelic richness (AR) and private allelic richness (PAR) adjusted for sample size (minimum sample size used for calculations = 24 genes). Standard deviations (s.d.) are provided for H*obs,* H*exp,* Mean ALoc, mean H*exp,* mean H*tot*, mean AR and mean PAR.

Sampling location(cal	**H*obs* ± s. d.**	**H*exp* ± s. d.**	**Mean ALoc ± s. d.**	**AR**	**PAR**
AT1	0.65 ± 0.21	0.73 ± 0.15	7.78 ± 4.38	6.71	0.58
AT2	0.67 ± 0.15	0.78 ± 0.14	10.89 ± 4.65	8.04	0.94
AT4a	0.54 ± 0.31	0.57 ± 0.20	5.44 ± 5.27	4.55	1.44
AT4b	0.54 ± 0.32	0.50 ± 0.27	5.11 ± 4.37	4.35	0.68
AT5	0.72 ± 0.18	0.74 ± 0.17	10.11 ± 5.51	7.30	1.16
AT10a	0.63 ± 0.19	0.72 ± 0.19	7.44 ± 3.91	7.04	0.99
AT11	0.70 ± 0.12	0.74 ± 0.12	6.67 ± 2.12	5.88	0.49
AT12	0.61 ± 0.23	0.65 ± 0.25	8.33 ± 5.12	6.02	0.36
AT13	0.63 ± 0.23	0.72 ± 0.24	11.11 ± 5.06	8.23	0.98
AT14	0.54 ± 0.08	0.59 ± 0.10	4.22 ± 1.86	3.33	0.58
AT23	0.75 ± 0.11	0.81 ± 0.15	12.00 ± 5.07	9.24	1.34
Mean ± s. d.	0.63 ± 0.19	0.68 ± 0.19	8.10 ± 4.30	6.43 ± 1.81	0.87 ± 0.36
Mean H*tot* ± s.d.	0.84 ± 0.13				

Table 4. Partitioning of molecular variance within *A. thalictrifolia;* global index of fixation across loci based on F_{ST} (Weir & Cockerham, 1984) and G_{ST} (Nei, 1973a); global estimate of population differentiation based on D_{est} (Jost, 2008). P-values (F_{ST}) and 95% C.I. (G_{ST} and D_{est}) within brackets are based on 1000 permutations.

Component of variation	***df***	***SS***	**Variance component**	**% Variation**	F_{ST}	G_{ST}	D_{est}
Among populations	10	379.49	0.67	20.61	0.21*	0.20 * (95% C.I.= 0.1988 - 0.2044)	0.61* (95% C.I.= 0.5984 - 0.6300)
Within populations	579	1477.11	2.56	79.39			

*df degrees of freedom, SS sum of squares, * p<0.01*

Table 5. Pairwise G_{ST} (above diagonal) and Jost's D_{est} (below) values estimated for every pair of sampled locations (p-values < 0.05 over 1000 bootstrapped matrices after Bonferroni correction; Mantel test between the 2 triangular matrices: r = 0.76, p-value = 9.999e-05).

	AT1	**AT2**	**AT4a**	**AT4b**	**AT5**	**AT10a**	**AT11**	**AT12**	**AT13**	**AT14**	**AT23**
AT1	-	0.03	0.12	0.15	0.11	0.05	0.08	0.10	0.09	0.13	0.07
AT2	0.28	-	0.11	0.15	0.09	0.05	0.07	0.10	0.10	0.11	0.06
AT4a	0.58	0.57	-	0.12	0.18	0.08	0.15	0.16	0.17	0.18	0.13
AT4b	0.61	0.62	0.36	-	0.21	0.12	0.16	0.17	0.18	0.23	0.14
AT5	0.66	0.61	0.80	0.81	-	0.12	0.07	0.14	0.07	0.18	0.05
AT10a	0.31	0.37	0.44	0.53	0.75	-	0.08	0.13	0.10	0.10	0.07
AT11	0.61	0.60	0.67	0.69	0.55	0.55	-	0.10	0.05	0.14	0.03
AT12	0.61	0.59	0.65	0.64	0.69	0.62	0.59	-	0.11	0.17	0.08
AT13	0.60	0.65	0.69	0.68	0.60	0.58	0.38	0.62	-	0.15	0.03
AT14	0.63	0.62	0.60	0.69	0.76	0.48	0.67	0.62	0.68	-	0.13
At23	0.57	0.56	0.70	0.63	0.55	0.50	0.38	0.57	0.26	0.66	-

Table 6. Correlation coefficients for Mantel tests executed on the whole distribution range and within the clusters identified by bayesian STRUCTURE analyses for K = 3.

Variable	$G_{ST}/(1-G_{ST})$				$D_{est}/(1-D_{est})$			
	ALL	**VE**	**LE**	**PU-MI**	**ALL**	**VE**	**LE**	**PU-MI**
Straight-line distance (km)	0.66 *	0.23	-0.23	0.84	0.66 *	0.32	0.49	0.87
Residuals	0.75	0.43	0.75	0.99	-0.28	0.22	0.75	0.99

** p<0.01; ALL= complete data-set; VE=Valvestino; AM= Val d'Ampola; PU-MI=Val Pubregno-Valle S. Michele.*

b) Population structure

The AMOVA results (Table 4) revealed 79.4 % and 20.6 % of variation respectively within and among sampling locations. Both F_{ST} and G_{ST} values averaged among locations were moderately high and statistically significant (F_{ST} = 0.21, p-value = 0.000; G_{ST} = 0.20, p-value = 0.001, 95% C.I.= 0.1988 - 0.2044), while overall D_{est} was at least 3 times higher (D_{est} = 0.61, p-value = 0.001, 95% C.I.= 0.5984 - 0.6300).

A square matrix comparing pairwise G_{ST} and D_{est} values is shown in Table 5. G_{ST} between pairs of sampling locations ranged from 0.03 (AT1/AT2) to 0.23 (AT4b/AT14), while, D_{est} values ranged from 0.28 (AT1/AT2) to 0.81 (AT5/AT4b). Consistently, also pairwise D_{est} estimates were found to be at least 3 times higher than pairwise G_{ST}. A strong and highly significant correlation was revealed by the Mantel test between the 2 measures (r = 0.76 , p-value = 9.999e-05). The AIC selected the nonlinear equation as the best model to fit the data (lnL = -131.8796 *vs* -41.84355), as also confirmed by the ANOVA analyses (p-value < 2.2e-16). Figure 3C shows a plot of the relationship between D_{est} and G_{ST}.

Results of Mantel tests estimating migration-drift equilibrium are summarized in Table 6. A significant pattern of IBD between genetic and geographic distances was found at the global scale among all sampling locations of *A. thalictrifolia*, both for linearised G_{ST} and D_{est} (Figure 2A-B). Nevertheless, the null hypothesis was rejected because no significant positive association was found between the degree of scatter and the geographic distance for neither G_{ST} nor D_{est}. Similarly, no significant positive correlation was found within the cluster Valvestino (VE) as identified by STRUCTURE (see below), neither between genetic and geographic distances or between residuals and geographic distances (see Table 6 and scatterplots of Figure 2C-D). This suggests that the null hypothesis of gene flow - drift equilibrium is rejected also in this case. However, while in VE the number of pairwise comparisons was sufficient to show a clear pattern, the statistical power of the analyses within the 2 remaining clusters was limited by the low number of populations, so that presence of IBD could not be thoroughly verified.

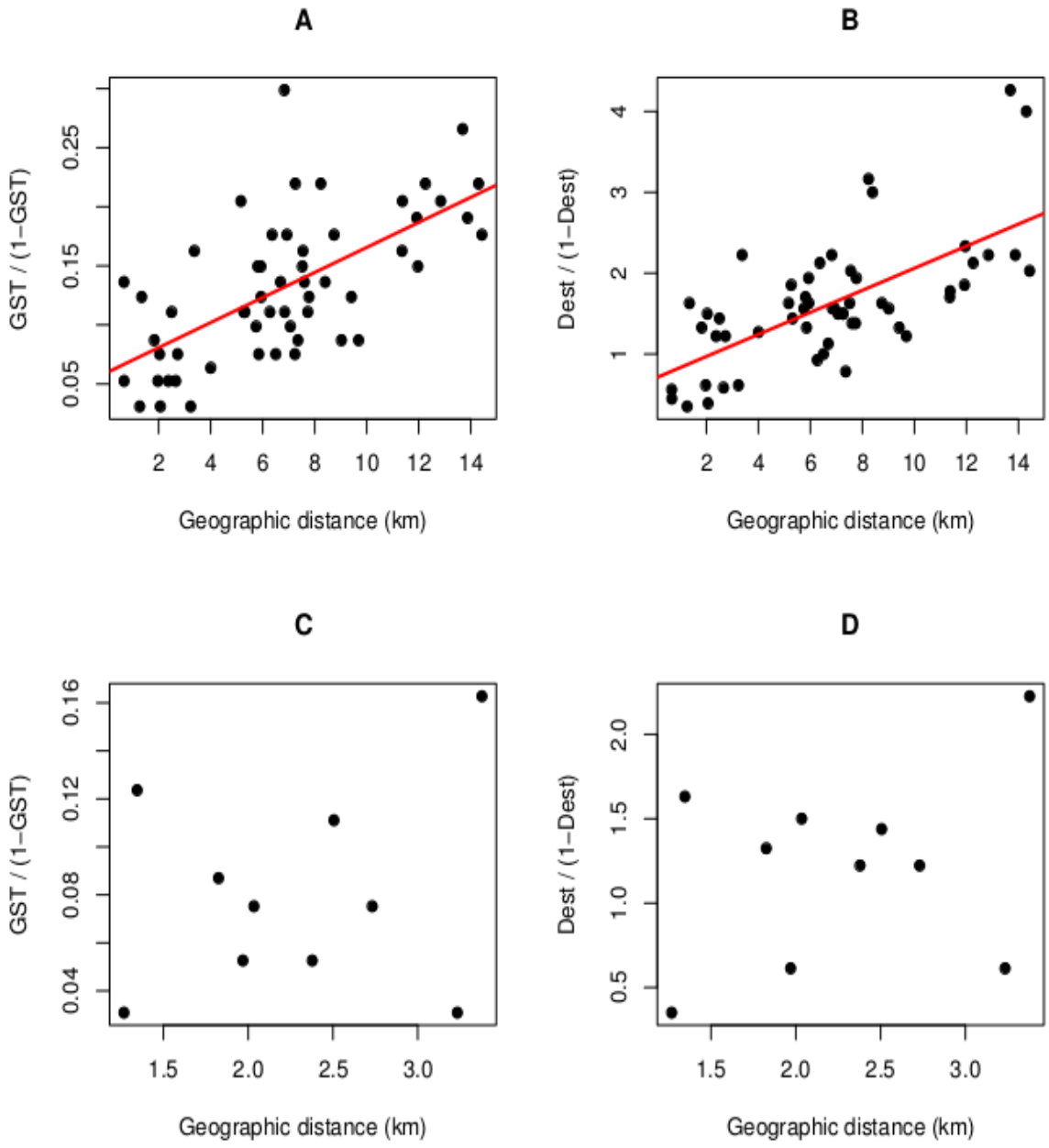


Figure 2. Relationships between genetic distances (G_{ST} and D_{est} estimates) and straight-line geographic distances among pairwise populations for **A-B.** the complete data-set; **C-D.** VE, one cluster identified by bayesian STRUCTURE analyses for K = 3.

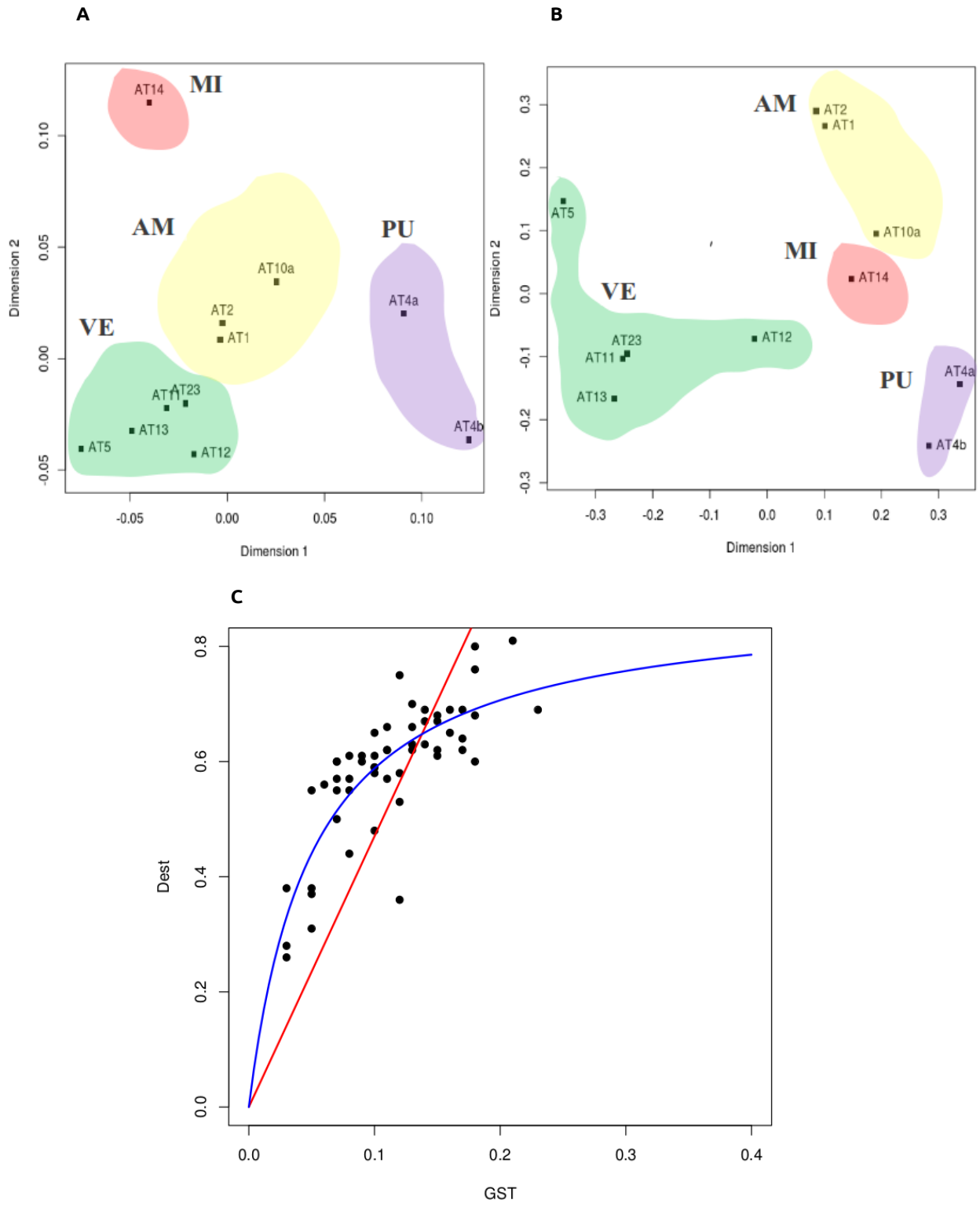


Figure 3. **A.** CMDS plot based on pairwise G_{ST} values. **B.** CMDS plot based on pairwise D_{est} values. Populations labels, identifications of the valleys and colours refer to Figure 1 and Table 1. **C.** Scatter plot of pairwise D_{est} versus pairwise G_{ST} estimates. Red line, interpolation of the linear model described by: $y = a*x$; blue line, interpolation of a nonlinear model described by: $y = a*x/(1+b*x)$.

Although the CMDS plots showed a rather diffuse structure, some groups could be identified, which differed qualitatively between G_{ST} and D_{est} data sets. Indeed, the CMDS plot built on the G_{ST} matrix (Figure 3A) grouped together the populations located in VE (AT5, AT11, AT12, AT13, AT23) and Val d'Ampola (AM; including AT1, AT2, AT10a), while placing the populations of Val Pubregno (PU; AT4a, AT4b) on the opposite part of the first axis (32.96 % of total variance). The second axis (17.67 %) separated population AT14 located in Valle S. Michele (MI). Alternatively, the plot of Jost's D_{est} (Figure 3B) clustered AM and MI, and slightly separated the populations of VE, with AT3 placed in the middle between the 2 groupings on the first axis (46.91 % of the total variance). Finally, the second axis (28.78 %) distinguished the pairs of populations from PU.

STRUCTURE analyses indicated 3 possible most likely values of K (Evanno, 2005) in the following order of importance: 19, 3 and 11. The subdivision in 19 groups has little biological meaning, as the choice for the best K should ideally aim for the smallest value of K capturing the major structure in the data (Pritchard et al., 2000). Moreover, only 5 retrieved clusters could be considered as well defined ($0.7 \leq p \leq 0.9$), while the remainder were characterized by very low posterior probabilities ($p \leq 0.1$) for almost the whole set of individuals included. On the basis of these considerations, we excluded K = 19 from further interpretations. On the other hand, K = 3 resulted in 3 clusters largely corresponding to the main valleys where the endemic is distributed (see barplot of Figure 4A). The first group included locations belonging to the geographic region VE, namely AT5, AT11, AT13 and AT23. The other population representative of this area is AT12, which resulted to be highly admixed among the 3 clusters (41 %, 38 % and 20 % belonging to AM, VE and PU+MI respectively). Within the first cluster, 59.3 % of all individuals were assigned to this group with $p \geq 0.9$ by the Q-matrix produced by CLUMPP. The remaining 41.7 % belonged mainly to population AT5, which was slightly admixed with the second cluster (mean p = 0.75, Figure 4A). The latter included populations located in AM (AT1, AT2 and AT10a) with assignment of 91.2 % of individuals with $p \geq 0.9$. Populations of the third cluster were located in PU (AT4a and AT4b) and MI (AT14). Within this cluster, 63.1 % of all individuals were assigned with $p \geq 0.9$, corresponding to representatives from AT4a and AT4b. The remainder 36.9 % presented moderate admixture with AM (AT14, mean p = 0.78). With regards to K = 11, the clustering results distinguished every sampling location (Figure 4B). The exceptions were represented by AT13 and AT23, which could not be unambiguously assigned to one group, and by AT4a and AT4b, which were identified as a single population.

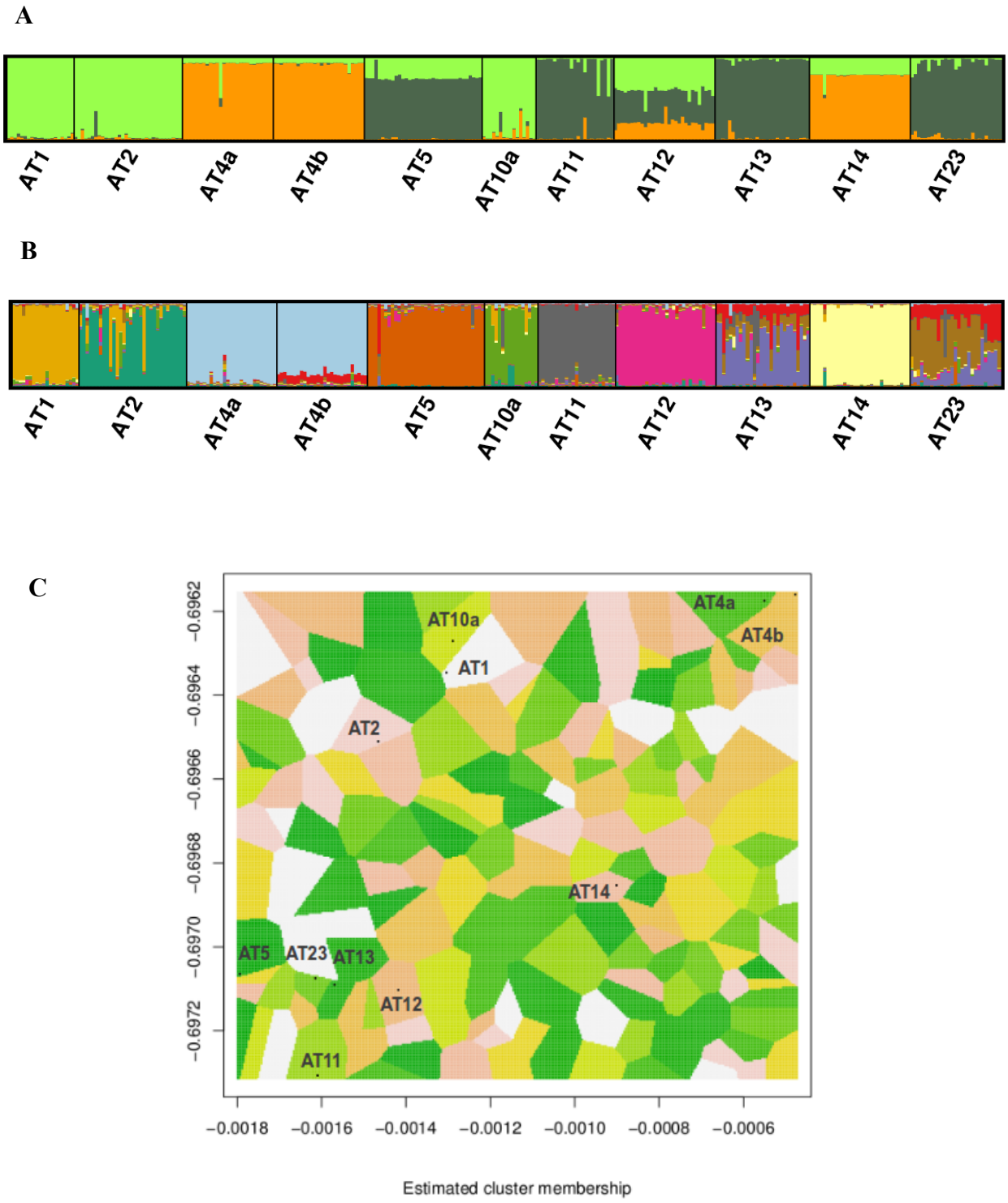


Figure 4. A-B. Bar plot showing the assignment of individuals to clusters and their levels of admixture assessed with STRUCTURE. The length of colored bars represents the fractional assignment of individuals for a K = 3 (**A**) or a K = 11 (**B**) population model. **C.** Map of estimated population membership produced by GENELAND for K = 11 clusters, the maximum a posteriori estimate of K. Populations labels refer to Table 1.

GENELAND MCMC analyses assuming an F-model distinguished 11 genetic pools in 11 independent runs which corresponded exactly to the 11 sampling locations. The remaining 9 runs gave K = 12 as the best K, but one cluster resulted a ghost population. This was disregarded for further interpretations following author's recommendations (Guillot, 2008). Therefore, all 20 runs showed an optimal structure with 11 populations. The mean posterior probability (PP) of simulated parameters along MCMC simulation for each of the 20 independent runs was calculated and runs were sorted by decreasing average posterior density following Corander et al. (2003). In the best run, all populations appeared clearly distinguished, with a membership coefficient of PP = 1. On the other hand, MCMC simulations run under the D-model showed both K=7 (7 runs) and K=8 (13 runs) as the most likely number of clusters. Between these 2 alternative values, the run with 8 clusters had the highest average PP. However, population memberships of individuals were split through the 8 groups with no possibility to unambiguously assign an individual to one cluster ($p \leq 0.9$). Hence, it was concluded that F-model best fitted the real data and only results assuming correlated allele frequencies were retained (Figure 4C).

