Multi-species phylogeographic investigations in closely-related taxa and mutualistic plant-insect systems

Thèse présentée à la Faculté des Sciences
Institut de Biologie
Université de Neuchâtel

Pour l'obtention du grade de docteur ès sciences

par

Yann Triponez

Acceptée sur proposition du jury:

Dr Nadir Alvarez, co-directeur de thèse

Prof. Martine Rahier, co-directrice de thèse

Dr Denis Michez (UMH, Mons, Belgium), rapporteur

Dr Bertrand Schatz (CEFE, CNRS Montpellier, France), rapporteur

Prof. Thomas Schmitt (Universität Trier, Germany), rapporteur

Soutenue le 20 Mai 2010

Université de Neuchâtel

2010

Université
de Neuchâtel unine

FACULTE DES SCIENCES
Secrétariat-Décanat de la faculté
- Rue Emile-Argand 11
- CP 158
- CH-2009 Neuchâtel

IMPRIMATUR POUR LA THESE

Multi-species phylogeographic investigations in closely-related taxa and mutualistic plant-insect systems

Yann TRIPONEZ

UNIVERSITE DE NEUCHATEL

FACULTE DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel,
sur le rapport des membres du jury

Mme M. Rahier (co-directrice de thèse),
MM. N. Alvarez (co-directeur de thèse), T. Schmitt (Université de Trier, D),
B. Schatz (Montpellier, F) et D. Michez (Mons, B)

autorise l'impression de la présente thèse.

Neuchâtel, le 16 août 2010

Le doyen :
F. Kessler

Felix Kessler

■ Téléphone : +41 32 718 21 00 ■ E-mail : secretariat.sciences@unine.ch ■ www.unine.ch/sciences

TABLE OF CONTENTS

SUMMARY

Phylogeography is the study of historical processes responsible for the contemporary geographic distributions of individuals and gene lineages. Studying species in space and time combining ecology and genetics allows disentangling patterns at the boundary between biogeography, population genetics and evolutionary ecology. As until now in Europe (but also in a very large portion of the World) many organisms have been considered by single-species phylogeographic approaches, it is time to develop a multi-species approach in the nascent field of comparative phylogeography, either considering closely-related taxa or organisms involved in tight ecological interactions. In this study, the fate of several species spanning a large variety of insect and plant systems is investigated. Among them are found alpine leaf-beetles, temperate widespread orchids, and mutualistic oil-producing flowers and oil-collecting bees. The methods used to fulfil the bio- and phylogeographic analyses in order to unravel past histories of these associate organisms state-of-the-art methods in genetics and ecology, by combining gene sequencing and genome fingerprinting, with ecological descriptors and niche modelling. Main results obtained confirm several paradigms addressed in the field of classical phylogeography but also bring new insights into the broad picture of phylogeography of European organisms. Thanks to our multi-species approach, the biogeographic patterns of a particular species can be studied in a global framework, including results obtained for related taxa or ecologically associated partners. In addition, the methods used here allow inferring further conclusions and perspectives in other fields: for instance, evaluation of genetic diversity indices useful in conservation biology, or analytical elements about species delimitation applicable to taxonomy, represent some of the supplementary (but also complementary) topics addressed in this thesis.

Keywords: Phylogeography, Europe, Alpine leaf-beetles, deceptive orchids, oil-collection mutualism, DNA sequencing, AFLP fingerprinting, ecological niches modeling, genetic lineages structure, comparative phylogeography, species delineation, conservation genetics.

RESUME

La phylogéographie se définit comme l'étude des processus historiques responsables de la distribution géographique actuelle des individus et des lignées génétiques. Etudier les espèces dans l'espace et dans le temps sur la base d'approches combinant la génétique et l'écologie représente un fantastique défi pour la biologie de l'évolution. En Europe, mais aussi ailleurs dans le monde, de telles approches ont été appliquées à de nombreux organismes, mais jusqu'ici la plupart de ces études n'ont toutefois été faites qu'au niveau d'espèces uniques. Le moment est donc venu de développer des études regroupant plusieurs espèces à la fois dans le domaine nouveau de la phylogéographie comparée, en utilisant par exemple des taxons très proches, ou alors des organismes interdépendants au sein de fortes interactions écologiques. Les espèces sélectionnées pour ce travail couvrent une large variété de groupes de plantes et d'insectes, tels que des coléoptères alpins, des orchidées de régions tempérées, ainsi que certaines espèces de plantes productrices d'huiles et d'abeilles mutualistes spécialisées dans la récolte de cette huile. Les méthodes d'analyses bio- et phylogéographiques permettant de dévoiler l'histoire des organismes étudiés associent des techniques de pointe telles que le séquençage génétique et le screening génomique, avec des descriptions et des modélisations de niches écologiques. Les principaux résultats obtenus confirment certains paradigmes mis en évidence jusqu'à présent, mais ils amènent également des notions nouvelles. Grâce à une approche globale, combinant les histoires évolutives de plusieurs lignées, les patrons biogéographiques d'une espèce en particulier peuvent être mieux compris en se basant sur les résultats obtenus parallèlement pour des taxons proches, ou pour des partenaires associés écologiquement. Les méthodes utilisées ici ont également permis de tirer des conclusions et d'entrevoir certaines perspectives au-delà de la phylogéographie ou de l'écologie évolutive : des considérations sur la diversité génétique, utiles en biologie de la conservation, ou des problèmes liés à la délimitation des espèces, applicables à la taxonomie, sont quelques uns des domaines additionnels (mais aussi complémentaires) que cette thèse aborde.

Mots-clés: Phylogéographie, Europe, coléoptères alpins, orchidées « trompeuses », mutualisme de pollinisation, séquençage d'ADN, caractérisation d'AFLP, modélisation de niches écologiques, structuration de lignées génétiques, phylogéographie comparée, concept d'espèces, génétique de la conservation.

REMERCIEMENTS / ACKNOWLEDGMENTS

Pour commencer je désire remercier mes deux co-directeurs de thèse, Martine Rahier et Nadir Alvarez, qui m'ont permis chacun à leur façon de concrétiser cette thèse en bonne et due forme et dans les meilleurs délais. Une belle aventure bien évidemment scientifique, mais aussi et surtout humaine, qui me restera à jamais en mémoire. Alors merci à vous deux d'avoir cru en moi et de m'avoir offert l'opportunité d'effectuer cette thèse au sein de vos projets de recherches respectifs et de ce(s) laboratoire(s), entouré d'une sacrée équipe…

Et dans cette équipe sans hésiter je débute par dire merci à Anahí ! Autant qu'il a cru en moi, Nadir a eu l'excellent idée d'avoir besoin de nous deux pour venir à bout se son projet POMAPII… C'est grâce à lui qu'ensemble nous avons pu partir à la découverte de cette Europe incroyable, des fjords norvégiens aux dolines balkaniques et des plaines hongroises aux sierras espagnoles. Un périple de plus de 100'000 km qui fut une chance incroyable de découvrir notre bon vieux continent, ainsi qu'une expérience de vie unique. Merci pour ta (presque) constante bonne humeur, ton enthousiasme et ton énergie inépuisable, ainsi que ton aide scientifique à maintes reprises lors des ces dernières années.

Je tiens ici aussi à remercier les étudiants de master qui ont partagé nos intérêts au sein de ce même projet : Makiala, Natacha, Robin, David, et plus particulièrement Lucas qui a participé plus directement à cette thèse via le sujet son travail de master. Tout ça sans oublier l'aide précieuse d'Anouk, notre laborante préférée et efficace qui nous a tous drillés au mieux en labo !

Hors de cette mini-équipe de recherche, ce sont des collègues de bureau, devenus des amis très chers que j'aimerai remercier maintenant. Tout d'abord Vio et Nico. J'ai partagé avec grande joie votre bureau pendant plus de deux ans, et passé du rire aux larmes en votre compagnie. Malgré la distance, c'est toujours un bonheur que de vous retrouver avec vos moitiés respectif(ve)s, en chair et en os, ou via skype. Je vous remercie pour tous ces moments ensemble, et aussi pour votre aide ! Puis merci aussi à Ivan, plus qu'un collègue, un colloc' pour quelques années, et avec toi beaucoup de bons moments autant à l'appart qu'au labo. Vous tous avez en très grande partie contribué à la bonne ambiance et à ma motivation de revenir chaque jour avec le sourire au travail.

Finalement merci à Cristina, Marie-Eve, Russ, Greg, Sarah, Claudia, Christelle, Elvira, Georg, Marco, Betty, Ted et tous les autres membres des labos pour ces nombreuses années en votre compagnie. Et sans oublier bien sûr nos secrétaires et bibliothécaires adorées : Brigitte, Natacha et Josiane.

Pour compléter la partie scientifique, je tiens à remercier les trois membres du jury qui ont pris du temps pour évaluer ma thèse et ont porté leur intérêt scientifique et soutien d'experts concernant ma recherche : Thomas Schmitt, Bertrand Schatz et Denis Michez. Je remercie également Loïc Pellissier, Christophe Randin et Antoine Guisan à l'UNIL pour leur aide en modélisation écologique. Un chaleureux merci aux aides ponctuelles partout en Europe lors du travail de terrain: N. Petrou, D. Průša, J. Ježek, L. Desprès, M. Gibernau, G. Scopece, F. Conti, B. Lumbierres, M. Ronikier, M. Puscas, H. Hipperson, K. Tan, G. Blaich, A. Grill, D. Kharazishvili, B. Dolinar, et ceux que j'oublie…Finalement, merci à mes divers « sponsors » : le Fonds National Suisse pour la recherche, l'Association Suisse des Sciences Naturelles, la Fondation Suisse pour les Orchidées (et particulièrement Samuel Sprunger), la Fondation Wütrich / Mathey-Dupraz, le Parc National des Grands Causses et la Société Française d'Orchidophilie.

Avec tous les gens cités jusqu'ici, il est désormais clair qu'une thèse ne s'effectue pas tout seul… Et c'est ici surtout toutes les personnes bienfaisantes au moral en dehors du labo que j'aimerai remercier énormément.

Pour commencer merci à mes parents, Huguette et André sans qui je ne serai pas là, et certainement jamais arrivé à ce point. Ainsi je leur dédie cette thèse et ces quatre années de travail. Un très grand merci bien sûr à tout le reste de ma famille répartie aux quatre coins de notre beau pays.

Merci ensuite à Géraldine. Merci simplement pour tout ce que tu es et pour ce que tu m'apportes. Je t'aime, et je me réjouis d'effectuer cette belle aventure de la vie à deux à tes côtés, pour très longtemps. Et merci à toute la famille Dayer pour leur gentillesse et leur soutien : Hildegard, Jean-Marc, Johanna, André, Coralie, et … !

Merci enfin à la « success team » des biologistes et à tou(te)s les fidèles ami(e)s: Chantal, Judith, Caro, Séverine, Fabi, Virginie, Aline, Stéphanie, Barbara, Zully, Elo, Gaëlle, Corine, Vivie, Erica, Saskia, Joelle, Alex, Adri, Fabien, Claude, Raghav, Yvan, Blaise, René, Lulu, Ju, Stéphane, Urs, Tof, Steph, Dave.

GENERAL INTRODUCTION

1. Studying the history of species through space and time: European phylogeography as a case study

Over geologic time, climatic changes have continuously modified landscapes and molded diversity on our planet. Such climatic modifications are essentially due to interactions between orbital and tectonic forces, generally referred as the Milankovic cycles (Hays et al. 1976). Our climate began to cool about 60 Mya, and since the arctic ice cap was formed about 3 Mya, orbital oscillations involving regular eccentricities in the Earth's orbit around the sun appear to have been the main driving forces modifying climate. These periodic major climatic modifications through the Pleistocene (last 2-3 My) along with current Global Change, are key forces shaping genetic diversity at the intra- and inter-specific levels, leading to migration, extinction, isolation and speciation of populations through every cycle. Pleistocene oscillations were therefore a fundamental factor shaping the biogeographic history and genetic structure of species (Hewitt 1996). Tropical regions, even if less affected by these climatic oscillations, endured repeated shifts in the respective setting and range of tropical forests, savannas and deserts (Fig. 1). But it is under temperate latitudes that Pleistocene oscillations most dramatically modified ecosystems' composition and species richness (Jansson 2003), when landscapes largely get adapted to the fluctuating cycles between glacial and temperate conditions (see Fig. 2 for the particular example of Europe).

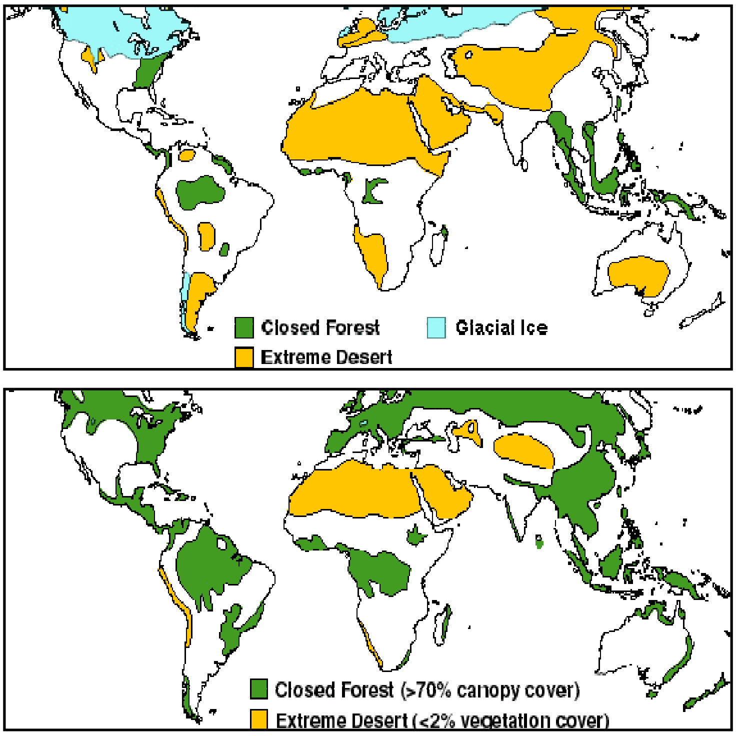


Figure 1: Global coverage of forest and desert at last glacial maximum (above) and actual (below). *Source: http://www.lakepowell.net/sciencecenter /paleoclimate.htm*

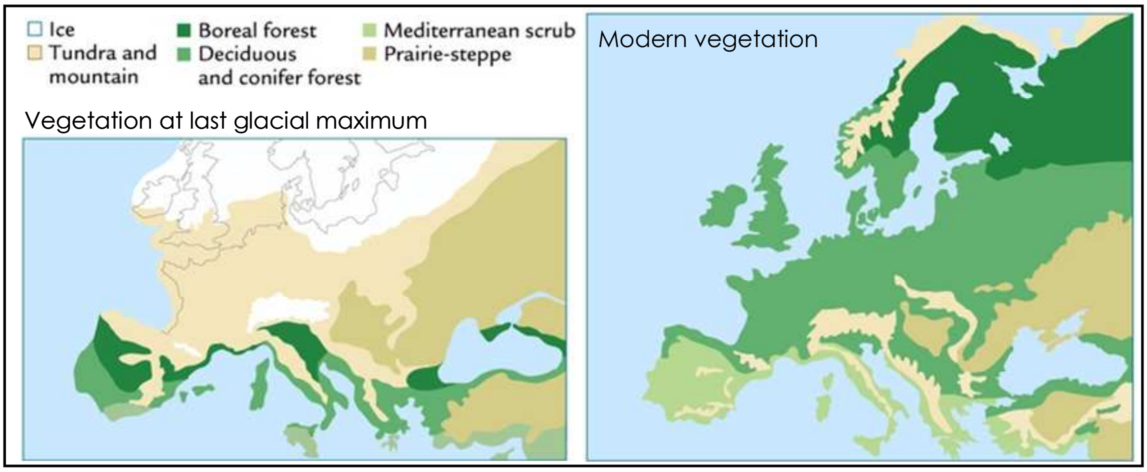


Figure 2: European vegetation modification from last glacial period (left) to actual state (right). *Source: WF Ruddiman, "Earth's Climate, Past and Future" 2nd Edition (modified)*

As a consequence, temperate terrestrial species episodically underwent migrations to the south when climate chilled and to the north when climate heated up again, following the shifts of their suitable ecological conditions. In the Nearctic and eastern Palearctic, such migrations could occur in a fairly continuous way, since only few geographic barriers obstruct the movement of populations. This is not the case of the West-Palearctic region (that can be broadly seen as the European continent), in which massive physical barriers impede dispersal of organisms. Main European barriers are represented by the east–west orientation of the Pyrenees, the Alps and the Carpathians, as well as by the Mediterranean Sea that acts as an insuperable border for many species (Svenning 2003). Because migration to warmer places in the south during cold periods was stopped by the Mediterranean Sea, many European species adapted to temperate conditions were thus confined to a few restricted refugia during glacial maxima, when did not definitively disappear. As another consequence for western Palearctic biota, the extinction rates might have been greater than in other regions at same latitudes, because some species might not have been able to reach appropriate ecological conditions. The previous scenario is rather contradictory for species adapted to colder environments. Such species may have undergone very different biogeographical histories from temperate species (Varga & Schmitt 2008), as climatic oscillations have affected them the other way round. Indeed, fossil evidence suggests that in the past, such species could have been more broadly distributed than today. In warm interglacial periods (like the present) these species are then restricted to remote environments, because their refugia are zones at high altitudes and latitudes that are mainly extensively covered by ice-shields during the glacial periods (but currently available).

Thanks to numerous studies performed on European plants, vertebrates, insects and other invertebrates, and to the rapid advance in molecular techniques, the main phylogeographic patterns (i.e. refugia for both mountain and temperate species, principal re-colonization pathways and lineages' contact zones) are currently well described for the western Palearctic (Stewart et al. 2010). Recognized refugia during Last Glacial Maximum (LGM) for temperate species located in the peninsulas of southern Europe (Fig. 3) are progressively more precisely defined (Médail & Diadema 2009). Even ancient refugia probably isolated since the Pliocene have been found in several animal species in Iberian and Balkan peninsulas (e.g. Paulo et al. 2001). Concerning species adapted to colder environments, several studies show the existence of central and peripheral refugia within the main European mountain ranges (several examples are reviewed in Schönswetter et al. 2005 and Schmitt 2009). It has been proved that re-

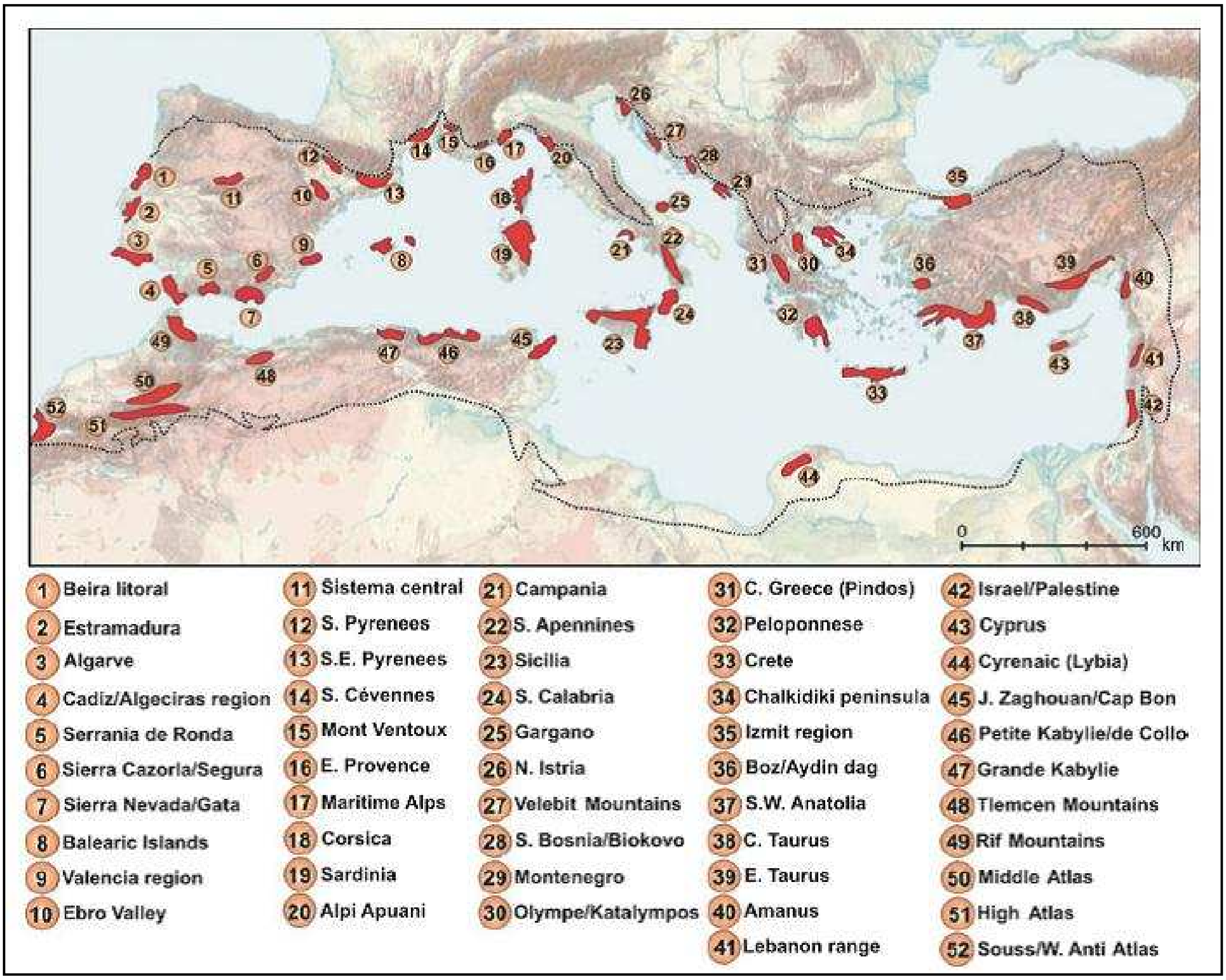


Figure 3: Fine-scale geographical localization of putative Mediterranean refugia. *Source: Médail & Diadema 2009*

colonization routes could be drastically different, depending on the respective ecologies and dispersal abilities of studied organisms to colonize suitable areas (Alvarez et al. 2009), but also because the location of refugia varies according to each species' ecology (Schönswetter et al. 2005). Based on fossil and genetic evidence, three major paradigm patterns (see Fig. 4) of European postglacial re-colonization for temperate terrestrial species (vertebrates, insects and plants) have been described (Hewitt 2000, 2001, Taberlet and Cheddadi 2002). As a consequence, two populations of a particular species may nowadays be very close geographically, and exhibit a very ancient divergence. Conversely, current geographically distant populations could show very little genetic difference. After re-colonization from distinct refugia, cryptic suture zones can emerge between diverging genomes. Whereas some populations interbreed in suture zones, others do not, and two lineages of the same species may not interbreed over several glacial cycles. These movements will have restructured the genomes of species and produced the genetic and species diversity found today in Europe.

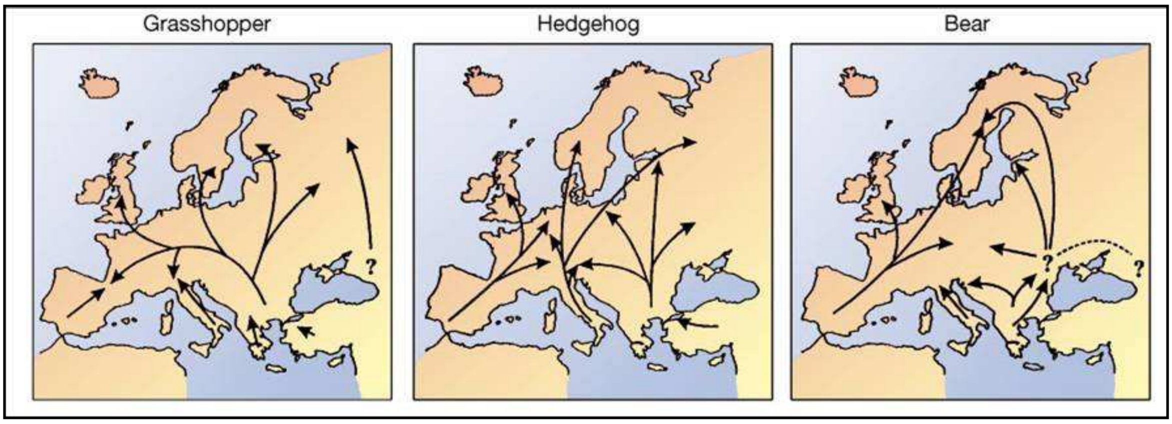


Figure 4: Three paradigm patterns of postglacial re-colonization from southern European refugia, represented by the grasshopper *Chorthippus parallelus*, the hedgehog *Erinaceus europeus* and the bear *Ursus arctos*. Each pattern illustrates the respective influence of southern peninsulas in the actual composition of genetic lineages across the continent. *Source: Hewitt 2000*

2. Investigating the evolutionary history of closely-related lineages

Delimitating taxa in closely related organisms and trying to explain their biogeographical history have long been important objectives of systematic biology. With the emergence of modern molecular methods over the two last decades, the classical species' concept mostly based on distinguishable morphological and/or ecological characteristics was sometimes challenged by incongruent phylogenetic views on species or species boundaries. In several recent phylogeographic studies, the authors had to handle intricate scenarios of extinction and re-colonization through the Quaternary in order to explain the current spatial genetic structure of populations (e.g. Spooner & Ritchie 2006, Böhme et al. 2007, Milankov et al. 2008). Sometimes very high levels of intra-specific differentiation were found regionally in a few populations. On the opposite, only tiny genetic distances were occasionally detected among several species well recognized under morphological or ecological concepts. Consequently many species were rather categorized as species complexes or species groups. Such closely related organisms are expected to regularly experience genetic exchanges with their relatives (via hybridization and progressive introgression). Even if current important geographical distance can impede such genetic events, we must consider the recurrent Pleistocene movements of populations and try to find out if genetic connections have occurred, for instance, when closely related species survived altogether in shared refugia.