3.5. Discussion

a) Population structure and implications for conservation

Analyses performed in the present study confirmed the presence of a strong population structure within *A. thalictrifolia*, a narrow endemic with fragmented distribution and specialized ecological niche in the heterogeneous Alpine landscape. Our results yield new evidence on the strong effects exerted by discrete geographic barriers in shaping the genetic diversity within this Aquilegia alpine taxon, and provide insight on the driving forces acting on a restricted genetic pool within the intricate background of the European scenario (Fior et al., 2013).

The genetic variability of *A. thalictrifolia* seems to be hierarchically organized in 2 main levels. First, STRUCTURE analyses identified 3 major clusters, corresponding to the valleys of the species distribution range: VE, AM and the area including PU + MI (see Figure 1). Mantel tests suggest that the limited connectivity among the 4 four valleys is one of the principal drivers for its genetic structuring. In fact, despite the finding of a significant association between genetic and geographic distances, the residuals of their linear regression were scattered with respect to geographic straight line distances, thus confirming the predominant role of genetic drift over migration (Hutchinson & Templeton, 1999). Along with orographic discontinuities, the habitat in which *A. thalictrifolia* occurs is a plausible factor that plays a role in limiting homogenization of allele frequencies within the species. Forests of Norway spruce and beech, that typically surround the rock ledges hosting *A. thalictrifolia*, are known to interfere with long-range pollinators movements. In fact, pollen dispersal in European *Aquilegia* relies uniquely on insects, mostly hymenoptpers, lepidopters and dipters which are more effective in dispersing pollen in open land than in closed forests (Kamm, Gugerli, Rotach, Edwards, & Holderegger, 2010; Kreyer, Oed, Walther-Hellwig, & Frankl, 2004). Particularly, the foraging area of a species of *Bombus*, one of the most common pollinators of *Aquilegia* in Europe (Martinell et al., 2011; Medrano, 2006; Lavergne, Debussche, & Thompson, 2005; Macior, 1966), was proved to be greatly affected by forest patches (Diaz-Forero et al., 2011), and lepidopters are more impeded by lower light intensity in forests with respect to meadows (Ross, Matter & Roland, 2005). Furthermore, similarly to other *Aquilegia* species (Martinell et al., 2011;

Castellanos et al., 2011; Yang & Hodges, 2010; Brunet & Sweet, 2006; Herlihy & Eckert, 2002), *A. thalictrifolia* is not expected to be constrained to outcrossing and a proportion of self-fertilization is expected to reinforce population structuring. In this context, SSR-based paternity determination would be a powerful tool to estimate to what extent gene flow via pollen contributes to population connectivity.

On the other hand, STRUCTURE analyses suggest that a certain degree of gene flow is maintained especially between geographically close populations. In a recent study on endemic *Aquilegia* species from Sardinia, Garrido et al. (2012) mentioned the possibility of seed dispersal on longer distances by herbivore endozoochory (Manzano et al., 2005; Manzano and Malo, 2006). Consistently, Martinell et al. (2011) found traces of goat predation of fruits in the endemic Iberian *A. paui*. Following these reports, the substantial predation of the columbine follicles recorded in combination with the presence of ungulate faeces in several populations of *A. thalictrifolia* (Lega M., pers. obs.) could be interpreted as a sign of endozoochory also for this species. The high glandulosity that characterizes *A. thalictrifolia* could also allow epizoochory (Sorensen, 1986).

The second level of genetic structuring is represented by the single sampling localities, which were all distinguished by GENELAND and almost completely defined by STRUCTURE analyses. Similarly, both the pairwise G_{ST} and D_{est} estimates depicted a clear and substantial differentiation between sampling locations, and a significant level of among-population variance was retrieved by the AMOVA analysis. For instance, Mantel tests performed on populations belonging to VE cluster showed a pattern that well approximated the case III of the Hutchinson & Templeton model (1999). This describes the scenario where genetic drift is historically more influential than gene flow. It is noteworthy that this pattern was detected even for such short geographic distances (from 1.3 to 3.4 km). The remaining STRUCTURE clusters included too few populations to provide any reliable pattern as inferred from pairwise comparisons.

The strong population genetic structure detected in *A. thalictrifolia* suggests that in situ preservation actions should in principle target every population, as the disappearance of even only one of these implies the loss of a unique allele pool. In the paragraph below some indications will be given for conservation of target populations deserving particular attention based on their genetic diversity. With regards to *ex situ* conservation, the partition of genetic variance identified by the AMOVA recognizes importance both to within and among population differences, thus indicating that numerous seeds should be collected for germplasm banks in all the studied populations.

b) Population diversity and implications for conservation

Overall values of observed and expected heterozigosity in *A. thalictrifolia* populations were rather high (mean *Hexp* ± s.d. = 0.68 ± 0.19), both absolutely and if compared to the values detected in other *Aquilegia* species with similar pollination biology, like the endemics of Sardinia [mean Nei's gene diversity (1987) ± s.d. = 0.0349 ± 0.13, AFLPs, Garrido et al., 2012] or the Iberian *A. paui* (mean *Hexp* ± s.d. = 0.006 ± 0.01, allozymes, Martinell, López-Pujol, Bosch & Blanché, 2010). However, this result is at odds with several aspects of the biology of the species, including the small census and number of recorded populations and the possible contraction of population sizes due to regression of suitable habitat (Lega M., pers. obs.; T. Abeli, pers. com.; Bonomi, Castellani & Longo, 2008), the low levels of gene flow and the self-compatibility in *Aquilegia*. These elements are expected to strengthen the effect of demographic, environmental and genetic stochasticity, which increase inbreeding depression and extinction risk (Frankham et al., 2002).

One possible reason for this discrepancy is that observed populations constitute only a fraction of bigger demes, which were not exhaustively sampled, and several individuals might occur on unreachable rock ledges well above the main body of the population.

Second, more populations are likely to exist besides those currently recognized and sampled in this study, as indicated by past herbarium specimens collected at different locations (C. Castellani, unpublished master's thesis; herbarium codes: FI, K, HBBS, PAD, TR). Therefore, the estimate of the heterozigosity achieved in the present study could be modified when the contribution of some "hidden" genetic diversity coming from additional populations or groups of individuals connected with the genotyped populations is included in the analyses.

Third, populations may have undergone a very recent demographic decline that is not yet evident in heterozigosity levels. Indeed, allelic diversity decreases faster than heterozigosity after a reduction in population size when loci evolve under the IAM model (Maruyama & Fuerst, 1985). Assessment of the existence of a temporary heterozigosity excess relative to allele diversity in *A. thalictrifolia* would be needed to check for recent population bottlenecks, based on simulation of the coalescent process conditioned on the observed number of alleles for every microsatellite locus (e.g. Piry, Luikart and Cornuet, 1999). Hypothetical loss of heterozigosity due to population bottlenecks is

also likely buffered by the the existence of a persistent soil seed bank, which was in fact shown in some *Aquilegia* endemics of Sardinia (Mattana et al., 2012). Additionally, *A. thalictrifolia* is a long-lived plant, with a generation time of approximately 12 years (Bonomi, Castellani & Longo, 2008) which may reduce the loss of alleles caused by genetic drift (Young, Boyle, & Brown, 1996; Honnay & Jacquemyn, 2007) consequential to population bottlenecks.

Finally, introgression with *A. einseleana* should not be excluded as a further source of genetic diversity for *A. thalictrifolia*. These mountain species have both fragmented distribution centred in the Eastern Alps, with population ranges that partially overlap and give rise to putatively hybrid zones (e.g. Valsugana, TN; Tramonti di Sopra, PN; see Chapter 4). Nonetheless, the unique morphological characters and the very peculiar ecological niche of *A. thalictrifolia* that earned it the current status of species probably indicate reproductive isolation of the endemic in the core distribution range. Further genetic studies including populations of *A. einseleana* and other neighbouring alpine *Aquilegia* taxa are necessary to disentangle the taxonomic relationships existing between the 2 taxa, and more generally, the evolution of the *Aquilegia* diversity in the Alpine system. With this regards, Chapter 4 is entirely dedicated to the investigation of the genetic contribution of different populations with respect to taxonomic boundaries, in order to further address the driving forces of speciation in the complex European scenario.

Although the high levels of heterozigosity observed in *A. thalictrifolia* populations suggest that the endemic is not at immediate risk of extinction (Spielman, Brook & Frankham, 2004), this statement must be interpreted with caution. Indeed, recent studies on mating-systems in *A. formosa* and *A. pubescens* demonstrated that, despite outcrossing rate was higher than 50 %, a considerable part of it was ascribable to biparental inbreeding, with total selection acting against inbred progeny in the early life stages (Yang & Hodges, 2010). This implies that an important number of descendants are routinely eliminated from populations before flowering, with a subsequent drastic reduction of the effective population size and thus higher sensitivity to genetic, environmental and demographic stochasticity (see above).

Therefore, some important information to prevent the extinction process of the strict endemic *A. thalictrifolia* can be obtained at the population level, since genetic structure analyses recognized the status of demes to each sampled location. For example, populations AT4a, AT4b and AT14 showed the lowest levels of H*exp,* mean number of ALoc and AR. These populations belonged to a unique genetic cluster identified by STRUCTURE, PU+MI, which seems to be the most genetically

impoverished and thus endangered complex of valleys. Moreover, all 3 populations are located along the borders of forest roads, whose construction caused a fragmentation of the original wood *continuum*. As a consequence, the increased exposure to sun radiations possibly altered the primary ecological niche of *A. thalictrifolia*, usually characterized by shady and wet substrate underneath calcareous rocks continuously exposed to water leakage. Indeed, a past botanical report of a probably extinct population was made in Val del Singol, not far from MI, and numerous vouchers collected in MI at the beginning of XXth century (Castellani, unpublished master's thesis), proved that other populations than AT14 were living in that area and probably disappeared only recently. Interestingly, population AT4a showed the highest level of private allelic richness of the whole data-set, possibly because of its small population dimensions, which made it more vulnerable to genetic drift and random fixation of alleles. The very high private allelic richness of AT4a is particularly striking considering that this population has a straight line distance of only 700 meters with the neighbouring AT4b in the same valley, without any explicit physical barrier between the populations. Therefore, particular attention should be deserved to the in-situ conservation of the only 3 known populations of the cluster PU+MI. On the other side, populations AT2 and AT23 were not only the most heterozygote but also those with the highest number of ALoc, the highest AR and the second highest levels of PAR after AT4a. In this light, populations AT2 and AT23 represent important sources of genetic diversity for the whole endemism and should deserve conservation priority (Petit, El Mousadik & Pons, 1998). Coherently, AT23 is also the biggest and most largely distributed population (Lega, pers. field obs.) with respect to the other sampled populations, occupying as usual the bases of big blocks of rocks scattered into a very large area of forest. On the contrary, although formerly signalled by botanists as very big and productive, nowadays AT2 appears quite small, with a few flowering individuals, growing along the borders of a secondary road and suffering from the regular grass mowing that prevent flowering, seed ripening and dispersal. The discrepancy between the high genetic diversity and the small census of AT2 could be explained by an incomplete sampling of a bigger deme. Finally, the uniqueness and importance of populations AT2 and AT23 is also represented by their being representatives of the AM and VE clusters respectively.

c) Methodological considerations and evolutionary insights: G_{ST} VS D_{est}

In our study, we compared a traditional index of fixation, G_{ST}, with a more recent index of allele differentiation, Jost's D_{est}, in order to quantify and interpret the different aspects of population structure provided by each estimate.

Remarkably higher values of Jost's D_{est} than G_{ST} were found, both across populations and in single pairwise comparisons. Considering the rather high levels of within population expected heterozigosity (mean *Hexp* across populations = 0.68 ± 0.19 s.d.), these results are in line with Jost (2008) and similar to other recent studies applying both statistics (Ensing et al., 2011; Kuss et al., 2011; Vik et al., 2010). Although a high and significant positive correlation between the 2 measures was found, it is important to remember that this is not a necessity, and different scenarios may occur, for example when *Hexp* and within population variation are very low, because the 2 indexes do not estimate the same quantity (Thomas Städler, pers. com.). In fact, important differences could be observed between D_{est} than G_{ST}.

First, different properties of the 2 measures were visualized by a nonlinear relationship between them in Figure 3C. In this plot, rather similar to the one produced by Raeymaekers et al. (2012), Jost's D_{est} increased rapidly for low values of G_{ST}, but slowed down for higher levels of G_{ST}. In other words: when gene flow and within population diversity are high, this index is more sensitive to population structure than G_{ST}, because it is less influenced by gene flow and drift. For the same reason, when the effect of drift is higher, migration between groups decreases and within population diversity is low, G_{ST} performs better in detecting population structure (Raeymaekers et al., 2012; Jost, 2009).

Second, the 2 estimates identified different population structures, which were well depicted by the CMDS visualizations. Indeed, if the matrix of D_{est} values clustered together populations from AM and MI and separated VE on the first dimension of the CMDS plot, the parallel plot on G_{ST} values tended to divide AM and MI, while finding some continuity between VE and AM. Populations AT4a and AT4b from AM, instead, were isolated from the rest in both CMDS representations. An interesting example of an integrated approach that took in consideration and interpreted the results from both measures is represented by the recent work of Raeymaekers et al. (2012). The authors demonstrated both empirically and through simulations that on short time scales G_{ST} provided a

more correct picture of genetic structure as recently shaped by migration and drift, while D_{est}, for being much slower than G_{ST} in reaching equilibrium, still kept trace of the previous pleistocenic colonization history. If the interpretation of Raeymaekers et al. (2012) holds in our case study, this would imply that the picture provided by Jost's D_{est} reflects, on short time scales (from late Pleistocene on), a more ancient structuring of populations which is not masked by contemporary demographic events. Thus, the clustering of populations from AM to MI could identify a past continuum between the 2 valleys that is not existing anymore. A series of populations could have been present in the past between these 2 neighbouring valleys, which assured a stepping-stone mode of population connectivity. Indeed, some old herbarium specimens of *A. thalictrifolia* were found to be collected in 2 locations at the entry of Val Gaton, not far from Mount Tremalzo, approximately in-between AM and MI. These populations have not been found again at present (Castellani, unpublished master's thesis), indicating a possible limitation to contemporary gene flow between the 2 valleys. This was well depicted by G_{ST}, which is more sensitive than D_{est} to recent migration and drift and indicated 2 different clusters for the 2 valleys. Another main difference between the 2 indexes was that G_{ST} aggregated populations of VE and AM, while D_{est} separated VE from the remainder of the valleys. In this case, G_{ST} apparently revealed a contemporary or recent exchange of migrants between the 2 valleys. Indeed, some other herbarium specimens collected between 1872 and 1929 witness the past presence of *A. thalictrifolia* on Mount Tombea and Mount Caplone (Castellani, unpublished master's thesis). Importantly, these mountains occupy an intermediate position between AM and VE. Some of the populations identified by past botanists were never found again, while 2 of them were visited recently, and another population was recently recovered near Bondone, not far from Monte Tombea (Cristina Castellani, unpublished master's thesis). For economy of time, these particular locations were not included in the present study, but they are very likely to represent, or they have been until recently, part of an active corridor of gene flow from VE to AM and viceversa. Migration between VE and AM may also be indicated by detection of some admixture between the respective STRUCTURE clusters. Nevertheless, as STRUCTURE algorithm does not model the demographic history, it is not possible from its output to discriminate real admixture from shared polymorphism. A rigorous verification of the presence of admixture between the neighbouring valleys and thus, indirectly, also of the soundness of the pattern provided by G_{ST} and D_{est} may come from the application of Approximate Bayesian Computation (ABC; Beaumont, Zhang, & Balding, 2002), using an higher number of SSRs. Indeed, ABC proved to be a reliable tool to distinguish between these 2 alternative scenarios (Sousa, Beaumont, Fernandes, Coelho, &

Chikhi, 2012).

On the other hand, the presence of a distinct cluster of VE populations opposed to the rest showed by D_{est} may provide some hints about the more ancient post-glacial history of this endemic. During the last Quaternary glaciation cycle (Würmian), several areas north-west of lake Garda served as glacial refugia on calcareous bedrock (Schönswetter et al., 2005) from the numerous ice tongues that branched off from the Alpine icecap and occupied southern peripheral valleys. In particular, the mountain chain Tremalzo-Tombea was placed exactly in the middle of 2 main valley ice-shields: Valle del Sarca and its branch Valle di Ledro at east and Valle del Chiese at west (Avanzini, 1999). The highest elevated areas and the lateral slopes of this chain remained uncovered by ice and therefore represented important nunataks and peripheral glacial refugia (*sensu* Holderegger & Thiel-Egenter, 2009) for the survival of numerous species of plants. As a confirmation, most of the endemics and local endemics of the whole Eastern Alps were recovered in the southern calcareous Alps between Lake Como and Lake Garda (Tribsch, 2004), including the area of Tremalzo-Tombea. This mountain chain itself encompasses the highest number of endemics of Trentino Alto-Adige (Prosser, 1999; but see also Mount Baldo, in Chapter 2). Therefore, one can suppose a post-glacial differentiation of the endemic *A. thalictrifolia*, possibly from some generic pre-pleistocenic *Aquilegia* populations that took refuge in the ice-free areas of the Tremalzo-Tombea mountains. Reasonably, the plant could have remained isolated for thousands of years accumulating enough mutations to be considered a different species from a putative pre-pleistocenic *Aquilegia*, similar to what Bastida et al. (2010) hypothesized for the endemic European columbines grouping the species with narrow distributions in the Pyrenees, Betic Mountains, Alps, Apennines and Balkans. During the last post-glacial period, the newly formed taxon could have gradually recolonized again the surrounding valleys where the ice retreated. Coming back to the results of pairwise Jost's D_{est} estimates, the distinctiveness of VE populations may suggest the existence of a refugium on Mount Tombea in the neighbourhood of VE. The highest levels of allelic richness and private allelic richness of VE populations with respect to the rest, both averaged on the single sampling locations and aggregated on the base of STRUCTURE clusters (K=3; data not shown) suggest that the populations which survived in the hypothetical refugia on Mount Tombea where numerous and of bigger dimensions than the remainder of the populations, and thus less influenced by genetic drift. As a confirmation, VE is also the valley where most of the known populations of *A. thalictrifolia* are growing, and the one characterized by the highest concentration of past botanical reports for the endemic (Castellani, unpublished master's thesis).

The 2 remaining clusters identified by D_{est} may suggest that other different and independent glacial refugia for *Aquilegia* were possibly present, likely somewhere on the upper parts of Tremalzo mountains. Indeed, this mountain chain is placed in-between the complex of the three valleys where *A. thalictrifolia* is distributed. Moreover, AM, MI and PU clustered together on the first axis in the D_{est} results, although PU was separated on the second dimension. The data in our hands do not allow to say whether the contemporary distribution range of *A. thalictrifolia* was derived by vicariance, by 1 recolonization event or by different ones, and from which refugia this possible recolonization started. In the hypothesis of vicariance, or of multiple recolonization events, the independent nuclei could have developed similar morphological traits in virtue of a parallel ecological adaptation to the same very specific ecological niche. None of these hypotheses can be rejected a priori, and only a coalescent approach should ideally shed light on the ancestral origin of the endemic and indicate if and which one of the 2 proposed hypotheses is more likely (e.g. Afzal-Rafii & Dodd, 2007). Anyway, the comparison between Mantel tests on the whole distribution range based on D_{est} and G_{ST} matrixes showed absence of IBD in both cases, thus possibly indicating that the mountains chains were important physical barriers in the past and still they play a similar role today.

Comparison of the overall relative values of G_{ST} and D_{est} potentially sheds light on other interesting aspects of the evolutionary history that shaped population structure. Indeed, when both estimates are large (> 0.15) like in the current case, Leng & Zhang (2011) simulations indicated three different possibilities: a strong population differentiation with very weak gene flow ($Nm<1$); a very small population size ($N \leq 100$); a low mutation rate ($\leq 10^{-4}$) coupled with a very ancient population differentiation. The first hypothesis seems to be the most likely one based on Bayesian assignment tests, AMOVA analyses and the geographical distribution of populations within the different valleys. However, the second possibility cannot be rejected, as field observations suggest that some populations were characterized by less than 100 flowering individuals (Thomas Abeli, personal communication). With regards to the third scenario, a low mutation rate is not likely because of the dinucleotidic nature of the SSRs and high number of alleles detected on average per locus. Coalescent analyses will be however necessary to reach a firmer conclusion about what scenario best applies to *A. thalictrifolia* populations.

To summarize, population structure was quantified by comparing 2 different measures of fixation (G_{ST}) and allelic differentiation (D_{est}), encouraged by a growing body of empirical and theoretical studies (Hedrick, 2005; Meirmans, 2006; Jost, 2008, 2009; Ryman & Leimar, 2009; Gerlach et al., 2010; Meirmans & Hedrick, 2011; Leng and Zhang, 2011; Raeymaekers et al., 2012). Here, the 2

types of measures were considered to answer different research questions, and an integrative approach for inferring the evolutionary processes that influence population structure was applied following Raeymaekers et al. (2012) and Leng & Zhang (2011), against Whitlock's reasoning (2011) that Jost's D_{est} has no evolutionary meaning. The contrast analysis of G_{ST} and D_{est} on our data-set allowed to gain deeper insights into possible short term processes of *A. thalictrifolia* populations at different temporal scales. Indeed, Jost's D_{est} and Nei's G_{ST} recognized distinct population structures which possibly reflected the more ancient colonization history and the recent demographic processes, respectively (Raeymaekers et al., 2012). More generally, the present study confirms the usefulness of combining together measures of genetic differentiation and fixation to unravel population structure, as already advocated by the meta-analyses of Heller & Siegismund (2009) and Meirmans & Hedrick (2011).

3.6. Conclusions

This work represents the first study of population genetic structure and diversity applied to an alpine endemic taxon of the model genus *Aquilegia* in Europe, which represents a textbook example of a very rapid and recent radiation through the Northern Hemisphere. The majority of the European taxa occur in the Alpine system, but the processes regulating genetic differentiation in this heterogeneous landscape are still widely unexplored. We used microsatellites to study population genetic structure and diversity of *Aquilegia thalictrifolia* Schott & Kotschy, an alpine endemic distinguished by a high ecological specificity and fragmented distribution. The relative influences of gene flow and neutral genetic drift were analysed to understand how these evolutionary processes shaped genetic structure. Moreover, an analytical comparison of the results obtained by applying a measure of fixation (G_{ST}) and population differentiation (D_{est}) was performed to characterize different aspects of population genetic structure. Despite its endemic distribution, *A. thalictrifolia* shows a considerable spatial genetic structuring of populations as indicated by bayesian assignment analyses, and in fact, genetic drift was proved to be historically more influential than gene flow. The retrieved pattern suggests that natural barriers like mountain ridges and the ecological niche could act as barriers to migration, thus favouring among population differentiation of the endemic. Despite the predominance of genetic drift, overall high levels of heterozigosity were found. G_{ST} and D_{est} showed different population genetic patterns, and the distinct properties of these measures provide insights into post-glacial history and more recent demographic events respectively. Implications for the conservation of the alpine stenoendemic are discussed in the light of the results obtained.