Despite little studied until now, such species groups and complexes are expected to be rather common, and maybe even more frequent than single and well distinguished species whatever criteria are considered. Moreover, as phylogeographic studies generally consider large geographic areas, the probability is high to encounter isolated relatives, close to the focal species (even if restricted to a particular region). One might actually have sampled a species group, but misleadingly regards it as one single taxon during analyses. Considering only one species, or one sub-group, within a whole group of closely related species might lead to an incomplete understanding of the group's biogeographic history. It is now important to consider the full range of taxa included within a particular species group and establish their phylogeography taking into account these close genetic relationships. The phylogeographic information from each of the members might bring the necessary complementary information in order to optimize the comprehension of the group's history and evolution.

3. Investigating the phylogeography of ecological interactions

It can be expected that climatic oscillations also strongly influence ecological interactions and species co-evolution. Up to now studies have almost exclusively addressed the phylogeographies of single species, or of closely related members within a species group. The next step is therefore to understand how glaciations cycles molded the pattern of specific interactions among some of these European organisms. Interactions generally take place when a species strongly depends on another, as in the specific cases of antagonism and mutualism. Up to now, one single study comparing the intra-specific spatial genetic structures of mutualistic organisms was performed on a very special case of digestion mutualism between the South African carnivorous plant *Roridula* and its associated bugs *Pameridea* (Anderson et al. 2004). In Europe, a few specific interactions between plants and insects exist within the category of pollination mutualisms. However in such mutualisms, exclusive one-to-one relationships (that are very uncommon in general) are even probably absent from temperate Europe. Hence, the best examples of very close mutualistic interactions under our latitudes comprise more than two actors. They generally include a few effective mutualistic species that, under particular conditions, keep interacting with other associated organisms either neutral or even parasitic for the whole interaction (e.g. the *Trollius* / *Chiastocheta* nursery mutualism; Despres & Jaeger 1999).

Due to the Pleistocene history of the western Palearctic, these mutualisms are more threatened during every glacial cycle than independent organisms, as the presence of a species is strongly linked to its partner's history (and vice versa). As a preliminary hypothesis, we could expect to detect at least one shared refugia for all members of the partnership. Indeed, if the previous assessment is not verified, it would mean that current mutualisms are very young, a very unlikely theory in view of the substantial adaptations (mainly morphological but also ecological) that evolved in the actors of a given specific interaction. Other expectations could, for instance, suppose identical re-colonization pathways, even if here one should also consider the dispersal abilities of each individual species. Finally, by comparing the respective spatial genetic structures of mutualistic species, we would be able to identify if glaciations cycles strongly influenced the pattern of specific interactions, and how. But such a comparative approach would also give us a better idea about the strength of a particular mutualism to overcome climate's regular perturbations, during past ice ages, and that additionally may be informative in the perspective of future climate changes.

Optimally investigating the three main topics addressed above (phylogeography, closely-related lineages and ecological interactions) would have been impossible without considering several types of organisms. The same kind of research carried out on different European species permits to treat either vegetal or animal kingdoms, temperate or alpine organisms, closely-related or well defined species, and finally independent, antagonistic or mutualistic interactions. All case-study plants and insects I investigated during my PhD are presented in details below. A few keywords are given in order to quickly introduce the topics targeted through the corresponding organism(s).

4. An alpine leaf-beetle species group as first case-study

Keywords: Large-scale phylogeography, species group, species concept, alpine habitat

In **chapter I** we attest for the efficiency of simple gene sequencing in resolving two main biological questions in an alpine leaf-beetle species group. First, the species concept of the recently evolved *O. speciosa* / *O. alpestris* species group is widely discussed. Based on the comparison of our results with classical consideration of described species, we address the incongruence between morphological and phylogenetical species concept concerning the three taxa *O. speciosa*, *O. ganglbaueri* and *O. alpestris*. Second, main phylogeographic patterns of this group of sister leaf-beetles are described, with an emphasis on key-refugia zones in Europe, genetic links among mountains ranges and crossroads of lineages in the Alps as well as neighboring massifs. Both aspects (species concept and phylogeography) proved the importance of combining complementary genetic information from different genomic regions (here mitochondrial and nuclear DNA), in order to resolve the evolutionary history of such recently diverged taxa.

4.1 The leaf-beetle genus *Oreina*

The phytophagous genus *Oreina* Chevrolat, 1837 (Coleoptera, Chrysomelidae) synonymous of *Chrysochloa* Hope, 1840 or *Orina* Weise, 1884, and including 28 species (Kippenberg 1994, 2008; pers. comm.), is found throughout the mountain regions of Europe, with some incursions into the lowlands, and western Siberia. Most species are oligophagous on either Apiaceae or Asteraceae. Four species feed on both families and have been considered as polyphagous, whereas monophagy seems scarce but does exist (*O. gloriosa* on *Peucedanum ostruthium* for instance) (Jolivet et al. 1986). The reproductive biology of *Oreina* represents a case of continuum between oviparity and viviparity (Dobler et al. 1996, Dobler & Rowell-Rahier 1996). In the same genus we find (i) species that are oviparous (e.g. *O. elongata*), (ii) species that incubate the eggs inside the mother and lay eggs containing embryos at different stages of development (facultative viviparity as for *O. luctuosa*), and (iii) species that give birth to completely developed larvae which have apparently been nourished inside the mother (e.g. *O. speciosa*, *O. alpestris* or *O. gloriosa*). *Oreina* species are generally considered to be non-flying leaf-beetles despite possessing completely developed red-coloured wings. Even though some cases of flight have been recorded in a few species, these are usually rare and seem to be more an exception than the general rule.

Chemical defence, which is a widespread strategy in the Chrysomelidae, is common to all *Oreina* species except one (*O. melanocephala*, Dobler et al. 1996). Adults use active secretion via exocrine glands situated on the pronotum and the elytra, whereas larvae stock chemicals in their body. All species feeding on Apiaceae synthesise de novo cardenolides from ubiquitous plant sterols. A few *Oreina* species that feed on Asteraceae secrete both cardenolides and pyrrolizidine alkaloid N-oxides (PAs) which are sequestered from their food plants. Finally, one species (*O. cacaliae*) secretes only PAs and has lost the capacity to produce cardenolides (Pasteels et al. 1992, Rowell-Rahier & Pasteels 1994, Dobler et al. 1996).

Several species are morphologically variable at the inter-, as well as the intra-population level and may present colour polymorphism. Consequently they have been divided into a number of subspecies and forms, some of which are restricted to limited geographic areas. Caryological studies show that the formula 2n = 24 chromosomes is uniform in the genus (Petitpierre 1999). Due to their morphological similarities, identification of species is often difficult. The safest characteristics for determination are the male genitalia, which are well documented and illustrated (Freude et al. 1966), but that may also vary along a morphological continuum within the same species.

4.2 The *Oreina speciosa* / *Oreina alpestris* species group

This species group is composed of two closely related species that proved to share many morphological and genetic features within the genus (Dobler et al. 1996). Considering their respective large European distributions (see below) as well as alpine species ecologies, we expected this system to reveal interesting patterns of isolation and migration across the different European mountain ranges, thus providing an ideal case-study within the framework of alpine phylogeography applied to a species' group.

Oreina speciosa (Linnaeus 1767) measures 9-13 mm. with prothorax and elytra coloration ranging from green to blue, violet or black, sometimes with red or orange bands (see Fig. 5). Color patterns are very variable in this species, at the inter-population level as well as within populations. Females are generally a little larger than males. In the field, some morphological external criteria (shapes of maxillary palps, prosternum and anal sternite) help in recognizing *O.*

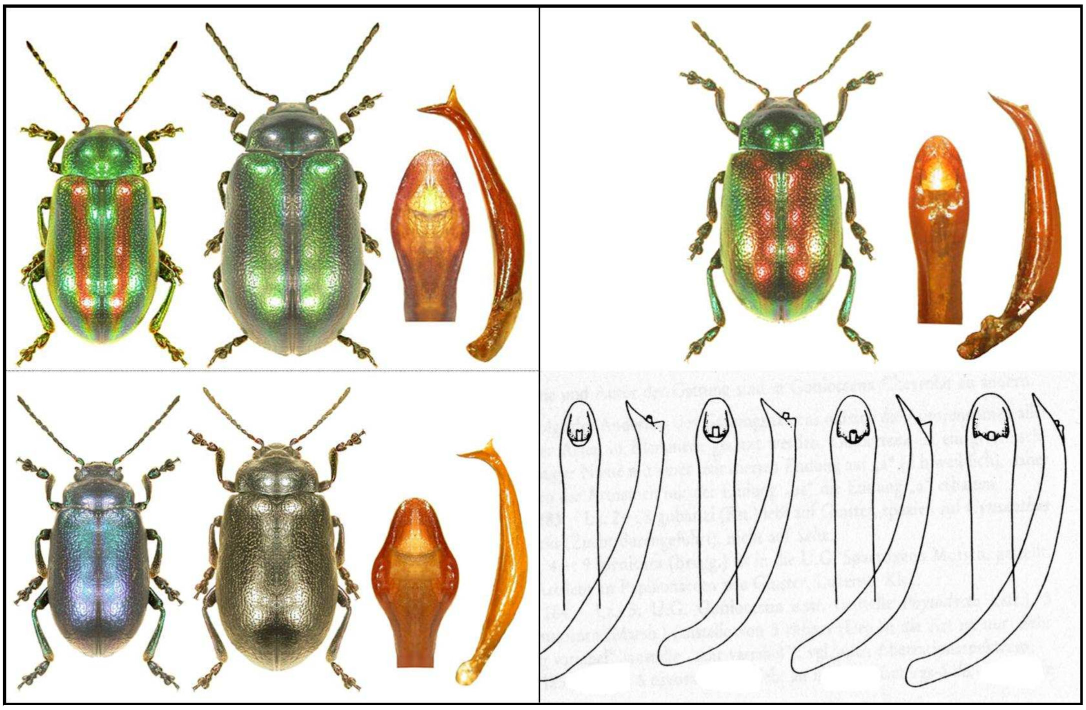


Figure 5: Shape, coloration and genitalic morphology of O. speciosa ssp speciosa (above left), O. speciosa ssp ganglbaueri (below left) and O. alpestris (right), also showing the genitalic variation within this species. *Source: http://www.biol.uni.wroc.*

speciosa from other *Oreina* species. However, this discriminating trait easily leads to confusion when compared to its sister species *O. alpestris*. *Oreina alpestris* (Schummel 1843) is a 7-13 mm. long species showing also highly variable coloration (pers. obs.). However, in the middle of Europe, individuals within a population seem to be homogenously coloured (Freude et al. 1966).

In addition of being highly similar in terms of coloration and morphology, the two species share many other characteristics including life cycle, ecology and distribution (see below) and can often be found together in sympatric populations. In such cases, determination of the correct taxon relies mostly on examination of aedeagi. Even if those are more similar between both species than compared to other *Oreina*, each species possesses characteristic male genitalia (Fig. 5). Thus in most cases and at least in males, this criterion makes the determination between *O. speciosa* and *O. alpestris* reliable. *Oreina speciosa* is identified by one single type of genitalia. In contrast, *O. alpestris* represents one case of leaf-beetle in which the shape of aedeagi varies (as presented on Fig. 5 right down). Consequently, extreme types of genitalia belonging to *O. alpestris* resemble the type of *O. speciosa*.

Oreina speciosa is distributed across most of the mountain regions of Europe (see Fig. 6). It is found throughout the Alps, but also in the mountains of the Balkan Peninsula as far as Montenegro, in the Vosges, in the French Massif Central and in the Jura. As many other *Oreina*,

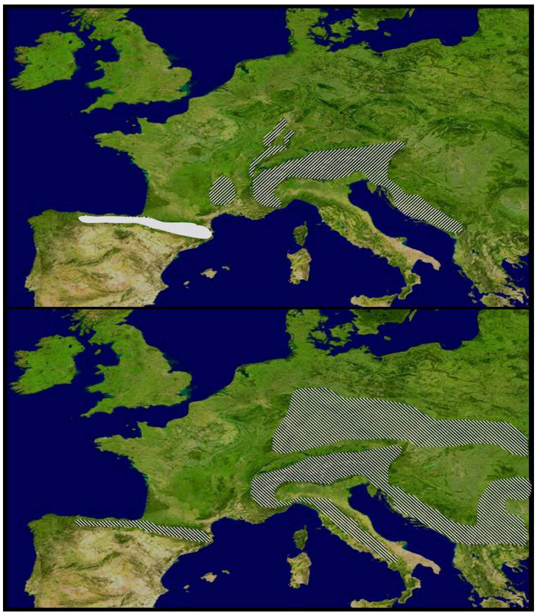


Figure 6: Distribution area of *O. speciosa* (above) and *O. alpestris* (below). The white area on the upper map represents the range of *O. speciosa* ssp. *ganglbaueri*.

O. speciosa is divided into subspecies with distinct distribution areas. The subspecies *O. speciosa* ssp. *ganglbaueri* Bechyné 1958, distributed in the Pyrenees and Cantabrics mountains demonstrates a discussed status as it is sometimes considered as a distinct species (*O. ganglbaueri* Jakob 1953), notably on the basis of its larger size and different type of aedeagus (also larger, see Fig. 5). The European distribution of *O. alpestris* mostly covers that of *O. speciosa*, but is larger extending eastwards to the Carpathian Mountains, south-east to Macedonia and Bulgaria and southwards into Italy (Abruzzi). Here again the forms of different regions are split into subspecies.

Oreina speciosa and *O. alpestris* are oligophagous, feeding on several species of Apiaceae. Although *O. alpestris* presents a slightly smaller number of suitable host-plants, the following plants are reported as the main host plants for the species group throughout Europe (Jolivet et al. 1986, Freude et al. 1966): *Aegopodium podagraria*, *Angelica sylvestris*, *A. razulsi*, *Anthriscus* sp., *Apium nodiflorum*, *Chaerophyllum aromaticum*, *Ch. hirsutum*, *Ch. villarsii*, *Heracleum sphondylium*, *H. pyrenaicum*, *Laserpitium latifolium*, *Meum athamanticum*, *Peucedanum ostruthium*, but also *Seseli libanotis* (pers. obs.). Typical habitats containing these host plants can be found at altitudes from about 1000 to 2000 m. a.s.l. either beside or in the understorey of fir, larch or beech forests, but also in megaphorbs, typical mountain cold and humid grasslands with high vegetation. As both species share most host plants, they logically can be found in sympatry in many populations.

Larvae of both species are darkly colored during the first instar and become light brown to whitish at the second instar. After preliminary observations on *O. speciosa*, we supposed that the life cycle extends over two years (unpublished data). In the first year, the viviparous females lay larvae that grow until the end of summer. These probably bury in the soil to overwinter as incompletely developed larvae. In the spring of the second year the incompletely developed larvae finish their development on host plants and metamorphose during summer. The pupae directly hatch as adults that do not reproduce during their first summer.

5. A temperate orchid species group as second case-study

Keywords: Large-scale phylogeography, species group, evolutionary significant units, conservation biology, endemics lineages, temperate habitat

Chapters II and III report a second example of lineages interactions within members of a species group, but in a temperate climate context. Both chapters are dedicated to the Fly orchid group, composed of three recognized species (either widespread or endemic) that share most of their genetic components and are thus described on the basis of morphologic and ecological features. **Chapter II** represents a necessary preface to further phylogeographic work, and establishes the genetic identity and circumscription of each species (that was not clearly defined until now) with both gene sequencing and fine-scale AFLP genotyping. In addition, the same genetic data allow determining patterns of genetic diversity at the level of the species group. Results attest for unexpected high values of diversity in other regions than those enclosing endemics and for rare cpDNA haplotypes occurring in some easternmost populations. In addition, the study shows that endemics are sharing similar haplotypes as the widespread species. The discussion is focused on defining evolutionary significant units in the fly orchid group and on identifying several areas of interest for further conservation strategies. **Chapter III** establishes general phylogeographic patterns for the whole species group (in view of previous results revealing strong genetic similarity), using AFLP data combined to ecological niche modelling at the last glacial maximum (LGM). The paper discusses, in the light of the combined approach, the results obtained notably concerning two particular endemic genetic lineages. Generally accounting for a large admixture within populations, the discussion shows that these endemic lineages could either relate to the Quaternary history of the group, or to biotic and abiotic ecological properties of newly emerged endemic species. The conservation and taxonomical status of endemic species are then further examined.

5.1 The European genus *Ophrys*

The genus *Ophrys* L., placed in the subfamily Orchidoideae, tribe Orchideae, subtribe Orchidinae, was first described by Carl von Linné (1753, Species Plantarum: 945). This Euro-Mediterranean genus occurs primarily around the Mediterranean Basin. The northern border of the distribution lies in central Scandinavia, and the southern border is formed by the Sahara desert in northern Africa. The most western occurrence is reported from the Macaronesian

Islands, and the most eastern observations are from Iran. Areas of high species' diversity are found in southern Italy and in the Aegean region (Soliva et al. 2001). Even if most species described are localized and/or endemic to small islands, peninsulas or even isolated mountains ranges, some other present a larger European distribution, mostly linked to the Mediterranean or Atlantico-Mediterranean climatic regions (e.g. *O. bombyliflora*, *O. lutea*, *O. tenthredinifera*). Only few species present a distribution centered on middle Europe. However, these are reaching northernmost borders of the continent, as far as Scandinavia and British Isles (e.g. *O. apifera*, *O. insectifera*, *O. fuciflora*, *O. sphegodes*).

Ophrys are especially famous for being sexually deceptive orchids that attract pollinators (male Hymenoptera mostly from the Apoidea and Sphecoidea superfamilies) with odorant and morphological clues. After the first mention by Pouyanne (1917), many studies have reported similarities between the female insect and the plant flower mimic (e.g. Kullenberg 1961, Borg-Karlson 1990, Hermosilla et al. 1999). Some of these are morphological such as pilosity, shape and size of labellum (e.g. Kullenberg 1961, Vereecken 2009). Other similarities are odorant, as shown by volatile substances implicated in the attraction of males resembling those emitted by female bees or wasps (Borg-Karlson et al. 1987, Ayasse et al. 2001, Schiestl 2005). Pollination is hence ensured by these lured males during pseudocopulation events with fake females. But as a consequence that male pollinators may learn to recognize such traps, variability in floral signals is important. As pollinators learn more quickly to avoid common deceptive morphs than rare ones, morphological variability may be maintained by negative frequency-dependant selection (Ayasse et al. 2000, Schiestl 2005).

Due to morphological variability, species number and nomenclature within the genus (mostly based on morphological studies) has evolved very quickly during the last century (Vela et al. 2007). At the subgeneric level, *Ophrys* species have been since long divided into two distinct sections with different pollination biology. Section *Pseudophrys* (presumably the basal *Ophrys* species) present an abdominal pollination type (head-down position of the pollinator during pseudocopulation), whereas species of *Euophrys* section show a cephalic pollination type (head-up position of pollinator). This partition has however proved to be non-supported phylogenetically, as section *Euophrys* may be paraphyletic (Bateman et al. 2003, Devey et al. 2008). Pollinator's specificity has also been the main motive for species description, assuming the basic postulation that each *Ophrys* species possesses its own specific, dedicated pollinator. From authors such as Sundermann (1980) who recognized only 16 species plus 34 subspecies, to recent orchidologists who described as many as 252 species forming 32 complexes (Delforge

2001), the taxonomy of *Ophrys* can be considered as confusing (Soliva et al. 2001). Studies using optimized molecular methods to resolve the detailed phylogeny and species' delimitation of this monophyletic genus arose only in the very beginning of the 21st century (Soliva et al. 2001, Bateman et al. 2003, Devey et al. 2008). The most recent by Devey et al. (2008) generally concluded that "the genus has been substantially over-divided at the species level". So there is an actual strong conflict between "splitter" orchidologists considering clues such as fine-scale morphology, pollination biology and/or geography, and those basing their assumptions on molecular methods, tending more for a reunification into aggregates or groups (i.e., they can be considered as "lumpers").

Figure 7: The three members of the *Ophrys insectifera* species group: *Ophrys insectifera* (left), *O. aymoninii* (centre) and *O. subinsectifera* (right). *Pictures Y.Triponez*

5.2 The *Ophrys insectifera* group, or Fly orchid group

The *Ophrys insectifera* group consists in three morphologically closely related species (Fig. 7), all of them morphologically and genetically very distinct from the other species of the genus (Delforge 2001, Soliva et al. 2001, Devey et al. 2008). In the past, the three species were for long considered as a single one (*O. insectifera* s.l.) with some authors distinguishing the three different ssp. *insectifera*, *aymoninii*, *subinsectifera* (and some others, today rather considered as varieties). Currently *O. aymoninii* and *O. subinsectifera* are accepted as distinct species from *O. insectifera*. One of the crucial reasons for this splitting is the different type of pollinator observed for each previous subspecies of *O. insectifera*. *Ophrys insectifera* L., the representative species of the group, is considered as one of the most famous *Ophrys* species and the one with the largest distribution (Wolff 1951, Delforge 2001) (Fig. 8). Several independent data assess that its pollination is regularly assured by *Argogorytes mystaceus* males (Hymenoptera, Sphecidae): 1) many observations of flower's visits and pseudocopulation events are documented (e.g. Wolff 1950, Kullenberg 1951, 1961, and personal observation, see Fig. 9); 2) *O. insectifera* emits a characteristic series of aliphatic hydrocarbons that are very similar to that produced in the

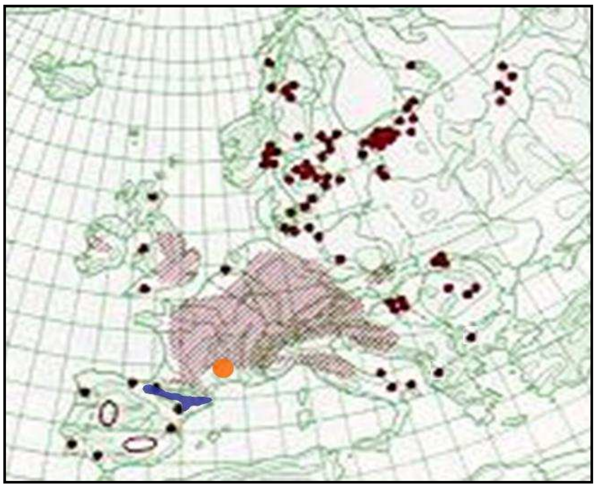


Figure 8: Distribution of *O. insectifera* (in red and dots), *O. aymoninii* (in orange) and *O. subinsectifera* (in blue). *Source: http://linnaeus.nrm.se (modified)*

Figure 9: Two males (!) *A. mystaceus* pseudocopulating on *O. insectifera*, Cudham, UK. *Picture Y. Triponez*

Dufour's gland of *Argogorytes* spp. females (Borg-Karlson et al. 1987); 3) some of these hydrocarbons release strong electroantennogram responses when tested one by one on *Argogorytes* male antennae (Ågren & Borg-Karlson 1984); 4) extractions of *O. insectifera* aliphatic alcohols showed high degree of attraction for *Argogorytes* spp. when tested into the field (Borg-Karlson et al 1993). The second species of the group, *O. aymoninii* BREISTROFFER is restricted to a tiny distribution zone covering the karstic southern French Massif Central (Breistroffer 1981, Delforge 1983). Compared to *O. insectifera*, *O. aymoninii* has got flowers with a wider and yellow edged labella and green-yellowish petals. It has been for long considered as a subspecies of *O. insectifera* (*O. insectifera* ssp *aymoninii*) but was elevated to the rank of species after reported observations of pseudocopulation events by the bee *Andrena combinata* (Hymenoptera, Andrenidae). Borg-Karlson et al. (1993) also showed that *A. combinata* was strongly attracted to aliphatic alcohols emitted by *O. aymoninii*. The last species of the group, *O. subinsectifera* HERMOSILLA & SABANDO, is a Spanish vicariant growing in the south-western foothills of the Pyrenees (Amardeilh 1996, Hermosilla & Sabando 1996). Here as well, it looks quite alike *O. insectifera*, differing from the latter by smaller flowers, reduced lateral lobes of labella and petals with a greenish tip. Only recently, the pollinator of *O. subinsectifera* has been described by Hermosilla et al. (1999) as the sawfly *Sterictiphora gastrica* (Hymenoptera, Argidae).

6. A European pollination mutualism as third case-study

Keywords: Large-scale phylogeography, comparative phylogeography, pollination mutualism, temperate habitat

The approach combining AFLP-based spatial genetic structure and past ecological niches modeling demonstrates once again to be powerful in revealing crucial phylogeographic features in **Chapter IV**. Here the particular case of a mutualistic interaction between plants and insects is investigated towards a comparative phylogeography approach. Similarities and differences between phylogeographic patterns are compared throughout the western Palearctic in the different partners involved in this mutualism: the oil-producing plant *Lysimachia vulgaris* and its oil-collecting bees pollinators *Macropis* sp. This attempt to characterize conjointly the biogeographic features of different mutualistic organisms is the first of its kind done in Europe. As initially hypothesized, the comparative phylogeographic approach resulted in an important overlapping of potential refugia during LGM between the plant and its pollinators. However the current spatial genetic structure does not allow concluding that re-colonization pathways were similar. These results are discussed in view of the different dispersal abilities and ecological preferences of the actors. More conclusions could also be drawn in parallel with a complementary paper (**Appendix 1**) that brings to light new insights in the comprehension of the evolutionary ecology of the two main *Macropis* pollinators in Europe.

6.1 The oil-producing plant *Lysimachia vulgaris* (yellow loosestrife)

Lysimachia was traditionally placed in tribe Lysimachieae within the Primulaceae family (Pax and Knuth, 1905). Only recently it was demonstrated that *Lysimachia* and the other members of tribe Lysimachieae belonged to the Myrsinaceae (Källersjö et al. 2000), a rather large family (35 genera and about 1000 species) from the order Ericales. The genus *Lysimachia* includes 191 species found mainly in temperate and subtropical areas of the Northern Hemisphere (Hao et al. 2004), with a few species spanning the Neotropical region and Africa. China shows the highest concentration of species, with 122 *Lysimachia* species inventoried out of which 110 are endemic (Chen & Hu 1979). Depending on authors, mainly Chen & Hu (1979) for Eurasia and Ray (1956) for the New World, *Lysimachia* is subdivided into eight subgenera (Hao et al. 2004). Data sequencing from either nuclear or chloroplastic genomes showed that *Lysimachia* is clearly paraphyletic (Hao et al. 2004, Anderberg et al. 2007).

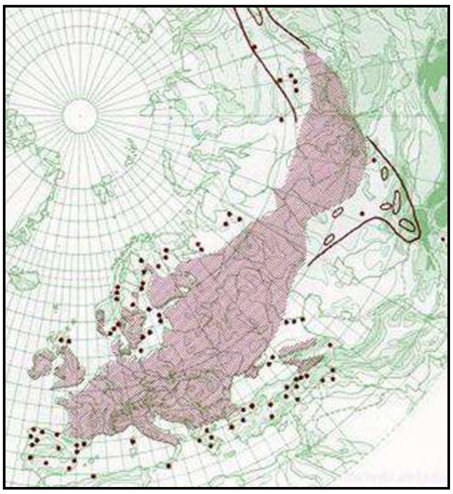


Figure 8: Distribution of *L. vulgaris* in Eurasia. *Source:* ***http://linnaeus.nrm.se***

Among the 13 species of *Lysimachia* present in Europe (out of which ten are native), *L. vulgaris* L. is the most frequently found and the one with the largest distribution. *Lysimachia vulgaris* is a rhizomatous or stoloniferous perennial plant native of temperate Eurasia. It is a 50-150 cm tall plant with erect panicles of conspicuous yellow flowers blooming from June to September. It is found growing wild throughout most of Europe and Asia (as shown in Fig. 10), especially in moist areas, grasslands and near streams and waterways. Plants from the easternmost of its distribution (N-E China and Japan) are recognized as *L. vulgaris* var. *davurica* (also called *L. davurica* LEDEB.). *Lysimachia vulgaris* was naturalized in Canada and in the North of the USA (Britton & Brown 1913) where it is regarded as a serious weed pest. Rare records from the East Coast of Australia (Kodela 2006) and Westland of New Zealand's South Island (Sykes 1981) are attributed to introductions from var. *davurica*.

Lysimachia vulgaris represents one particular example of the estimated 1500 to 1800 species of oil-offering flowers. This special pollination syndrome is known from 11 different plant families worldwide, although mainly occurring in the tropics and subtropics. Several papers and reviews previously argued for oil rewards to have evolved in angiosperms many times independently (e.g. Buchmann 1987, Vogel 1988, Rasmussen & Olesen 2000). Only very recently phylogenetical evidences confirmed the supposed independent evolutions of oil as a pollinator reward, with at least 28 different events (Renner & Schaefer 2010). In the Holarctic region, oil-offering flowers are all included into the Myrsinaceae family and into the genus *Lysimachia*. Vogel (1986, 1988) found evidence of floral oil in at least 75 species of *Lysimachia*. Here the oil is produced by elaiophores (oil secreting trichomes) located at the basal part of petals and anther's filaments (Simpson et al. 1983). All oil-producing *Lysimachia* are pollinated by females of the oil-collecting bee genus *Macropis* (Michez & Patiny 2005; see below for further details). Arguments in favor of this tight interaction in Europe are exemplified by Vogel (1976) followed later by Hoffmann (2005). These authors noticed that seed set *of L. vulgaris* with potential self-pollination capacity is considerably reduced without insect visitation. Moreover in *L. vulgaris*, more than 90% of the visits were from the oil-collecting monolectic solitary bee *M. europaea* (at least in their respective location of research). The rest of the visits were from small syrphids, which are not responsible for pollination and do not represent alternative efficient pollinators.

6.2 The oil-collecting bees *Macropis* sp.

The bee genus *Macropis* Panzer 1809 (Hymenoptera, Mellitidae) is one of the 14 genera included in Mellitidae (Michener 2000). The 16 species of the genus are subdivided into three subgenera (*Macropis* s. str., *Paramacropis* Popov & Guiglia 1936 and *Sinomacropis* Michener 1981) and are distributed in the Holarctic region only (Michez & Patiny 2005). All *Macropis* species are tightly associated with *Lysimachia* flowers (Cane et al. 1983, Vogel 1976, Michez & Patiny 2005). Females of all bee species display morphological adaptations to oil collecting. The *Lysimachia* oil is used by *Macropis* females for cell walls coating as well as for larval provisions when mixed with the pollen.

Three species of *Macropis* can be found in the western Palearctic ecozone: *Macropis europaea* Warncke 1973, *M. frivaldszkyi* Mocsary 1878 and *M. fulvipes* Fabricius 1805 (Michez & Patiny 2005). Among them, *M. frivaldszkyi* shows the smallest distribution, limited to Balkans, Anatolia and Syria. *Macropis europaea* has a medio-european distribution (Fig. 11), even if it was

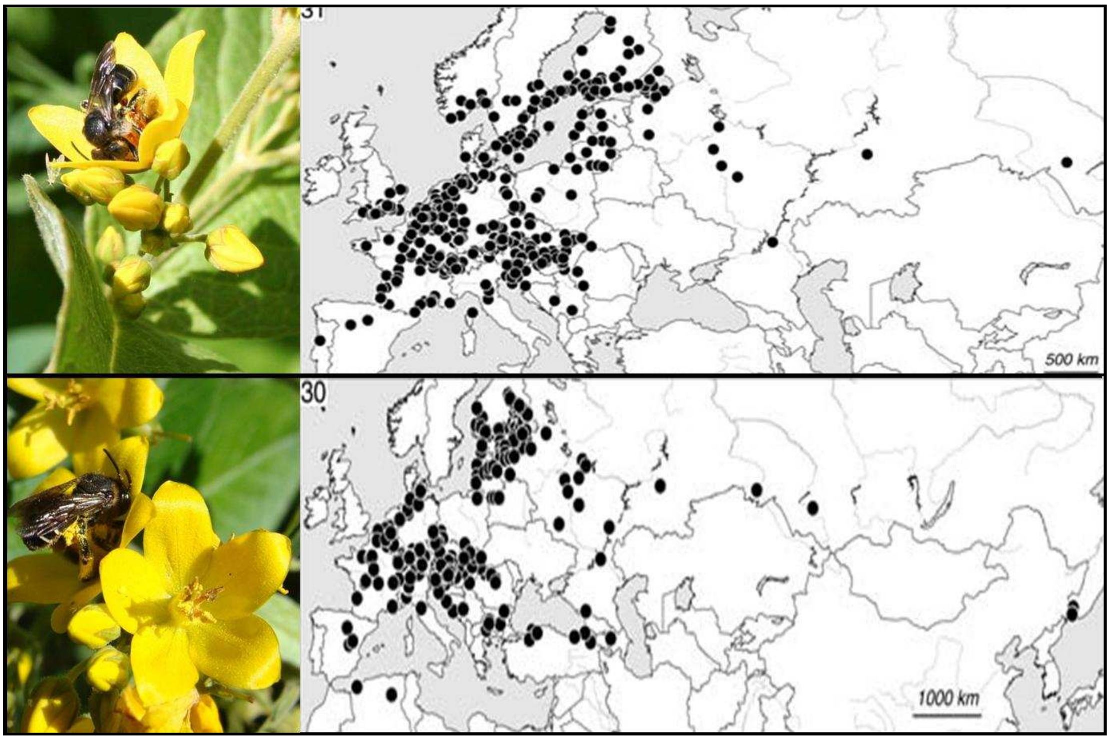


Figure 9: Typical oil-collection behaviour for *M. europaea* (above) and *M. fulvipes* (below) on *L. vulgaris*. Respective distribution ranges are displayed on the right. *Source: Michez & Patiny 2005. Pictures L. Bassin and Y. Triponez*

recorded further east (Barnoul, Russia). To the north, *M. europaea* reaches 65°N in Finland. It is the only species distributed in the British Islands and in the Scandinavian Peninsula (i.e. Sweden and Norway). To the south, its distribution extends in Spain, in Italy and in the Balkans as far as Macedonia and Bulgaria, but is absent from Anatolia and Caucasus. *Macropis fulvipes* has a distribution comparable to that of *M. europaea* although presenting some differences. First, *M. fulvipes* is absent from the British Isles and from Scandinavia (except Finland). Second it is widely distributed in Anatolia and Caucasus. Finally, it was also recorded from Maghreb (Algeria) and from far-eastern Russia (Ussuriysk).

In Europe, *M. europaea* and *M. fulvipes* can be found together in sympatric populations (Celary 2004, Pekkarinen et al. 2003, pers. obs.). Both species display a high level of similarity in their morphology, ecology and floral preferences. A comparative study conducted in southern Poland (Celary 2004) showed that all features concerning the nest's architecture and the larval stages differ almost only in size (slightly bigger for *M. fulvipes*). The most obvious difference between the two species stands in their respective phenologies, with a flight period delay of more or less ten days (earlier in *M. fulvipes* and later in *M. europaea*). This delay in the peak activity of both males and females from these two protandrous species was conjointly observed by different authors in different parts of Europe. Pekkarinen et al. (2003) hypothesized that such a delay in flight periods might be meaningful for resources partitioning in sympatric populations, especially in places where only one *Lysimachia* species is available. When looking at the floral preferences in the bees, all previous records have reported *L. vulgaris* to be the main (and maybe the only) resource of oil for *M. europaea*. Regarding *M. fulvipes*, the species was additionally observed on *L. nummularia* (sometimes presumed as its preferred plant species [Michez 2002]) and on *L. punctata*. In Anatolia and Caucasus, numerous populations of *M. fulvipes* were observed on this latter *Lysimachia* species (pers. obs.), which is naturally growing in mountainous areas around the Black Sea but currently largely cultivated (and sometimes naturalized) in central and western Europe.

References

Ågren L, Borg-Karlson AK. 1984. Responses of *Argogorytes* (Hymenoptera: Sphecidae) males to odor signals from *Ophrys insectifera* (Orchidaceae). Preliminary EAG and chemical investigation. *Nova Acta Regiae Societatis Scientiarum Upsaliensis* **Serie V:C:** 111-117.

Alvarez N, Thiel-Egenter C, Tribsch A, Holderegger R, Manel S, Schonswetter P, Taberlet P, Brodbeck S, Gaudeul M, Gielly L, Kupfer P, Mansion G, Negrini R, Paun O, Pellecchia M, Rioux D, Schupfer F, Van Loo M, Winkler M, Gugerli F. 2009. History or ecology? Substrate type as a major driver of patial genetic structure in Alpine plants. *Ecology Letters* **12:** 632-640.

Amardeilh J-P. 1996. *Ophrys subinsectifera* Hermosilla & Sabando. Une nouvelle orchidée du nord de l'Espagne. *L'Orchidophile* **123:** 149-154.

Anderberg AA, Manns U, Källersjö M. 2007. Phylogeny and floral evolution of the *Lysimachieae* (Ericales, Myrsinaceae): evidence from ndhF sequence data. *Willdenowia* **37:** 407-421.

Anderson B, Olivieri I, Lourmas M, Stewart BA. 2004. Comparative population genetic structures and local adaptation of two mutualists. *Evolution* **58:** 1730-1747.

Ayasse M, Paxton RJ, Tengo J. 2001. Mating behavior and chemical communication in the order Hymenoptera. *Annual Review of Entomology* **46:** 31-78.

Ayasse M, Schiestl FP, Paulus HF, Lofstedt C, Hansson B, Ibarra F, Francke W. 2000. Evolution of reproductive strategies in the sexually deceptive orchid *Ophrys sphegodes*: How does flower-specific variation of odor signals influence reproductive success? *Evolution* **54:** 1995-2006.

Bateman RM, Hollingsworth PM, Preston J, Yi-Bo L, Pridgeon AM, Chase MW. 2003. Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). *Botanical Journal of the Linnean Society* **142:** 1-40.

Böhme MU, Fritz U, Kotenko T, Dzukic G, Ljubisavljevic K, Tzankov N, Berendonk TU. 2007. Phylogeography and cryptic variation within the *Lacerta viridis* complex (Lacertidae, Reptilia). *Zoologica Scripta* **36:** 119-131.

Borg-Karlson AK. 1990. Chemical and ethological studies of pollination in the genus *Ophrys* (Orchidaceae). *Phytochemistry* **29:** 1359-1387.

Borg-Karlson AK, Bergstrom G, Kullenberg B. 1987. Chemical basis for the relationship between *Ophrys* orchids and their pollinators. 2. Volatile compounds of *O. insectifera* and *O. speculum* as insect mimetic attractants excitants. *Chemica Scripta* **27:** 303-&.

Borg-Karlson AK, Groth I, Agren L, Kullenberg B. 1993. Form-specific fragances from *Ophrys insectifera* L. (Orchidaceae) attract species of different pollinator genera. Evidence of sympatric speciation? *Chemoecology* **4:** 39-45.

Breistroffer M. 1981. Notes succinctes sur quelques équivalences nomenclaturales d'espèces céveno-caussenardes et description d'une sous-espèce nouvelle d'Orchidée. *Bulletin de la société botanique de France. Lettres botaniques* **128:** 69-72.

Britton NL, Brown A. 1913. An illustrated flora of the northern United States, Canada and the British Possessions. Vol. 2: 711 *USDA-NRCS PLANTS Database*.

Buchmann SL. 1987. The ecology of oil flowers and their bees. *Annual Review of Ecology and Systematics* **18:** 343-369.

Cane JH, Eickwort GC, Wesley FR, Spielholz J. 1983. Foraging, grooming and mate-seeking behaviors of *Macropis nuda* (Hymenoptera, Melittidae) and use of *Lysimachia ciliata* (Primulaceae) oils in larval provisions and cell linings. *American Midland Naturalist* **110:** 257-264.

Celary W. 2004. A comparative study on the biology of *Macropis fulvipes* (Fabricius, 1804) and *Macropis europaea* Warncke, 1973 (Hymenoptera : Apoidea : Melittidae). *Folia Biologica-Krakow* **52:** 81-85.

Chen FH, Hu CM. 1979. Taxonomic and phytogeographic studies on Chinese species of *Lysimachia*. *Acta Phytotaxonomica Sinica* **17:** 21-53.

Delforge P. 1983. Remarques sur *Ophrys insectifera* subsp. *aymoninii* Breistroffer et description d'un hybride nouveau de cette sous-espèce: *Ophrys insectifera* nsubsp. *tytecaeana* Delforge. *L'Orchidophile* **55:** 307-312.

Delforge P. 2001. *Guide des orchidées d'Europe, d'Afrique du Nord et du Proche-Orient, 2nd ed.* Delachaux et Niestlé: Paris, France.

Despres L, Jaeger N. 1999. Evolution of oviposition strategies and speciation in the globeflower flies *Chiastocheta* spp. (Anthomyiidae). *Journal of Evolutionary Biology* **12:** 822-831.

Devey DS, Bateman RM, Fay MF, Hawkins JA. 2008. Friends or relatives? Phylogenetics and species delimitation in the controversial European orchid Genus *Ophrys*. *Annals of Botany* **101:** 385-402.

Dobler S, Mardulyn P, Pasteels JM, Rowell-Rahier M. 1996. Host-plant switches and the evolution of chemical defense and life history in the leaf beetle genus *Oreina*. *Evolution* **50:** 2373-2386.

Dobler S, Rowell-Rahier M. 1996. Reproductive biology of viviparous and oviparous species of the leaf beetle genus *Oreina*. *Entomologia Experimentalis Et Applicata* **80:** 375-388.

Freude H, Harde KW, Lohse GA. 1966. Die Käfer Mitteleuropas. In: Evers G, ed. Krefeld, Germany.

Hao G, Yuan YM, Hu CM, Ge XJ, Zhao NX. 2004. Molecular phylogeny of *Lysimachia* (Myrsinaceae) based on chloroplast trnL-F and nuclear ribosomal ITS sequences. *Molecular Phylogenetics and Evolution* **31:** 323-339.

Hays JD, Imbrie J, Shackleton NJ. 1976. Variations in Earth's orbit: Pacemaker of ice ages. *Science* **194:** 1121-1132.

Hermosilla CE, Amardeilh JP, Soca R. 1999. *Sterictiphora furcata* Villers, pollinisateur d'*Ophrys subinsectifera* Hermosilla & Sabando. *L'Orchidophile* **139:** 247-254.

Hermosilla CE, Sabando J. 1996. Notas sobre orquideas: 2. *Estud. Mus. Cienc. Nat. Alava* **10:** 119-140.

Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58:** 247-276.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907-913.

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology* **10:** 537-549.

Hoffmann F. 2005. Biodiversity and pollination. Flowering plants and flower-visiting insects in agricultural and semi-natural landscapes. University of Groningen.

Jansson R. 2003. Global patterns in endemism explained by past climatic change. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270:** 583-590.

Jolivet P, Petitpierre E, Daccordi M. 1986. Les plantes-hôtes des *Chrysomelidae.* Quelques nouvelles précisions et additions (*Coleoptera*). *Nouvelle Revue Entomologique* **3:** 341-357.

Källersjö M, Bergqvist G, Anderberg AA. 2000. Generic realignment in primuloid families of the Ericales s.l.: A phylogenetic analysis based on DNA sequences from three chloroplast genes and morphology. *American Journal of Botany* **87:** 1325-1341.

Kippenberg H. 1994. Familie Chrysomelidae. In: Lohse GA LW, ed. *Die Käfer Mitteleuropas, 3. Supplementband*. Krefeld, Germany: Goecke & Evers. 65-83.

Kippenberg H. 2008. Revision der Untergattung *Protorina* Weise der Gattung *Oreina* Chevrolat (Coleoptera: Chrysomelidae: Chrysomelinae). *Koleopterologische Rundschau* **78:** 367-418.

Kodela PG. 2006. *Lysimachia* (Myrsinaceae) in New South Wales. *Telopea* **11:** 147-154.

Kullenberg B. 1951. *Ophrys insectifera* L. et les insectes. *Oikos* **3:** 53-70.

Kullenberg B. 1961. Studies on *Ophrys* pollination. *Zoologiska Bidrag Fran Uppsala* **34:** 1-340.

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

Michener CD. 2000. *The bees of the world*. The Johns Hopkinds University Press: Baltimore.

Michez D. 2002. Monographie systématique, biogéographique et écologique des Melittidae (Hymenoptera, Apoidea) de l'Ancien Monde – Premières données et premières analyses.

Michez D, Patiny S. 2005. World revision of the oil-collecting bee genus *Macropis* Panzer 1809 (Hymenoptera : Apoidea : Melittidae) with a description of a new species from Laos. *Annales De La Societe Entomologique De France* **41:** 15-28.

Milankov V, Stahls G, Stamenkovic J, Vujic A. 2008. Genetic diversity of populations of *Merodon aureus* and *M. cinereus* species complexes (Diptera, Syrphidae): integrative taxonomy and implications for conservation priorities on the Balkan Peninsula. *Conservation Genetics* **9:** 1125-1137.

Pasteels JM, Eggenberger F, Rowell-Rahier M, Ehmke A, Hartmann T. 1992. Chemical defense in Chrysomelid leaf beetles: storage of host-derived pyrrolizidine alkaloids versus *de novo* synthesized cardenolides. *Naturwissenschaften* **79:** 521-523.

Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA. 2001. The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London Series B-Biological Sciences* **268:** 1625-1630.

Pax F, Knuth R. 1905. Primulaceae. In: Engler A, ed. *Das Pflanzenreich 22 (IV, 237)*. Leipzig.

Pekkarinen A, Berg O, Calabuig I, Janzon LA, Luig J. 2003. Distribution and co-existence of the *Macropis* species and their cleptoparasite *Epeoloides coecutiens* (Fabr.) in NW Europe (Hymenoptera : Apoidea, Melittidae and Apidae). *Entomologica Fennica* **14:** 53-59.

Petitpierre E. 1999. The cytogenetics and cytotaxonomy of *Chrysolina* Mots. and *Oreina* Chevr. (Coleoptera, Chrysomelidae, Chrysomelinae). *Hereditas* **131:** 55-62.

Pouyanne M. 1917. La fécondation des *Ophrys* par les insectes. *Bulletin de la Société d'Histoire Naturelle de l'Afrique du Nord* **8:** 6-7.

Rasmussen C, Olesen JM. 2000. Oil flowers and oil-collecting bees. *Scandinavian Association for Pollination Ecology honours Knut Fegri* **39:** 23-31.

Ray JD. 1956. The Genus *Lysimachia* in the New World. In: Press TUoI, ed. *Illinois Biological Monographs, 24 (3–4)*. Urbana, IL.

Renner SS, Schaefer H. 2010. The evolution and loss of oil-offering flowers: new insights from dated phylogenies for angiosperms and bees. *Philosophical Transactions of the Royal Society B-Biological Sciences* **365:** 423-435.

Rowell-Rahier M, Pasteels JM. 1994. A comparison between allozyme data and phenotypic distances from defensive secretion in *Oreina* leaf-beetles (Chrysomelinae). *Journal of Evolutionary Biology* **7:** 489-500.

Schiestl FP. 2005. On the success of a swindle: pollination by deception in orchids. *Naturwissenschaften* **92:** 255-264.

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

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.

Simpson BB, Neff JL, Seigler DS. 1983. Floral biology and floral rewards of *Lysimachia* (Primulaceae). *American Midland Naturalist* **110:** 249-256.

Soliva M, Kocyan A, Widmer A. 2001. Molecular phylogenetics of the sexually deceptive orchid genus *Ophrys* (Orchidaceae) based on nuclear and chloroplast DNA sequences. *Molecular Phylogenetics and Evolution* **20:** 78-88.

Spooner LJ, Ritchie MG. 2006. An unusual phylogeography in the bushcricket *Ephippiger ephippiger* from Southern France. *Heredity* **97:** 398-408.

Stewart JR, Lister AM, Barnes I, Dalen L. 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B-Biological Sciences* **277:** 661-671.

Sundermann H. 1980. *Europäische und mediterrane Orchideen. Eine Bestimmungsflora mit Berücksichtigung der Ökologie. Ed. 3.*: Hildesheim, Germany.

Svenning JC. 2003. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters* **6:** 646-653.

Sykes WR. 1981. Checklist of dicotyledons naturalised in New Zealand 11. Apocynales, Campanulales, Ericales, Gentianales, Loganiales, Plantaginales, Primulales, and Rubiales. *New Zealand Journal of Botany* **19:** 319-325.

Taberlet P, Cheddadi R. 2002. Quaternary refugia and persistence of biodiversity. *Science* **297:** 2009-2010.

Varga ZS, Schmitt T. 2008. Types of oreal and oreotundral disjunctions in the western Palearctic. *Biological Journal of the Linnean Society* **93:** 415-430.