4. *AQUILEGIA* SPECIES IN THE ALPS: A BROADER PERSPECTIVE

4.1. Introduction

a) Case study

The rapid and recent radiation of European *Aquilegia* species renders the identification of interspecific differentiation a very difficult task (Fior et al., 2013; Bastida et al., 2010), even when using highly informative markers like AFLPs (Garrido et al., 2012). While *A. thalictrifolia* (see Chapter 3) is an example of taxon clearly distinguishable on an ecological and morphological basis, within the remaining European taxa of the genus a great phenotypic variety is distributed in the heterogeneous environments that characterize the Alpine System (*sensu* Ozenda, 2009). This favours continuous taxonomic revisions on the base of a few, often unstable, morphological characters, which may simply represent ecotypic variants of formerly described species.

Here, we used highly polymorphic microsatellite markers to extend the previous study on *A. thalictrifolia* (see Chapter 3) to other species of *Aquilegia* that represent very well the geographical and morphological complexity of the Alpine System. We selected the above mentioned taxa using the following criteria: i) solid taxonomic classification based on morphology and ecology, or, alternatively, ambiguous taxonomic classification due to presence of some intermediate phenotypic traits between species or to putatively diagnostic traits that recently led to taxonomic revisions; ii) high number of populations per taxon, when possible; iii) clearly defined and confined distribution range, in order to achieve an exhaustive sampling.

Moreover, the selected species possess peculiar complementary characteristics which may reveal very useful to gain more insights into the evolutionary processes that possibly shaped the differentiation process.

A. bertolonii Schott (1853, Verh. K. K. Zool.-Bot. Ges. Wien, 3 : 127) is an endemic species of the Apuan Alps, with some dozens of populations mostly distributed on calcareous bedrock on steep slopes and stabilised screes, in sun-exposed and dry mountain tops. Until recently, *A. bertolonii*

included also disjunct populations on the Maritime and Julian Alps, which now have been classified as different species based on subtle morphological differences: *A. iulia* Nardi (Nardi, 2011) and *A. reuterii* Boss (Gismondi, 1950).

A. einseleana Schultz (1848, Arch. Fl. Fr. Allem. 135) is a widespread species characterized by populations separated by large geographic distances, ranging from Rhaetic to Austrian and Slovenian Alps. It grows on saxicolous and calcareous substrate in gorges and scree canals with different levels of soil rockiness, depth, moisture and sun exposure.

Finally, *A. thalictrifolia*, as already mentioned in Chapter 3, is an endemic of Judicarian Alps, with approximately 22 known populations located in perennially humid environments, on limestone substrate at the base of rocky cliffs, usually shaded by surrounding forest. The attribution to this species of 3 additional populations disjuncted from the core distribution area is dubious as their individuals carry a *continuum* of hybrid morphological traits between *A. thalictrifolia* and *A. einseleana*. Therefore, 2 of these populations were sampled and included in the present study. Moreover, an intermediate form between these 2 taxa in the Judicarian Alps has been recently described as *A. vestinae* (Pfenninger & Moser, 2002).

While *A. einseleana* and *A. thalictrifolia* distribution ranges partly overlap (see map of Figure 1), *A. bertolonii* populations extend over a very circumscribed and isolated area, so that absence of current gene flow between the endemic of Apuan Alps and the remainder of the above mentioned taxa can be expected. Moreover, Fior et al. (2013), in their phylogeny of the genus *Aquilegia* demonstrated a possible 2 step origin of the European columbines, with a first colonization event from Eurasia bringing to Europe the ancestors of the modern *A. einseleana* and *A. bertolonii* and a second, more recent, migration wave likely giving origin to the remainder of the European columbine species, included *A. thalictrifolia*.

b) Research aims

Here, an extended and comprehensive sampling compared to that of Fior et al. (2013) was carried out on *A. einseleana* and *A. bertolonii*, and marginally on *A. julia*, *A. reuterii* and *A. vestinae*, complementing the previous extensive sampling effectuated on *A. thalictrifolia*. A total of 32 locations were sampled and 826 individuals were collected that exhaustively represented the distribution range of the 6 taxa, and population genetic structures and diversities were assessed to explore the possible evolutionary processes governing the diversification of these European columbines. More in detail, we asked: (I) if these alpine *Aquilegia* taxa defined on a morpho-ecological basis are also genetically distinguishable and if admixture zones or ancestral shared polymorphism exist among them; (II) if different patterns of genetic diversity and population differentiation apply to species of columbines in the Alps; (III) if some elements exist suggesting that populations with morphologically intermediate phenotypes between *A. thalictrifolia* and *A. einseleana* may be hybrids between these 2 species; (IV) which are the conservation implications for the alpine *Aquilegia* endemics according to their genetic diversity and structure.

4.2. Materials and methods

a) Sampling

A total of 531 individuals in addition to the 295 *A. thalictrifolia* individuals previously collected (Table 1, Chapter 3) were sampled from 21 different locations representing mainly *A. bertolonii* and *A. einseleana* distribution ranges (see map of Figure 1). At the time of sampling, *A. bertolonii* s.l. was considered as a long-established single taxonomic unit (Pignatti, 1982) characterized by morphological variation among the Apuan, Maritime and Julian Alps; however, recent floristic accounts suggest the recognition of different species for the 3 regions, namely *A. bertolonii* s. str. (1), *A. reuterii* (2) and *A. iulia* (3) (Pignatti, in press). Following this classification, our sampling includes 176 individuals for *A. bertolonii* s. str. from 6 locations in the core distribution area of Apuan Alps, 41 individuals from 2 populations of A. *iulia* and 20 individuals from 1 population of *A. reuterii*. *A. einseleana* is represented in this study by 201 individuals from 9 populations across its range, covering the easternmost and westernmost extremes, respectively in Slovenian and Swiss Alps. With regards to *A. thalictrifolia*, 71 new individuals representing potential hybrids with *A. einseleana* were collected in 2 different locations of Eastern Pre-Alps. Moreover, 22 individuals from one additional population of the partly sympatric *A. vestinae* Pfenninger & Moser were also sampled. Fresh and young leaves from minimum 16 to maximum 43 individuals per location were sampled for each taxon in small patches throughout the area, choosing individuals spaced 1-3 meters within the same patch in a 2-dimensional scheme. Approximate population length and width, patch length, distance between patches and between individuals were recorded, together with information concerning the growing substrate of each single plant (rock, gravel, soil), humidity of the substrate, and any other detail concerning the health status of the plant (presence of pathogens, herbivores predation, etc.; data not shown). Detailed information about sampling locations, populations and number of individuals sampled per location is given in Table 1.

A

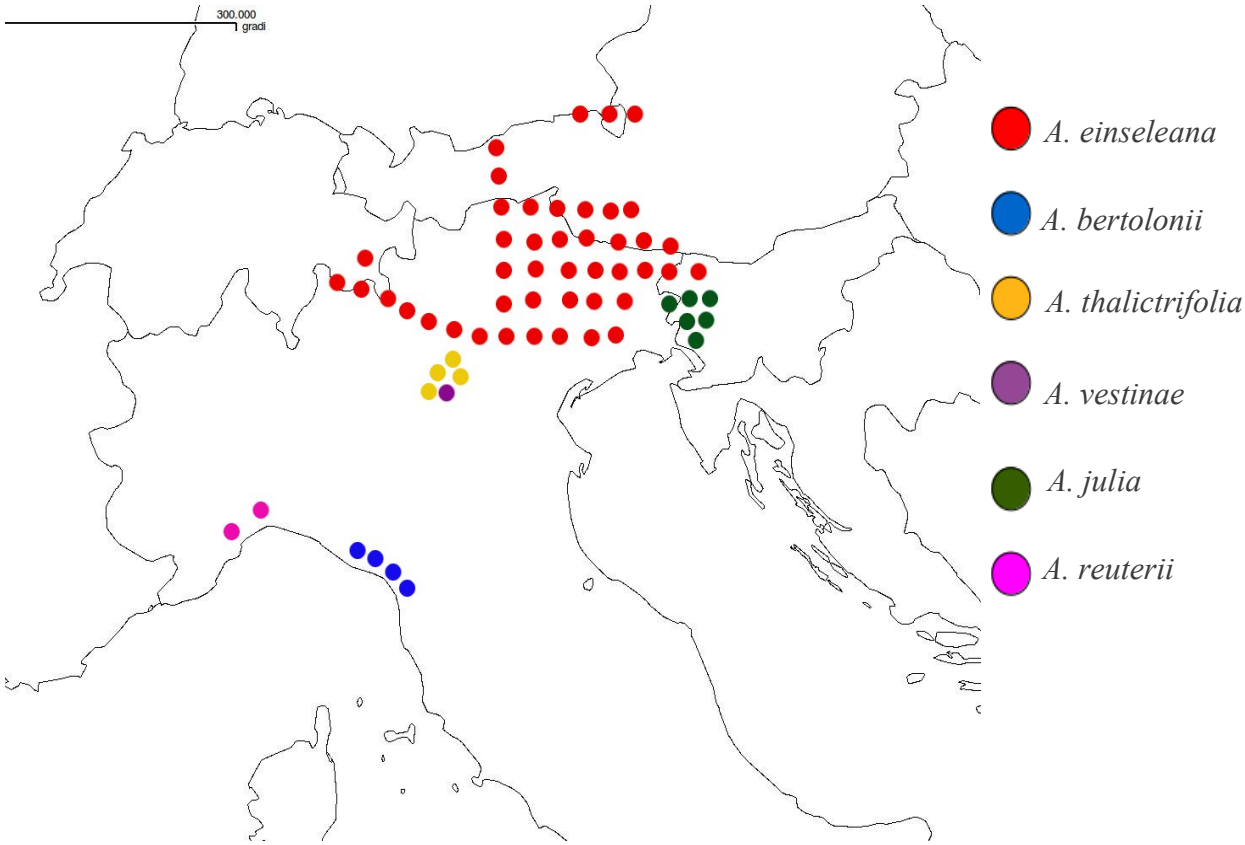
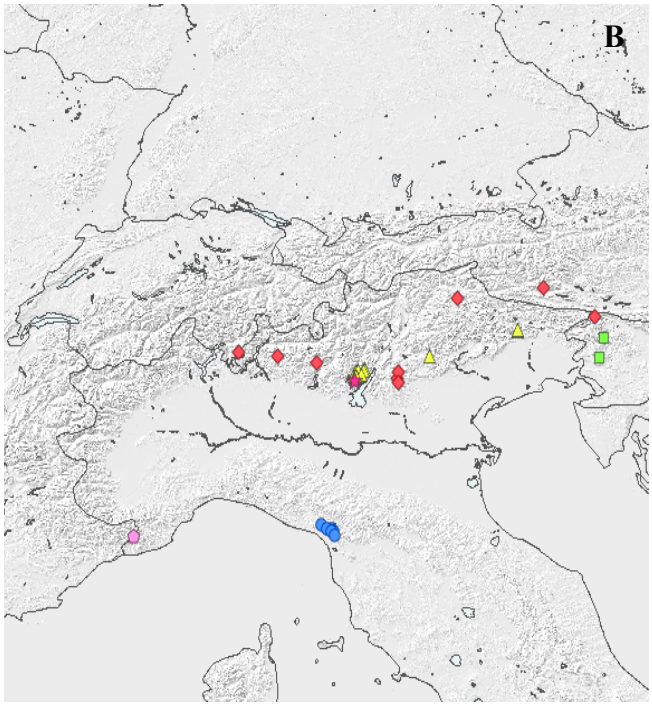


Figure 1. **A.** Main distribution ranges of the *Aquilegia* species under study, from Atlas Flora Europaea (http://www.luomus.fi/english/botany/afe/index.htm).
Coloured circles identify the different species: red = *A. einseleana*; blue = *A. bertolonii*; yellow = *A. thalictrifolia*; violet = *A. vestinae*; rose = *A. reuterii*; green = *A. julia*. **B.** Main distribution areas of *Aquilegia* sampling locations under study. Colours refer to **A.**



Maps kindly provided by USGS (2004), Shuttle Radar Topography Mission, Global Land Cover Facility, University of Maryland, College Park, Maryland, USA;
http://thematicmapping.org/downloads/world_borders.php provided by Bjorn Sandvik;
VMAP0 data, NGA,USA,
http://geoengine.nga.mil/geospatial/SW_TOOLS/NIMAMUSE/webinter/rast_roam.html.

Table 1. Country, sampling location, population identification and number of individuals sampled per location for *A. bertolonii*, *A. einseleana*, *A. iulia*, *A. reuterii*, *A. thalictrifolia* and *A. vestinae*. Information concerning *A. thalictrifolia* sampling locations is shown in Table 1, Chapter 3. Voucher specimens were deposited in the herbarium collection of Museo Civico di Rovereto (ROV) for future reference.

Taxon	**Country**	**Location**	**Population no.**	**No. samples**
A. bertolonii	Italy	Monte Sumbra, Alpi Apuane (MS)	AB1	20
	Italy	Pizzo delle Saette, Alpi Apuane (LU)	AB2	19
	Italy	Foce della Pianza, Alpi Apuane (MS)	AB3	40
	Italy	Passo del Vestito, Alpi Apuane (MS)	AB4	43
	Italy	Passo Croce, Alpi Apuane (LU)	AB5	28
	Italy	Monte Nota-Procinto, Alpi Apuane (LU)	AB6	26
A. julia	Italy	Zeleni Rob, Trnovski Gozd Plateau	AJ7	20
	Italy	Mt Crna Prst / Rodoviza	AJ8	21
A. einseleana	Italy	Valle delle Prigioni (TN)	AE1	18
	Italy	Bocchetta di Salmurano (SO)	AE2	23
	Italy	Monte Pasubio (TN)	AE3	20
	Italy	S. Vigilio, Val Badia (BZ)	AE4	25
	Austria	Nikolsdorf	AE6	25
	Italy	Passo della Presolana (BG)	AE7	22
	Italy	Passo di Campogrosso (VI)	AE9	24
	Switzerland	Denti della Vecchia	AE10	25
	Slovenia	Velika Pisnica, near Kranjska Gora	AE11	19
A. reuterii	Italy	Mt. Pietravecchia (IM)	AR1	20
A. thalictrifolia	Italy	Puele, Valsugana (TN)	AT6	37
	Italy	Tramonti di Sopra (PN)	ATPN	34
A. vestinae	Italy	Val Vestino (BS), *locus classicus*	AV1	22

Picture 1. **A.** *A. bertolonii*, flower, AB3. **B**. *A. bertolonii*, growing site, AB4. **C.** *A. vestinae*, sub-adult individual on watering rocks, AV1. **D.** *A. einseleana*, growing site, AE6. **E.** *A. einseleana*, sub-adult individual, AE7. **F.** *A. einseleana*, flower, AE9. Note the different length of nectar spurs compared to *A. bertolonii*. Pictures by Margherita Lega.

b) DNA extraction

Total genomic DNA for *A. bertolonii*, *A. julia, A. einseleana*, *A. reuterii* and *A. vestinae* was isolated from dried or frozen leaves using the CTAB extraction method of Doyle & Doyle (1987). DNA extractions for *A. thalictrifolia* hybrid populations were carried out following the same procedure already used for the other populations of the same taxon (see Chapter 3, page 59). Extracted DNA was quantified on 1.0 % agarose gel stained with ethidium bromide and then diluted to approximately 5 ng/µl for PCR amplifications.

c) SSR genotyping

Microsatellite genotyping and scoring of SSR profiles were performed applying the same procedure described for *A. thalictrifolia* in Material and Methods of Chapter 3.

4.3. Data analysis

a) Descriptive statistics

All descriptive statistics applied here refer to Chapter 3, Data analysis.

b) Population structure

The scoring of microsatellite profiles revealed the same non-stepwise pattern of mutation observed for *A. thalictrifolia* (see Chapter 3, Data analysis) to be more common. Thus, also in this case we applied the infinite allele model (IAM; Kimura and Crow, 1964).

Population structure for the whole data-set of *Aquilegia* taxa was investigated using 4 complementary approaches.

- Nei's chord distances (D_A; Nei, Tajima & Tateno, 1983) between all pairs of populations were calculated in MICROSATELLITE ANALYZER 4.0 (Dieringer and Schlötterer, 2002). This distance measure was chosen because when mutation rate is high, like in the case of SSR markers, it increases with time reaching its maximum value of 1.0 very quickly if compared with the standard genetic distance of Nei (1972, 1978), D_S, and the Weir & Cockerham (1984) analogue of F_{ST}, θ (Kalinowski, 2002). Moreover, computer simulations on a SSR data set (Takezaki & Nei, 1996) demonstrated that D_A, together with Cavalli-Sforza and Edward's chord distance (D_C; 1967), were the best distance measures with respect to D_R (Rogers, 1972), D_m (Nei, 1973b), D_S, F_{ST} (Latter, 1972) and X^2 (Sanghvi, 1953) in obtaining the true tree topology when varying heterozigosity levels, number of generations, mutation rate, number of loci and presence or absence of populations bottlenecks. Particularly, D_A yielded a better linear relationship with time than D_C. Additionally, an empirical study on a data set with 783 SSRs from different human populations confirmed that D_A was the most accurate measure for phylogenetic reconstruction, mainly due to its very low coefficients of variation with regards to the other distance indexes (Takezaki

& Nei, 2008). The resulting distance matrix was subsequently used to build a Neighbor Joining (Saitou & Nei, 1987) phylogram in PAUP 4.0b10 (Swofford 2002). Support for branches was based on 1000 replicate distance measures constructed in MICROSATELLITE ANALYZER by permuting genotypes among populations.

- A multivariate PCO-MC analysis among populations (Reeves & Richards, 2009) was executed based on the same genetic distances as above. The clustering procedure followed the authors' recommendations (available at: http://lamar.colostate.edu/~reevesp/PCOMC/PCOMC.html; see also Chapter 2).
- A non-spatial Bayesian assignment method was adopted on the whole data-set to display the number of genetic clusters and the level of admixture among them, using STRUCTURE v2.3.3 (Pritchard et al., 2000), with the same parameter settings as in Chapter 3, for K = 1 to K = 40. As this software unravels only the upper hierarchical structuring of populations, to achieve a finer resolution the same analysis was repeated within each of the inferred clusters using the same set of parameters as above, and a range of K values proportional to the number of sampling locations present in each group.
- Hierarchical analyses of molecular variance (Excoffier et al., 1992) were performed in accordance with the mutation model assumed (IAM). First of all, an AMOVA was carried out without assuming any taxonomic prior. In other words, the input consisted in a unique group containing the 32 sampling locations defined as different samples. Two other AMOVAs were then executed on the same matrix considering groups corresponding to recognized taxa: *A. thalictrifolia, A. einseleana* and *A. bertolonii* first, and then all 6 taxa including *A. vestinae*, *A. reuterii* and *A. julia.* Moreover, 2 separate AMOVAs were computed on *A. einseleana* and *A. bertolonii* populations. Finally, one additional AMOVAs included the 2 major clusters identified by STRUCTURE analysis (see Results). Significance of group partitioning was tested against alternative random distribution of individuals among groups through 10000 random permutations. Differentiation between pairs of populations was also assessed in ARLEQUIN calculating pairwise F_{ST} values. Overall D_{est} values averaged across loci and

populations were also calculated for *A. einseleana* and *A. bertolonii*, for which a considerable number of sampling locations was available, and compared with the overall D_{est} calculated for *A. thalictrifolia.* Refer to Chapter 3 for details of D_{est} calculation.

4.4. Results

a) Descriptive statistics

Although the 11 *A. thalictrifolia* populations located in the core distribution area of the endemics were already analysed and discussed singularly in Chapter 3, for sake of completeness in the compared analysis, the following results will refer to the general data-set including also those populations.

The data-set contained 12.4 % of missing data averaged across loci. However, 3 loci (*25.6–16*, *11-3*, *7–27.2*) were responsible alone for 64.95 % of the missing information, while the 6 remaining loci only accounted for 36.15 % of it. Indeed, the average percentage of missing data carried by these 6 loci taken alone was much lower (6.2%). Therefore, in order to verify if the missing information of loci *25.6–16*, *11-3* and *7–27.2* significantly affected the main results, bayesian, phylogenetic and multivariate analyses were also carried out for the partial data-set excluding loci *25.6–16*, *11-3* and *7–27.2* (see below). Negative and migration controls behaved similarly as for *A. thalictrifolia* samples (Chapter 3).

LD resulted significant among 2 pairs of loci in 1 *A. julia* sampling location (AJ7) and one pair of loci in 2 *A. einseleana* locations (AE6 and AE9). LD was also detected for 3 pairs of loci in the *A. reuterii* population under study (and 2 pairs of loci in 2 *A. thalictrifolia* populations, see Chapter 3). No significant LD was found in the remaining data-set. Ninety-eight LD tests out of 1152 produced no information because of excess of missing data (Locus *7–27.2* was the most concerned, with 68 tests presenting missing information). As the *25.3–33* / *11–3* pair of loci was the most affected by LD (4 locations out of 32), the respective nucleotide sequences were blasted with the available sequences of the *A. coerulea* Goldsmith genome (www.phytozome.net), in order to verify whether the 2 loci mapped close on the same chromosome. Since the *A. coerulea* genome is still unassembled, the blast search is performed on the separate scaffolds. The query showed both *25.3–33* and *11–3* microsatellites to be located on scaffold 2 (e-value= 2.5e-34 and 1.2e-39 respectively), thus proving the physical proximity of the 2 loci. However, as there was no consistent evidence of LD across all sampled locations, this pair of loci was retained in the microsatellite matrix.

Six of 288 probability-tests showed significant departures from HWE proportions for locus *50-21* in

locations AE7, Ae10, AT2 and AT6; locus *7–27.2* in location AT12; locus *11-3* in location AT4a. Moreover, heterozygote deficit compared with HWE expectations was significantly detected for 13 out of 288 tests. Locus *50–21* showed homozygote excess in locations AB2, AE7, Ae10, AT2, AT6 and AT13; locus *50–7* in locations AB2, AB5, AT12, ATPN; locus *11-3*, *25.3–33* and *10–15* in locations AT4a, AB4 and AE10 respectively. On the other hand, significant heterozygote excess was found in AB7 at locus *50–21* and in AT4a at locus *25.6–16.* Null alleles at moderate frequencies ($0.16 < p < 0.26$) according to Howes et al. (2006) were distinguished in 5 loci (*50–21*, *50–7*, *25.3–33*, *11–3*, *10–15*) and 10 locations. The remaining data set revealed rare null allele frequencies ($p \leq 0.15$) for all loci within all 32 locations. To summarize, locus *50–21* exhibited the highest number of homozygote and heterozygote excess and null alleles frequencies, followed by *50–7* and *11–3*. Nevertheless, as departures from HWE in these loci were concerning only a few populations and the mean population frequencies of null alleles were rare ($p < 0.08$), these loci were considered unlikely to heavily bias the analyses results and thus the entire set of loci was examined.