Vela E, Tirard A, Renucci M, Suehs CM, Provost E. 2007. Floral chemical signatures in the genus *Ophrys* L. (Orchidaceae): A preliminary test of a new tool for taxonomy and evolution. *Plant Molecular Biology Reporter* **25:** 83-97.

Vereecken NJ. 2009. Pollinator-mediated selection, reproductive isolation and the evolution of floral traits in the genus *Ophrys* (Orchidaceae). Free University of Brussels.

Vogel S. 1976. *Lysimachia* : Ölblumen des Holartiks. *Naturwissenschaften* **63:** 44-45.

Vogel S. 1986. Ölblumen and Ölsammelnde Bienen: Zweite Folge *Lysimachia* und *Macropis*. In: Verl. FS, ed. *Akademie der Wissenschaftenund der Literatur Mainz*. Stuttgart, Germany.

Vogel S. 1988. The oil-bee oil-flower relationships: parallelism and other aspects of their evolution in space and time. *Zeitschrift Fur Zoologische Systematik Und Evolutionsforschung* **26:** 341-362.

Wolff T. 1950. Pollination and Fertilization of the fly ophrys, *Ophrys insectifera* L. in Allindelille Fredskov, Denmark. *Oikos* **2:** 20-59.

Wolff T. 1951. Ecological Investigations on the fly ophrys, *Ophrys insectifera* L in Allindelille Fredskov, Denmark. *Oikos* **3:** 71-97.

CHAPTER I

Discordances between phylogenetic and morphological patterns of alpine leaf beetles attest for intricate history of lineages in postglacial Europe

Yann Triponez, Matthias Borer, Russell E. Naisbit, Martine Rahier and Nadir Alvarez

(Version formatted according to *Molecular Ecology*)

Neuchâtel, August 2010

Title Page

1. Title: Discordances between phylogenetic and morphological patterns of alpine leaf beetles attest for intricate history of lineages in postglacial Europe

2. Author's names: Y. Triponez[a,*], M. Borer[a], R. E. Naisbit[a,b], M. Rahier[a] and N. Alvarez[a]

3. Full postal addresses:

[a] Laboratory of Evolutionary Entomology, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, CH-2009 Neuchâtel, Switzerland.

[b] Unit of Ecology and Evolution, Department of Biology, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland.

[*] Corresponding author

4. Keywords: Phylogeography, Chrysomelidae, species concept, European mountains

5. Corresponding author:

Yann Triponez, Laboratory of Evolutionary Entomology, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, CH-2009 Neuchâtel, Switzerland.

E-mail address: yann.triponez@unine.ch

Fax number: +41 32 718 21 03

6. Running title: Phylogeography of *Oreina* alpine leaf beetles

Abstract

Delimitating taxa in closely related organisms and trying to explain their biogeographical history have long been important objectives of systematic biology. The development of biomolecular tools over the last two decades brought noticeable advances in these fields, but often resulted in discrepancies between new results and previous evaluations. Species concepts in particular have proved controversial in the face of conflicts between classical morphological species concepts (MSC) and more recent phylogenetic species concepts (PSC) based on DNA sequences. The present study aims at testing the validity of the morphological species concept (MSC) used to circumscribe three sister species of the *Oreina* genus, *O. speciosa*, *O. ganglbaueri* and *O. alpestris* (Coleoptera, Chrysomelidae), as well as describing the phylogeographic patterns of these European non-flying leaf beetles. Based on more than 700 sequences from one nuclear and three mitochondrial genes, we depict the phylogenetic relationships between individuals from 63 sites sampled across the main European mountain systems. Our results demonstrate incompatibility between MSC and PSC, especially considering *O. speciosa* and *O. alpestris*, but reveal the existence of several genetically and morphologically distinct lineages, notably the Pyrenean *O. ganglbaueri*. The phylogeography of the group shows parallels with recent studies, but also novel patterns in diversity and genetic links between the European mountains, demonstrating the importance of the Alps as a crossroads of genetic lineages in alpine insects.

Introduction

Long before the application of molecular techniques based on DNA, the current distribution areas of many European species were already supposed to be the result of Pleistocene climatic oscillations that begun about 2Mya (Coope 1994, Hewitt 2000, Schmitt 2007). During glacial periods, species went extinct over large parts of their range, dispersed to new locations or survived in refugia. When the climate warmed again, they recolonized the previously ice-covered areas. In the last decade, sequence-based phylogeographies revealed that the biogeographic histories of temperate and arctic-alpine species were likely to be different (e.g. Hewitt 2001, Schmitt et al. 2005). Temperate species survived during cold periods in southern refugia in Iberia, Italy, the Balkans and Caucasus region, repopulating northern Europe after the Last Glacial Maximum and creating hybrid zones in the areas where the expanding genomes met (Hewitt 2000). In contrast, alpine species or those with an arctic-alpine disjunctive distribution must have been widespread during glacial maxima and probably covered cold steppes over most of Europe, whereas they were restricted to remote environments during interglacial stages like the present (Hewitt 2004). As a consequence, some species with formerly large distributions continue to persist only as relict species in relatively small areas.

Large scale phylogeographies of the European fauna have been investigated in several groups, such as vertebrates (e.g. Taberlet et al. 1998, Böhme et al. 2007, Salomone et al. 2007, Sotiropoulos et al. 2007) and insects (e.g., crickets [Hewitt 2001], butterflies [Schmitt et al. 2006, Haubrich & Schmitt 2007, Habel et al. 2008, Espeland et al. 2007], hoverflies [Milankov et al. 2008], mayflies [Williams et al. 2006], and beetles [Gomez-Zurita & Vogler 2003, Cardoso & Vogler 2005, Cardoso et al. 2009]). Unexpectedly, high levels of differentiation among populations and regions were found in several of these studies, and their authors had to deal with intricate scenarios of extinction and recolonization through the Quaternary to explain the current spatial genetic structure of populations. As a consequence, many species were considered as species complexes encompassing high levels of genetic, ecological and morphological variation. Interestingly, nearly all published phylogeographic studies among European taxa involved organisms with substantial migratory abilities, foreseeing rather clearly their recent postglacial re-colonization patterns. However, one might ask whether or not patterns of recolonization should be so clear in species highly limited in their dispersal abilities, as for instance non-flying insects. A rational expectation in such organisms is a marked pattern of isolation by distance, with disjunctive areas of the distribution being highly isolated genetically from each other.

Here, we aim to tackle this problem by presenting the phylogeography of three sister species of Alpine leaf beetles within the genus *Oreina* Chevrolat, 1837 (Coleoptera, Chrysomelidae). This genus includes 28 species (Kippenberg 1994, 2008 and pers. comm.) described following a morphological species concept (MSC) based mainly on male genitalia and the shape of some other external body parts. All morphological species are found throughout the mountain regions of Europe, with some incursions into the lowlands, and western Siberia. Most are oligophagous on either Apiaceae or Asteraceae (Jolivet et al. 1986). *Oreina* are generally considered to be non-flying despite possessing completely developed red-coloured wings, and the recorded cases of flight in a few species are usually rare (Kalberer et al. 2004). As a consequence, studies based on allozyme and microsatellite data suggest that populations are highly isolated (Knoll & Rowell-Rahier 1998, Margraf et al. 2007). Here, we treat the three sister taxa *O. speciosa*, *O. ganglbaueri* and *O. alpestris*, among which *O. alpestris* shows the widest distributional range over European mountains (from the Pyrenees to the Carpathians), followed by *O. speciosa* (spanning the Alps, the Jura, the Massif Central and northern Balkans) and the endemic *O. ganglbaueri* (restricted to the Pyrenean-Cantabric range). The three taxa share many characteristics such as their general external morphology, highly variable coloration and autogenous cardenolide production (Dobler et al. 1996, Triponez et al. 2007). They finally show predilection for similar habitat and host plants (all are oligophagous on several species of Apiaceae) and therefore can be often found living together in sympatric populations in the Alps and Pyrenees.

In this study, our objectives are as follows:

1) to determine whether taxa defined using the MSC coincide with major monophyletic clades, or in other words, with a cladistic approach based on a general phylogenetic species concept (PSC) for these three taxa.
2) to examine the main phylogeographic patterns of these leaf beetles in comparison with general trends in the alpine phylogeography of Europe, while considering how their limited dispersal abilities could have affected their spatial genetic structure.

Material and Methods

Sampling

Sites within the main European alpine massifs were visited during summers 2005, 2006 and 2007. In total, the three sister species were collected in 63 sites (see Table 1), comprising mostly male specimens, with a total of 211 collected specimens. All sampled insects were determined, mostly on the basis of male genitalia shape and with the help of some external morphological features for females, following Kippenberg (1994, 2008). Outgroups were sampled among four closely-related species according to Hsiao & Pasteels (1999), namely *O. bifrons*, *O. cacaliae*, *O. gloriosa* and *O. liturata*.

DNA extraction, PCR amplification and cycle sequencing

Total genomic DNA was extracted from an average of three individuals per population, using the DNeasy® Tissue Kit (Qiagen, Hilden, Germany). Extraction was performed using 4–6 legs of each individual. One nuclear region and three mtDNA regions were amplified using the following primers: partial internal transcribed spacer (*ITS2*) region (ITS3 and ITS4 from Gomez-Zurita & Vogler, 2003), *16s* ribosomal RNA (LR-N-13398 and LR-J-12883 from Simon et al. 1994), partial cytochrome oxidase I (*COI*) (C1-J-1751 and C1-N-2191 from Simon et al. 1994) and partial cytochrome oxidase II (*COII*) (modTL2-J-3037 and modC2-N-3661 from Mardulyn et al. 1997). Amplification was carried out in a standard 30 µl PCR reaction including: 3 µl of 10X PCR buffer (Promega, Madison, WI, USA), 3 µl of a $MgCl_2$ solution (25 mM), 3 µl of dNTPs (1.5 mM), 0.5 µl of forward and reverse primers (10 mM), 0.3 µl of Taq DNA polymerase (Promega, Madison, WI, USA), 3 µl of extracted DNA, all made up to 30 µl with purified MilliQ water. The PCR reactions were run in a TGradient thermocycler (Biometra, Goettingen, Germany) with the following program: initial denaturation at 93°C for 1 min 30 s; 35 cycles comprising denaturation steps at 93°C for 1 min 30 s, annealing steps at 45°C (*16s rRNA*, *COI*) or at 53°C (*COII*, *ITS2*) for 1 min, extension steps at 72°C for 2 min; and final extension at 72°C for 8 min. The PCR product purification and sequencing was carried out by Macrogen (Seoul, South Korea). Sequencing was performed with both forward and reverse primers under BigDyeTM terminator cycling conditions, purifying the reacted products by using ethanol precipitation, and running them using an Automatic Sequencer 3730XL (Applied Biosystem, Foster City, USA).

Sequence alignment

Sequences (forward and reverse) were manually corrected and assembled using the software CHROMAS PRO 1.34 (Technelysium, Helensvale, Australia). Alignments of *ITS2* and *16s rRNA* were carried out using CLUSTALW Multiple Alignment (Thompson et al. 1997) within the software BIOEDIT 7.0.5.3 (distributed by Tom Hall, North Carolina State University, Raleigh, USA), followed by minor manual correction. For *COI* and *COII*, alignment was trivial as all sequenced fragments were of the same size. For each defined partition, the best-fit substitution model was selected using MrAIC.pl 1.4.3 (Nylander 2004) based on the Akaike information criterion (AIC; Akaike 1974). The three mtDNA partitions were shown to be congruent using the mILD test (Planet & Sarkar 2005) and a supermatrix comprising *16s rRNA*, *COI* and *COII* was built. In contrast, the mILD analysis revealed that *ITS2* was incongruent with the mtDNA regions and it was treated separately.

Phylogenetic reconstructions

To estimate phylogenetic relationships among individuals, both the mtDNA supermatrix and the *ITS2* alignment were analyzed using parsimony ratchet (Nixon, 1999) as implemented in PAUPrat (Sikes & Lewis, 2001). Based on recommendations by Nixon (1999), ten independent searches were performed with 200 iterations and 15% of the parsimony informative characters perturbed using PAUP* version 4.0b10 (Swofford, 2003). The shortest equally most parsimonious trees were combined to produce a majority-rule consensus tree. Node support was determined by computing Bremer support values as implemented in TREEROT.V3 (Sorenson & Franzosa 2007) and using PAUP* version 4.0b10 (Swofford, 2003).

Bayesian analyses (Nylander et al. 2004) were also performed separately for the *ITS2* region and the mtDNA supermatrix (treating the three mtDNA regions as separate partitions), using MRBAYES version 3.1 (Huelsenbeck & Ronquist 2001), with substitution models as estimated by

MRAIC (Nylander et al. 2004) and four estimated alpha categories for the gamma term (Yang 1994). Four simultaneous Monte Carlo Markov Chains were run for 10^8 generations in two independent runs, saving a tree every 1000 generations. Convergence of the MCMC runs was tested by computing the Potential Scale Reduction Factor (Gelman & Rubin 1992) criterion as implemented in MRBAYES, and by determining the Effective Sample Size using TRACER 1.4.1 (Rambaut & Drummond 2008). Accordingly, the burn-in period was set to $3x10^7$ generations until stationarity in the likelihood value was established among the runs, so that 30000 sample points were discarded. The last 70000 trees were used to calculate the half-compatible topology (*i.e.,* majority-rule) and the Bayesian posterior probability (BPP) at each node. Bremer support values were also determined on the half-compatible topology, as implemented in TREEROT.V3 and using PAUP* version 4.0b10.

To confront phylogenetic hypotheses with a network approach, mtDNA and *ITS2* haplotype networks were performed using statistical parsimony networks (SPN) as implemented in TCS 1.21 (Clement et al. 2000). Analyses were carried out by applying a 99% connection limit and gaps were treated either as “missing” for mtDNA or as “fifth state” for *ITS2* (due to a lower level of polymorphism but high number of indels in this later region). To be taken into consideration, we assumed that a network should comprise at least two ingroup samples.

Spatial genetic structure

Based on the topologies and node supports obtained both for mtDNA and *ITS2* phylogenetic analyses, supported clades were defined and then displayed on geographical maps using ARCGIS 9.1 (ESRI, Redlands, CA, USA), by representing each population as a pie-chart showing the number of samples from each clade.

Results

Morphological determination

The classical determination key for *Oreina* species from Kippenberg (1994, 2008) uses male genitalia shape to discern *O. speciosa*, *O. ganglbaueri* and five subtypes for *O. alpestris* (defined along a morphological continuum). In total, we could only distinguish in this study four discrete genitalia types, represented in Fig. 1 together with their geographical distributions. The "*ganglbaueri*" type is easily diagnosable by being much larger. Interestingly the type "*alpestris* β" has an intermediate morphology between "*alpestris* α" and *O. speciosa*. Except in two populations, one in Abruzzi and one in the northern Carpathians, types "*alpestris* α" and "*alpestris* β" were never found together. Moreover, type "*alpestris* β" appears to be restricted to the edges of the distribution and was never found in the Alps. In contrast, all *O. alpestris* populations sympatric with *O. speciosa* (in other words, Alpine populations) always showed an "*alpestris* α" morphology.

Sequence polymorphism

Amplification of mtDNA regions was successful for 185 specimens, whereas *ITS2* was successfully amplified for 184 specimens (with 158 specimens amplified for both). The genetic analysis yielded a total of 1442 bp for the three mtDNA regions: 513 bp for *16s rRNA* (20 parsimony-informative sites among 35 polymorphic sites), 345 bp for partial *COI* (54 parsimony-informative sites among 68 polymorphic sites), and 584 bp for partial *COII* (63 parsimony-informative sites among 85 polymorphic sites). The total length of the nuclear *ITS2* region was 611 bp (43 parsimony-informative sites among 73 polymorphic sites). All sequences were deposited in GenBank (accession numbers GQ392138-GQ392886). The best-fit substitution models for each of the four regions were as follows: Hasegawa–Kishino–Yano (HKY) model with gamma parameter for *ITS2* and *16s rRNA*; General Time Reversible (GTR) model with gamma parameter and a proportion of invariable sites for *COI* and *COII*.

MtDNA phylogenetic analyses

The MP analysis yielded a well-resolved mtDNA majority-rule consensus tree, based on 125 equally parsimonious trees (650 steps; CI=0.55; RI=0.89). *Oreina speciosa, O. ganglbaueri* and *O. alpestris* clustered together with a Bremer support of 2 and support values ranged from 0 to 14 for clades within the ingroup (see Fig. 2). Among the three species, only *O. ganglbaueri* was monophyletic, with *O. alpestris* and *O. speciosa* being mutually paraphyletic. The half-compatible consensus tree obtained through Bayesian inference analyses largely confirmed the topology obtained by MP (see supplementary file #1). Overall, BPP values were relatively high, confirming the monophyly of the clade comprising the three ingroup taxa with very high support (BPP=0.98), as well as the monophyly of *O. ganglbaueri* (BPP=1). Bremer indexes on the Bayesian topology supported the ingroup with a value of 3, and ranged between 0 and 14 within the ingroup, even if basal clusters were slightly less well supported than in the MP analysis. When comparing the broad-scale topologies produced by MP and Bayesian Inference, only two inconsistencies were apparent: *Oreina ganglbaueri* diverged one step earlier in the MP analysis than in the Bayesian Inference analysis, and a small group comprising five *O. speciosa* individuals from northern Italy (Dolomites and Piemonte regions) fell in two different clades depending on the reconstruction method. Since the two methods yielded highly similar results, we discuss our results based on the MP topology due to its higher level of Bremer support in basal nodes. However, in order to define entities that were supported by both methods, we consider only clades supported by Bremer indexes ≥ 1 and BPP ≥ 0.8. Globally, the phylogeny consists of seven major clades, referred to as M1 to M7 (see Fig. 2), each of them showing marked geographic structuring (see Fig. 3 and supplementary file #2, in which the distribution of each clade is considered independently and shown in parallel with the topology of one of the most parsimonious trees). Figure 6a illustrates the results of the 12 independent networks obtained with the SNP haplotype analysis of the mtDNA matrix. These networks are highly similar with both MP and Bayesian topologies, but with the following two slight differences. First, although clades M1, M5 and a large proportion of M7 are retrieved as segregated entities, they are linked in one single large network. Second, the high level of diversity within clade M6 as attested by the high Bremer support of subclades in the MP topology is confirmed since it is split further into six independent networks (plus two single individuals disconnected from any network).

ITS2 phylogenetic analyses

The MP analysis yielded an *ITS2* majority-rule consensus tree based on 2009 equally parsimonious trees (78 steps; CI=0.85; RI=0.93). The resulting topology was, however, much less well-resolved than the mtDNA topology for the ingroup taxa and most samples were embedded in polytomies. Bremer supports ranged from 1 to 2 for clades within the ingroup, and *O. speciosa*, *O. ganglbaueri* and *O. alpestris* clustered together with a Bremer support of 3 (see Fig. 4). Again, only *O. ganglbaueri* was monophyletic among the three species, with *O. alpestris* and *O. speciosa* being mutually paraphyletic. The half-compatible consensus tree obtained through Bayesian inference analyses confirmed the topology obtained by MP (data not shown), although the splitting of major clades was not identical. BPP values were much lower than in the mtDNA analysis, but the monophyly of the clade comprising the three taxa was confirmed with high support (BPP=1.00), as was the monophyly of *O. ganglbaueri* (BPP=0.98). Bremer indexes on the Bayesian topology also supported the ingroup with a value of 3, and ranged between 1 and 2 within the ingroup. In order to be as consistent as possible with the mtDNA analyses, only clades supported by Bremer indexes ≥ 2 and BPP ≥ 0.95 will be discussed in the light of the MP topology (*NB*. threshold levels were increased compared to the mtDNA analyses in order to remove biases related to the much lower level of polymorphism within the *ITS2* dataset; *i.e.*, the *ITS2* tree was more than eight times shorter than the mtDNA tree). Overall, only two clusters, referred to as N1 and N2, were well defined, both nested in a wide polytomy N0 (see Fig. 4 and Fig. 5). The SNP haplotype analysis (Fig. 6b) resulted in only two independent networks. Compared to MP analysis, the first network is identical to clade N2. The second one connects all other samples (N0 + N1), in which N1 is a segregated "clan" (*sensu* Wilkinson et al. 2007) within the unrooted network.

Comparison between morphological and phylogenetic data

In order to discuss the nature of inconsistencies among morphological, mtDNA and *ITS2* data, we provide a contingency table (Table 2) containing information from unambiguously determined specimens (excluding females and specimens with missing genetic information).

- Only the group composed of specimens with *O. ganglbaueri* morphology (five specimens) showed absolute congruence between morphological, mtDNA and *ITS2* patterns (respectively only clustered into M2 and N1).
- Within the *ITS2* clade N2, more than 80% of the samples demonstrated an "alpestris α" morphology and conversely, more than 80% of all specimens characterized by an "alpestris α" morphology were found within the N2 clade; among them, about one third exclusively composed the most basal mtDNA

clade, M1 (distributed in the Balkanic region as well as in South-eastern Austria), whereas the other two thirds were distributed within the more widespread mtDNA clades M3, M4 and M5 (which also contain specimens with an *O. speciosa* morphology). The remaining specimens in the N2 cluster showed an “alpestris β” morphology.

- The *ITS2* polytomy N0 represents all other “alpestris β”, together with a few samples with an “alpestris α” morphology, and 100% of the “speciosa” samples. Among the “speciosa” specimens, 60% form the mtDNA clade M7. The other 40% were distributed across mtDNA clades M3, M4, M5 and particularly M6. The remaining specimens with an “alpestris α” morphology clustered in clades M3 and M6, whereas most beetles with an “alpestris β” morphology clustered into clade M6.
- Among the mtDNA clades, M1, M2 and M7 contain one single morphotype in each. M3 and M4 include “alpestris α” morphology together with “speciosa”. M5 and M6 clades are finally the two only mtDNA clades with beetles representing three different morphologies: “alpestris α”, “alpestris β” and “speciosa”.

Discussion

Incompatibility between MSC and PSC

Based on the simple criterion of the shape of male genitalia (and on a few external criteria for females), the species circumscription of each taxon studied here seemed quite evident for classical coleopterologists, with the exception of *O. alpestris* (formerly known as *O. variabilis*, see Kippenberg 1994, Dobler et al. 1996) which showed some local variation and which was split into five to eight subspecies, depending on the authors. However, based on our phylogenetic approach, species delimitation is not so clear cut.

Only *O. ganglbaueri* can be objectively considered as a segregated entity, both from morphological and phylogenetic points of view. It can be seen as a micro-endemic restricted to the central valleys of the Pyrenean range, which has remained isolated for a sufficient length of time for the development of barriers to hybridization that impede introgression of *O. ganglbaueri* with other specimens from *O. alpestris*, even in sympatry (as in population SP1 for instance); the male genitalia of *O. ganglbaueri* are by far the largest in size among the three taxa, a feature that could act as a physical pre-zygotic barrier. Interestingly, the *O. ganglbaueri* clade (M2) splits very early in the mtDNA phylogeny, attesting to the ancient history of this lineage in this mountain range.

The consistency between morphology and phylogeny is also strong in the most basal clade of the mtDNA phylogeny (M1), uniquely composed of *O. alpestris* specimens from the Balkans and Eastern Austria, with an “alpestris α” morphology and belonging to the N2 nuclear clade. This argues in favor of a very ancient history of this lineage as well (similar to the case of *O. ganglbaueri*) that may have survived long enough in allopatry in the Balkanic region to have developed reproductive barriers despite its current sympatry with *O. speciosa* in Austria (as in population VAL for instance). This well-circumscribed entity contrasts with the high level of morphological and phylogenetic variation found in other *O. alpestris* samples from across the European mountain ranges. Such heterogeneity could be explained by the following two antithetic hypotheses:

- first, the sympatric status of *O. alpestris* with *O. speciosa* (in the Alps) and *O. ganglbaueri* (in the Pyrenean range) might have created an opportunity for selection due to outbreeding depression to act to reduce hybridization and accelerate the development of pre-zygotic barriers. It is well known that genitalia are among the fastest evolving characters in insects, and that selection on these structures may even increase speciation rates (Polihronakis 2009). As a consequence, it might have led to independent evolution of similar homoplasic features of the male genitalia in disparate populations that were then grouped by taxonomists;
- second, the wide range of morphological and phylogenetic variation could simply be the result of several independent events of introgression with other lineages.

It is, however, not possible to favor one or the other hypothesis based on our dataset, since genome-wide markers and experimental crosses would be required to resolve the question. Nonetheless, our analyses clearly highlight the paraphyletic status of *O. alpestris*. In addition to the specimens found in the Balkans (M1), several *O. alpestris* with an “alpestris α” morphology come from the southern and central Alps (M3-M4) and northern Alps plus Carpathians (M5). The other *O. alpestris* specimens (from the Apennines and the Pyrenean-Cantabric range), which cluster into clade M6 together with *O. speciosa* specimens, mostly have an “alpestris β” morphology. Therefore, their genitalic structures (more similar to those from *O. speciosa*) as well as their phylogenetic position (very close to *O. speciosa*) suggest that they are intermediate between the two nominal species.