All 9 microsatellites were moderately polymorphic across all taxa (Table 3), with 8 alleles per locus on average, ranging from a minimum of 1 (AJ8, locus *50–21*; AT14, locus *10–15*) to a maximum of 22 alleles per locus (AB4 and AT5, locus *7–27.2*). The associated standard deviation was rather high (mean s.d. across loci and populations = 2.73), mostly due to locus *7–27.2* which displayed a much higher number of alleles than the remainder of the loci (14 versus 7 mean number of alleles per locus respectively). Expected heterozygosities under HWE were rather high across all loci and sampling locations (Mean H*exp* = 0.71 ± 0.16 s.d.; see Table Figure 3), with similar values for observed heterozygosities. The taxon showing the highest level of H*exp* was *A. bertolonii* (mean H*exp* = 0.80 ± 0.12 s.d.), while the most homozygote was *A. thalictrifolia* (mean H*exp* = 0.70 ± 0.17 s.d., potential hybrids included). Mean H*exp* were found to be slightly lower than mean H*tot* in all taxa, with *A. einseleana* and *A. thalictrifolia* showing the highest deviation. AR and PAR calculated on a sample size of 13 genes ranged from 2.88 to 7.78 and from 0.10 to 0.73 respectively. *A. bertolonii* presented the highest AR averaged across sampling locations (mean = 6.07 ± 0.57 s.d), followed by *A. thalictrifolia* (mean = 5.17 ± 1.34 s.d) and *A. einseleana* (mean = 4.93 ± 0.77 s.d).

On the other hand, PAR averaged across locations was very similar between *A. einseleana* (mean = 0.32 ± 0.16 s.d) and *A. thalictrifolia* (mean = 0.31 ± 0.18 s.d), while *A. bertolonii* presented the lowest level of PAR (mean = 0.23 ± 0.09 s.d).

Table 3. Genetic diversity estimates for each sampling location of *Aquilegia* taxa. Mean observed heterozygosity (H*obs*), Nei's (1978) unbiased expected heterozygosity (H*exp*), mean H*exp* and mean H*tot* across locations; mean number of alleles per locus (ALoc); allelic richness (AR) and private allelic richness (PAR) adjusted for sample size (minimum sample size for calculations = 13 genes). Standard deviations (s.d.) are given in parentheses.

Sampling location	**H*obs* ± s. d.**	**H*exp* ± s. d.**	**Mean ALoc ± s. d.**	**AR**	**PAR**
AB1	0.72 ± 0.13	0.76 ± 0.11	8.78 ± 5.09	5,55	0,35
AB2	0.69± 0.16	0.81 ± 0.14	9.22 ± 3.31	6,37	0,15
AB3	0.72± 0.17	0.75 ± 0.11	9.22 ± 4.49	5,27	0,11
AB4	0.75 ± 0.17	0.81 ± 0.16	12.56 ± 6.19	6,59	0,24
AB5	0.73 ± 0.19	0.84 ± 0.12	9.89 ± 3.82	6,65	0,21
AB6	0.74 ± 0.16	0.80 ± 0.09	9.00 ± 3.32	5,99	0,32
	Mean H*exp* ± s. d.	0.80 ± 0.12	**Mean H*tot* ± s. d.**	0.84 ± 0.12	
AJ7	0.58 ± 0.25	0.62 ± 0.24	5.00 ± 2.55	3,99	0,26
AJ8	0.45 ± 0.17	0.54 ± 0.20	3.75 ± 1.75	3,02	0,11
AE1	0.77 ± 0.15	0.78 ± 0.12	7.78 ± 3.83	5,52	0,41
AE2	0.59 ± 0.21	0.63 ± 0.14	5.00 ± 1.94	3,72	0,17
AE3	0.82 ± 0.14	0.80 ± 0.08	8.22 ± 3.07	5,79	0,24
AE4	0.73 ± 0.12	0.75 ± 0.12	8.11 ± 4.49	5,22	0,58
AE6	0.62 ± 0.25	0.68 ± 0.23	6.44 ± 2.51	4,68	0,32
AE7	0.65 ± 0.21	0.71 ± 0.07	5.78 ± 1.48	4,41	0,44
AE9	0.75 ± 0.20	0.76 ± 0.15	9.11 ± 3.41	5,63	0,45
AE10	0.67 ± 0.20	0.78 ± 0.09	8.56 ± 3.36	5,48	0,15
AE11	0.62 ± 0.28	0.61 ± 0.22	4.78 ± 2.05	3,94	0,11
	Mean H*exp* ± s. d.	0.72 ± 0.14	**Mean H*tot* ± s. d.**	0.90 ± 0.05	
AR1	0.56 ± 0.16	0.61 ± 0.15	4.22 ± 1.64	3,68	0,33
AT1	0.65 ± 0.21	0.73 ± 0.15	7.78 ± 4.38	5,21	0,21
AT2	0.67 ± 0.15	0.78 ± 0.14	10.89 ± 4.65	6,08	0,1
AT4a	0.54 ± 0.31	0.57 ± 0.20	5.44 ± 5.27	3,68	0,73
AT4b	0.54 ± 0.32	0.50 ± 0.27	5.11 ± 4.37	3,52	0,35
AT5	0.72 ± 0.16	0.74 ± 0.17	10.11 ± 5.51	5,56	0,54
AT6	0.69 ± 0.17	0.74 ± 0.09	7.56 ± 3.36	4,91	0,15
AT10a	0.63 ± 0.19	0.72 ± 0.19	7.44 ± 3.91	5,4	0,18
AT11	0.70 ± 0.12	0.74 ± 0.12	6.67 ± 2.12	4,88	0,23
AT12	0.61 ± 0.23	0.65 ± 0.25	8.33 ± 5.12	4,65	0,23
AT13	0.63 ± 0.23	0.72 ± 0.24	11.11 ± 5.06	5,97	0,39
AT14	0.54 ± 0.08	0.59 ± 0.10	4.63 ± 1.51	2,88	0,18
AT23	0.75 ± 0.11	0.81 ± 0.15	12.00 ± 5.07	6,69	0,37
ATPN	0.80 ± 0.09	0.88 ± 0.09	15.67 ± 7.28	7,78	0,36
	Mean H*exp* ± s. d.	0.70 ± 0.17	**Mean H*tot* ± s. d.**	0.87 ± 0.11	
AV1	0.63 ± 0.18	0.74 ± 0.19	9.22 ± 3.8	5,82	0,43

b) Population structure

The NJ phylogram and the PCO-MC plot based on 6 SSR loci (with the exclusion of loci *25.6–16*, *11-3* and *7–27.2*) showed similar outputs with respect to the complete data set (data not shown), but higher bootstrap support and higher stability values respectively. Thus, only results from this partial data set were retained and shown in Figure 3. The phylogram indicated three supported 'clans' (Wilkinson et al. 2007) that corresponded to currently recognized taxonomic entities, i.e. *A. bertolonii* s.str. (indicated in Figure 3 as b1-b6; BP = 76%), *A. julia* (b7-b8; BP = 93 %) and *A. thalictrifolia* (t1, t2, t4a, t4b, t5, t10a, t11, t12, t13, t14; t23; BP = 70%). With regards to *A. einseleana,* 3 clans, all poorly unsupported, were identified that reflected geographic distribution. In particular, a first clan was composed of populations of the eastern Pre-Alps in the Lessini Mountains (e1, e3, e9); a second clan included the more westerly located populations in the Orobie and Insubria Mountains (e2, e7, e10); a third clan comprehended populations of the Eastern range of the species distribution (e4, e6, e11), as well as the morphologically ambiguous populations of the Valsugana and Val Meduna (herein classified as *A. thalictrifolia:* t6 and tPN). Finally, *A. reuterii* from the Maritime Alps (r1) formed a clan with the geographically close *A. bertolonii*.

PCO-MC analysis (Figure 5B) exhibited only 2 stable clusters. One group incorporated all *A. thalictrifolia* locations, except for the putative hybrid populations (t6 and tPN), and *A. vestinae*. The other stable cluster encompassed all *A. bertolonii* locations (b1- b6). None of the 2 groups was statistically significant ($\alpha = 0.05$), but they yielded stability values ≥ 15 (23 and 40 respectively). On the other side, *A. einseleana* populations did not cluster together but were scattered across the space, with unstable and insignificant clusters reflecting geographic distribution, similar to the NJ phylogram. The first, second and third axis of the PCO-MC analysis explained respectively 17.1, 8.4 and 7.6 % of the total variance.

The most probable number of clusters identified by the STRUCTURE output on the whole data set with the method of Evanno et al. (2005) was K = 2 (see Figure 4A). The Q-individual matrix of assignment produced by CLUMPP reported 98.41% of individuals belonging to cluster I (cyan bars) with probability of assignment ≥ 0.90. On the other hand, individuals sampled in locations belonging to cluster II (orange bars) were partly admixed with the eastern cluster: only 83,48 % of individuals were assigned to cluster II with posterior probability ≥ 0.90, while the remaining 16.52

% presented a variable level of admixture, ranging from 11 to 78 %, and 2 probable immigrants were assigned to cluster I with p ≥ 0.90. *A. einseleana* presented the most admixed populations of the whole data-set (AE1, AE2, AE3, AE9, AE10).

The same analysis conducted in parallel on the partial data-set excluding those loci with most of the missing information (*25.6–16*, *11-3*, *7–27.2*) gave rather similar results (see Figure 4B). The most important difference between the 2 data-sets is represented by the lower level of admixture existing between the group *A. thalictrifolia* + *A. vestinae* (identified in Figure 4B with orange bars) and the western locations of *A. einseleana* (AE1, AE2, AE3, Ae7, A9 and AE10), with consequent more admixture of the latter with the alternative cluster containing *A. bertolonii*, the remainder of *A. einseleana* locations, *A. julia*, and *A. reuterii*. Given the similarity of results arising from the complete and partial data-sets, the missing information contained in loci *25.6–16*, *11-3* and *7–27.2* was not considered as a significant source of bias. Thus, the subsequent analyses of population structure were carried out on the complete data-set with 9 loci.

The hierarchical analysis conducted within cluster I allowed to separate *A. bertolonii* and *A. reuterii* from the remaining populations of Eastern Alps (best K = 2, see yellow and blue barplot in Figure 4D), coherently with geographical position (93.4 % of individuals assigned with p ≥ 0.90). Nevertheless, *A julia* (AJ7 and AJ8) and *A. thalictrifolia* hybrid populations (AT6, ATPN) presented a considerable number of genetic intermediates between the 2 sub-clusters, as highlighted in Figure 4D. With regards to cluster II, the more probable number of sub-clusters was also found to be K = 2 (see violet and green barplot in Figure 4C), mainly separating *A. thalictrifolia* sampling locations of VE (AT5, AT11, AT12, AT13 and AT23; 96 % of individuals belonging to violet sub-cluster with p ≥ 0.90) from 3 *A. einseleana* populations west of Garda Lake (AE1, AE3, AE9; 97 % of individuals belonging to green sub-cluster with p ≥ 0.90). On the other hand, the remaining populations of either taxon were found to be moderately to highly admixed between the sub-groups. In particular, the 3 *A. einseleana* populations from Western Alps (AE2, AE7, AE10) presented average values of membership coefficients to the "*A. einseleana*" green sub-group of 0.60 ± 0.00, 0.84 ± 0.03 and 0.62 ± 0.06 respectively. The remnant of *A. thalictrifolia* populations were found to be more admixed with *A. einseleana* populations (green colour) than with the geographically and morphologically closer *A. thalictrifolia* populations from VE (violet colour), with average levels of admixture with the first sub-group of 0.49 ± 0.02, 0.56 ± 0.14, 0.78 ± 0.06, 0.78 ± 0.04, 0.65 ± 0.16, 0.74 ± 0.01, for populations AT1, AT2, AT4a, AT4b, AT10a and AT14 respectively. Finally,

despite being morphologically more similar to *A. einseleana*, *A. vestinae* clustered mainly with *A. thalictrifolia* from VE, with an average coefficient membership to this sub-cluster of 0.78 ± 0.09.
The analyses of molecular variance performed on *A. thalictrifolia*, *A. bertolonii* and *A. einseleana* (Table 4) yielded very low amounts of variance among species (3.83 %) and much higher percentages of among populations and among individuals variation (17.37 % and 78.80 % respectively). The inclusion of *A. reuterii*, *A. julia* and *A. vestinae* resulted in analogous proportions of variance partitioning (Table 5). Similarly, the AMOVA on the clusters identified by STRUCTURE on the whole data set (9 SSR loci, Tables 6) revealed a very low amount of variance among clusters (2.53 %), a discrete amount of among locations within cluster variability (19.54 %), while within locations within cluster variance was the most explicative factor (77.93 %). Coherently with this partitioning of molecular variance, fixation index between clusters was very low and non significant (FCT = 0.026, p-value = 0.059), while FSC and FIS were higher and statistically significant (0.20, p-value = 0.000 and 0.22, p-value = 0.000 respectively). The same trend could be observed in the AMOVA performed without any *a priori* taxonomic delimitation (Table 6), where 21.43 % of the variability was among sampling locations and 78.57 % was within these groups, and in the analyses executed on the single a priori delimited taxa (Tables 7-8; see Table 4 in Chapter 3 for AMOVA of *A. thalictrifolia* without hybrids). Values of overall F_{ST} averaged across locations for *A. einseleana* and *A. bertolonii* were significant but very low, while D_{est} estimates were both statistically significant and considerably higher than F_{ST} (Tables 7-8, Table 4 in Chapter 3).

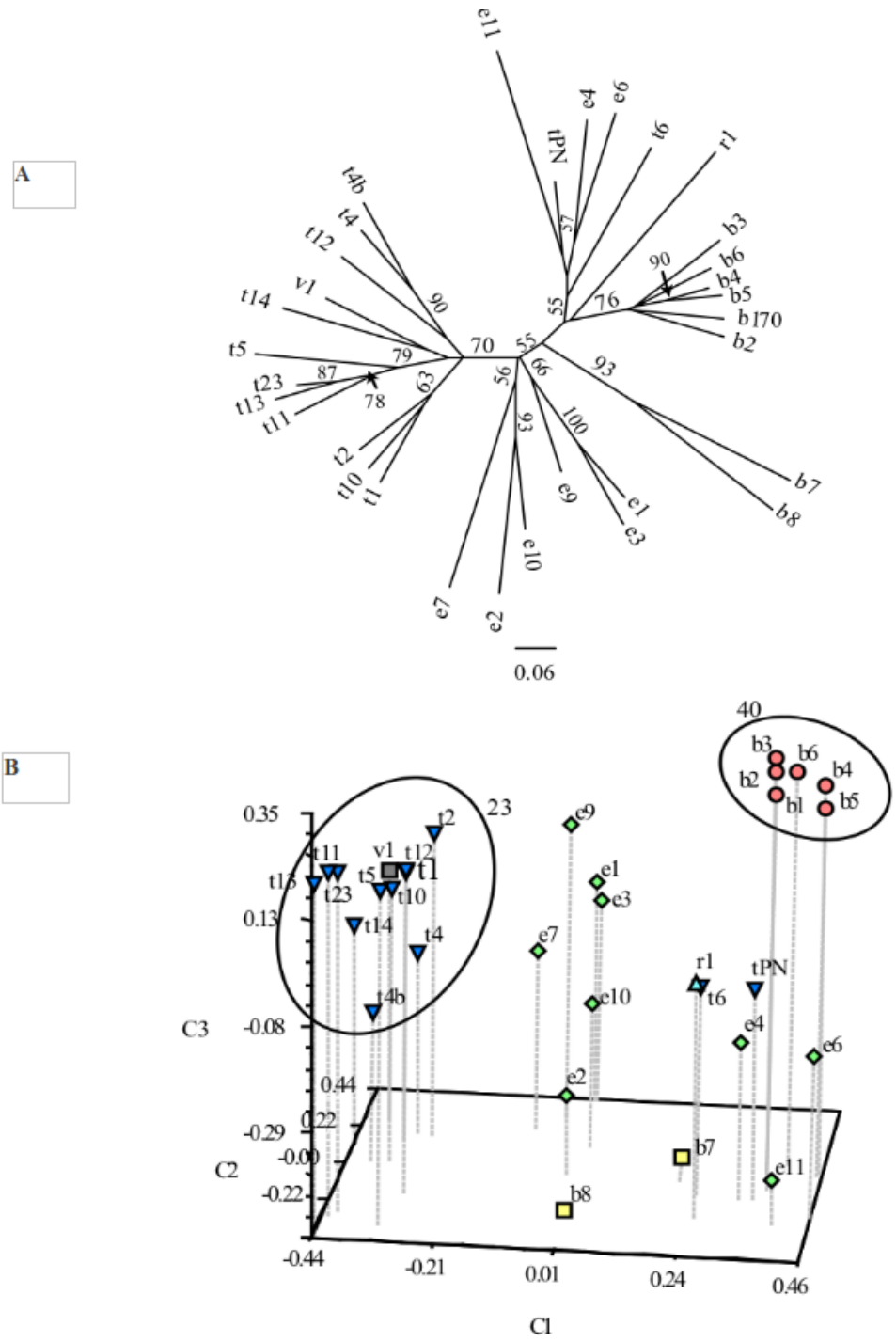


Figure 3. **A**: NJ phylogram based on Nei's chord distance among populations. Number at branches are bootstrap values from 1000 replicates. **B**: PCO-MC based on Nei's chord distance among populations. Retrieved clusters are shown, and numbers indicate stability values. No cluster was statistically significant (p-value = 0.05). b 1-b6= *A. bertolonii;* b7- b8 = *A. julia*; e = *A. einseleana*; t = *A. thalictrifolia*; r = *A. reuterii*; v = *A. vestinae*. All the analyses were performed on the partial data set of 6 SSR loci (loci *25.6–16*, *11-3* and *7–27.2* were excluded).

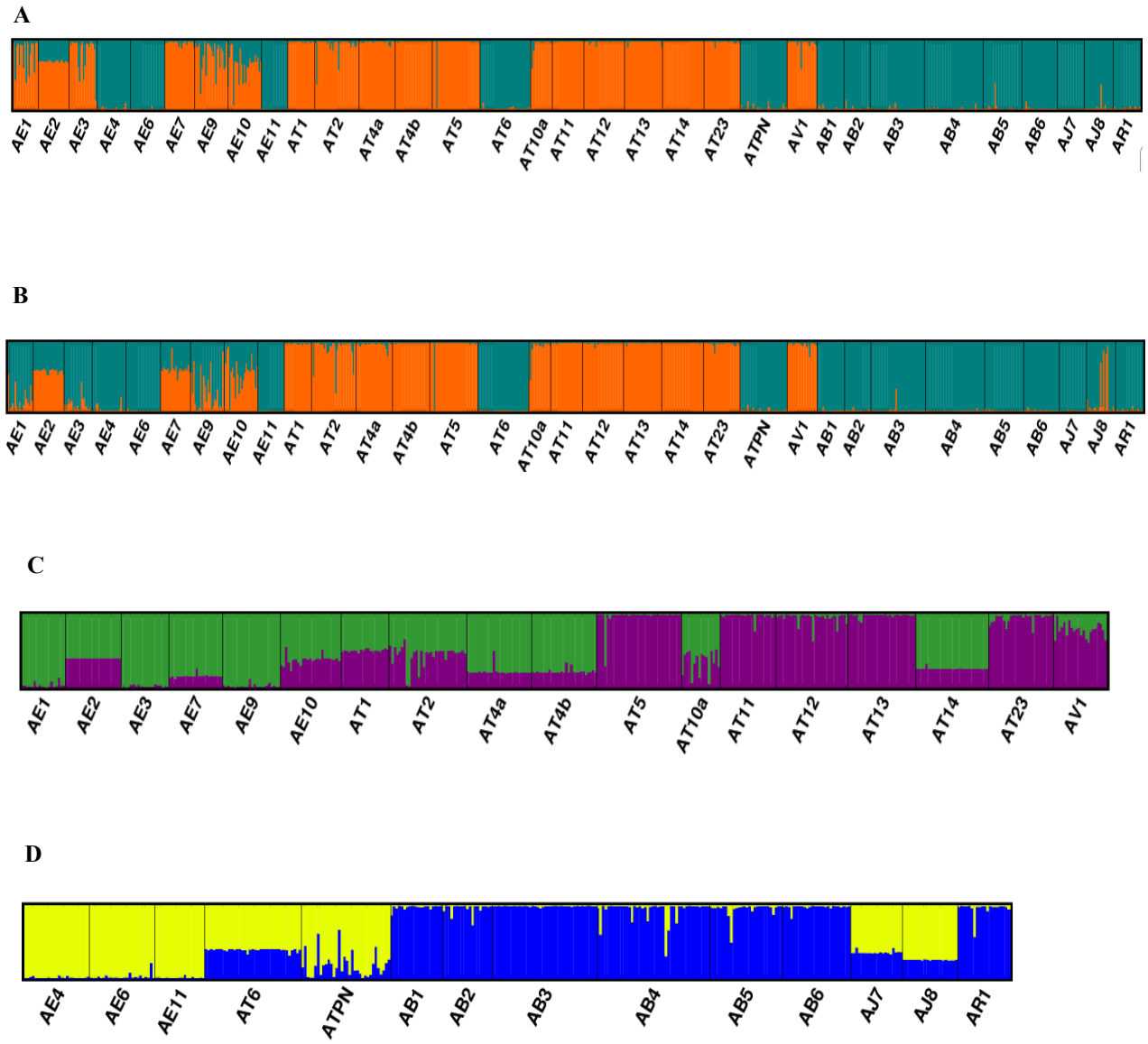


Figure 4. A-B. Bar plots showing the assignment of individuals to clusters and their levels of admixture assessed with a hierarchical analysis in STRUCTURE. The length of coloured bars represents the fractional assignment of individuals for K = 2. Sampling locations labels refer to Table 1. **A.** Complete data-set with 9 SSR loci. **B.** Partial data-set with 6 SSR loci (loci *25.6–16*, *11-3*, *7–27.2* were excluded). **C.** Subset of western populations based on orange coloured bars in **A**. Exclusion of the admixed *A. einseleana* populations (AE1-2-3-7-9-10) produces the barplot of Figure 4A for K=3 in Chapter 3. **D.** Subset of eastern populations based on cyan coloured bars in **A**.

Table 4. Partitioning of molecular variance among and within *A. thalictrifolia, A. einseleana* and *A. bertolonii,* species following the most recent taxonomic delimitations at the species level (see Materials & Methods for details).

Component of variation	*df*	*SS*	Variance component	% Variation	Fixation index	Significance
Among species	2	187.718	0.13	3.83	0.05*	0.000±0.000
Among locations within species	25	822.674	0.57	17.37	0.19¶	0.000±0.000
Among individuals within locations	1458	3801.479	2.61	78.80	0.22$	0.000±0.000

df degrees of freedom, SS sum of squares

** FCT, fixation index among species ¶ FSC; fixation index among location within species; $ FIS, fixation index among individuals within locations*

Table 5. Partitioning of molecular variance among and within *A. thalictrifolia, A. einseleana, A. bertolonii, A. julia and A. reuterii* following the most recent taxonomic delimitations at the species level (see Materials & Methods for details).