Regarding specimens with a strict *O. speciosa* morphology, their undefined position in the *ITS2* phylogeny, as well as the later branching of the mtDNA clades in which they fall (together with specimens showing other morphologies) argue for an incomplete development of barriers to reproduction with other lineages. Nonetheless, the existence of one major mtDNA clade, M7, strictly composed of central European *O. speciosa* could represent evidence for a long history of this specific lineage in the Alps.

Although it is not possible to unequivocally interpret the history of the different lineages and morphotypes addressed in this study, our results globally suggest that rather than three discrete species, we are dealing with what is generally called a species complex, more precisely resembling a cladospecies-metaspecies (see Donoghue 1985 or de Queiroz & Donoghue, 1988). Reproductive barriers seem to be incomplete, and despite their low dispersal ability, *Oreina* lineages within these taxa tend to intermix where they meet within the Alps. Only lineages that have been isolated for a long time in remote mountain ranges (i.e., *O. ganglbaueri* in the Pyrenees and the M1 clade of *O. alpestris* in the Balkans) seem to have developed barriers to hybridization that maintain their status as non-introgressed entities. As a consequence, the MSC used in order to describe and recognize *Oreina* taxa might be biased in using fast evolving characters such as male genitalia as the primordial criteria for species delimitation. In such cases, consideration of other morphological clues compatible with analyses of character evolution based on DNA sequence data should be preferred as proposed by Huber (2003).

Taxonomic considerations

At the species level, both MSC and PSC match for *O. ganglbaueri* (M2 and N1 are monophyletic and exclusive), even if this result might be tempered by the *ITS2* haplotype network, in which *O.*

ganglbaueri is retrieved as a “clan” and not as an independent network. In contrast, these two concepts are not compatible at all for the morphological entities *O. speciosa* and *O. alpestris*. A general phylogenetic approach would then redistribute them as follows: 1) Evaluating the *ITS2* topology, the phylogenetic species *O. alpestris* could be retrieved on the basis of the N2 clade that includes only specimens from the Alps, Carpathians, Black Forest and Balkans. The other specimens (all embedded into the N0 polytomy) could be regarded as the phylogenetic metaspecies *O. speciosa**. 2) Evaluating the mtDNA topology and haplotype network, at least four distinct entities (M1, M3, M4 and M5-7) could be regarded as four phylogenetic species. Among the different lineages identified by the mtDNA analysis (clades and/or networks), several fit with a series of subtaxa described on the basis of morphological features and/or distribution patterns (see Kippenberg 1994, 2008). For instance, the monophyletic mtDNA clades M6b and M6c correspond to the micro-endemics *O. alpestris* ssp. *marsicana* (Luig.) from the Appenines and *O. alpestris* ssp. *nigrina* (Suffr.) from the Pyrenees respectively. Other entities only partly match described subtaxa, or at least show a larger distribution than previously thought. This is notably the case for M6d, which corresponds to *O. speciosa* ssp. *pretiosa* (Suffr.) from the Swiss Jura (but expanding into the Massif Central and the Swiss western Alps according to our results) or M1, which might correspond to *O. alpestris* ssp. *balcanica* (Wse.) described from Macedonia and Bulgaria (ranging from Serbia and Montenegro to southern Austria as shown here). Similarly, a partial match is found with clade M5, which could be an assemblage of three described subspecies of *O. alpestris*: ssp. *variabilis* (Wse.) (from the Alps and Black Forest), ssp. *banatica* (Wse.) (from the central Carpathians and Transsylvania) and ssp. *alpestris* s. str. (Schumm.) (from the northern Carpathians and Sudets).

Phylogeography of the species complex

Setting aside the conflicting circumscription of taxa, some noticeable patterns from the clades obtained (particularly with the mtDNA topology) are informative about the phylogeography of this species complex. European refugia classically considered during Pleistocene climatic oscillations can be identified here. Two of the most basal mtDNA lineages (M1 and M2) comprise beetles from the Pyrenees and from the Balkanic range, respectively (see Fig. 2, Fig. 3 and supplementary file #2), two areas known for having played the role of Pleistocene refugia. The M5 clade includes pure populations in the Carpathians and the Black Forest, which may be related to putative Northern refugia (Stewart & Lister 2001). When considering the Alps, the picture is more intricate although some pure populations of the M3 and M4 lineages (eg, western Dolomites [IT2, IT3], Adamello [IT5], Swiss Inn valley [SUS]) could represent persistent refugia within the Alps, with range expansion leaving just a few representatives at their eastern and western edges. As identified by the mtDNA haplotype network analysis, other micro-endemic lineages confirm the role played by different massifs as persisting refugia, such as Liguria, the Maritime Alps, Abruzzi or the Pyrenees. But our results also address new areas of importance, such as the central Jura and the northern Swiss Alps, in which similar processes of isolation followed by genetic divergence occurred during the Pleistocene, thereby increasing the intricacy of the phylogeographical history of these alpine leaf beetles.

Genetic links between the main mountain ranges are apparent in our data. Some of these provide additional examples of previously reported biogeographical and genetical links reviewed by Schmitt (2009). For instance, some northern Alpine populations harbour individuals with mtDNA originating in the Carpathians (M5 clade). At the western edge of the distribution, specimens from the Pyrenees, Massif Central and Apennines (M6a,b,c) as well as those from the western /

southern Alps, Massif Central and Jura (M6d,e) are also connected. Contrastingly, other patterns obtained here bring novelty to the phylogeographic knowledge of the European mountains fauna. This is notably the case of the genetic connection between the Alps and Balkans (for which data was unavailable until now for insects according to Schmitt 2009): two mtDNA clades (M1 and M7) provide strong evidence for relationship between 1) the eastern Alps and southern Balkans (Serbia and Montenegro) and 2) the Alps (as a whole) and eastern Croatia. In addition we should note that the obvious genetic connexion between Black Forest and Carpathians, as demonstrated by M5 clade, represents a relatively unusual pattern (T. Schmitt pers. comm.).

European high mountain taxa have proved once again to be moulded by multiple genetic influences especially when considering massifs such as the Pyrenees, the Alps, but also the western Balkans. Interestingly however, in contrast to what has been found in Trichoptera (Pauls et al. 2006) as well as in Angiosperms (e.g. Mraz et al. 2007, Ronikier et al. 2008), no genetic differentiation could be detected (either with mtDNA or nDNA) within the Carpathians, despite samples spanning the chain North to South. This phylogenetic homogeneity in the easternmost range of the distribution area contrasts with the pattern found in narrow middle-high mountain systems such as the Massif Central or the Jura. This later region actually shows, together with the Alps, the highest level of admixture of distantly-related mtDNA haplotypes. Our results therefore reinforce the status of the Alps and Jura as a crossroads for alpine insect species in Europe.

References

Akaike H (1974) New Look at Statistical-Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

Bohme MU, Fritz U, Kotenko T, *et al.* (2007) Phylogeography and cryptic variation within the

Lacerta viridis complex (Lacertidae, Reptilia). *Zoologica Scripta*, **36**, 119-131.

Cardoso A, Serrano A, Vogler AP (2009) Morphological and molecular variation in tiger beetles of the *Cicindela hybrida* complex: is an 'integrative taxonomy' possible? *Molecular Ecology*, **18**, 648-664.

Cardoso A, Vogler AP (2005) DNA taxonomy, phylogeny and Pleistocene diversification of the *Cicindela hybrida* species group (Coleoptera : Cicindelidae). *Molecular Ecology*, **14**, 3531-3546.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1659.

Coope GR (1994) The response of insect faunas to glacial-interglacial climatic fluctuations. *Philosophical Transactions: Biological Sciences*, **344**, 19-26.

de Queiroz K, Donoghue MJ (1988) Phylogenetic systematics and the species problem. *Cladistics*, **4**, 317-338.

Dobler S, Mardulyn P, Pasteels JM, Rowell-Rahier M (1996) Host-plant switches and the evolution of chemical defense and life history in the leaf beetle genus *Oreina*. *Evolution*, **50**, 2373-2386.

Donoghue MJ (1985) A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist*, **88**, 172-181.

Espeland M, Aagaard K, Balstad T, Hindar K (2007) Ecomorphological and genetic divergence between lowland and montane forms of the *Pieris napi* species complex (Pieridae, Lepidoptera). *Biological Journal of the Linnean Society*, **92**, 727-745.

Freude H, Harde KW, Lohse GA (1966) Die Käfer Mitteleuropas (ed. Evers and Goecke), Krefeld, Germany.

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457-511.

Gomez-Zurita J, Vogler AP (2003) Incongruent nuclear and mitochondrial phylogeographic patterns in the *Timarcha goettingensis* species complex (Coleoptera, Chrysomelidae). *Journal of Evolutionary Biology*, **16**, 833-843.

Habel JC, Meyer M, El Mousadik A, Schmitt T (2008) Africa goes Europe: The complete phylogeography of the marbled white butterfly species complex *Melanargia galathea* / *M. lachesis* (Lepidoptera: Satyridae). *Organisms Diversity & Evolution*, **8**, 121-129.

Haubrich K, Schmitt T (2007) Cryptic differentiation in alpine-endemic, high-altitude butterflies reveals down-slope glacial refugia. *Molecular Ecology*, **16**, 3643-3658.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology*, **10**, 537-549.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **359**, 183-195.

Hsiao TH, Pasteels JM (1999) Evolution of host-plant affiliation and chemical defense in *Chrysolina* - *Oreina* leaf beetles as revealed by mtDNA phylogenies. In: *Advances in Chrysomelidae Biology* (ed. Cox ML), pp. 321-342, Backhuys, Leiden.

Huber BA (2003) Rapid evolution and species-specificity of arthropod genitalia: fact or artifact ? *Organisms Diversity & Evolution*, **3**, 63-71

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754-755.

Jolivet P, Petitpierre E, Daccordi M (1986) Les plantes-hôtes des *Chrysomelidae.* Quelques nouvelles précisions et additions (*Coleoptera*). *Nouvelle Revue Entomologique*, **3**, 341-357.

Kalberer NM, Turlings TCJ, Rahier M (2005) An alternative hibernation strategy involving sun-exposed 'hotspots', dispersal by flight, and host plant finding by olfaction in an alpine leaf beetle. *Entomologia Experimentalis et Applicata,* **114**, 189-196.

Kippenberg H (1994) Familie Chrysomelidae. In: *Die Käfer Mitteleuropas, 3. Supplementband* (ed. Lohse GA and Lucht WH), pp. 65-83. Krefeld, Germany.

Kippenberg H (2008) Revision der Untergattung *Protorina* WEISE der Gattung *Oreina* CHEVROLAT (Coleoptera: Chrysomelidae: Chrysomelinae). *Koleopterologische Rundschau,* **78**, 367-418.

Knoll S, Rowell-Rahier M (1998) Distribution of genetic variance and isolation by distance in two leaf beetle species: *Oreina cacaliae* and *Oreina speciosissima*. *Heredity,* **81**, 412-421.

Mardulyn P, Milinkovitch MC, Pasteels JM (1997) Phylogenetic analyses of DNA and allozyme data suggest that *Gonioctena* leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history of host-plant family shifts. *Systematic Biology,* **46**, 722-747.

Margraf N, Verdon A, Rahier M, Naisbit RE (2007) Glacial survival and local adaptation in an alpine leaf beetle. *Molecular Ecology,* **16**, 2333-2343.

Milankov V, Stahls G, Stamenkovic J, Vujic A (2008) Genetic diversity of populations of *Merodon aureus* and *M. cinereus* species complexes (Diptera, Syrphidae): integrative taxonomy and implications for conservation priorities on the Balkan Peninsula. *Conservation Genetics,* **9**, 1125-1137.

Mraz P, Gaudeul M, Rioux D, *et al.* (2007) Genetic structure of *Hypochaeris uniflora*

(Asteraceae) suggests vicariance in the Carpathians and rapid post-glacial colonization of the Alps from an eastern Alpine refugium. *Journal of Biogeography*, **34**, 2100-2114.

Nixon KC (1999) The Parsimony Ratchet, a new method for rapid parsimony analysis. *Cladistics-the International Journal of the Willi Hennig Society*, **15**, 407-414.

Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL (2004) Bayesian phylogenetic analysis of combined data. *Systematic Biology*, **53**, 47-67.

Pauls SU, Lumbsch HT, Haase P (2006) Phylogeography of the montane caddisfly *Drusus discolor*: evidence for multiple refugia and periglacial survival. *Molecular Ecology*, **15**, 2153-2169.

Planet PJ, Sarkar IN (2005) mILD: a tool for constructing and analyzing matrices of pairwise phylogenetic character incongruence tests. *Bioinformatics*, **21**, 4423-4424.

Polihronakis M (2009) Hierarchical comparative analysis of genetic and genitalic geographical structure: testing patterns of male and female genital evolution in the scarab beetle *Phyllophaga hirticula* (Coleoptera: Scarabaeidae). *Biological Journal of the Linnean Society,* **96**, 135-149.

Rambaut A, Drummond AJ (2008) Tracer v1.4.1.

Ronikier M, Cieslak E, Korbecka G (2008) High genetic differentiation in the alpine plant *Campanula alpina* Jacq. (Campanulaceae): evidence for glacial survival in several Carpathian regions and long-term isolation between the Carpathians and the Alps. *Molecular Ecology,* **17**, 1763–1775.

Rowell-Rahier M, Pasteels JM (1994) A comparison between allozyme data and phenotypic distances from defensive secretion in *Oreina* leaf-beetles (Chrysomelinae). *Journal of Evolutionary Biology*, **7**, 489-500.

Salomone N, Vignoli V, Frati F, Bernini F (2007) Species boundaries and phylogeography of the "*Euscorpius carpathicus* complex" (Scorpiones : Euscorpiidae) in Italy. *Molecular Phylogenetics and Evolution*, **43**, 502-514.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology,* **4**, 13.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology,* **6**.

Schmitt T, Hewitt GM, Muller P (2006) Disjunct distributions during glacial and interglacial periods in mountain butterflies: *Erebia epiphron* as an example. *Journal of Evolutionary Biology*, **19**, 108-113.

Schmitt T, Rober S, Seitz A (2005) Is the last glaciation the only relevant event for the present genetic population structure of the meadow brown butterfly *Maniola jurtina* (Lepidoptera : Nymphalidae)? *Biological Journal of the Linnean Society*, **85**, 419-431.

Sikes DS, Lewis PO (2001) Software manual for PAUPRat: A tool to implement Parsimony Ratchet searches using PAUP*, University of Connecticut, Storrs, CT.

Simon C, Frati F, Beckenbach A, *et al.* (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Annals of the Entomological Society of America,* **87**, 651-701.

Sorenson MD, Franzosa EA (2007) TreeRot, version 3. Boston University, Boston, MA.

Sotiropoulos K, Eleftherakos K, Dzukic G, *et al.* (2007) Phylogeny and biogeography of the alpine newt *Mesotriton alpestris* (Salamandridae, Caudata), inferred from mtDNA sequences. *Molecular Phylogenetics and Evolution,* **45**, 211-226.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota.

Trends in Ecology & Evolution, **16**, 608-613.

Swofford DL (2003) *PAUP*. Phylogenetic analysis using (*and other methods). Version 4.0b10* Sunderland, MA: Sinauer Associates.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology,* **7**, 453-464.

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

Triponez Y, Naisbit RE, Jean-Denis JB, Rahier M, Alvarez N (2007) Genetic and environmental sources of variation in the autogenous chemical defense of a leaf beetle. *Journal of Chemical Ecology,* **33**, 2011-2024.

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

Williams HC, Ormerod SJ, Bruford MW (2006) Molecular systematics and phylogeography of the cryptic species complex *Baetis rhodani* (Ephemeroptera, Baetidae). *Molecular Phylogenetics and Evolution*, **40**, 370-382.

Yang ZH (1994) Maximum-Likelihood phylogenetic estimation from DNA-sequences with variable rates over sites-approximate methods. *Journal of Molecular Evolution,* **39**, 306-314.

Acknowledgments

This research was made possible thanks to the University of Neuchâtel and the *National Centre of Competence in Research* (NCCR “Plant Survival”). The authors are very grateful to Y. Borcard for species determination and drawings of genitalia morphotypes, and to N. Villard and

F. Schüpfer for laboratory technical assistance. We also wish to thank our precious field helpers: R. Arnoux, V. Baštić, A. Favre, A. Grill, N. Lamon and N. Margraf.

Figure Legends

Figure 1. Drawings of the four genitalia morphotypes examined in this study with their respective spatial distributions. A. type *ganglbaueri*; B. type *speciosa*; C. type *alpestris* β; D. type *alpestris* α.

Figure 2. Majority-rule consensus tree of the MP analysis for the three mtDNA regions *COI*, *COII* and *16s rRNA* in a total evidence approach. Label names comprise the population code (see Table 1), the code of the individual specimen (a, b, c, d, e, f, g), followed by the genitalic morphology (ALPα, ALPβ, GAN, SPE) for male specimens. In female specimens (highlighted with a "*"), the global body shape and morphology allowed us to distinguish the three species, coded by ALP, GAN and SPE (but with no way to separate ALPα and ALPβ). Bremer support values are indicated on each node. Main clades are defined with codes from M1 to M7. The group illustrated with a "§" switched position between the MP and the Bayesian inference analyses (see supplementary material #1).

Figure 3. Geographical distribution of each mtDNA clade identified in the MP analysis. Colors refer to clades identified in Figure 2 and codes correspond to populations detailed in Table 1.

Figure 4. Majority-rule consensus tree (equal branch-length cladogram) of the MP analysis for the nuclear region *ITS2*. Label names are identical to Figure 2. Bremer support values are indicated on each node. The two main clades are defined with codes N1 and N2. Specimens that do not fall in a well-supported clade (N1 or N2) are considered as belonging to group N0.

Figure 5. Geographical distribution of each nuclear clade identified in the Maximum Parsimony analysis. Colors refer to N1 and N2 clades, as identified in Figure 4. Samples in white correspond to the loosely defined N0 group.

Figure 6. Schematic representation of the mtDNA (a) and *ITS2* (b) haplotype network analyses. Each box represents an independent network. In each box is indicated the correspondence with phylogenetic clades (as shown in Figures 2 and 4), the geographical origin and the number of samples.

Tables

Table 1. Geographical details of the sampled populations. The names of the collected morphological species are also indicated.

POP.	RANGE	COUNTRY	LOCATION	ALT.	GPS	SPECIES
AI1	ALPS	Switzerland	Brulisau	1600	N 47°17'04.08" E 009°29'05.90"	O. alpestris / O. speciosa
AUT1		Austria	Arlbergpass	1354	N 47°08'03.80" E 010°12'20.90"	O. speciosa
AUT2		Austria	Mittersill (Gamsblick)	1317	N 47°11'25.20" E 012°28'37.50"	O. speciosa
AUT3		Austria	Badgastein (Stubnerkugel)	1876	N 47°06'35.40" E 013°07'36.20"	O. alpestris
AUT5		Austria	Koralpe	1564	N 46°48'30.50" E 014°56'34.10"	O. speciosa
BE1		Switzerland	Kandersteg	1314	N 46°28'21.10" E 007°39'23.00"	O. alpestris / O. speciosa
DOL2		Italy	Passo Rolle	1969	N 46°17'46.00" E 011°46'56.00"	O. speciosa
DOL3		Italy	Passo Duran	1477	N 46°18'45.20" E 012°05'29.20"	O. speciosa
FR11		France	Col du Galibier	1999	N 45°05'07.40" E 006°26'19.10"	O. speciosa
FR13		France	Ailefroide	1700	N 44°53'41.50" E 006°26'44.70"	O. speciosa
FR14		France	Abries	1886	N 44°48'49.10" E 006°58'28.30"	O. speciosa
FR16		France	Saint-Martin Vesubie	1795	N 44°06'37.10" E 007°18'41.10"	O. speciosa
GL1		Switzerland	Schwanden	1500	N 46°57'28.30" E 009°05'55.90"	O. speciosa
GR1		Switzerland	Tschiertschen	1400	N 46°48'42.00" E 009°36'40.00"	O. alpestris / O. speciosa
GRH		Austria	Grosser Hengst	1615	N 47°27'6.15" E 014°25'50.86"	O. speciosa
IT1		Italy	Col du Petit St-Bernard	1996	N 45°42'12.00" E 006°52'29.40"	O. speciosa
IT2		Italy	Passo Pian delle Fugazze	1296	N 45°44'47.30" E 011°09'32.30"	O. alpestris
IT3		Italy	Monte Baldo	1247	N 45°47'03.60" E 010°52'14.00"	O. speciosa
IT4		Italy	Val Daone	1319	N 46°01'19.80" E 010°30'47.60"	O. speciosa
IT5		Italy	Passo del Tonale	1784	N 46°15'53.40" E 010°36'46.80"	O. speciosa
IT8		Italy	Terme di Valdieri	1419	N 44°12'10.30" E 007°16'13.60"	O. speciosa
IT9		Italy	Crissolo (Piano del Re)	1342	N 44°41'59.00" E 007°09'11.20"	O. speciosa
IT11		Italy	Breuil-Cervinia	2149	N 45°55'40.10" E 007°37'53.30"	O. speciosa
IT12		Italy	Macugnaga	1343	N 45°57'58.50" E 007°56'14.00"	O. speciosa
LOT		Switzerland	Lotschental	1888	N 46°26'21.67" E 007°52'23.46"	O. speciosa
SLO2		Slovenia	Predmeja	1142	N 45°55’56.70” E 013°50’31.20”	O. speciosa
SLO8		Slovenia	Logarska Dolina	1394	N 46°22’09.50” E 014°35’04.70”	O. speciosa
SLO9		Slovenia	Dom na komni	1261	N 46°16’58.10” E 013°47’13.40”	O. speciosa
SLO10		Slovenia	Vrsic pass	1387	N 46°25’29.00” E 013°44’34.60’	O. speciosa
SUR		Switzerland	Sur	1515	N 46°31'16.40" E 009°37'27.00"	O. speciosa
SUS		Switzerland	Susch	1486	N 46°44'50.13" E 010°4'28.92"	O. speciosa
VAL		Austria	Valentinshalm	1540	N 46°36'29.48" E 012°57'18.89"	O. alpestris / O. speciosa
VD1		Switzerland	Col des Mosses	1843	N 46°23'13.93" E 007°07'48.60"	O. alpestris / O. speciosa
VS2		Switzerland	Sanetsch	1680	N 46°18'26.66" E 007°20'07.46"	O. speciosa
VS4		Switzerland	Saas Almagell	1620	N 46°06'15.76" E 007°56'57.93"	O. speciosa
VS5		Switzerland	Les Haudères	1436	N 46°04'52.00" E 007°30'18.10"	O. speciosa
VS6		Switzerland	Emosson	1944	N 46°03'55.30" E 006°55'43.90"	O. speciosa
VS7		Switzerland	Chandolin	2000	N 46°14'41.85" E 007°36'10.01"	O. speciosa
VS8		Switzerland	La Fouly	1571	N 45°56'10.20" E 007°05'36.10"	O. speciosa

Table 1. continued

POP.	RANGE	COUNTRY	LOCATION	ALT.	GPS	SPECIES
ABZ1	**Apennines**	Italy	Vado di Sole	1650	N 42°23'51.70" E 013°47'19.40"	O. alpestris
ABZ2		Italy	Sarnano1	1333	N 42°58'57.40" E 013°15'00.20"	O. alpestris
ABZ5		Italy	Sarnano2	1431	N 43°01'27.90" E 013°13'25.20"	O. alpestris
IT7		Italy	Passo del Penice	1141	N 44°47'26.20" E 009°18'12.70"	O. alpestris
SW3	**Black Forest**	Germany	Zastler	1068	N 47°54'13.40" E 007°58'58.00"	O. alpestris
EUR	**Cantabrics**	Spain	Pico de Europa	1295	N 43°07'40.44" W004°52'38.82"	O. alpestris
CZ4	**Carpathians**	Czech Republic	Karlova Studanka	1267	N 50°04'13.30" E 017°14'42.30"	O. alpestris
PL2		Poland	Zakopane	1016	N 49°16'31.80" E 019°51'12.10"	O. alpestris
RO1		Romania	Monti Rodnei	1111	N 47°35'54.90" E 024°55'21.20"	O. alpestris
RO2		Romania	Sinaia	1386	N 45°21'26.50" E 025°31'05.70"	O. alpestris
HR2	**Dinarics**	Croatia	Risnjak	1402	N 45°25'39.50" E 014°37'19.40"	O. speciosa
HR4		Croatia	Sjeverni Velebit	1422	N 44°48'28.00" E 014°58'14.70"	O. speciosa
MON2		Montenegro	Savnik (Slatina)	1363	N 42°59'55.50" E 019°09'58.00"	O. alpestris
SER1		Serbia	Kopaonik	1700	N 43°20'32.40" E 020°46'01.90"	O. alpestris
FR1	**Jura**	France	Cret de la Neige	1715	N 46°15'04.96" E 005°55'28.54"	O. speciosa
JU1		Switzerland	Undervelier	550	N 47°17'57.06" E 007°13'24.32"	O. speciosa
NE1		Switzerland	Motiers	760	N 46°54'09.63" E 006°36'59.38"	O. speciosa
SO1		Switzerland	Weissenstein	1274	N 47°15'03.40" E 007°29'07.70"	O. speciosa
FR8	**Massif Central**	France	Puy Mary	1550	N 45°06'40.90" E 002°40'51.80"	O. speciosa
FR9		France	Puy de Dome	1292	N 45°46'07.70" E 002°57'33.70"	O. speciosa
FR3	**Pyrenees**	France	Vernet-les-Bains	1737	N 42°30'00.20" E 002°24'30.40"	O. alpestris
FR4		France	Col de Port	1334	N 42°53'39.10" E 001°26'50.40"	O. alpestris
FR5		France	Col de Peyresourde	1548	N 42°47'50.00" E 000°27'10.60"	O. ganglbaueri
SP1		Spain	Salardu	1636	N 42°39'57.90" E 000°55'06.70"	O. alpestris / O. ganglbaueri