Component of variation	*df*	*SS*	Variance component	% Variation	Fixation index	Significance
Among species	5	325.720	0.15	4.59	0.05*	0.000±0.000
Among locations within species	26	863.006	0.59	17.85	0.19¶	0.000±0.000
Among individuals within locations	1620	4128.124	2.55	77.56	0.22$	0.000±0.000

df degrees of freedom, SS sum of squares

** FCT, fixation index among species ¶ FSC; fixation index among location within species; $ FIS, fixation index among individuals within locations* FCT, ¶ FSC, $ FST*

Table 6. Partitioning of molecular variance among and within clusters identified by STRUCTURE analyses on the complete data set with 9 SSR loci.

Component of variation	df	SS	Variance component	% Variation	Fixation index	Significance
Among clusters	1	57.41	0.083	2.53	0.03 *	0.059 ± 0.007
Among locations within clusters	16	549.72	0.64	19.54	0.20 ¶	0.000+-0.000
Among individuals within locations	880	2249.52	2.55	77.93	0.22 $	0.000+-0.000

df degrees of freedom, SS sum of squares

** FCT, fixation index among clusters ¶ FSC; fixation index among location within clusters; $FIS, fixation index among individuals within locations*

Table 7. Partitioning of molecular variance within *Aquilegia einseleana*.

Component of variation	*df*	*SS*	Variance component	% Variation	F_{ST}	D_{est}
Among locations	8	199.158	0.51720	22.38	0.02*	0.76* (95 % C.I. = 0.72 - 0.79)
Among individuals within locations	192	351.732	0.03786	1.64		

*df degrees of freedom, SS sum of squares, * p<0.01*

Table 8. Partitioning of molecular variance within *Aquilegia bertolonii* (Apuan Alps).

Component of variation	*df*	*SS*	Variance component	% Variation	F_{ST}	D_{est}
Among locations	5	64.23	0.18	2.6	0.06*	0.34* (95 % C.I. = 0.31 - 0.37)
Among individuals within locations	346	922.90	2.67	93.78		

*df degrees of freedom, SS sum of squares, * p<0.01*

4.5. Discussion

a) Genetic differentiation processes in six alpine European *Aquilegia*

AMOVA analyses, F_{ST} and D_{est} values indicated that only a very low fraction of the genetic variation is due to differences among *A. einseleana*, *A. thalictrifolia* and *A. bertolonii*, whereas most of the variation is among individuals and sampling locations within taxa. Inclusion of *A. julia*, *A. reuterii* and *A. vestinae* did not significantly change these results. The weak differentiation among taxa defined on a morphological and ecological basis was hypothesized to be partly explainable by a very recent process of diversification of the genus *Aquilegia* in Europe (Fior et al., 2013; Bastida et al., 2010), and it is in accordance with the polytomy detected at the base of the European clade in the recent *Aquilegia* phylogeny based on plastome data (Fior et al., 2013). Nevertheless, as vegetative traits in European columbines were demonstrated to have more evolutionary potential than floral traits (Castellanos et al., 2011), likely because of strong selective pressures correlated with different habitats (Bastida et al., 2010), it is possible that the genes under selection for these traits will uncover a much higher genetic differentiation among taxa compared to neutral loci. This finding would be in accordance with a mechanism of ecological speciation, where populations become reproductively isolated as a consequence of ecologically-based divergent natural selection, as opposed to a mutation-order speciation in which reproductive isolation arises by chance under similar selective pressures (Schluter, 2009). Ecological speciation can act even with a certain amount of gene flow, when this force does not influence the loci under adaptive selection (e.g. Schluter & Conte 2009). In the case of incipient speciation like the one of European *Aquilegia*, the genotypic diversification among species could be detectable only in a few genomic islands controlling the traits under selection, or in more extended genomic regions including also loci in linkage with these traits (Michel et al. 2010), while the remainder of the genome would still remain undifferentiated. Prolonged periods of reproductive isolation, maintained by the minor fitness of

hybrids on the adaptive traits in case of sympatric speciation, or by geographical confinement in case of allopatric speciation, will allow differentiation at the whole genome level, neutral loci included. For example, Cooper at al. (2010) found that the interbreeding *A. formosa* and *A. pubescens* of North America were indistinguishable at 9 nuclear neutral loci, despite the presence of clear differences in floral characters, and considered this evidence as a facilitation for a future genome-wide scan aimed at finding highly differentiated loci under selection in *Aquilegia*. In Europe, a high genetic similarity measured by 7 allozymes was found between the endemic *A. paui* and the widespread, sympatric congener *A. vulgaris*, in spite of important ecological and morphological differences (Martinell et al., 2010). Lavergne (2003) compared the endemic *A. viscosa* with *A. vulgaris* and concluded as well that ecological differentiation was mainly responsible for the maintenance of distinction between the 2 taxa. Moreover, Garrido et al. (2012) studied the genetic diversity and structuring of 3 endemic columbines of Sardinia (*A. barbaricina*, *A. nuragica* and *A. nugorensis*) and found out very low levels of genetic diversity across species with an almost null effect of population size on diversity. Therefore, they hypothesized that this unsubstantial diversity was consequent to local adaptation of the plants to their very restrained and peculiar habitats, like bare rocks and waterfront meadows. Indeed, Mattana et al. (2012) studied the autoecology of *A. barbaricina* and *A. nugoresis* dissecting their phenological trends, seed dispersal period and germination requirements, and demonstrated that the 2 endemics are rigorously adapted to their microhabitats. Finally, another evidence for ecological speciation was brought on *A. vulgaris* and *A. pyrenaica* subspecies by Alcantara et al. (2010), which indicated divergent selection acting on inflorescence height and number of flowers per inflorescence between habitats with different soil rockiness. Divergent selection related to elevation gradients was also proved to operate on the number of leaves per plant, but this trait apparently influenced only intra-specific variability. Thus, a process of ecological speciation with diversifying selection performing on specific traits is very likely to have characterized (or being actively characterizing) also the *Aquilegia* taxa analysed in this study.

Although a low proportion of among species molecular variance suggested that probably neutral loci have not yet accumulated the same level of differentiation with regards to the loci under selection, a rather high distinctiveness of the endemics *A. bertolonii*, *A. thalictrifolia* and *A. julia* was indicated by phylogenetic and multivariate analyses, and partly, also by bayesian assignment tests. First, *A. bertolonii* was clearly recognized as a separate lineage both under the phylogenetic

(Donoghue 1985; de Queiroz & Donoghue, 1988) and genotypic clustering (Mallet, 1995) criteria, showing a robust clan in the Neighbour Joining phylogram, and the most stable cluster in PCO-MC analysis. Hierarchical bayesian analyses (Figure 4D) confirmed the uniqueness of the endemic of Apuan Alps which was assigned a cluster shared exclusively with the geographically proximal population of *A. reuterii*. The relative ancient origin of *A. bertolonii* (approximately 2.5 Mya, Fior et al., 2013) within the European clade could partly explain the clear distinctiveness of this endemic at the neutral level, as the evolutionary forces of mutation and drift had more time to randomly operate on the genome enhancing diversification. Moreover, geographical isolation of *A. bertolonii* in the glacial refugia of Apuan Alps (Médail & Diadema, 2009) seemingly reinforced the speciation process. It is also interesting to note that while weak phenotypic characters distinguish *A. bertolonii* from *A. julia* and *A. reuterii*, a discrete differentiation is present at the genotypic level between the first endemic and the other 2, reflecting independent evolutionary origins (Fior et al., 2013).

Second, a statistically robust clan was recognized for *A. thalictrifolia* by Neighbour Joining phylogram, thus confirming it as a separate lineage according to the phylogenetic species concept. Agreement with the genotypic clustering criterion was also indicated by PCO-MC analyses, which grouped all *A. thalictrifolia* sampling locations in one stable cluster. STRUCTURE analyses, if on one side confirmed the uniqueness of *A. thalictrifolia* with regards to the rest, on the other side allowed a certain level of introgression of *A. thalictrifolia* within *A. einseleana* populations (stronger in the complete data set than in the partial one with 6 SSR loci), which will be discussed below. The prolonged geographic isolation of *A. thalictrifolia* in the glacial refugia of the Tremalzo-Tombea mountain ridges (see Chapter 3) seems a plausible explanation for the diversification pattern found on neutral genes, which is further confirmed by the preponderant role played by genetic drift over gene flow in shaping genetic population structure (see Chapter 3). Instead, the younger age of *A. thalictrifolia* relative to *A. bertolonii* (approximately 1.8 Mya; Fior et al., 2013) apparently excludes a preponderant role of time in the differentiation process.

Third, *A. julia* was considered a separate lineage only according to the phylogenetic species criterion, with the highest statistical support of all clans in the NJ phylogram. On the contrary, the sampled populations did not cluster together in a stable cluster of PCO-MC, nor were distinguished by STRUCTURE even for K=3 in hierarchical analysis (data not shown). Thus, if on one hand the taxonomic status of independent species recently proposed by Nardi (2011) for *A. julia* with respect to *A. bertolonii* seems to be confirmed, on the other hand the genotypes of the Slovenian columbine

populations present important similarities with some *A. einseleana* and hybrid populations (AT6 and ATPN) of the nearby Eastern Alps. This junction of genetic lineages from south-eastern Alps and western Balkan mountains is part of a more general series of connections widely observed at the subspecies and species level in the hilly refugial areas among adjacent mountain systems of Europe (Schmidtt, 2009).

Finally, *A. reuterii*, the taxon assimilated to *A. bertolonii* in previous taxonomies, was only recognized as a different lineage by the Neighbor Joining phylogram, but a distinct cluster was retrieved for this taxon based on chloroplast data (Fior et al., 2013), thus indicating an active process of radiation of this taxon located in the glacial refugium of Maritime Alps (Médail & Diadema, 2009).

If *A. bertolonii* and *A. thalictrifolia* could be distinguished as independent lineages at the phylogenetic and genotypic clustering level, and *A. julia* together with *A. reuterii* were separated according to the last criterion, the same is not true for *A. einseleana*, whose populations belonged to unsupported clans and unstable clusters that reflected geographic distribution. The grouping of the eastern populations of *A. einseleana* and the admixture of its western populations with the cluster including *A. bertolonii* (Figure 4A-B), could be indicative of the common origin of these 2 species, which in fact fall within the same clade in *Aquilegia* phylogeny and diverged much earlier relatively to the remainder of the European columbines (Fior et al., 2013). Post glacial secondary contacts, hybridization and reticulate evolution with neighbouring populations along the late Quaternary recolonization routes from the respective glacial refugia (Greimler, Park & Schneeweiss, 2011; Petit et al., 2003) could have partly confounded this original pattern. Consequently, hybrid populations originated like the ones found in the Friulian plain (ATPN) or in Valsugana (AT6), and a certain extent of gene flow is present (or was recently present) with some populations of Balkan Alps (see above). Likewise, the observed admixture between *A. einseleana* and *A. thalictrifolia* in the western cluster could indicate an ongoing process of hybridization between neighbouring populations of the 2 taxa, whose distribution range partly overlap, and/or past crossing events. The highest allelic richness and private allelic richness characterizing *A. einseleana* as a whole with regards to the remainder of the species seems to validate this possibility, as the areas where different lineages admixed were proved to be melting spots of genetic diversity, where most of the variation coming from different refugia is concentrated (Petit et al., 2003).

A. vestinae as well could not be distinguished as a different species by neutral loci and was

assimilated to *A. thalictrifolia*. Pfenniger & Moser (2002) described this species by a series of distinct morphological, ecological and isoenzymes characters, demonstrating that despite its occurrence within the distribution ranges of *A. einseleana* and *A. thalictrifolia*, this taxon was not a hybrid between the 2. Thus, an ongoing process of sympatric ecological speciation on traits under selection is not to be excluded for this columbine.

b) Implications for conservation

Confirmation of the species status for *A. thalictrifolia* in a broader data set including other species of geographically close columbines reinforce the conservation priority for the endemic already stated in Chapter 3. However, when *A. einseleana* populations of western Alps showing some admixture with *A. thalictrifolia* in the complete data set with 9 loci were taken together for a hierarchical analysis in STRUCTURE, the result is a discrete admixture between *A. einseleana* lineage with *A. thalictrifolia* populations of PU, AM and MI valleys (see Chapter 3 for labels referring to the valleys where *A. thalictrifolia* is distributed). Notwithstanding, in both data sets, the direction of gene flow was always from *A. thalictrifolia* to *A. einseleana*, suggesting a possible introgression of the first taxon into the second, and not the contrary. Moreover, STRUCTURE barplot produced on the partial data set (6 loci), showed much lower levels of admixture, thus generally confirming the distinctiveness of *A. thalictrifolia* from *A. einseleana*. The taxon presented the lowest levels of average heterozigosity across all congeners, but the absolute level of heterozigosity remains nevertheless high (see Chapter 3), and values of allelic richness and private allelic richness are quite similar to its congeners. The population of Val Meduna which was previously classified as *A. thalictrifolia* (ATPN; see Costalonga et al., 2006), together with that of Valsugana (AT6) are probably to be considered hybrids between the endemic and some populations of *A. einseleana* of eastern Alps, and therefore may not deserve an immediate priority for conservation.

The strong distinctiveness of *A. bertolonii* from Apuan Alps with regards to the remaining taxa confirms the validity of the recent taxonomic revisions that recognized to this group of populations

the species rank, while considering *A. reuterii* and *A. julia* as different entities. Thus, the IUCN world level classification of "lower risk" stated for *A. bertolonii* (Buord et al., 2011) needs to be updated, as it considered a much more extended distribution range and higher number of existing populations for *A. bertolonii*, ranging from south-eastern France to northern and central Italy and the Slovenian Alps. Moreover, the endemic grows on calcareous bedrock, often in the proximity of active marble pits (e.g. populations AB3, AB4) which heavily impacted the habitat of the plant within the Regional Park of Apuan Alps. Therefore, more attention should be paid to the active *in situ* protection, at least circumscribing the present populations in a visible way in order to partly limit habitat destruction.

4.6. Conclusions

The present work shed some light upon the evolutionary histories of 6 European *Aquilegia* taxa distributed in the alpine landscape encompassing south-eastern, Apuan and Maritime Italian Alps, together with Austrian and Slovenian Alps.

These taxa were chosen among the whole complex of European columbine species because they possessed clear morphological and ecological features, and circumscribed distribution ranges not far from one to another.

Despite this morpho-ecological distinctiveness, analyses of molecular variance showed low levels of among taxa differentiation, thus probably indicating that a very recent and still active process of ecological speciation is operating on some traits under divergent selection, which it is not yet widespread at the whole genome level.

Nevertheless, *A. bertolonii*, *A. thalictrifolia* and partly *A. julia* and *A. reuterii* could be distinguished on the basis of phylogenetic and/or multivariate analyses, thus suggesting an active role of geographic isolation within glacial refugia in enhancing diversification at neutral loci, and confirming the relatively more ancient origin of *A. bertolonii* demonstrated by a recent *Aquilegia* phylogeny (Fior et al., 2013).

On the other hand, a more complex history apparently applies for *A. einseleana*, where the shared ancient origin with *A. bertolonii* seems to be partly confounded by successive secondary contact and hybridization events on the post glacial recolonization routes.

Conservation implications for the endemics are discussed in the light of the results obtained. Indeed, the conservation priority for *A. thalictrifolia* already stated in Chapter 3 was strengthen by its distinctiveness from the remainder of the taxa. On the contrary, the 2 morphologically ambiguous populations were suggested to be hybrids between *A. thalictrifolia* and *A. einseleana*. Genetic analyses showed that *A. bertolonii* from Apuan Alps could be very clearly distinguished from the rest, thus deserving immediate conservation actions and the updating of IUCN red lists classification. With regards to the remaining taxa, more populations should be sampled and analysed for an accurate evaluation of their state of the risk.

5. GENERAL CONCLUSIONS

5.1. Towards conservation: integrating taxonomy, genetics and ecology

Despite the increasing availability of advanced DNA-based technologies, large population samples and sophisticated algorithms for data analysis, conservation geneticists of plants and animals are continuously faced with the difficult problem of converting their quantitative data to practical and effective actions to preserve biodiversity at small scales (Vernesi et al., 2008).

Here, an attempt was made to produce reliable scientific data while keeping in mind the need for a working application of these data in the field. The first important contribution to fill this gap, was the adoption of a unified species concept (USC; de Queiroz, 2007) applied to species delimitation of potentially endangered taxa. An increasing attention is dedicated to species delimitation applied to conservation of biodiversity (Wiens, 2007). Indeed, the species represents one of the most important units of comparison in several if not all domains of biological sciences, and also the most readily recognizable unit of biodiversity. The USC approach represents in this context a solution to the never ending species concept controversy that is gaining increasing consensus in the scientific community. In the case of *B. repanda* complex, the continuum of morphological differences and similarities among the populations across the taxon distribution range led to ambiguous taxonomic classifications at the subspecies level in the past. Application of the USC proved that BRG and BRB have acquired multiple properties that satisfy the phylogenetic criteria of monophyly and diagnosability, as well as the genotypic cluster criterion. For this reason, a taxonomic recombination classifying them as different species was proposed. Elevating BRG and BRB to the rank of species in virtue of their evolutionary distinctiveness both from each other and from the remainder of the complex means that these taxa will have a concrete possibility to be considered in future conservation planning. The next step will be to produce an IUCN state of the risk classification for BRB, in order to put it officially in the list of threatened species, like it was recently made for BRG (Branca et al., 2011). The added value of the IUCN risk assessment for BRB lies in the possibility to integrate available ecological information with genetic diversity and genetic structure data of

natural populations, which is still not the common practice in IUCN risk assessments (http://www.iucnredlist.org/technical-documents/categories-and-criteria).

Another important point of interest from a practical conservation perspective is represented by the study of the genetic structure and diversity of taxa living within ancient refugia. As mentioned above, BRB and *A. thalictrifolia* populations grow on the mountains ridges of mount Baldo and mounts Tremalzo-Tombea, which are well known calcareous refugia of south-eastern Alps; moreover, BRG distribution range extends on the lowlands of Friulan eastern Alps which were ice-free during the last glacial maximum; finally, *A. bertolonii* and *A. reuterii* are located in glacial refugia of Apuan and Maritime Alps, respectively. If we think about refugia not only in terms of Quaternary glacial and interglacial periods, but broadly as habitats where living organisms shelter during long term adverse conditions and from which they expand when these environmental conditions become suitable again, the identification and protection of refugia acquires a strategic importance to reduce the negative impacts of contemporary man-induced climate change (Keppel et al., 2012; Ashcroft, 2010; Médail & Diadema, 2009).

One could argue that with increasing global warming several pleistocenic refugia won't maintain their role as *in situ* refugia, and also that numerous species won't probably have sufficient dispersal capability to rapidly reach alternative *ex situ* refugia outside their distribution range. Nevertheless, as mentioned in the Introduction, some microclimates can be produced by fine-scale topographic complexity that are completely decoupled from regional climate, and thus may persist longer, even with increasing global warming (Dobrowski, 2011).

As already specified in Chapter 3, *A. thalictrifolia* ecological niche is defined by a very peculiar microclimate, where water dropping keeps constant moisture on the calcareous bedrock and overhanging ledges may protect the growth site from strong thermic excursions. Although this niche seems to have been partly altered by drought (Bonomi, Castellani & Longo, 2008), it still remains intact for most of the populations. Therefore, the growing sites of the endemic could represent persistent microrefugia (Rull, 2009) from ongoing climate warming which are scattered over the alpine landscape of south-eastern Alps.

Identification and protection of these microrefugia is of strategic importance not only for *A. thalictrifolia* but also for other endangered species presently associated to this ecological niche, like the endemics *Physoplexis comosa* (L.) Schur (*Campanulaceae*), protected at regional and national

level (IUCN: Lower Risk) and *Saxifraga aracnoidea* Sternb. (*Saxifragaceae*), protected at regional level (IUCN: Lower Risk), as well as for species that may benefit from them in future as *ex situ* refugia.

Data-loggers measuring spatio-temporal variability in microclimate represent in this context very useful tools for describing this kind of microrefugia (Keppel et al., 2012). Two of them were put under the ground in the growing site of populations AT12 and AT14, with temperature being measured every 2 hours for the whole year since 2005. Castellani (unpublished master's thesis) processed the temperature data of population AT12 from November 2005 to May 2006 and found that winter average temperatures ranged from 2°C to 5°C, thus indicating that the soil never froze. A comparative analysis of temperature values among different years will be carried out in the near future, thus contributing to characterize at fine scale these microrefugia.

5.2. Limits of the study

The study performed on the endemic angiosperm taxa of eastern Alps is largely incomplete, with several possible questions still left to be answered.

First, a thorough landscape genetic approach was not carried out. Storfer (2007) defined landscape genetics as "research that explicitly quantifies the effects of landscape composition, configuration and matrix quality on gene flow and spatial genetic variation". However, in our case, landscape genetic analyses were confined uniquely to assignment tests with spatial priors and isolation-by-distance tests. Storfer et al. (2010), reviewed the current state of the art in this discipline and observed that while these kind of analyses were the most popular in this context, integration with multivariate models including landscape variables led to more complete and realistic explanations of population spatial genetic structuring.

For example, from a conservation point of view, it would be interesting to evaluate if a differential effect of contemporary and historic landscape features exists on gene flow, in order to quantify the species sensitivity to human-induced landscape alterations. This could reveal particularly useful in the case of BRG, whose distribution range was heavily degraded during the last 30 years by anthropogenic activities like agriculture, modifications of water streams and soil pollution. If the genetic structure of the taxon results to be more correlated to the current landscape heterogeneity than to past landscape features, the taxon is likely to be very sensitive to habitat degradation, while viceversa is true if a higher correlation is found between current population structure and past landscape.

Moreover, no landscape investigation was done on small spatial scales to detect further genetic structuring, for example through spatial autocorrelation analyses within populations. With regards to *A. thalictrifolia*, if no IBD was found both on the whole distribution range and within the STRUCTURE clusters, this does not mean that no correlation between genetic and geographic distances can be present at smaller scales. Seed dispersion mainly by gravity or small water streams, the possibility of self-fertilization, the conspicuous percentage of biparental inbreeding found in the genus (Yang & Hodges, 2010), and the possible clonality by ramet formation from secondary rhizomes (Lega, personal observation, to be verified thoroughly) are important elements that

suggest spatial autocorrelation over small distances. If IBD is not observed in *A. thalictrifolia* even at finer scales, but cryptic genetic structures were found, this could be explained by presence of barriers which are less evident than the physical impediments to gene flow characterizing large spatial scales (mountain ridges and forests). Given the high specificity of the ecological niche in *A. thalictrifolia*, unsuitability of natural habitat could for example represent an important barrier to gene flow for the endemics, causing fine and cryptic breaks in genetic structure (Storfer, 2007; 2010).

Finally, a more complete analysis to identify the possible impediments to gene flow in *A. thalictrifolia* should be carried out even on higher spatial scales, for example finding out which particular matrix of physical distances correlates better with the matrix of genetic distances when applying Mantel tests.