Table 2. Contingency table showing the number of individuals cross-classified by morphological categories, mtDNA and nuclear phylogenetic clades.

		mtDNA clustering							
***ITS2* clustering**	**genitalic morphology**	M1	M2	M3	M4	M5	M6	M7	**Total**
N0	alpestris α			1			5		**6**
	alpestris β						15		**15**
	speciosa			9	11	2	19	61	**102**
Total N0				10	11	2	39	61	**123**
N1	ganglbaueri		5						**5**
Total N1			5						**5**
N2	alpestris α	8		1	6	10			**25**
	alpestris β					5			**5**
Total N2		8		1	6	15			**30**
Total		**8**	**5**	**11**	**17**	**17**	**39**	**61**	**158**

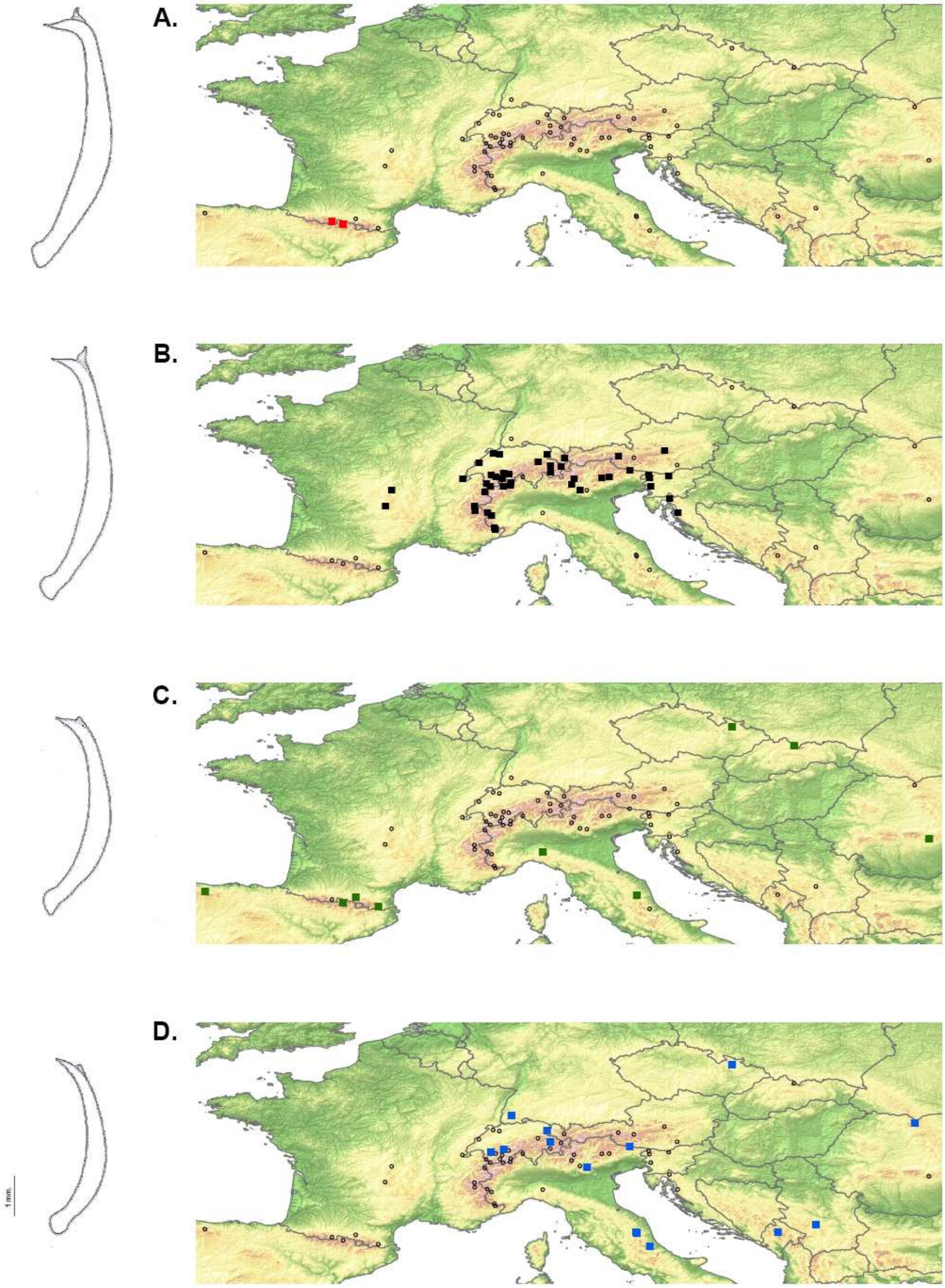


Figure 1: Drawings of the four genitalia morphotypes examined in this study with their respective spatial distributions. A. type *ganglbaueri*; B. type *speciosa*; C. type *alpestris* β; D. type *alpestris* α.

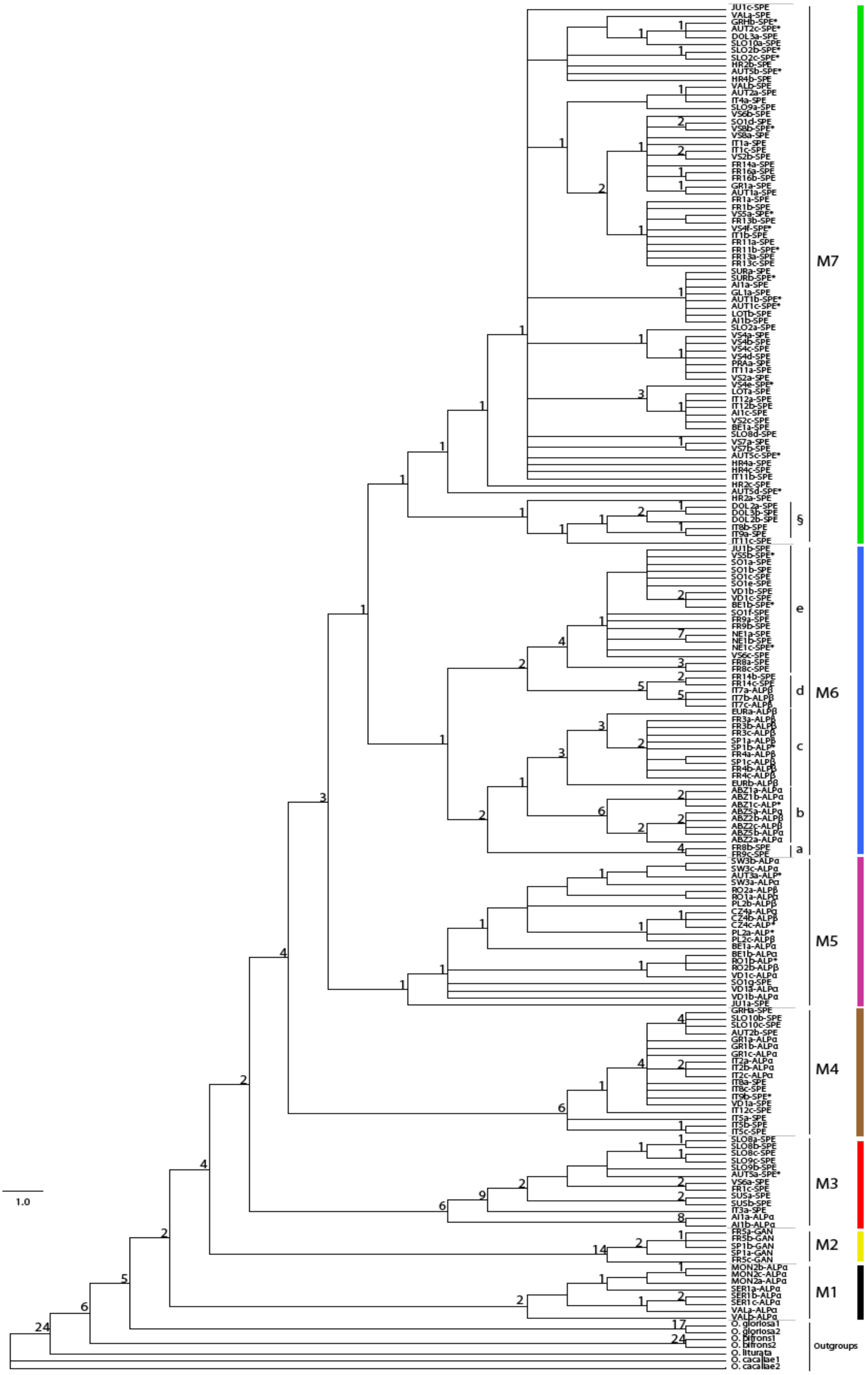


Figure 2: Majority-rule consensus tree of the MP analysis for the three mtDNA regions *COI*, *COII* and *16s rRNA* (see complete text above for details in this figure)

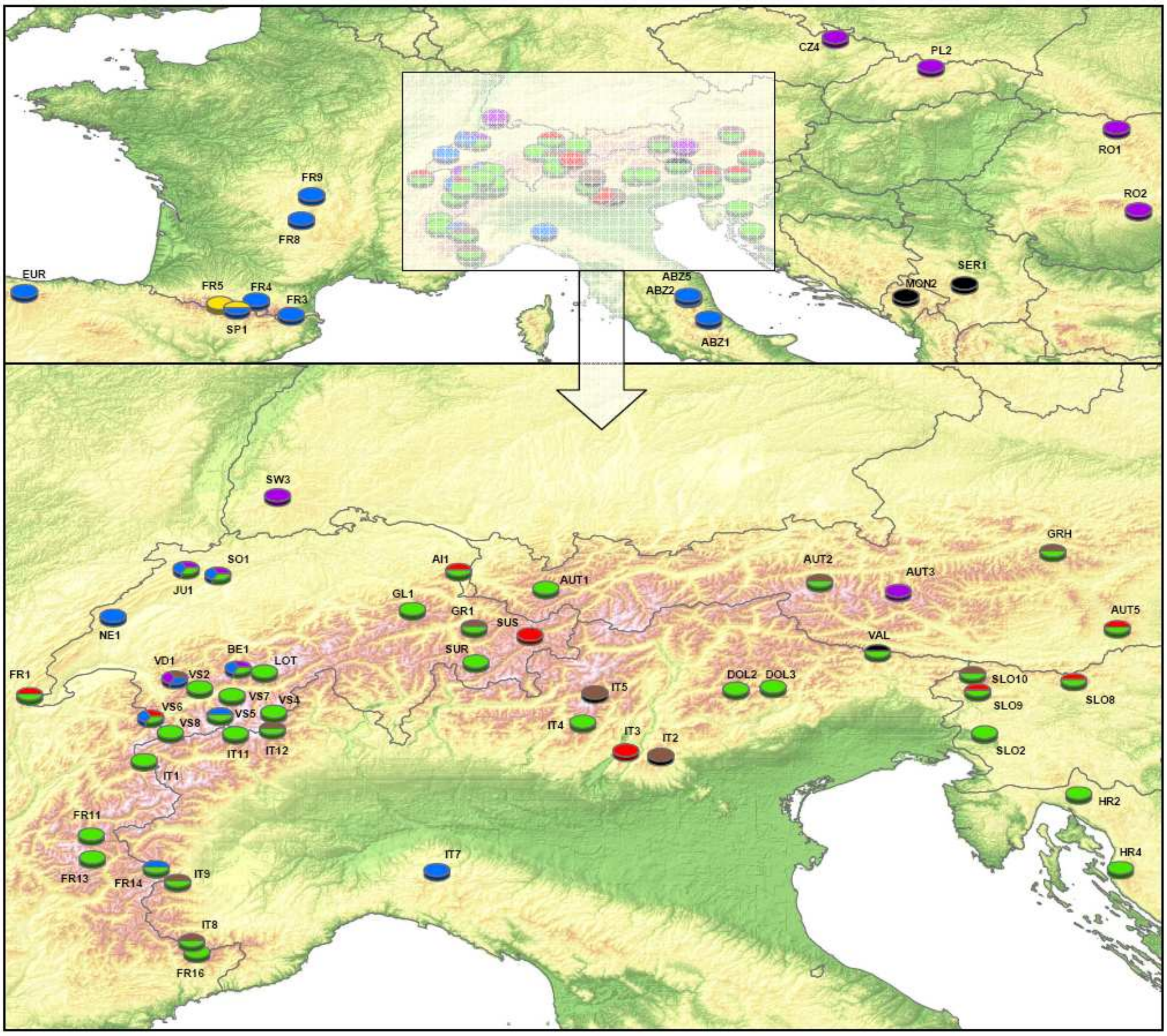


Figure 3: Geographical distribution of each mtDNA clade identified in the MP analysis. Colors refer to clades identified in Figure 2 and codes correspond to populations detailed in Table 1.

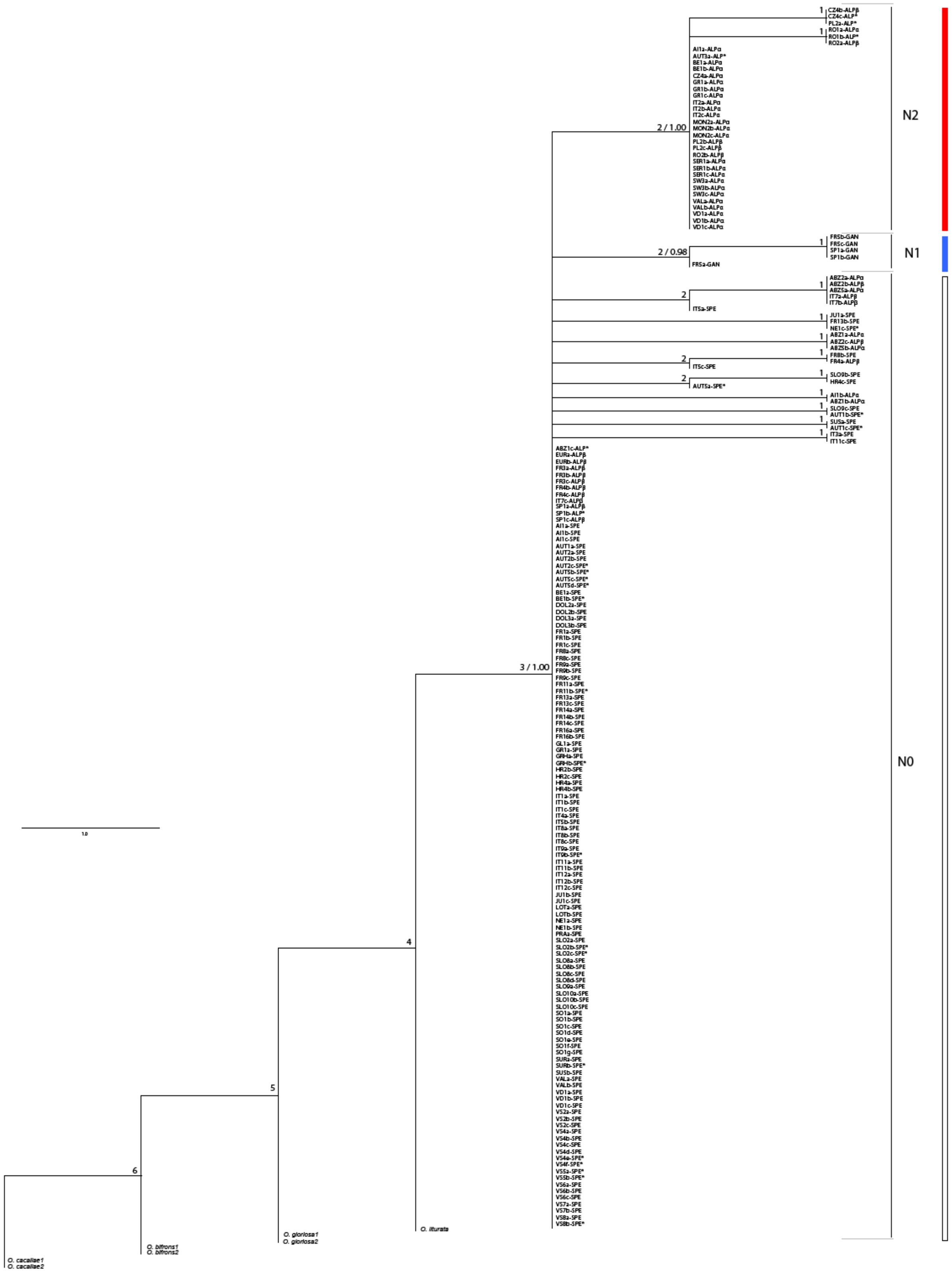


Figure 4: Majority-rule consensus tree (equal branch-length cladogram) of the MP analysis for the nuclear region ITS2. Label names are identical to Figure 2. Bremer support values are indicated on each node. The two main clades are defined with codes N1 and N2. Specimens that do not fall in a well-supported clade (N1 or N2) are considered as belonging to group N0.

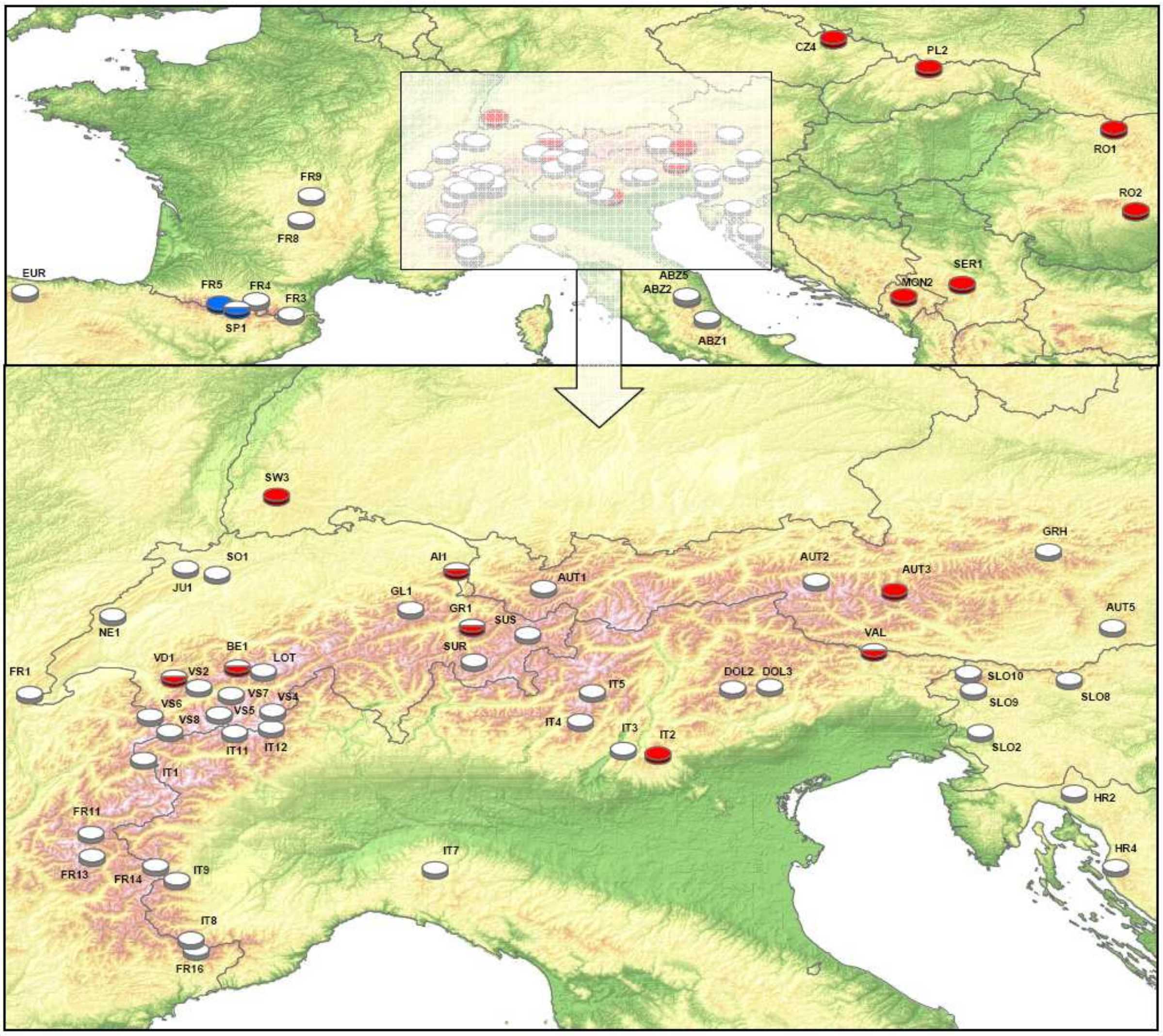


Figure 5: Geographical distribution of each nuclear clade identified in the Maximum Parsimony analysis. Colors refer to N1 and N2 clades, as identified in Figure 4. Samples in white correspond to the loosely defined N0 group.