Another important limit of the present work is represented by the adoption of different kinds of markers for genetic analyses of the 2 studies (Chapters 2 and 3), which does not allow a true comparability among taxa sharing a similar geographic distribution range in south-eastern Alps. The molecular studies on BRB and BRG, and *Aquilegia* species on the other side, gave very different results in terms of population genetic structuring and diversity. Indeed, the first 2 endemics revealed rather low levels of within-taxon genetic diversity and population differentiation, and a strong between-taxa distinctiveness. On the other side, the columbine species, and notably *A. thalictrifolia*, were rather structured and highly diverse internally, but a much weaker differentiation among taxa was observed, with clear evidence of admixture in certain cases.

One could question if these differences are mainly due to the distinct kind of molecular markers used for the 2 genera (AFLP and SSR respectively), or if other factors were more influential. However, it is not possible to answer this question with the data in our hands, as this depends, as mentioned above, on a parallel analysis including both SSR and AFLPs on the data sets. Here, only one kind of marker per plant was employed.

Despite the intrinsic limit of the present work, some predictions about the possible degree of correlation expected between the 2 markers within each taxon can be done based on the existing information. Indeed, Mariette et al. (2002), measured through simulations the trade-off existing between AFLPs and SSRs, and demonstrated that a low correlation is expected in 3 main cases: low heterogeneity among populations, for example when gene flow is high and populations are big;

heterogeneity of the genome and/or a few loci sampled throughout the genome; recent origin of populations, characterized by nonequilibrium conditions between drift, migration and mutation. Here, low population heterogeneity seemed to characterize both BRG and BRB, as demonstrated by AMOVA and assignment tests, thus suggesting that a microsatellite study would give different results. However, Mariette et al. (2002) demonstrated also that, irrespective of the evolutionary scenario, the correlations between microsatellites and AFLP genome scans were augmented with the number of sampled SSRs, and 4 to 10 times as many loci were necessary for dominant markers compared to codominant ones to reach the same level of precision. Thus, in this case, the SSR genomic sampling effort should approach to the above indicated proportions. As 628 AFLP bands were analysed for Brassica, ideally at least 60 SSRs should be used.

On the other side, a good correlation between markers may be found in *A. thalictrifolia* data set, because of the strong population heterogeneity detected (high values of fixation and differentiation indexes, recognition of different clusters corresponding to sampling locations by bayesian algorithms). Nevertheless, rejection of IBD highlighted that populations were not in migration-drift equilibrium, and because drift had more importance than gene flow, a high genome heterogeneity is expected; additionally, only 9 microsatellite loci were analysed. Therefore, one could argue that, despite the high level of population structuring, a parallel study with AFLP markers on *A. thalictrifolia* may possibly give different outcomes from the present one.

The above reasoning implies that the different results arisen from *Brassica* and *Aquilegia* data sets could be partly due to the kind of marker used. Even in the hypothetical case that one was able to quantify the proportion of variance explained by the marker effect, it would be anyway very hard to partition the remaining variance components. Indeed, several factors are likely to have contributed to the present genetic diversity and structure of the taxa under study, and disentangling all of them appears rather difficult.

First of all, although BRB and *A. thalictrifolia* evolved in a similar mountain system during the Pleistocene, BRG apparently followed a different evolutionary path. In fact, the first 2 taxa are located in similar glacial refugia of south-eastern calcareous Alps not far one to each other (Tribsch, 2004; Schönswetter et al., 2005). Both of them have survived to the last glacial maximum within these refugia, probably simply performing altitudinal shifts in response to climate adverse conditions, without any further expansion over long distances afterwards, similar to many genetic lineages of the smaller European high mountain systems (Schmitt, 2007). On the contrary, BRG is

part of the parasteppic and gravel plain flora of the Friulian Plain just outside the eastern Alps, which remained largely unglaciated during the Pleistocene and which now is the only lowland region in the study area of Tribsch (2004) rich in endemisms. Second, BRB and BRG probably originated in very different historical moments with respect to *A. thalictrifolia*. Phylogenetic molecular dating at the genus level suggested an approximate age for the radiation between the genus *Brassica* and *Arabidopsis* based on the NADH subunit 4 at 14–20 Myr (Yang et al., 1999), and the analysis of chalcone synthase and alcohol dehydrogenase loci produced an average distance between the 2 clades of 24 Myr (Koch et al., 2000). On the contrary, the genus *Aquilegia* was indicated by Fior et al. (2013) to have originated much more recently, about 6.85 Myr, with the radiation of the European clade including *A. thalictrifolia* dated 1.83 Myr.

Third, the three angiosperms have rather different reproductive modes: while *Aquilegia* is partly self-crossing, even if with a conspicuous contribution of biparental inbreeding (Yang & Hodges, 2010), *B. repanda* is probably mainly outcrossing and possibly self-incompatible, as suggested by studies on other *Brassica* species (e.g. Zhang et al., 2011).

6. BIBLIOGRAPHY

Afzal-Rafii Z & Dodd RH. 2007. Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in western Europe. *Molecular Ecology.* 16:723–736.

Agapow P, Bininda-Emonds OR, Crandall KA, Gittleman JL, Mace GM, Marshall JC, Purvis A. **2004**. The impact of species concept on biodiversity studies. *The Quarterly Review of Biology*. 79: 161-179.

Akeroyd, J. 1993. *Aquilegia*. In T. Tutin, V. Heywood, N. Burges, D. Valentine, S. Walters, and D. Webb [eds.], Flora Europaea I: Psilotaceae to Platanaceae. Pp: 287–290. Cambridge University Press, Cambridge.

Alcantara JM, Bastida J.M. & Rey PJ. 2010. Linking divergent selection on vegetative traits to environmental variation and phenotypic diversification in the Iberian columbines (*Aquilegia*). *Journal of evolutionary biology* 23: 1218–1233.

Al-Shehbaz IA, Beilstein MA, Kellogg EA. **2006**. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Plant Systematics and Evolution*. 259: 89-120.

Alvarez N, Thiel-Egenter C, Holderegger R, Tribsch A et al. 2009. History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecology Letters*, 12,

632–640.

Andersson L. 1990. The driving force: species concepts and ecology. *Taxon*. 39:375-382.

Ashcroft MB. 2010. Identifying refugia from climate change. *Journal of Biogeography*, 37, 1407–1413.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden*. 82, 247–277.

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J. 2000. Microsatellites can be misleading: an empirical and simulation study. *Evolution.* 54, 1414–1422.

Barrett CF, Freudenstein JV. **2011**. An integrative approach to delimiting species in a rare but widespread mycoheterotrophic orchid. *Molecular Ecology*. 20(13): 2771-2786.

Bastida JM, Alcantara JM, Rey PJ, Vargas P & Herrera CM. 2010. Extended phylogeny of *Aquilegia*: the biogeographical and ecological patterns of two simultaneous but contrasting radiations. *Plant Systematics And Evolution.* 284: 171–185.

Beilstein MA, Al-Shehbaz IA, Kellogg EA. **2006**. Brassicaceae phylogeny and trichome evolution. *American Journal of Botany*. 93: 607-619.

Bellusci F, Musacchio A, Palermo AM, Pellegrino G. **2010**. Exploring the history of *Serapias politisii* (Orchidaceae), a narrow endemic on the opposite coasts of the Otranto channel: insights from molecular investigations. *Botanical Journal of the Linnean Society*. 162: 572-580.

Benjamini Y & Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*. 57 (1): 289–300.

Bilz M, Kell SP, Maxted N, Lansdown RV. 2011. *European Red List of Vascular Plants*. Luxembourg: Publications Office of the European Union. European Commission.

Bivand RS, Pebesma EJ & Gomez-Rubio V. 2008. *Applied spatial data analysis with R*. Springer, NY., http://www.asdar-book.org/

Bonin A, Bellemain E, Bronken EP, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*. 13: 3261-3273.

Bonomi C, Castellani C & Longo M. 2008. *Aquilegia thalictrifolia* Schott & Kotschy. Flora da conservare: implementazione delle categorie e dei criteri IUCN (2001) per la redazione di nuove Liste Rosse. *Informatore Botanico Italiano*, 40/1: 58-61.

Branca F, Donnini D, Dulloo ME & Kell SP. 2011. *Brassica glabrescens*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. Available at: http://www.iucnredlist.org/details/161859/0.

Brunet J & Sweet HR. 2006b. The maintenance of selfing in a population of the Rocky Mountain columbine. *International Journal of Plant Sciences*. 167:213–219.

Brunet J, Larson-Rabin Z & Stewart CM. 2012. The Distribution of Genetic Diversity Within and Among Populations of the Rocky Mountain Columbine: The Impact of Gene Flow, Pollinators,

and Mating System. *International Journal of Plant Sciences*.173(5):484-494.

Buord S, Gargano D, Gigot G, Jogan N & Montagnani C. 2011. *Aquilegia bertolonii*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. Available at: www.iucnredlist.org.

Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA, Baillie, JEM et al. 2010. Global Biodiversity: Indicators of Recent Declines. *Science.* 328: 1164-1168.

Canestrelli D, Sacco F, Nascetti G. 2012. On glacial refugia, genetic diversity, and microevolutionary processes: deep phylogeographical structure in the endemic newt. *Biological Journal of the Linnean Society*, 105:42–55.

Carreras-Carbonell J, Macpherson E, Pascual M .2006. Population structure within and between subspecies of the Mediterranean triplefin fish *Tripterygion delaisi* revealed by highly polymorphic microsatellite loci. *Molecular Ecology.* 15, 3527–3539.

Casado-Amezúa P, Goffredo S, Templado J, & Machordom A. 2012. Genetic assessment of population structure and connectivity in the threatened Mediterranean coral *Astroides calycularis* (Scleractinia, Dendrophylliidae) at different spatial scales. *Molecular ecology.* 21(15):3671–85.

Castellani, C. Ecologia dei semi di *Aquilegia thalictrifolia* Schott et Kotschy e loro longevità in banca del germoplasma. Tesi di laurea specialistica in Scienze della Natura, Percorso in Gestione Sostenibile del Patrimonio Naturalistico. Università degli Studi di Pavia. AA 2005-2006.

Castellanos MC, Alcantara JM, Rey PJ, & Bastida JM. 2011. Intra- population comparison of vegetative and floral trait heritabilities estimated from molecular markers in wild *Aquilegia* populations. *Molecular Ecology.* 20: 3513–3524.

Cavalli-Sforza LL & Edwards AW. 1967. Phylogenetic analysis. Models and estimation procedures. *The American Journal of Human Genetics*. 19: 233–257.

Chung M, Gelembiuk G, Givnish TJ. **2004**. Population genetics and phylogeography of endangered *Oxytropis campestris* var. *chartacea* and relatives: arctic-alpine disjuncts in eastern North America. *Molecular Ecology*. 13: 3657-3673.

Clausen J, Keck DD, Hiesey WM. 1945. *Experimental studies on the nature of species. II. Plant evolution through amphiploidy and autoploidy, with examples from Madiinae.* Carnegie Inst.Wash. Publ. 564:1–174.

Conti F, Manzi, A, Pedrotti F. 1992. *Libro rosso delle piante d'Italia*. TIPAR Poligrafica Editrice, Roma.

Conti F, Manzi A, Pedrotti F. 1997. *Liste Rosse Regionali delle Piante d'Italia*. WWF Italia, Società Botanica Italiana, Camerino.

Cooper EA, JB Whittall, SA Hodges, M Nordborg. 2010. Genetic variation at nuclear loci fails to distinguish two morphologically distinct species of *Aquilegia*. *PLoS ONE.* 5:e8655

Corander JC, Waldmann P, Sillanpää MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics.* 163:367–374.

Coumou D & Rahmstorf S. 2012. A decade of weather extremes. *Nature Climate Change*. 2(7): 491-496.

Coyne JA & Orr HA, **2004**. *Speciation*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Cracraft J. **1983**. Species concepts and speciation analysis. *Current Ornithology*. 1: 159–187.

Davis JI, Nixon KC. **1992**. Populations, genetic variation, and the delimitation of phylogenetic species. *Systematic Biology*. 41: 421-435.

Dawson TP, Jackson ST, House JI, Prentice IC, Mace, GM. 2011. Beyond predictions: biodiversity conservation in a changing climate. *Science.* 332, 53.

Debussche M, Thompson JD. **2002**. Morphological differentiation among closely related species with disjunct distributions: a case study of Mediterranean *Cyclamen* L. subgen. *Psilanthum* Schwarz (Primulaceae). *Botanical Journal of the Linnean Society*. 139: 133-144.

Dempster AP, Laird NM, Rubin, DB. 1977. Maximum Likelihood from incomplete data via the *EM* algorithm (with discussion). *Journal of the Royal Statistical Society.* 39(1): 1-38.

de Queiroz K, Donoghue MJ. **1988**. Phylogenetic systematics and the species problem. *Cladistics*. 4: 317–338.

de Queiroz K, **1998**. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: Howard DJ , Berlocher SH, eds. *Endless forms: Species and speciation*. Oxford University Press, New York, 57–75.

de Queiroz K. **2007**. Species concepts and species delimitation. *Systematic Biology*. 56: 879-886.

de Queiroz K. **2011**. Branches in the lines of descent: Charles Darwin and the evolution of the species concept. *Biological Journal of the Linnean Society*. 103: 19-35.

DeSalle R & Amato G. 2004. The expansion of conservation genetics. *Nature Reviews–Genetics* 5: 702–712.

Després L, Gielly L, Taberlet P. **2003**. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. *Molecular Phylogenetics and Evolution*. 27: 185-196.

Dieringer D, Schlötterer C. 2003. Microsatellite analyser (MSA): a plat-form independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3, 167–169.

Dobrowski SZ. 2011. A climatic basis for microrefugia: the influence of terrain on climate. *Global Change Biology.* 17, 1022–1035.

Dewoody J, Nason JD, Hipkins VD. 2006. Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* 6, (4): 951-957.

Diaz-Forero I, Kuusemets V, Maend M, et al. 2011. Effects of forest habitats on the local abundance of bumblebee species: a landscape-scale study. *Baltic Forestry*.17(2): 235-242.

Dobzhansky, T. 1950. Mendelian populations and their evolution. *The American Naturalist*. 84:401–418.

Dobzhansky T. 1970. *Genetics of the evolutionary process*. Columbia University Press, New York.

Donoghue MJ. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist*. 88: 172– 181.

Doyle JJ & Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bullettin.* 19: 11-15.

Dray S & Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*. 22(4): 1-20.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A. 2010. *Geneious v5.1.* Available from http://www.geneious.com.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology.* 7: 214.

Earl DA & von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources.* 2: 359-361.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research.* 32(5): 1792-1797.

Ekblom R, Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*. 107:1–15.

Ensing D, Prodöhl PA, McGinnity P, Boylan P, O'Maoiléidigh N, Crozier WW. 2011. Complex pattern of genetic structuring in the Atlantic salmon (*Salmo salar* L.) of the River Foyle system in northwest Ireland: disentangling the evolutionary signal from population stochasticity. *Ecology & Evolution.* 1(3):359-72.

Erschbamer B, Kiebacher T, Mallaun M & Unterluggauer P. 2009. Short-term signals of climate change along an altitudinal gradient in the South Alps. *Plant Ecology*. 202:79–89

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

Excoffier L, Lischer H. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564-567.

Falush D, Stephens M, Pritchard JK. **2003**. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 164: 1567-1587.

Fior S, Li M, Oxelman B, Viola R, Hodges SA, Ometto L, Varotto C. 2013. Spatiotemporal reconstruction of the *Aquilegia* rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions. *New Phytologist*. In press.

Fiori A. 1923-1929. *Nuova flora analitica d'Italia*. Edagricole, Firenze.

Frankham R, Ballou JD, Briscoe DA, **2002**. *Introduction to conservation genetics*. Cambridge: Cambridge Univ. Press.

French GC, Hollingsworth PM, Silverside AJ, Ennos,RA. 2008. Genetics, taxonomy and the conservation of British *Euphrasia*. *Conservation Genetics*. 9: 1547-1562.

Garcia-Pereira MJ, Caballero A, Quesada H. 2010. Evaluating the relationship between evolutionary divergence and phylogenetic accuracy in AFLP data sets. *Molecular Biology and Evolution.* 27: 988–1000.

Garrido JL, Fenu G, Mattana E, Bacchetta G. 2012. Spatial genetic structure of *Aquilegia* taxa endemic to the island of Sardinia. *Annals of Botany.*109: 953–964.

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.

Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P. 2010. Calculations of population differentiation based on G_{ST} and D: forget G_{ST} but not all of statistics! *Molecular Ecology.*19, 3845-3852.

Gismondi A. 1950. *Aquilegia*. In: Gismonid A., editor. *Prospetto della Flora Ligustica*. Genova (Italy): Scia. Pp. 410-411.

Gómez-Campo C, **1993**. *Brassica*. In: Castroviejo S , Aedo C, Gómez-Campo C, Laínz M, Montserrat P, Morales R, Muñoz Garmendia F, Nieto Feliner G, Rico E, Talavera S, Villar L, eds.

Flora Iberica. Madrid, 362-384.

Gómez-Campo C & Martínez Laborde J. **1998**. Reajustes taxonómicos y nomenclaturales en la tribu Brassiceae (Cruciferae). *Anales del Jardín Botánico de Madrid*. 56(2): 379-380.

Gómez-Campo C, **1999**. Taxonomy. In: Gómez-Campo C, eds. *Biology of Brassica coenospecies*. Amsterdam, 3-32.

Gómez-Campo C. **2003**. The genus *Guenthera* Andr. in Bess. (Brassicaceae, Brassiceae). *Anales del Jardin Botanico de Madrid*. 60: 301-307. **Greimler J, Park J-M, Schneeweiss H. 2011**. Gentianella (Gentianaceae): a model taxon for evolution in the Alps. *Taxon 60: 427–435.*

Gregorius HR. 1987. The relationship between the concepts of genetic diversity and differentiation. *Theoretical and Applied Genetics*. 74:397–401.

Greuter W, Burdet HM, Long G. **1986**. Med-Checklist, 3 vols. Genève & Berlin.

Gruber S, Pekrun C, Claupein W. 2004. Seed persistence of oilseed rape (*Brassica napus*): variation in transgenic and conventionally bred cultivars. *Journal of Agricultural Science*. 142:29–40.

Guillot G, Estoup A, Mortier F, Cosson JF. 2005. A spatial statistical model for landscape genetics. *Genetics*. 170: 1261-1280.

Guillot G, Mortier F, Estoup A. 2005. Geneland: a computer package for landscape genetics. *Molecular Ecology*

Notes. 5: 712–715.

Guillot G. 2008. Inference of structure in subdivided populations at low levels of geneticdifferentiation – the correlated allele frequencies model revisited. *Bioinformatics*. 24, 2222–2228.

Guillot G & Santos F. **2010**. Using AFLP markers and the Geneland program for the inference of population genetic structure. *Molecular Ecology Resources*. 10: 1082-1084.

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

Hagen M, Wikelski M, Kissling WD. **2011**. Space use of bumblebees (*Bombus* spp.) revealed by radio-tracking. *PLoS ONE*. 6(5): e19997.

Haldane, JBS.1954. An exact test for randomness of mating. *Journal of Genetics* 52: 631-635.

Hamrick JL & Godt JW. **1996**. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London Biological Sciences*. 351: 1291-1298.

Hartl, DL. 2000. *A Primer of Population Genetics* (3rd edn). Sinauer Associates, Inc., Sunderland, MA. Pp. 221.

Hartl DL, Clark AG, **2007**. *Principles of population genetics*. Sinauer Associates, Inc. Publishers, fourth edition.

Hedrick PW. 2005. A standardized genetic differentiation measure. *Evolution.* 59, 1633–1638.

Heller R & Siegismund H. 2009. Relationship between three measures of genetic differentiation G_{ST}, D_{est} and G'_{ST} : how wrong have we been? *Molecular Ecology.* 18, 2080–2083.

Herlihy CR & CG Eckert. 2002. Genetic cost of reproductive assurance in a self-fertilizing plant. *Nature.* 416:320–323.

Hewitt GM. 2011. *Mediterranean peninsulas – the evolution of hotspots*. In: Zachos FE, Habel JC, eds. Biodiversity hotspots. Amsterdam: Springer.

Hey J. 2006. On the failure of modern species concepts. *Trends in ecology & evolution. 21*(8), 447–50.

Heywood VH & Akeroyd JR, **1993**. *Brassica* L. In: Tutin TG , Burges NA, Chater AO, Edmonson JR, Heywood VH, Moore DM, Valentine DH, Walters SM, Webb DA, eds. *Flora europaea, ed. 2*. Cambridge, etc, 405-409.

Hodges SA, & Arnold M. 1994a. Columbines - a Geographically Widespread Species Flock. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 91: 5129–5132.

Hodges SA, & Arnold M. 1994b. Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens*. P*roceedings Of The National Academy Of Sciences Of The United States of America.* 91(7):2493–2496.

Hodges SA, & Arnold M. 1995. Spurring plant diversification: Are floral nectar spurs a key innovation? *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 262: 343–127

348.

Holderegger R, Thiel-Egenter C. 2009. A discussion of different types of glacial refugia used in mountain biogeography and phylogeography. *Journal of Biogeography*. 36:476–480

Honnay O & Jacquemyn H. 2007. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology*. 21, 823–831.

Honnay O, Bossuyt B, Jacquemyn H, Shimono A, Uchiyama K. 2008. Can the seedbank maintain the genetic variation in the above ground plant population? *Oikos.* 117, 1–5.

Howes BJ, Lindsay B, Lougheed SC. 2006. Phylogeography of a temperate lizard. *Molecular Phylogenetics and Evolution.* 40, 183–194.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science.* 294(5550): 2310-2314.

Hutchinson DW, Templeton AR. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution.* 53: 1898–1914.

Jakobsson M & Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* 23:1801–1806.

Jalas J, Suominen J, Lampinen, R. 1996. *Atlas Florae Europaeae. Distribution of Vascular Plants in Europe. 11. Cruciferae (Ricotia to Raphanus)*. Helsinki, 252-256.

Jost L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology.* 17, 4015–4026.

Jost L. 2009. D vs. G_{ST} : response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology.* 18: 2088–2091.

Kalinowski ST. 2002. Evolutionary and statistical properties of three genetic distances. *Molecular Ecology.* 11:1263 – 1273

Kalinowski ST. 2004. Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conservation Genetics*. 5:539–543.

Kalinowski ST. 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes.* 5, 187–189.

Kamm U, Gugerli F, Rotach P, Edwards P, Holderegger, R. 2010. Open areas in a landscape enhance pollen-mediated gene flow of a tree species: evidence from northern Switzerland. *Landscape Ecology*. 25:903–911

Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AGT, Hopper SD, Franklin SE. 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography.* 21, 393–404.

Kholodova MV. 2009. Comparative phylogeography: molecular methods, ecological interpretation. *Molecular Biology.* 43, 847–854.

Kimura M, Crow JF. 1964. The number of alleles that can be maintained in a finite population. *Genetics*. 49:725–738.

Kimura M. & Weiss GH. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics.* 49: 561–576.

Klein C, Wilson K, Watts M, Stein J, Berry S, Carwardine J, Stafford Smith M, Mackey B & Possingham H. 2009. Incorporating ecological and evolutionary processes into continental-scale conservation planning. *Ecological Applications.* 19(1): 206-217.

Koch M, Haubold B, Mitchell-Olds T. 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Ara- bidopsis, Arabis and related genera (Brassicaceae). *Molecular Biology and Evolution*. 17:1483–1498.

Koopman WJM, Wissemann V, Cock KD, Huylenbroeck JV, Riek JD, Sabatino GJH, Visser D, Vosman B, Ritz CM, Maes B, Werlemark G, Nybom H, Debener T, Linde M, Smulders MJM. **2008**. AFLP markers as a tool to reconstruct complex relationships: a case study in *Rosa* (Rosaceae). *American Journal of Botany*. 95: 353-366.

Koopman WJM. 2005. Phylogenetic signal in AFLP data sets. *Systematic biology.* 54(2): 197-217.

Kramer AT & Havens K. 2009. Plant conservation genetics in a changing world. *Trends in Plant Science*. 14: 599–607.

Kreivi M, Rautiainen P, Aspi J, Hyva M. **2005**. Genetic structure and gene flow in an endangered perennial grass, *Arctophila fulva* var. *pendulina*. *Conservation Genetics*. 6: 683-696.

Kreyer D, Oed A, Walther-Hellwig K, Frankl R. 2004. Are forests potential landscape barriers for foraging bumble- bees? Landscape scale experiments with *Bombus terrestris* agg. and *Bombus pascuorum* (Hymenoptera, Apidae). *Biological Conservation*. 116:111–118.

Kuss P, Armbrustera GFJ, Ægisdóttira HH, Scheepensa JF, Stöcklin J. 2011. Spatial genetic structure of *Campanula thyrsoides* across the European Alps: Indications for glaciation-driven allopatric subspeciation. *Perspectives in Plant Ecology Evolution and Systematics.* 13(2):101-110.

Latter BDH. 1972. Selection in finite populations with multiple alleles. 111. Genetic divergence with centripetal selection and mutation. *Genetics.* 70:475-490.

Lavergne S. 2003. *Les espèces végétales rares ont-elles des caractéristiques écologiques et biologiques qui leur sont propres? Applications à la conservation de la flore en Languedoc-Roussillon*. Académie de Montpellier. École Nationale Supérieure Agronomique de Montpellier, Montpellier.

Lavergne S, Debussche M & Thompson JD. 2005. Limitations on reproductive success in endemic *Aquilegia viscosa* (Ranunculaceae) relative to its widespread congener *Aquilegia vulgaris*: the interplay of herbivory and pollination. *Oecologia.* 142(2):212–220.

Lega M, Fior S, Prosser F, Bertolli A, Li M, Varotto C. 2012. Application of the unified species concept reveals distinct lineages for disjunct endemics of the *Brassica repanda* (Brassicaceae) complex. *Biological Journal of the Linnean Society*. 106: 482–497.

Leng L, Zhang DX. 2011. Measuring population differentiation using G_{ST} or D? A simulation study with microsatellite DNA markers under a finite island model and nonequilibrium conditions. *Molecular Ecology.* 20, 2494–2509

Leroy SAG & Arpe K. 2007. Glacial refugia for summer- green trees in Europe and south-west Asia as proposed by ECHAM3 time-slice atmospheric model simulations. *Journal of Biogeography.* 34, 2115–2128.

Li WH. 1976. Distribution of nucleotide differences between 2 randomly chosen cistrons in a subdivided population - finite island model. *Theoretical Population Biology.* 10(3): 303-308.

Loveless M, Hamrick JL. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*. 15:65–95.

Luzzani F. 1932. *Aggiunte alla flora della Val del Chiese e dintorni*. Studi Trentini di Scienze Naturali 10, 3-25.

Macior LW. 1966. Foraging behavior of *Bombus* (Hymenoptera: Apidae) in relation of *Aquilegia* pollination. *American Journal of Botany*. 53:302–309.

Malécot G. 1948. *Les Mathématiques de l'Hérédié*. Masson, Paris.

Mallet J. **1995**. A species definition for the modern synthesis. *Trends in Ecology and Evolution*. 10: 294–299.

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution.*18, 189–197.

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

Cancer Research. 27, 209– 220.

Manzano P, Malo JE, Peco B. 2005. Sheep gut passage and survival of Mediterranean shrub seeds. *Seed Sciences Research.*15: 21–28.

Manzano P & Malo JE. 2006. Extreme long-distance seed dispersal via sheep. *Frontiers in Ecology and Environment.* 4: 244–248.

Mariette S, Le Corre V, Austerlitz F, Kremer A. 2002. Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Molecular Ecology.* 11, 1145–1156.

Martinell MC, López-Pujol J, Bosch M, Blanché C. **2010**. Low genetic variability in the rare, recently differentiated *Aquilegia paui* (*Ranunculaceae*). *Biochemical Systematics and Ecology.* 38:390–397.

Martinell CM, Blanché ARC & Bosch M. 2011. Shift towards autogamy in the extremely narrow endemic *Aquilegia paui* and comparison with its widespread close relative *A. vulgaris* (Ranunculaceae). *Plant Systematics and Evolution*. 295:73–82.

Maruyama T & Fuerst PA. 1985. Population bottlenecks and non-equilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics.* 111:675–689.

Mattana E, Daws MI, Fenu G, Bacchetta G. 2012. Adaptation to habitat in *Aquilegia* species endemic to Sardinia (Italy): Seed dispersal, germination and persistence in the soil. *Plant Biosystems*.146(2): 374-383.

Mayr E, **1942**. *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. New York: Columbia University Press.

Médail F & Diadema, K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*. 36, 1333–1345.

Medrano M, Castellanos MC, Herrera CM. 2006. Comparative floral and vegetative differentiation between two European *Aquilegia* taxa along a narrow contact zone. *Plant Systematics and Evolution*. 262:209–224.

Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution.* 60:2399–2402.

Meirmans PG & Hedrick PW. 2011. Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources.* 11: 5–18.

Meirmans PG, Goudet J, Gaggiotti OE. **2011**. Ecology and life history affect different aspects of the population structure of 27 high-alpine plants. *Molecular Ecology. 20*(15):3144–55.

Meudt HM & Clarke AC. 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science*. 12(3):106-117.

Meudt HM, Lockhart PJ, Bryant D. **2009**. Species delimitation and phylogeny of a New Zealand plant species radiation. *BMC Evolutionary Biology*. 9:111.

Michalakis Y, Excoffier L. 1996. A generic estimation of population subdivision using distances

between alleles, with special reference for microsatellite loci. *Genetics*. 142:1061–1064.

Michel AP, Sima S, Powella THQ, Taylora MS, Nosilb P, Federa JL. 2010. Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of the United States of America.* 107(2):19724–9729.

Morin PA, Manaster C, Mesnick SL, Holland R. 2009. Normalization and binning of historical and multi-source microsatellite data: overcoming the problems of allele size shift with allelogram. *Molecular ecology resources.* 9(6): 1451-5.

Mummenhoff K, Franzke A, Koch M. **1997a**. Molecular data reveal convergence in fruit characters used in the classification of *Thlaspi* s.l. (Brassicaceae). *Botanical Journal of the Linnean Society*. 125: 183–99.

Munier, DJ, Brittan KL & Lanini, WT. 2012. Seed bank persistence of genetically modified canola in California. *Environmental Science and Pollution Research.* 2281-2284.

Munz PA. 1946. *Aquilegia: The cultivated and wild columbines*. Gentes Herbarium. 7:1–150.

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

Nardi E. 2011. Diagnoses aquilegiarum novarum in Europa crescentium. *Webbia*, 66(2): 231-232.

Nei M. 1972. Genetic distance between populations. *American Naturalist.* 106:283–292.

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

Nei M. 1973a. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA*. 70:3321–3323.

Nei M. 1973b. *The theory and estimation of genetic distance*. Pp. 45-54. In: Genetic Structure ofPopulations, edited by N. E. MORTON. University Press of Hawaii, Honolulu.

Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*. 19:153–170.

Nei M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Newman D & Pilson D. 1997. Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution.* 51(2):354–362.

Nixon KC & Wheeler QD. **1990**. An amplification of the phylogenetic species concept. *Cladistics*. 6: 211–223.

O'Reilly CL, Cowan RS & Hawkins JA. 2007. Amplified fragment length polymorphism genetic fingerprinting challenges the taxonomic status of the near-endemic species. *Botanical Journal of the Linnean Society*, 467–476.

Osborne JL, Carreck NL, Swain JL, Knight ME, Goulson D, Hale RJ, Sanderson RA. **2008**. Bumblebee flight distances in relation to the forage landscape. *Journal of Animal Ecology*. 77: 406-415.

Ozenda P. 2009. On the genesis of the plant population in the Alps: New or critical aspects. *Comptes Rendus Biologies.* 332: 1092–1103.

Pampanini R. 1909. Intorno a due *Aquilegia* della flora italiana. *Nuovo Giornale Botanico Italiano.* 16: 5-22.

Parolo G & Rossi G. 2008. Upward migration of vascular plants following a climate warming trend in the Alps. *Basic and Applied Ecology.* 9:100–107.

Pauli H, Gottfried M, Dullinger S, Abdaladze O, Akhalkatsi M, Alonso JLB, Coldea G, Dick J, Erschbamer B, Calzado RF et al. 2012. Recent Plant Diversity Changes on Europe's Mountain Summits. *Science.* 336(6079):353-355.

Pearson RG. 2006. Climate change and the migration capacity of species. T*rends in Ecology and Evolution.* 21:111–113.

Pebesma EJ & Bivand RS. 2005. Classes and methods for spatial data in R. *R News.* 5(2), http://cran.r-project.org/doc/Rnews/

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M et al. 2003. Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science.* 300: 1563–1565.

Petit RJ, El Mousadik A, Pons P. **1998**. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*. 12:844–855.

Pfenninger A & Moser DM. 2002. Eine neue Aquilegia-Art aus den Judikarischen Alpen

(Valvestino, Prov. di Brescia, Italien): *Aquilegia vestinae* Pfenninger &D. M. Moser. *Candollea*. 57(2):317.

Pignatti S. 1982. Flora d'Italia. Edagricole.

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

Poldini L. **1973**. *Brassica glabrescens*, eine neue Art aus Nordost-Italien. *Giornale Botanico Italiano.* 107(4): 181-189.

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

Prosser F. 1998. *La distribuzione delle entità endemiche "strette" in Trentino alla luce delle più recenti esplorazioni floristiche*. Supplemento Annali Museo Civico Rovereto 14, 31-64.

Prosser F. 1999. *I fiori che hanno reso celebre il Monte Tombea*, in: Avanzini, M., Prosser, F., Zontini, G. (Eds.), *Tombea, giardino sulle Alpi*. CAI-SAT Sezione di Storo. Pp. 37-64.

Prosser F. 2001. *Lista Rossa della Flora del Trentino*. Osiride, Rovereto.

Prosser F & Bertolli A. **2007a**. A new subspecies of *Guenthera repanda* (Brassicaceae) from Mt Baldo (SE Prealps, Italy). *Willdenowia*. 37:191.

Prosser F & Bertolli A. **2007b**. Nuova combinazione in *Brassica* (Brassicaceae). *Annali Museo civico Rovereto, Sezione: Archeologia, Storia, Scienze naturali.* 22:295-297.

Raeymaekers, JAM, Lens, L,Van den Broeck, F, van Dongen, S & Volckaert, FAM. 2012. Quantifying population structure on short timescales. *Molecular Ecology*. 2(14): 3458-3473

Raymond M & Rousset F. 1995. An exact test for population differentiation. *Evolution.* 49:1280-1283.

R Core Team. 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

R Development Core Team, **2011**. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: http://www.R-project.org/.

Reed DH & Frankham R. **2003**. Correlation between fitness and genetic diversity. *Conservation Biology*. 17: 230-237.

Reeves PA & Richards CM. **2009**. Accurate inference of subtle population structure (and other genetic discontinuities) using principal coordinates. *PLoS ONE*. 4: e4269.

Reeves PA & Richards CM. **2011**. Species Delimitation under the General Lineage Concept: An Empirical Example Using Wild North American Hops (Cannabaceae: *Humulus lupulus*). *Systematic Biology*. 60: 45-59.

Reisch C, Kaiser AJ, Horn A, Poschlod, P. 2010. Genetic variation of *Mimetes hirtus* and

Mimetes fimbriifolius (Proteaceae) – a comparative analysis of two closely related fynbos species. *Plant biology.* 12, 537–544.

Ro KE, Keener CS & McPheron BA. 1997. Molecular phylogenetic study of the Ranunculaceae: utility of the nuclear 26S ribosomal DNA in inferring intrafamilial relationships. *Molecular phylogenetics and evolution. 8*(2):117–27.

Rogers JS. 1972. *Measures of genetic similarity and genetic distance.* Pp. 145-153. In: Studies in GeneticsVII. University of Texas Publication 7213, Austin, TX

Ronquist F & Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics.* 19(12): 1572-1574.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology*. 4:137–138

Ross JA, Matter SF, Roland J. 2005. Edge avoidance and movement of the butterfly *Parnassius smintheus* in matrix and non-matrix habitat. *Landscape Ecology.* 20, 127–135.

Ross KG, Gotzek D, Ascunce MS, Shoemaker DD. **2010**. Species delimitation: a case study in a problematic ant taxon. *Systematic Biology*. 59: 162-184.

Rousset F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103-106.

Ryman N & Leimar O. 2009. G_{ST} is still a useful measure of genetic differentiation – a comment on Jost's D. *Molecular Ecology.* 18, 2084–2087

Saccheri I, Kuussaari M, Kanare M, Vikman P, Fortelius W & Hanski I. 1998. Inbreeding and extinction in a butterfly metapopulation. *Nature*. 392:491–494.

Saitou N & Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 4: 406–425.

Sánchez-Teyer LF, Moreno-Salazar S, Esqueda M, Barraza A, Robert ML. **2009**. Genetic variability of wild *Agave angustifolia* populations based on AFLP: A basic study for conservation. *Journal of Arid Environments*. 73: 611-616.

Sanghvi LD.1953. Comparison of genetical and morphological methods for a study of biological differences. American Journal of Physical Anthropology. 11: 385-404.

Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH. **2010**. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual Review of Entomology*. 55: 421-438.

Schluter D & Conte, GL. 2009. Genetics and ecological speciation. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (1):9955–9962.

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science.* 323, 737–741.

Schmitt T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology.* 4:11.

Schmitt T. 2009. Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology.* 6:9.

Schönswetter P, Stehlik I, Holderegger R, Tribsch A. 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*. 14:3547–3555.

Schott HW & Kotschy CGT. 1853. Ueber Aquilegien. *Verhandlungen des zoologisch-botanischen Vereins*. 3:125-130.

Segarra-Moragues JG, Palop-Esteban M, Gonzalez-Candelas F, Catalán, P. 2005. On the verge of extinction: genetics of the critically endangered Iberian plant species, *Borderea chouardii* (Dioscoreaceae) and implications for conservation management. *Molecular Ecology*. 14: 969–982.

Simmons MP, Zhang L, Webb C, Müller K. 2007. A penalty of using anonymous dominant markers (AFLPs, ISSRs, and RAPDs) for phylogenetic inference. *Molecular Phylogenetics and Evolution.* 42: 528–542.

Smith RJ & Waldren S. 2009. Patterns of genetic variation in *Colchicum autumnale* L. and its conservation status in Ireland: a broader perspective on local plant conservation. *Conservation Genetics.* 11(4):1351–1361.

Soltis, PS & Gitzendanner, MA. 1999. Molecular Systematics and the Conservation of Rare Species. *Conservation Biology.* 13(3):471–483.

Sorensen, AE. 1986. Seed Dispersal by Adhesion. *Annual Review of Ecology and Systematics.*17: 443-463 .

Spielman D, Brook BW, Frankham R. **2004**. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States*

of America. 101:15261-15264.

Staden R, Beal KF, Bonfield JK, **1998**. The Staden Package. In: Misener S, Krawetz SA, eds. *Methods in Molecular Biology, Bioinformatics Methods and Protocols*. Humana Press Inc., Totowa, NJ.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson, Spear SF, Dezzani E, Delmelle E, Vierling L, Waits LP. 2007. Putting the 'landscape' in landscape genetics. *Heredity.* 98:128–142.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP. 2010. Landscape genetics: where are we now? *Molecular Ecology.* 19:3496–3514.

Swofford DL. 2003. *PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sinauer Associates: Sunderland, Massachusetts.

Takezaki, N & Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics.* 144: 389–399.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins D. **1997**. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*. 24: 4876-4882.

Thompson JD. **1999**. Population differentiation in mediterranean plants: insights into colonization history and the evolution and conservation of endemic species. *Heredity*. 82(3): 229-236.

Tribsch A & Schönswetter P. 2003. Patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon.* 52: 477–143

497.

Tribsch A. 2004. Areas of endemism of vascular plants in the Eastern Alps in relation to Pleistocene glaciation. Journal of Biogeography. 31:747–760.

Van Rossum F &Triest L. **2010**. Pollen dispersal in an insect-pollinated wet meadow herb along an urban river. *Landscape and Urban Planning*. 95(4) : 201-208.

Van Valen L. **1976**. Ecological species, multispecies, and oaks. *Taxon*. 25: 233–239.

Vernesi C, Bruford MW, Bertorelle G, Pecchioli E, Rizzoli A, Hauffe HC. 2008. Where's the conservation in conservation genetics? *Conservation Biology.* 22:802–804.

Vik U, Jørgensen MH, Kauserud H, Nordal I, Brysting AK. 2010. Microsatellite markers show decreasing diversity but unchanged level of clonality in *Dryas octopetala* (Rosaceae) with increasing latitude. *American Journal of Botany*, 97(6):988-97.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. **1995**. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*. 23: 4407– 4414.

Walter KS & Gillett HJ. 1998. *1997 IUCN Red List of Threatened Plants*. Compiled by the World Conservation Monitoring Centre. IUCN, Gland, Switzerland, and Cambridge, UK.

Warwick SI, Sauder CA. 2005. Phylogeny of tribe *Brassiceae* (Brassicaceae) based on chloroplast restriction site polymorphisms and nuclear ribosomal internal transcribed spacer and chloroplast trnL intron sequences. *Canadian Journal of Botany*. 83(5): 467–483.

Weir BS, Cockerham C. **1984**. Estimating F-statistics for the analysis of population structure. *Evolution*. 6: 1358–1370.

Weir BS, 1990b. *Genetic data analysis*. Sinauer Publ., Sunderland, Mass.

Whitlock MC. 2011. G'_{ST} and D do not replace F_{ST}. *Molecular Ecology.* 20:1083–1091.

Whittall JB, Medina-Marino A, Zimmer EA Hodges SA. 2006. Generating single-copy nuclear gene data for a recent adaptive radiation. *Molecular Phylogenetics And Evolution.* 39: 124–134.

Whittall JB & Hodges SA. 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature.* 447:706–U12.

Wiens JJ, Servedio MR. 2000. Species delimitation in systematics: inferring diagnostic differences between species. *Proceedings of the Royal Society of London | Biological Sciences*. 267: 631-636.

Wiens J.J. 2007. Species delimitation: new approaches for discovering diversity. *Systematic Biology.* 56:875–878.

Wilkinson M, O Mcinerney J, Hirt RP, Foster PG, Embley TM. 2007. Of clades and clans: terms for phylogenetic relationships in unrooted trees. *Trends in Ecology and Evolution.* 22(3): 114-115.

Wright, S. 1940. The statistical consequences of Mendelian heredity in relation to speciation. In: J. Huxley, ed. *The new systematics*. Oxford University Press, London. Pp:161–183.

Wright S. 1943. Isolation by distance. *Genetics.* 28, 114–138.

Wright S. 1951. The genetical structure of populations. *Annals of Eugenics.* 15, 323–354.

Wright S. 1978. *Variability Within and Among Natural Populations*. In: Evolution and the Genetics of Populations, Volume 4. University Of Chicago Press, Chicago, Illinois.

Yang YW, Lai KN, Tai PY & Li WH. 1999. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *Journal of Molecular Evolution*. 48:597–604.

Yang, JY, Counterman, BA, Eckert, CG, Hodges SA. 2005. Cross-species amplification of microsatellite loci in *Aquilegia* and *Semiaquilegia* (Ranunculaceae). *Molecular Ecology Notes* 5, (2): 317-320.

Yang JY & Hodges SA. 2010. Early inbreeding depression selects for high outcrossing rates in *Aquilegia formosa* and *Aquilegia pubescens*. *International Journal of Plant Sciences*. 171:860–871.

Young A, Boyle T, Brown T. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*. 11:413–418.

Zenari S, 1927. Sopra una nuova forma di "*Aquilegia thalictrifolia* (Schott et Kotschy) Nym.". *Nuovo Giornale Botanico Italiano.* 24: 201-203.

Zhang X, Yin D, Zhu W, Ma C, Fu T. 2011. Progress on characterization of self-incompatibility in *Brassica napus* L. *Euphytica*: 182(2): 147-155.

7. SUPPLEMENTARY MATERIALS

Chapter 2.

Figure A. MP strict consensus tree generated from the analysis of ITS sequences of the Brassiceae data set. Bootstrap values (> 50%) are reported above branches. *Brassica repanda* representatives included in this study are marked.

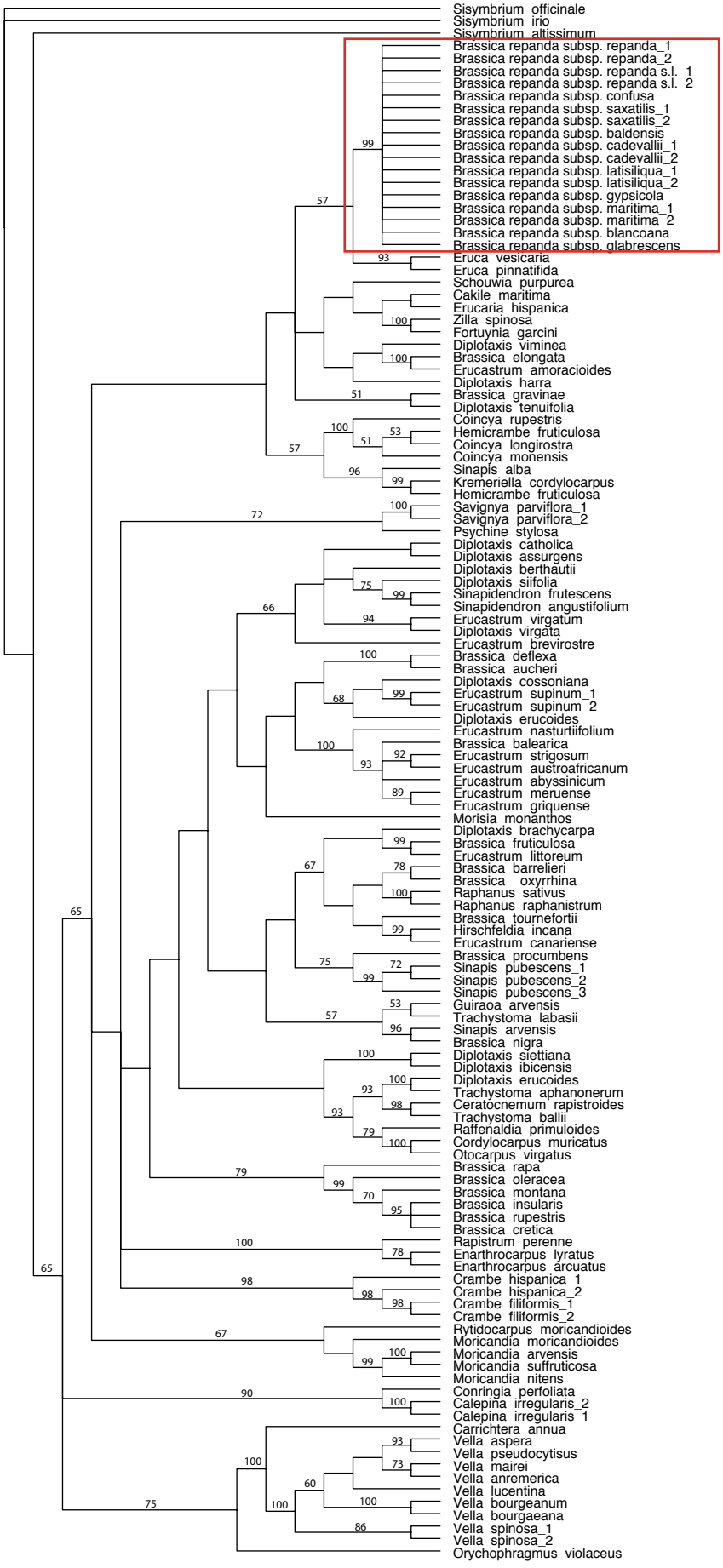
Figure B. Unrooted maximum clade credibility tree obtained from the BI analysis of AFLP data for the 47-accession matrix. Posterior probability values (>0.95) are shown above branches

.