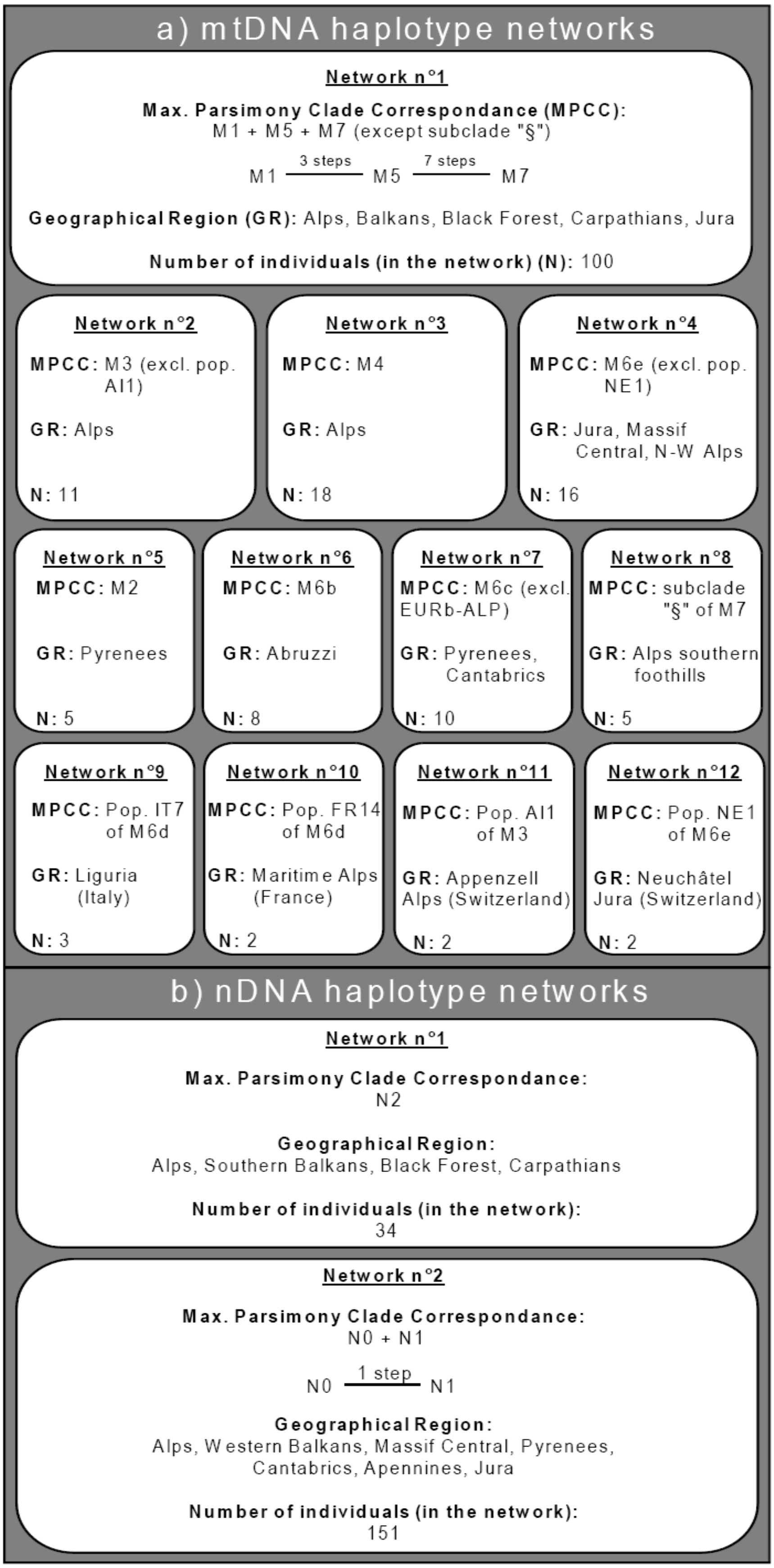
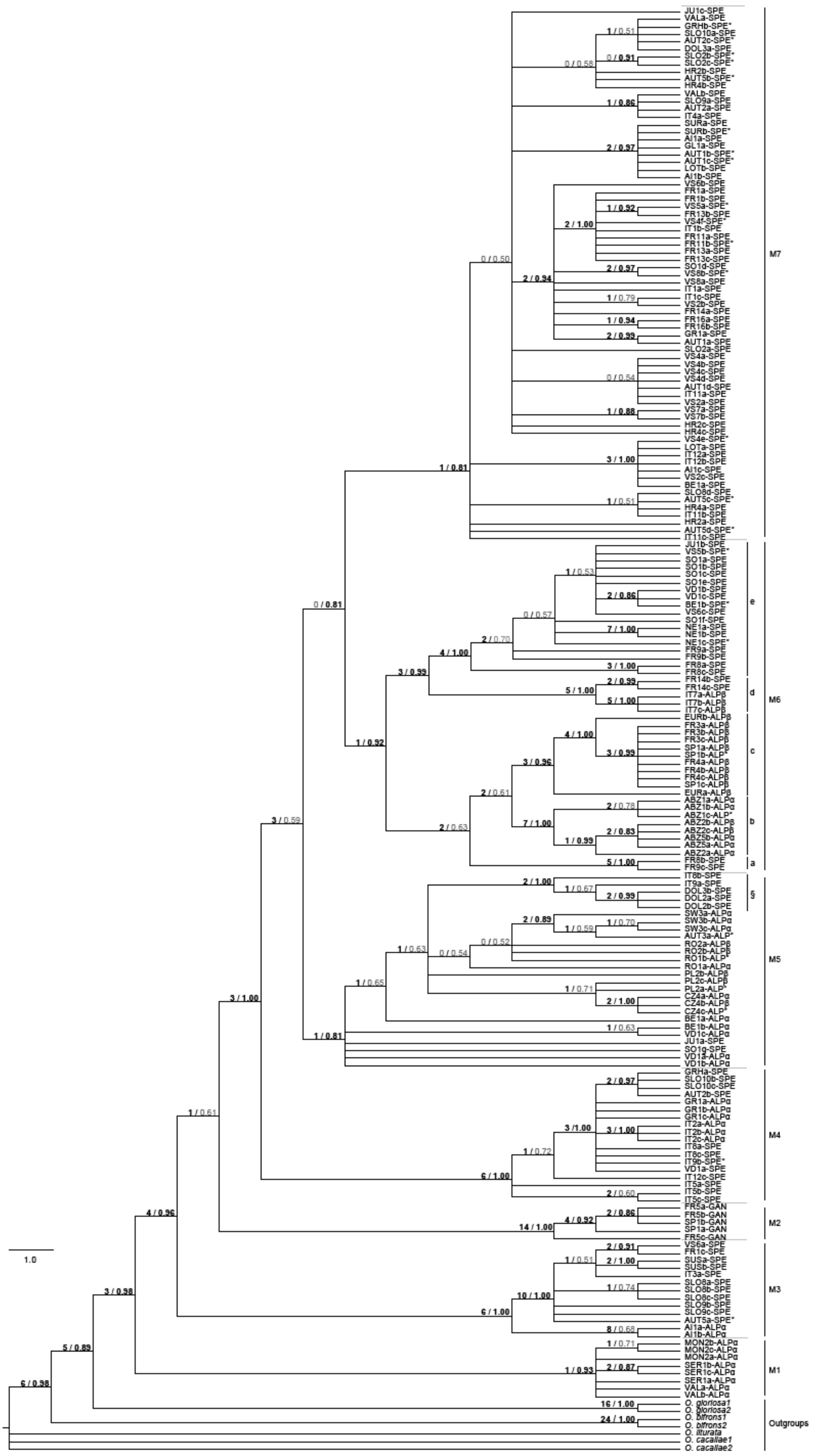
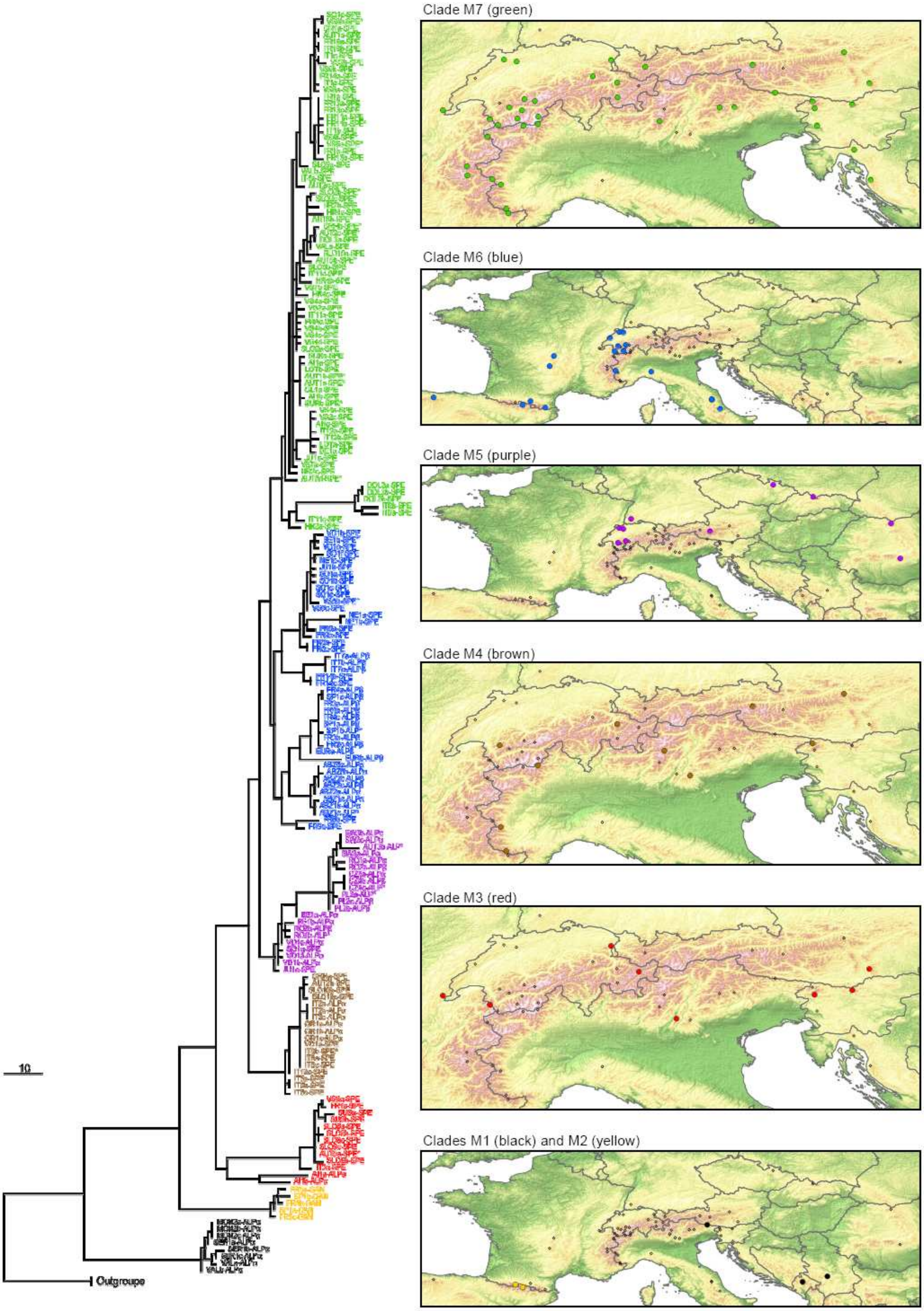


Figure 6: Schematic representation of the mtDNA (a) and *ITS2* (b) haplotype network analyses. Each box represents an independent network. In each box is indicated the correspondence with phylogenetic clades (as shown in Figures 2 and 4), the geographical origin and the number of samples.



Supplementary file #1: Half-compatible topology of the MrBayes analysis, with the corresponding Bremer support values and BPP on each node. Major clades from the MP analysis are shown to the right of the topology. The group indicated by § switched position between the MP and Bayesian analyses (see Fig.2).



Supplementary file #2: One of the 125 equally parsimonious trees, highlighting mtDNA clades M1 to M7, with their corresponding spatial distribution on the maps.

CHAPTER II

Hidden and visible diversity in the fly orchid (*Ophrys insectifera*): respective contributions in defining evolutionary significant units

Yann Triponez, Bertrand Schatz and Nadir Alvarez

(Version formatted according to *Biological Conservation*)

Neuchâtel, August 2010

Hidden and visible diversity in the fly orchid (*Ophrys insectifera*): respective contributions in defining evolutionary significant units

Triponez Y*, Schatz B[†] and Alvarez N[§]

* corresponding author: yann.triponez@unine.ch

Laboratoire d'Entomologie Evolutive, Institut de Biologie

Université de Neuchâtel, Rue Emile-Argand 11, cp 158

CH-2009 Neuchâtel, Switzerland

Phone +41 32 718 31 62 - Fax +41 32 718 30 01

[†] Centre d'Ecologie Fonctionnelle et Evolutive, CNRS Montpellier

UMR 5175, 1919 Route de Mende

F-34293 Montpellier cedex 5, France

Phone +33 04 67 61 33 00 - Fax +33 04 67 41 21 38

[§] Departement d'Ecologie et Evolution, Faculté de biologie et de médecine

Université de Lausanne, Biophore

CH-1015 Lausanne. Switzerland

Phone +41 21 692 41 60 - Fax +41 21 692 41 65

Short title: Genetic diversity within the *Ophrys insectifera* group

Number of words in abstract: 248

Number of words in the text (including literature cited): 5954

Abstract

Although pollination biology of *Ophrys* (Orchidaceae) has stimulated interests of both scientists and amateurs, genetic identity of most endemic species has not been clearly defined until now. Paradoxically several of these taxa are rare or even considered as threatened. The fly orchid group comprises the European widespread *O. insectifera* and its two sister species *O. subinsectifera* and *O. aymoninii*, endemics from northern Spain and southern France respectively. Each of these species possesses its own morphological and chemical characteristics, correlated to specific pollinator dependence. Extensive molecular data to evaluate their genetic identity and confirm species delimitations are however missing. Such data may allow determining patterns of genetic diversity, predicted to increase in the group's distribution zone where endemics occur.

Our analysis of fifty-nine populations results in only partial and weak genetic support of the phenotypic recognized species, even with fine-scale methods such as AFLP. Nei's genetic diversity, locally inferred from the whole dataset, unexpectedly shows highest values in other regions than those enclosing endemics. Finally, rare cpDNA haplotypes occurring in some easternmost populations attest for additional diversity hotspots, raising the level of unpredictability in genetic diversity patterns. As a first attempt for defining evolutionary significant units in the fly orchid group, we propose focusing conservation efforts in several areas of interest for either ecological or genetic reasons. Our study shows that, as probably for other orchid species, optimized conservation means should combine visible with hidden clues of diversity in order not to neglect important regions when developing conservation strategies.

Keywords

Ophrys insectifera, ESUs, endemics, genetic diversity, taxonomical status, conservation genetics

1. Introduction

Investigating the genetic diversity of taxa is fundamental to understand concurrent micro- and macro-evolutionary processes (Bickford et al., 2004) and to provide a relevant framework for conserving not just patterns but also processes at different biogeographical scales (Fady & Conord, 2009; Moritz, 2002). Despite the species level is the natural taxonomic rank forming the basis for conservation assessments at most spatio-temporal scales (Mace, 2004), its boundaries are often difficult to evaluate consistently (Petit and Excoffier, 2009). Consequently, conservation strategies should rely on a framework different from that of an unrealistic single species concept (for a review on the "species problem", see Hey, 2001). Evolutionary Significant Units (ESUs) might be an alternative for this issue in conservation biology (Moritz, 1994), and a consensus has recently been proposed (Hey et al., 2003). Accordingly, ESUs should be defined to maximize the potential for evolutionary success and therefore preserve adaptive diversity across the range of the taxon (Mace, 2004). Since it is often not possible to rely on a classical species concept in terrestrial plants, ESUs have been applied for conservation purposes in several hotspot biomes (e.g., Burke, 2005; Rouget et al., 2003), as well as in Mediterranean tree communities (Fortuna et al., 2009). Nonetheless, other groups might also benefit from such an approach, especially those in which species limitations are not obvious.

Orchidaceae is a highly diverse plant family, with an actual estimation of 25.000 species (Chase et al., 2003) distributed in a wide range of habitats over all continents. Despite a remarkable evolutionary success, the abundance of orchids is believed to have fallen to critical levels for half a century (Kull et al., 2006). In Europe, their decline started around 1950 when the strong anthropogenic pressure increased the impact and magnitude of modifications in land use. For instance, numerous European orchid species have suffered from the reduction in the area of dry grasslands and the increase in woodland floor shading (Dorland and Willems, 2002), as well as from a high rate of habitat fragmentation (Jacquemyn et al., 2007; Young et al., 1996). For these reasons, many conservation efforts are conducted for the protection of orchids. Because of the emblematic nature of Orchidaceae, some European taxa were largely oversplit and endemic forms were often upgraded at the species rank, although very few relevant synapomorphic features were described. As a consequence, an exaggerated number of orchid taxa was recognized in the recent decades (Pillon and Chase, 2007). This increasing number of recognized species may represent a serious obstacle to conservation, as some taxa may receive conservation attention, simply because of regional biases in the splitting level of taxonomical entities. While splitting not always relies on genetic assumptions, conservation of well-defined ESUs in European orchids represents an important challenge.

This statement is particularly true for species among *Ophrys* L., an emblematic 250 species-rich genus (representing approximately one third of the European orchid flora; Delforge, 2001). Within this genus famous for its pseudocopulatory pollination strategy involving Hymenoptera males (one among the different food-deceptive pollination syndromes occurring in approximately one-third of all Orchidaceae [Jersakova et al., 2006; Peakall, 2007]), most species have been described based on extreme fine-scale morphological criteria (most of the time leading to pollinators specificities). So far genomic sequencing generally failed to discriminate strong genetic structure among *Ophrys* species, as showed in previous phylogenetic studies based on sequence data (Bateman et al., 2003; Devey et al., 2008; Soliva et al., 2001). A recent genetically-based approach of a large number of *Ophrys* species, combining Amplified Fragments Length Polymorphisms (AFLPs) fingerprinting and gene sequencing, revealed that many species show only little (or even no) genetic variation among them (Devey et al., 2008). The authors therefore proposed to pool together *Ophrys* taxa into genetically well-established groups or aggregates. Therefore in *Ophrys*, where all species are theoretically inter-fertile but generally isolated by specific pollinators acting as a prezygotic barrier (Ågren and Borg-Karlson, 1984; Borg-Karlson et al., 1993; Kullenberg, 1961; Van der Cingel, 1995), considering ESUs for conservation purposes seems particularly relevant.

In the present study, we aim at defining realistic ESUs in the fly orchid group, applicable for conservation priorities. Our approach combines both ecologically and genetically-based clues using AFLP genome-fingerprinting in addition to classical sequencing. The fly orchid group includes three phenotypically distinct taxa: the widely distributed *O. insectifera* L., and the two endemics *O. aymoninii* BREISTROFFER and *O. subinsectifera* HERMOSILLA & SABANDO. This typical pattern of one (or a few) largely distributed species and several endemics is found in all other *Ophrys* groups as defined by Devey et al. (2008). The fly orchid group holds the advantage of being conceptually easier to handle than the other *Ophrys* aggregates, because it is not highly threatened and comprises only three distinctive species. Our objective is to shed light on two main questions: 1) do the two endemic species show genetic differences as suggested by their distinct phenotypes? 2) do regions of endemism comprise higher levels of genetic diversity (as expected in the south-western part of the fly orchid group's distribution, where endemics overlap with the widespread species)? Even if not strictly considered as an endangered taxa (see Methods section), results and conclusions drawn in the ESUs framework for this particular group might then be useful for sketching studies within other *Ophrys* aggregates, providing the required knowledge for establishing priorities in the preservation of this emblematic orchid genus.

Results and conclusions drawn in the ESUs framework for this particular group of species (even if not strictly considered as endangered taxa; see Methods section) might provide the required knowledge for establishing conservation priorities within other *Ophrys* aggregates, for the future preservation of this emblematic orchid genus.

2. Methods

2.1 Taxonomy, distribution, ecology and protection status of the fly orchid group

The most widespread species within the fly orchid group is *O. insectifera* L., whose distribution covers Europe, from Spain to the Balkans, and from Italy to southern Scandinavia, Baltic countries and British Isles (Delforge, 2001; Wolff, 1951). This species has the most widespread and the most northern distribution within the *Ophrys* genus (Delforge, 2001). It grows in heterogeneous habitats, from dry grasslands to marshes, or in the understorey of pine and temperate deciduous forests. Males of the wasp *Argogorytes mystaceus* (Hymenoptera, Sphecidae) are reported as the main pollinators of *O. insectifera* (Ågren & Borg-Karlson 1984; Borg-Karlson et al. 1993; Kullenberg, 1951, 1961; Kullenberg & Bergström 1976; Wolff, 1950; Triponez Y. and Schatz B. pers. obs.). With a much narrower distribution, *O. aymoninii* is restricted to a zone covering the karstic southern French Massif Central (Breistroffer, 1981; Delforge, 1983). Its main morphological difference with *O. insectifera* is found in the wider and shorter (in length) flowers with yellow-edged labella and green-yellowish petals. It is pollinated by the bee *Andrena combinata* (Hymenoptera, Andrenidae) (Borg-Karlson et al., 1993; Delforge, 2001; Vereecken, 2009). Finally, *O. subinsectifera* is a Spanish vicariant growing in the southern foothills of the Pyrenees (Amardeilh, 1996). As for *O. aymoninii*, it looks quite alike *O. insectifera*, differing from the latter by smaller flowers, reduced lateral lobes of labella and petals with a greenish tip. The pollinator of *O. subinsectifera* has been described as the sawfly *Sterictiphora gastrica* (Hymenoptera, Argidae) (Hermosilla et al., 1999). As a consequence of pollinator specificity, all three species emit well differentiated flower volatiles involved in the attraction of specific pollinators belonging to different genus (Borg-Karlson et al., 1993; Vereecken, 2009).

Members of the fly orchid group do not represent strictly speaking endangered taxa at the European scale and are not included in the general IUCN red list. However, *O. insectifera* is listed in several national red lists (e.g. in Bulgaria, Denmark, Estonia, Finland, Norway) with different threats level, and is also regionally protected (e.g. in Auvergne, France). *Ophrys aymoninii* is red-listed as vulnerable in France and only protected in the Aveyron department. Finally *O. subinsectifera* does not appear in the national Spanish red list. Actually, recent observations revealed that several populations of *O. aymoninii* and *O. subinsectifera* disappeared because of direct human impact in land use change leading to a progressive landscape closure (Schatz and Jacob, 2008).

2.2. Sampling

Leaves of *Ophrys insectifera, O. aymoninii* and *O. subinsectifera* were collected across most of its respective distribution, sometimes in sympatric populations within the distribution range of the two endemics (Table 1 and Fig. 2). The sampling was completed during flowering periods in springs 2007 and 2008 (from April to June depending on the regions). Within zones of endemism, we collected only plants showing morphological features typical of a defined species and discarded putative hybrids. Outgroups for phylogenetic analyses were collected when found in sympatry with specimens from the fly orchid groups and consisted of *O. araneola* REICH, *O. holoserica* (BURM.FIL) GREUTER, *O. lutea* (GOUAN) CAV. and *O. fusca* LINK. In order to minimize the impact of our sampling on the fitness of the plants, only one leaf per individual plant was cut. Fresh plant material was desiccated in silica gel according to Chase and Hills (1991).

2.3 DNA extraction, sequencing and alignment

Total genomic DNA was extracted from 10 mg of silica-dried leaf fragments using the DNeasy Plant Kit (Qiagen, Hilden, Germany). In a selection of 50 populations spanning the distributions of the three species within the *O. insectifera* group, two non-coding and fast-evolving cpDNA regions (Shaw, 2007) were amplified: 1) *ndhA* intron using primers ndhAx1 and ndhAx2 and 2) *psbJ-petA* intergenic spacer using primers psbJ and petA. Fragments were amplified using a standard 30 µl PCR mix including: 1.5 µl of extracted DNA, 3 µl of 10X PCR buffer (Promega, Madison, USA), 3 µl of MgCl2 solution (25 mM), 3 µl of dNTPs (1.5 mM), 1.5 µl of forward and reverse primers (Microsynth, Balgach, Switzerland), 0.3 µl of Taq DNA polymerase (Promega, Madison, USA), all made up to a final volume of 30 µl with purified MilliQ water. The PCR were run in a TGradient thermocycler (Biometra, Goettingen, Germany) using the following program for both cpDNA regions: initial denaturation of 1 min 30 s at 94℃, 35 cycles of elongation (35 s at 94℃, 1 min at 52℃, 45 s at 72℃), and a final elongation of 8 min at 72℃. The PCR products purification and sequencing was carried out by Macrogen Inc. (Seoul, South Korea). Sequencing was performed using forward primer *psbJ* for *psbJ-petA* region and reverse primer ndhAx2 for *ndhA* intron under BigDye TM terminator cycling conditions, purifying the products using ethanol precipitation and running them into an Automatic Sequencer 3730xl (Applied Biosystems, Foster City, USA). Sequences were manually corrected using the software Chromas Pro 1.34 (Technelysium, Helensvale, Australia).

2.4 Chloroplastic sequences alignment and haplotype network analysis.

Alignment was carried out for each chloroplastic region using ClustalW Multiple Alignment (Thompson et al., 1997) as implemented in the software BioEdit 7.0.5.3 (Hall, 1999), followed by minor manual corrections. Gaps were coded using FastGap 1.2 (Borchsenius, 2009), following the method of Simmons and Ochoterena (2000), and appended to the DNA matrix as a supplementary partition. Phylogenetic inference was performed according to Triponez et al. (2007). A haplotype network was then performed on the combined dataset of both regions using statistical parsimony network (SPN) as implemented in TCS 1.21 (Clement et al., 2000). Analysis was carried out by applying a 99% connection limit, with gaps treated as missing data.

2.5 AFLP procedure and analysis

AFLP analysis followed the classical method described by Vos et al. (1995), with slight modifications (detailed protocol available upon request to the first author). Reactions were conducted in 96-well plates, in which samples were randomly distributed. We used restriction enzymes *Eco*RI and *Mse*I for DNA digestion. A primer trial on a small number of samples from different species and origin was conducted using 16 primer combinations to identify pairs of selective primers that were repeatable and polymorphic enough within the fly orchid group. Each individual sample was fingerprinted with the two primer combinations *Eco*RI-ACAG/*Mse*I-CAA and *Eco*RI-ACAG/*Mse*I-CTG. Final selective PCR products were analysed using the GeneScan technology with an automated capillary-sequencer (ABI 3730XL, Applied Biosystems, Foster City, CA; service provided by Macrogen Inc., Seoul, South Korea). Resulting fluorescent AFLP patterns were scored using GeneMapper 3.7 (Applied Biosystems) with some changes in default parameters (maximum bin width=1 bp, light peak smoothing, peak threshold=200 RFU). We scored the presence or absence of each AFLP marker in each individual plant. The matrices of the two scored primer pairs were concatenated into one binary matrix where individuals and bands were stored as lines and columns, respectively. A modified Nei-Li distance matrix was computed using the RESTDIST program (PHYLIP 3.64 package; Felsenstein, 2009). Multivariate ordinations using the principal coordinate analysis (PCoA) approach were produced on this distance matrix to investigate genetic relationships among specimens and taxa using the software Ginkgo 1.5.8 (Bouxin, 2005). We further performed a Predictive Discriminant Analysis (PDA), which aims at allocating new cases (observations) to previously defined groups, assuming that a set of variables can predict group membership (Klecka, 1980). In order to account for the drawback of using the same set of samples both to define and predict group membership, we used the unbiased “Leave-One-Out” algorithm (Huberty, 1994; Hwang, 2001), as implemented in Ginkgo 1.5.8.