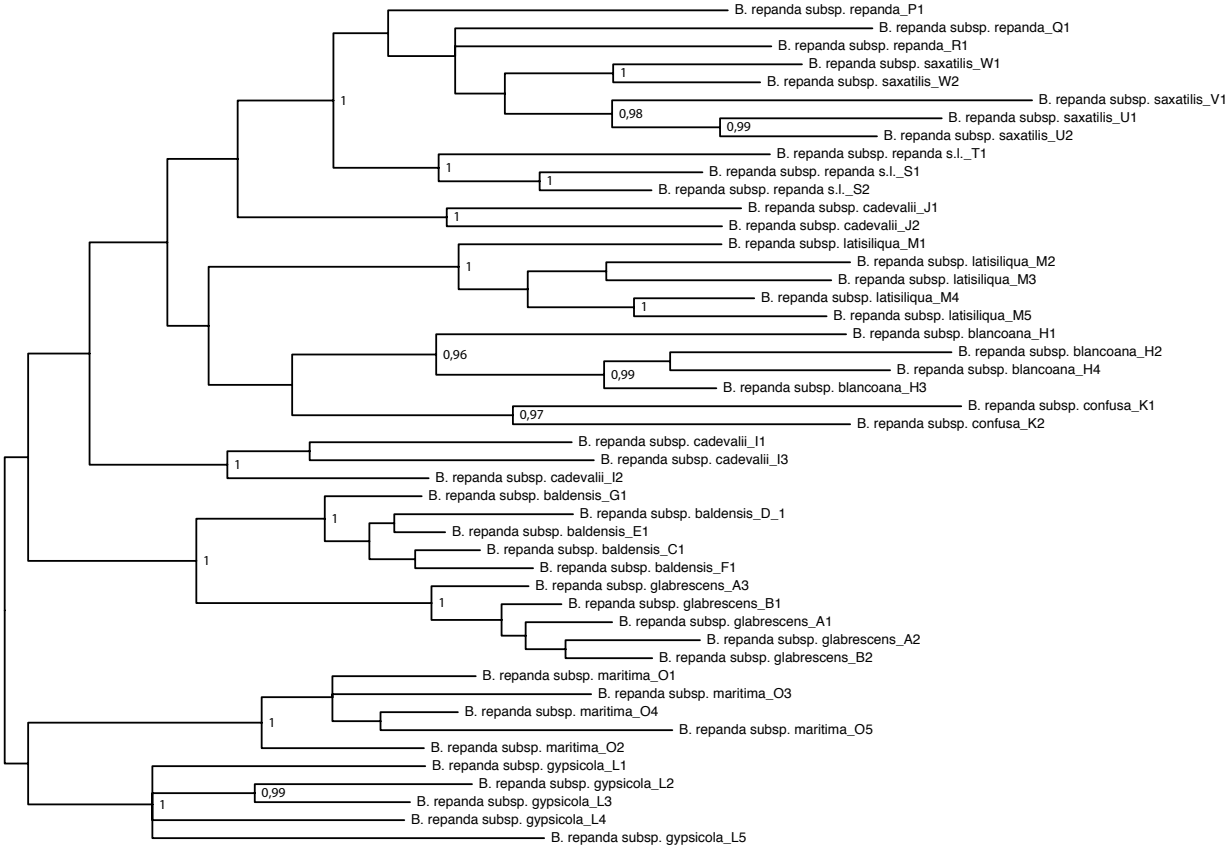
Figure C. Unrooted MP phylogram obtained for the AFLP matrix including all accessions. MP bootstrap support (>50%) and BI posterior probability (>0.95) values are shown above and below branches, respectively.

Figure D. PCO plot obtained from the 47-accession matrix including a proportionate number of samples for each subspecific entity. The first, second and third axis explained respectively 8.4, 7.0, 5.6 % of the total variance.

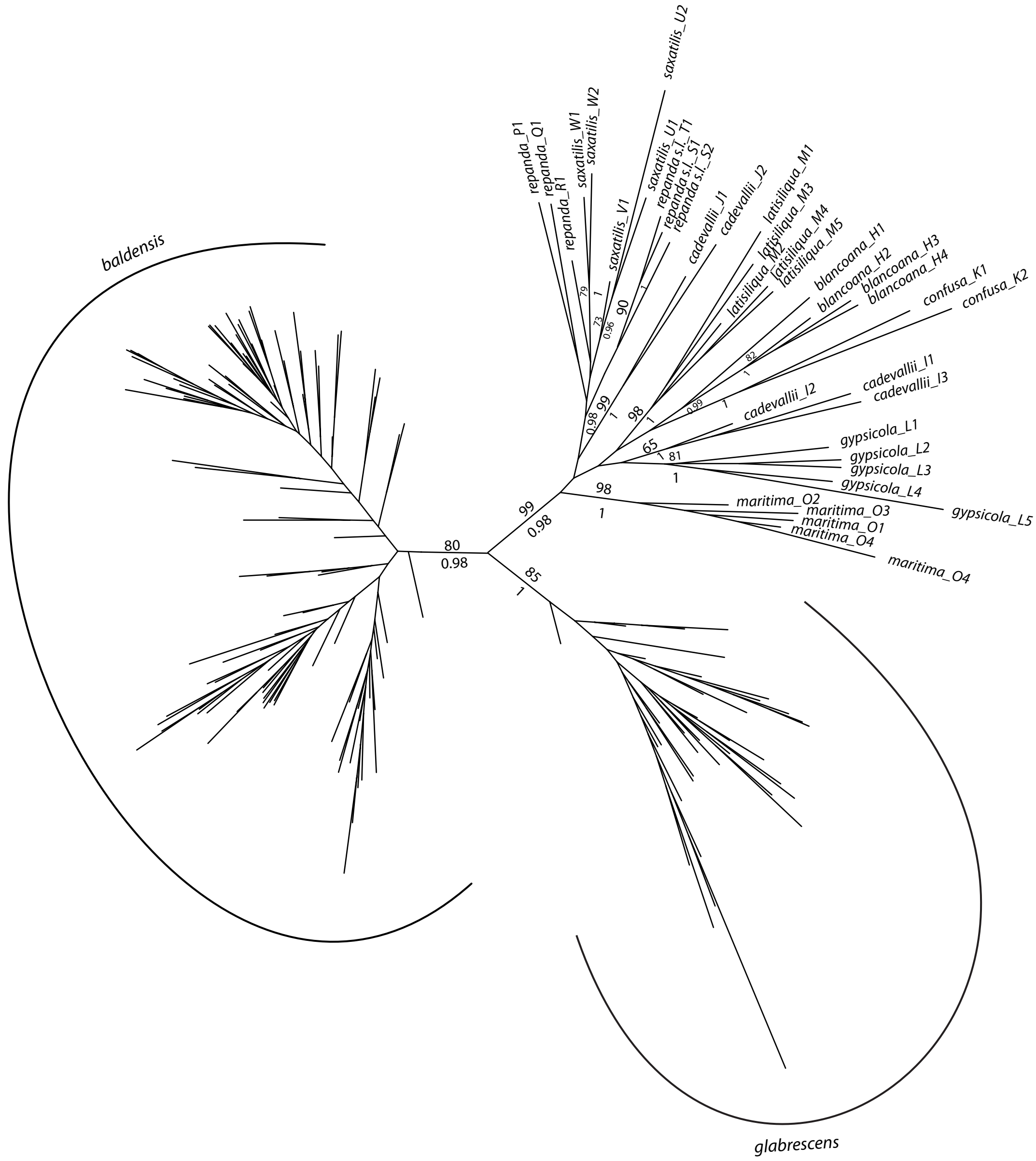
A


Sisymbrium officinale
Sisymbrium irio
Sisymbrium altissimum
Brassica repanda subsp. repanda_1
Brassica repanda subsp. repanda_2
Brassica repanda subsp. repanda s.l._1
Brassica repanda subsp. repanda s.l._2
Brassica repanda subsp. confusa
Brassica repanda subsp. saxatilis_1
Brassica repanda subsp. saxatilis_2
Brassica repanda subsp. baldensis
Brassica repanda subsp. cadevallii_1
Brassica repanda subsp. cadevallii_2
Brassica repanda subsp. latisiliqua_1
Brassica repanda subsp. latisiliqua_2
Brassica repanda subsp. gypsicola
Brassica repanda subsp. maritima_1
Brassica repanda subsp. maritima_2
Brassica repanda subsp. blancoana
Brassica repanda subsp. glabrescens
Eruca vesicaria
Eruca pinnatifida
Schouwia purpurea
Cakile maritima
Erucaria hispanica
Zilla spinosa
Fortuynia garcini
Diplotaxis viminea
Brassica elongata
Erucastrum amoracioides
Diplotaxis harra
Brassica gravinae
Diplotaxis tenuifolia
Coincya rupestris
Hemicrambe fruticulosa
Coincya longirostra
Coincya monensis
Sinapis alba
Kremeriella cordylocarpus
Hemicrambe fruticulosa
Savignya parviflora_1
Savignya parviflora_2
Psychine stylosa
Diplotaxis catholica
Diplotaxis assurgens
Diplotaxis berthautii
Diplotaxis siifolia
Sinapidendron frutescens
Sinapidendron angustifolium
Erucastrum virgatum
Diplotaxis virgata
Erucastrum brevirostre
Brassica deflexa
Brassica aucheri
Diplotaxis cossoniana
Erucastrum supinum_1
Erucastrum supinum_2
Diplotaxis erucoides
Erucastrum nasturtiifolium
Brassica balearica
Erucastrum strigosum
Erucastrum austroafricanum
Erucastrum abyssinicum
Erucastrum meruense
Erucastrum griquense
Morisia monanthos
Diplotaxis brachycarpa
Brassica fruticulosa
Erucastrum littoreum
Brassica barrelieri
Brassica oxyrrhina
Raphanus sativus
Raphanus raphanistrum
Brassica tournefortii
Hirschfeldia incana
Erucastrum canariense
Brassica procumbens
Sinapis pubescens_1
Sinapis pubescens_2
Sinapis pubescens_3
Guiraoa arvensis
Trachystoma labasii
Sinapis arvensis
Brassica nigra
Diplotaxis siettiana
Diplotaxis ibicensis
Diplotaxis erucoides
Trachystoma aphanonerum
Ceratocnemum rapistroides
Trachystoma ballii
Raffenaldia primuloides
Cordylocarpus muricatus
Otocarpus virgatus
Brassica rapa
Brassica oleracea
Brassica montana
Brassica insularis
Brassica rupestris
Brassica cretica
Rapistrum perenne
Enarthrocarpus lyratus
Enarthrocarpus arcuatus
Crambe hispanica_1
Crambe hispanica_2
Crambe filiformis_1
Crambe filiformis_2
Rytidocarpus moricandioides
Moricandia moricandioides
Moricandia arvensis
Moricandia suffruticosa
Moricandia nitens
Conringia perfoliata
Calepina irregularis_2
Calepina irregularis_1
Carrichtera annua
Vella aspera
Vella pseudocytisus
Vella mairei
Vella anremerica
Vella lucentina
Vella bourgeanum
Vella bourgaeana
Vella spinosa_1
Vella spinosa_2
Orychophragmus violaceus

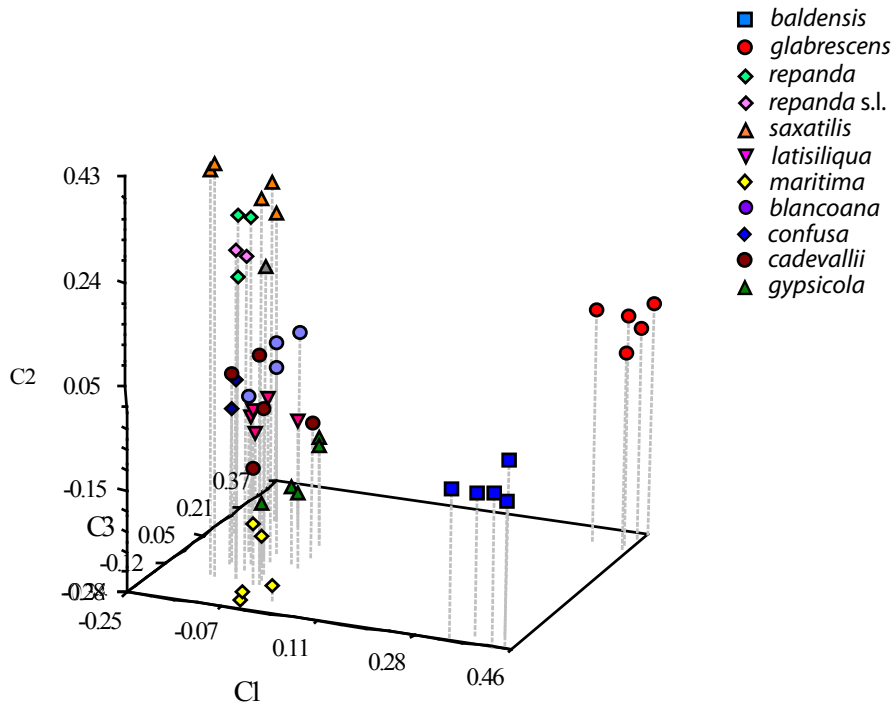
B


B. reppanda subsp. reppanda_P1
B. reppanda subsp. reppanda_Q1
B. reppanda subsp. reppanda_R1
B. reppanda subsp. saxatilis_W1
B. reppanda subsp. saxatilis_W2
B. reppanda subsp. saxatilis_V1
B. reppanda subsp. saxatilis_U1
B. reppanda subsp. saxatilis_U2
B. reppanda subsp. reppanda s.l._T1
B. reppanda subsp. reppanda s.l._S1
B. reppanda subsp. reppanda s.l._S2
B. reppanda subsp. cadevalii_J1
B. reppanda subsp. cadevalii_J2
B. reppanda subsp. latisiliqua_M1
B. reppanda subsp. latisiliqua_M2
B. reppanda subsp. latisiliqua_M3
B. reppanda subsp. latisiliqua_M4
B. reppanda subsp. latisiliqua_M5
B. reppanda subsp. blancoana_H1
B. reppanda subsp. blancoana_H2
B. reppanda subsp. blancoana_H4
B. reppanda subsp. blancoana_H3
B. reppanda subsp. confusa_K1
B. reppanda subsp. confusa_K2
B. reppanda subsp. cadevalii_I1
B. reppanda subsp. cadevalii_I3
B. reppanda subsp. cadevalii_I2
B. reppanda subsp. baldensis_G1
B. reppanda subsp. baldensis_D_1
B. reppanda subsp. baldensis_E1
B. reppanda subsp. baldensis_C1
B. reppanda subsp. baldensis_F1
B. reppanda subsp. glabrescens_A3
B. reppanda subsp. glabrescens_B1
B. reppanda subsp. glabrescens_A1
B. reppanda subsp. glabrescens_A2
B. reppanda subsp. glabrescens_B2
B. reppanda subsp. maritima_O1
B. reppanda subsp. maritima_O3
B. reppanda subsp. maritima_O4
B. reppanda subsp. maritima_O5
B. reppanda subsp. maritima_O2
B. reppanda subsp. gypsicola_L1
B. reppanda subsp. gypsicola_L2
B. reppanda subsp. gypsicola_L3
B. reppanda subsp. gypsicola_L4
B. reppanda subsp. gypsicola_L5
1
1
1
1
0,98
0,99
1
1
1
1
1
1
0,96
0,99
0,97
1
1
1
1
1
1
1
0,99
1

C


baldensis
glabrescens
saxatilis_U2
repanda_P1
repanda_Q1
repanda_R1
saxatilis_W1
saxatilis_W2
saxatilis_V1
saxatilis_U1
repanda s.l._T1
repanda s.l._S1
repanda s.l._S2
cadevallii_J1
cadevallii_J2
latisiliqua_M1
latisiliqua_M2
latisiliqua_M3
latisiliqua_M4
latisiliqua_M5
blancoana_H1
blancoana_H2
blancoana_H3
blancoana_H4
confusa_K1
confusa_K2
cadevallii_I1
cadevallii_I2
cadevallii_I3
gypsicola_L1
gypsicola_L2
gypsicola_L3
gypsicola_L4
gypsicola_L5
maritima_O2
maritima_O3
maritima_O1
maritima_O4
maritima_O4
80
0.98
99
0.98
85
1
98
1
65
81
1
99
0.98
98
1
90
0.96
73
79
82
0.99

D


baldensis
glabrescens
repanda
repanda s.l.
saxatilis
latisiliqua
maritima
blancoana
confusa
cadevallii
gypsicola
C2
0.43
0.24
0.05
-0.15
C3
0.37
0.21
0.05
-0.25
-0.07
0.11
0.28
0.46
C1

8. AKNOWLEDGEMENTS

Thanks to my supervisors Claudio Varotto and Stefano Leonardi, for trusting me and giving me the opportunity to do this PhD, providing their scientific and human support whenever needed.
I am particularly grateful to Simone Fior, for his continuous guidance, advice, encouragement and scientific lucidity throughout my doctoral work that helped me a lot to focus on my research, discover its sense, be curious and believe in myself.
I would also like to thank warmly Lino Ometto, for being a unique office neighbour, for his fundamental help with software issues, his competent advices, his contagious passion for nature and evolution, his cooperative way of doing research.
I express my gratitude to Enrico Barbaro for his deeply human and technical support during laboratory work. To Mingai Li, for helping me in organizing and planning my laboratory activities and for her technical advices. To Erika Stefani, Simone Larger and Daniela Nicolini for help with the sequencer.
Sincere thanks to Filippo Prosser and Alessio Bertolli for actively taking part to this research, with their amazing botanical competences, knowledge of the territory, and passion.
My warm and special thanks to all the people that helped me during field activities and shared their precious botanical knowledge with me: without them it would not have been possible to accomplish this work, and having such a fun while collecting plants! I thank Thomas Abeli, Gilberto Parolo, Costantino Bonomi, Bostjan Surina and his family, Andrej Podobnik, Igor Dakskobler, Renzo Vicentini, Stefania Dal Prà, Elena Bortoluzzi, Maddalena Bolognesi, Martina Fontanari, Giuseppe Menegoz, Lisa Angelini, Camille Pâtissier, Rocco Meloni, Federica Gilardelli.
Thanks also to Alberto Selvaggi, Andreas Hilpold, David Draper, Luc Garraud, and Noémie Fort for collection of those plants that I could not directly sample myself, and to Adriano Bruna for providing pictures of *B. glabrescens*.
My keen appreciation goes to Andrea Piotti, that was always available to discuss with me and give valuable advices on how to analyse and interpret my data, and precious R pipelines!

Helpful suggestions and materials also came from Cristina Castellani, Andreas Meraner, Andrea Gandolfi, Martina Bazzoli, Paolo Gratton, Anne Ghisla, Luca Delucchi, Markus Neteler, Marco Moretto, Alessandro Cestaro, Diego Micheletti, Nir Shahaf, Adam Konecny, Erica di Pierro, Cristiano Vernesi, Elena Mosca, Heidi Hauffe. Thanks to all of them.
My sincere gratitude goes then to all the people that I haven't mentioned yet, and that have been the simple happiness of my daily life in Fondazione Edmund Mach during my PhD, even "just" with a smile in a corridor, a chat on the train to S. Michele, or in the street to the canteen! Many thanks to Irene Criscuoli, Luisa Bresadola, Silvia Gazzani, Elisa Piaia, Laura Righetti, Michele Poli, Lorenzo Adami, Giorgio Maresi, Daniele Andreis, Valerio Mazzoni, Emanuela Kerschbamer, Thomas Roman, Chidananda Nagamangala, Emanuela Coller, Panagiotis Arapitsas, Andrea Campisano, Sara Castelletti, Sara Longhi, Emanuele de Paoli, Michela Troggio, Emanuele Eccel, Duccio Rocchini, Daniele Brazzale, Omar Rota Stabelli, Federico Ossi, Francesco Reyes, Florencia Ceppa, Noemi Herrero, Luca Cornetti, Wolters Pieter, Nadeesha Bandara, Lorena Herrera, Nada Surbanovski, Dan Sargent, Bo Wang, Qu Dong, and all those that I've forgot to put in the list!
I wish to express my cordial appreciation to Giorgio Bertorelle and Angelo Pavesi for their useful suggestions that helped to improve this thesis, and to 2 anonymous reviewers for their comments on parts of the manuscript.
I deeply thank my family, and all the wonderful friends that encouraged and supported me throughout this PhD experience, from everywhere...Trento-Piacenza-Nyamey-Kinshasa-Lille!

The Biological Journal of The Linnean Society, The Linnean Society of London and Blackwell Publishing Ltd (‘Wiley-Blackwell’) are acknowledged for allowing the use of parts of the already published Article and Abstract (DOI: 10.1111/j.1095-8312.2012.01887.x) in Chapter 2 of my thesis.
This research was funded by the Autonomous Province of Trento (Italy) within the ACE-SAP project ("Alpine Ecosystems in a Changing Environment: Biodiversity Sensitivity and Adaptive Potential"; regulation number 23, 12 June 2008, of the Servizio Università e Ricerca Scientifica) and held at Fondazione Edmund Mach (S. Michele all'Adige, Trento, Italy). Both institutions are acknowledged.

... and thanks to these amazing Mountains that always watch over me
giving me peace and good energy, everyday.

Picture by Camille Pâtissier