2.6 Genetic diversity analyses

An analysis of molecular variance (AMOVA) was performed with GenAlEx 6.3 (Peakall and Smouse, 2006) in order to estimate the proportion of genetic variation explained by the within- and among-populations levels, as well as by the among-species level. Percentages of polymorphic loci and Nei's unbiased expected heterozygosities (i.e., an analysis correcting for the number of specimens analysed per population) were also calculated with GenAlEx 6.3 for each population represented by at least two individuals. Values for Nei's unbiased heterozygosities were categorized into five classes and displayed geographically using ArcGIS 9.1 (ESRI).

3. Results

3.1 DNA sequencing

Amplification of chloroplatic regions was successful for 169 and 166 individuals for *ndhA* intron and *psbJ-petA* intergenic spacer, respectively. The alignment yielded a 1582 bp supermatrix for the two cpDNA regions: 821 bp for *ndhA* intron (one parsimony-informative site among 11 polymorphic sites in the ingroup) and 761 bp for *psbJ-petA* intergenic spacer (two parsimony-informative sites among two polymorphic sites in the ingroup). Our phylogenetic hypothesis based on the supermatrix comprising both *ndhA* intron and *psbJ-petA* distinguished with high support the monophyly of the fly orchid group, well segregated from all outgroup species (result not shown). However, genetic variation was very low within the 162 samples of the *O. insectifera* group (i.e., only three parsimony-informative sites among the two cpDNA regions). Figure 1 illustrates the results obtained with the SNP haplotype network analysis: five distinct haplotypes were detected, two frequent (i.e., hereafter "main") and three rare (i.e., hereafter "secondary"). Among the 45 populations represented by at least two sequenced individuals (out of 50 analysed in total), 37 presented one single haplotype (82%), 12 comprised two (27%), and one (i.e., the Polish population SZC) was even composed of three haplotypes. The two main haplotypes (i.e., main haplotypes A and B, hereafter MhA and MhB) were connected by two steps. Together they held more than 90% of all individuals analysed for the *O. insectifera* group, and did not circumscribe the three taxa *O. insectifera*, *O. aymoninii* or *O. subinsectifera*. Both main haplotypes showed large distributions across Europe (Fig. 2). While MhA occurred more frequently in Western Europe (and was even absent from the Balkans), MhB seemed mostly limited to middle and lower latitudes (with the notable exception of the Swedish population SJO). Only 14 individuals (9%) possessed rare haplotypes (i.e., secondary haplotypes A, B and C, hereafter ShA, ShB and ShC). Each one was connected by one single step to MhA. These rare haplotypes occurred mainly in Eastern Europe (Czech Republic, Poland, Croatia, Bulgaria and Greece, see Fig. 2) as well as in a Swiss population (PLA). They were found either in spatially distant populations (in Bulgaria and Switzerland for ShA), or in a limited area of the distribution (Balkans and Carpathians for ShB). Finally, ShC was unique to a single Czech population (OPO). Among the six populations presenting rare haplotypes, three (PLA, SZC and YAG) also comprised frequent haplotypes (see Fig. 2).

3.2 AFLP fringerprinting

The two primer combinations generated 325 AFLP markers, each of the 240 specimen (sampled from 65 populations) yielding between 155 and 194 alleles. The AMOVA performed on the complete AFLP dataset showed that the largest proportion (87%) of the genetic variation was explained by the within-population level, while the among-population (10%) and among–species (3%) levels accounted for a much lower proportion. This result was confirmed by the PCoA analysis that did not show any partitioning between species when considering the first three axes (explaining respectively 9.91%, 9.71% and 6.18% of total variance, i.e., 25.79% of cumulated variance; see Fig. 3A). Nevertheless, a slight species-effect was observed when projecting the fourth and seventh axes (explaining 5.82% and 4.18% of total variance, respectively) (Fig. 3B). Axis four partially discriminated about half of the *O. aymoninii* samples, here segregated from both *O. insectifera* and *O. subinsectifera*. Axis seven showed a slight trend for a demarcation of *O. subinsectifera*, even if all samples were still completely included into the wide *O. insectifera* scatter-plot. The “Leave-One-Out” PDA analysis resulted in low percentages of correct samples attribution, with an overall 80.42% matching in the re-assignment of all the specimens analysed. This result lies far beyond the 95% threshold, indicating that predicting the identity of a specimen, based on a model obtained with the *a priori* knowledge of the species to which it belongs, does not provide an accurate estimate of its species membership. There was however a trend for a better performance of the assignment model in *O. aymoninii* (87.50%) than in *O. insectifera* (76.87%) and *O. subinsectifera* (76.92%).

3.3 Genetic diversity

The percentage of polymorphic loci and the unbiased expected heterozygosity (hereafter referred as “PP” and “UHe”) ranged from 15.38% to 48.92% and from 0.085 to 0.187, respectively (Table 1 and Fig. 4). Average UHe genetic diversity measured over all loci and all populations was 0.141. Considering the low number of analysed individuals in several populations, diversity indices could be confidently deduced only from populations with a number of samples ≥ 5. Minimum values for both PP and UHe were located in the population BEL (31.4%, 0.130), and maximum values were found in populations VIL (48.9%) for PP and in the two populations THU and TRI (0.187) for UHe.

4. Discussion

4.1 The pitfall of endemic taxa: when phenotypes and genetics do not match

Even if phenotypic differences associated with distinct flower shapes and pollination biologies (Vereecken, 2009) justify the description of the two endemics *O. subinsectifera* and *O. aymoninii*, only weak genetic evidences compatible with such phenotypes could be addressed in our results. Although sequencing of highly variable cpDNA regions did well in recognizing strong genetic entities within *Ophrys* (i.e. the four outgroups; data not shown) and managed to detect five different haplotypes in the fly orchid group (see below), no characteristic haplotype for the two endemic taxa was detected. One could expect that AFLP fingerprinting would have performed better in identifying some structure in the genetic variation within the fly orchid, as demonstrated in recent studies of several *Ophrys* species' groups (Devey et al., 2009; Gögler et al., 2009; Schlüter et al., 2007) or of other taxonomically challenging genera of European orchids (e.g. Hedren et al., 2001; Pfeifer et al., 2009). In our study, this was however not the case: on the one hand, only 3% of the genetic variation was explained by the among-species level in the AMOVA; on the other hand, the PDA did not succeed in correctly re-assigning specimens to species. Nonetheless, as shown by the PCoA results on minor axes, *O. aymoninii* might show a slight trend towards genetic circumscription, as it partly segregated from the other two taxa (see Fig. 3); this trend is also illustrated by the PDA, in which *O. aymoninii* shows a higher re-assignment percentage than *O. insectifera* and *O. subinsectifera*.

From a general perspective, our results do however not attest for a strong genetic distinction between the two endemics and the more widely distributed *O. insectifera*. Different reasons might explain why the endemic taxa, well-segregated both ecologically and morphologically, did not correspond to strong genetic entities: one might consider incomplete sorting of ancestral lineages, among-taxa gene flow despite specialization on different pollinators, or a very recent adaptation and ongoing divergence. Whatever hypothesis is considered, further investigation combining ecological, morphological, functional and genetic clues through a macroecological approach (e.g. Kark et al., 2009; Pfeifer et al., 2009) will be necessary to fully understand the processes leading to the current evolutionary ecology patchwork in the fly orchid group. In view of the weak discrimination between taxa, we have chosen to further consider the three taxa as a single entity (equivalent to *O. insectifera* sensu lato) when comparing diversities among populations.

We first hypothesized that the highest levels of genetic diversity should be hosted in regions where *O. insectifera* sensu stricto is found in sympatry with one or the other endemic (i.e., in the South of France and in the North-Eastern part of Spain). Such a pattern was only partially confirmed by our results, with high genetic diversity values also found in the Southern Alps or in the Balkans for instance. There, despite smaller number of samples per population, the diversity values equalled or even outperformed those found in the south-western edge of the fly orchid group distribution (i.e. where zones of endemism are found). DNA sequencing also addressed the existence of rare cpDNA haplotypes in *O. insectifera* sensu stricto, mostly in the eastern part of its distribution. Both fingerprinting and sequencing results therefore underline the role played by central and Eastern Europe (highlighted in Fig. 4) in accounting for the overall diversity of the fly orchid aggregate. Although zones of endemism of *O. aymoninii* and *O. subinsectifera* should be considered with care in conservation priorities, one must keep in mind that other regions, even lacking the presence of phenotypically-defined endemics, also contribute to the genetic diversity of the fly orchid group.

4.2 Important ESUs for conservation purposes in the fly orchid group

In the light of our results, six remarkable zones (framed in Fig. 4) might be referred to as ESUs (*sensu* Crandall, 2000). Incorporating both ecological and genetic data, these zones are defined according to one (or several) of the following criteria: 1) presence of phenotypic endemic taxa, 2) high levels of genetic diversity and 3) occurrence of rare haplotypes.

Frames A and B (see Fig. 4) represent areas in which the widespread *O. insectifera* sensu stricto is sympatric with endemic taxa. Many populations within these two frames are associated with high levels of genetic diversity. South of France (i.e., frame A together with the population LAA that lies out of the distribution of *O. aymoninii*) presented an overall high level of genetic diversity. Particular attention should be paid to the populations within this region, especially because they harbour an emblematic narrow endemic of the French flora. The area where *O. subinsectifera* is found (i.e., frame B) showed medium to high diversity levels in populations (especially in the Spanish populations BER and COM). Here again, mostly motivated by the presence of an endemic species of the fly orchid group, populations of the southern foothills of Pyrenees merely deserve conservation interest. Southern and western Alpine regions (i.e., frame C), where high genetic diversity was detected both using AFLP genotyping (i.e., in populations THU and THO) and DNA sequencing (i.e., PLA hosts an additional rare haplotype), are also deserving preservation efforts. In Eastern Europe, the extreme south-eastern part of the Alps and northern Balkans (i.e., frame D) should draw particular attention in conservation management because despite fewer analysed samples in Slovenian and Croatian populations, genetic diversity indices were among the highest values throughout the whole distribution area. One rare haplotype detected in population GOR is an additional argument stressing the importance of this area. The opposite edge of Balkans (i.e., frame E) should also be considered as a conservation priority since it encompasses high levels of diversity, as well as a high proportion of rare haplotypes. Moreover, scarce inventoried populations from Bulgaria and Greece (Antonopoulos, 2009; Tsvetanov et al., 2005) are strongly isolated from the rest of the *O. insectifera* group's distribution area. This should be taken into consideration regarding the protection's status of these far-east *O. insectifera* representatives. Finally eastern populations (i.e., frame F) should not be forgotten by conservation managers: although demonstrating low values of expected heterozygosities (despite a possible bias related to particularly small sampling sizes), several rare chloroplastic haplotypes were detected by DNA sequencing, notably in population SZC, the single population in our study to harbour not less than three different haplotypes, and in population OPO, hosting a private cpDNA haplotype. Such cryptic genetic endemism could stimulate ecologists to investigate if there is coinciding variation in emitted scent and attracted pollinators. Discovering functionally differentiated species would certainly impact on the taxonomy of this *O. insectifera* group, as well as on conservation politics in the concerned countries (Kark et al., 2009).

In conclusion, the present study on the fly orchid group shows that, as probably in other orchid species, optimized conservation means should rely on both biological and genetic knowledge of the most consistent ESUs. Even if the biological apparent diversity of a species is highly variable, for instance in terms of morphology or ecology (Borg-Karlson et al., 1993; Delforge, 2001; Vereecken, 2009), our results demonstrate that considering only such criteria could fail in circumscribing the global diversity of a taxon. Ideally, complementary ecological and genetic approaches performed on the whole distribution of a taxon should be considered, in order to combine visible (e.g., morphology, pollination biology) and hidden (e.g., levels of expected heterozygosity, cryptic genetic endemism) clues of diversity. We hope that these guidelines will help in not neglecting important regions when developing further conservation strategies. The ESU concept, as defined in our study, might represent an alternative for future conservation perspectives in other *Ophrys* species, since the fiercely debated species' concept and questionable taxonomical oversplitting within the genus represent serious obstructions to efficient conservation.

5. Acknowledgments

We thank all scientists and keen amateurs who permitted to locate *Ophrys* populations across Europe: J.E. Arnold, A. Croce, C. Bernard, F. Dabonneville, B. Dolinar, P. Geniez, C.E.

Hermosilla, L. Jacob, A. Jacquet, N. Margraf, J. L. Menos, N. Petrou, D. Průša, G. Scopece, R. Souche, A. Soulié and many other people met here and there for punctual information. The authors are especially grateful to A. Espíndola for continuous help during field sessions and laboratory work, and to A. Sarr for support with AFLP analysis. This research was supported by the Swiss National Science Foundation (grant n° 310 0A0-116778/1), the Parc naturel regional des Grands Causses (grants n° 019792 and n° 040763) , the French Society of Orchidophily (SFO) and the Swiss Orchid Foundation at the herbarium Jeny Renz (Botanical Institute, Basel University).

6. Literature Cited

Ågren, L., Borg-Karlson, A., 1984. Responses of *Argogorytes* (Hymenoptera: Sphecidae) males to odor signals from *Ophrys insectifera* (Orchidaceae). Preliminary EAG and chemical investigation. Nova Acta Regiae Societatis Scientiarum Upsaliensis Serie V:C, 111-117.

Amardeilh, J.-P., 1996. *Ophrys subinsectifera* Hermosilla & Sabando. Une nouvelle orchidée du nord de l'Espagne. L'Orchidophile 123, 149-154.

Antonopoulos, Z., 2009. The bee orchids of Greece - The genus *Ophrys*. Mediterraneo Editions, Rethymno, Greece.

Bateman, R.M., Hollingsworth, P.M., Preston, J., Yi-Bo, L., Pridgeon, A.M., Chase, M.W., 2003. Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). Botanical Journal of the Linnean Society 142, 1-40.

Bickford, S.A., Laffan, S.W., Kok, R.P.J., Orthia, L.A., 2004. Spatial analysis of taxonomic and genetic patterns and their potential for understanding evolutionary histories. Journal of Biogeography 31, 1715-1733.

Borchsenius, F., 2009. FastGap 1.2, Department of Biological Sciences, University of Aarhus, Denmark.

Borg-Karlson, A.K., Groth, I., Agren, L., Kullenberg, B., 1993. Form-specific fragances from *Ophrys insectifera* L. (Orchidaceae) attract species of different pollinator genera. Evidence of sympatric speciation? Chemoecology 4, 39-45.

Bouxin, G., 2005. Ginkgo, a multivariate analysis package. Journal of Vegetation Science 16, 355-359.

Breistroffer, M., 1981. Notes succinctes sur quelques équivalences nomenclaturales d'espèces céveno-caussenardes et description d'une sous-espèce nouvelle d'Orchidée. Bulletin de la société botanique de France. Lettres botaniques 128, 69-72.

Burke, A., 2005. Endemic plants of the arid succulent karoo in Namibia: towards hypotheses for their evolution. Ecography 28, 171-180.

Chase, M.W., Cameron, K.M., Barrett, R.L., Freudenstein, J.V., 2003. DNA data and Orchidaceae systematics: a new phylogenetic classification. Natural History Publications Kota Kinabalu, Sabah.

Chase, M.W., Hills, H.H., 1991. Silica-Gel - an Ideal Material for Field Preservation of Leaf Samples for DNA Studies. Taxon 40, 215-220.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1659.

Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M., Wayne, R.K., 2000. Considering evolutionary processes in conservation biology. Trends in Ecology & Evolution 15, 290-295.

Delforge, P., 1983. Remarques sur *Ophrys insectifera* subsp. *aymoninii* Breistroffer et description d'un hybride nouveau de cette sous-espèce: *Ophrys insectifera* nsubsp. *tytecaeana* Delforge. L'Orchidophile 55, 307-312.

Delforge, P., 2001. Guide des orchidées d’Europe, d’Afrique du Nord et du Proche-Orient, 2nd ed. Delachaux et Niestlé, Paris, France.

Devey, D.S., Bateman, R.M., Fay, M.F., Hawkins, J.A., 2008. Friends or relatives?

Phylogenetics and species delimitation in the controversial European orchid Genus *Ophrys*. Annals of Botany 101, 385-402.

Devey, D.S., Bateman, R.M., Fay, M.F., Hawkins, J.A., 2009. Genetic structure and systematic relationships within the *Ophrys fuciflora* aggregate (Orchidaceae: Orchidinae): high diversity in Kent and a wind-induced discontinuity bisecting the Adriatic. Annals of Botany 104, 483-495.

Dorland, E., Willems, J.H., 2002. Light climate and plant performance of *Ophrys insectifera*; a four-year field experiment in The Netherlands (1998-2001), In Trends and Fluctuations and Underlying Mechanisms in Terrestrial Orchid Populations. eds P. Kindlmann, J.H. Willems, D.F. Whigham, pp. 225-238.

Fady, B., Conord, C., 2009. Macroecological patterns of species and genetic diversity in vascular plants of the Mediterranean basin. Diversity and Distributions 16, 53-64.

Felsenstein, J., 2009. PHYLIP (Phylogeny Inference Package) Version 3.64. Distributed by the author, Department of Genome Sciences. University of Washington, Seattle.

Fortuna, M.A., Albaladejo, R.G., Fernandez, L., Aparicio, A., Bascompte, J., 2009. Networks of spatial genetic variation across species. Proceedings of the National Academy of Sciences of the United States of America 106, 19044-19049.

Gögler, J., Stokl, J., Sramkova, A., Twele, R., Francke, W., Cozzolino, S., Cortis, P., Scrugli, A., Ayasse, M., 2009. Ménage à trois - two endemic species of deceptive orchids and one pollinator species. Evolution 63, 2222-2234.

Hall, T.A., 1999. BioEdit:a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp 41, 95-98.

Hedren, M., Fay, M.F., Chase, M.W., 2001. Amplified fragment length polymorphisms (AFLP) reveal details of polyploid evolution in *Dactylorhiza* (Orchidaceae). American Journal of Botany 88, 1868-1880.

Hermosilla, C.E., Amardeilh, J.P., Soca, R., 1999. *Sterictiphora furcata* Villers, pollinisateur d'*Ophrys subinsectifera* Hermosilla & Sabando. L'Orchidophile 139, 247-254.

Hey, J., 2001. The mind of the species problem. Trends in Ecology & Evolution 16, 326-329.

Hey, J., Waples, R.S., Arnold, M.L., Butlin, R.K., Harrison, R.G., 2003. Understanding and confronting species uncertainty in biology and conservation. Trends in Ecology & Evolution 18, 597-603.

Huberty, C.J., 1994. Applied Discriminant Analyses. Wiley and Sons, New York, USA.

Hwang, D.-Y., 2001. Issues in Predictive Discriminant Analysis: using and interpreting the Leave-One-Out Jacknife Method and the Improvement-Over-Change "I" Index Effect Size. Annual Meeting of the Southwest Educational Research Association. February 1-3, 2001. New Orleans, LA, USA.

Jacquemyn, H., Vandepitte, K., Brys, R., Honnay, O., Roldan-Ruiz, I., 2007. Fitness variation and genetic diversity in small, remnant populations of the food deceptive orchid *Orchis purpurea*. Biological Conservation 139, 203-210.

Jersakova, J., Johnson, S.D., Kindlmann, P., 2006. Mechanisms and evolution of deceptive pollination in orchids. Biological Reviews 81, 219-235.

Kark, S., Levin, N., Grantham, H.S., Possingham, H.P., 2009. Between-country collaboration and consideration of costs increase conservation planning efficiency in the Mediterranean Basin. Proceedings of the National Academy of Sciences of the United States of America 106, 15368-15373.

Klecka, W.R., 1980. Discriminant Analysis. Sage Publishing. Newbury Park, USA.

Kull, T., Kindlmann, P., Hutchings, M.J., Primack, R.B., 2006. Conservation biology of orchids: Introduction to the special issue. Biological Conservation 129, 1-3.

Kullenberg, B., 1951. *Ophrys insectifera* L. et les insectes. Oikos 3, 53-70.

Kullenberg, B., 1961. Studies on *Ophrys* pollination. Zoologiska Bidrag Fran Uppsala 34, 1-340.

Kullenberg, B., Bergstrom, G., 1976. Hymenoptera Aculeata males as pollinators of *Ophrys* orchids. Zoologica Scripta 5, 13-23.

Mace, G.M., 2004. The role of taxonomy in species conservation. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 359, 711-719.

Moritz, C., 1994. Defining evolutionarily-significant-units for conservation. Trends in Ecology & Evolution 9, 373-375.

Moritz, C., 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51, 238-254.

Peakall, R., 2007. Speciation in the Orchidaceae: confronting the challenges. Molecular Ecology 16, 2834-2837.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

Petit, R.J., Excoffier, L., 2009. Gene flow and species delimitation. Trends in Ecology & Evolution 24, 386-393.

Pfeifer, M., Schatz, B., Pico, F.X., Passalacqua, N.G., Fay, M.F., Carey, P.D., Jeltsch, F., 2009. Phylogeography and genetic structure of the orchid *Himantoglossum hircinum* (L.) Spreng. across its European central-marginal gradient. Journal of Biogeography 36, 2353-2365.

Pillon, Y., Chase, M.W., 2007. Taxonomic exaggeration and its effects on orchid conservation. Conservation Biology 21, 263-265.

Rouget, M., Cowling, R.M., Pressey, R.L., Richardson, D.M., 2003. Identifying spatial components of ecological and evolutionary processes for regional conservation planning in the Cape Floristic Region, South Africa. Diversity and Distributions 9, 191-210.

Schatz, B., Jacob, L., 2007. Enjeux pour la conservation des orchidées méditerranéennes. Journée scientifique au Parc naturel régional des Grands Causses, p. 88, Millau, France.

Schlüter, P.M., Ruas, P.M., Kohl, G., Ruas, C.F., Stuessy, T.F., Paulus, H.F., 2007. Reproductive isolation in the Aegean *Ophrys omegaifera* complex (Orchidaceae). Plant Systematics and Evolution 267, 105-119.

Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L., 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. American Journal of Botany 94, 275-288.

Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Systematic Biology 49, 369-381.

Soliva, M., Kocyan, A., Widmer, A., 2001. Molecular phylogenetics of the sexually deceptive orchid genus Ophrys (Orchidaceae) based on nuclear and chloroplast DNA sequences. Molecular Phylogenetics and Evolution 20, 78-88.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The clustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876-4882.

Triponez, Y., Naisbit, R.E., Jean-Denis, J.B., Rahier, M., Alvarez, N., 2007. Genetic and environmental sources of variation in the autogenous chemical defense of a leaf beetle. Journal of Chemical Ecology 33, 2011–2024.

Tsvetanov, T., Vladimirov, V., Petrova, A., 2005. New localities of *Ophrys insectifera* (Orchidaceae) in Bulgaria, In Proceedings of the Balkan Scientific Conference of Biology from 19th to 21nd May 2005. eds B. Gruev, M. Nikolova, A. Donev, Plovdiv, Bulgaria.

Van der Cingel, N.A., 1995. An Atlas of Orchid Pollination. European Orchids. A.A. Balkema, Rotterdam.

Vereecken, N.J., 2009. Pollinator-mediated selection, reproductive isolation and the evolution of floral traits in the genus *Ophrys* (Orchidaceae), In Faculty of Sciences, behavioural and evolutionary ecology. Free University of Brussels, Brussels, Belgium.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, 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.

Wolff, T., 1950. Pollination and Fertilization of the fly Ophrys, *Ophrys insectifera* L in Allindelille Fredskov, Denmark. Oikos 2, 20-59.

Wolff, T., 1951. Ecological Investigations on the fly Ophrys, *Ophrys insectifera* L in Allindelille Fredskov, Denmark. Oikos 3, 71-97.

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

7. Figure captions

Figure 1: Schematised results of the statistical parsimony network analysis. Each haplotype is represented by a different shape (reported on Fig. 2). The total number of samples (detailed for each phenotypic species) and the general distribution of the corresponding haplotype are given.

Figure 2: Map of sampled populations showing geographical localisation of haplotypes. Both main and secondary haplotypes are represented with similar symbols as in Fig. 1. Close-ups are provided for regions hosting endemic taxa. Populations represented by one single sequenced individual are marked with an asterisk.

Figure 3: PCoA plot of all genotyped specimens within the *O. insectifera* group. Each of the three phenotypic species is displayed on the first two axes (A) as well as on axes four and seven (B).

Figure 4: Pattern of genetic diversity for the *O. insectifera* group. Values of expected heterozygosities are showed according to five classes. The corresponding number of analysed individuals per population is also mentioned. Six regions (A-F) are framed for their importance in the diversity of the species' group, inferred both by AFLP and sequence analyses (see discussion). Populations of the two endemics (in frames A and B) are marked with an asterisk when they also contained *O. insectifera* samples.

8. Tables

Table 1: Indications of the geographical localisation of investigated populations (listed by country) and information regarding genetic diversity indices: number of analysed individuals (N), genetic diversity value (UHe) and percentage of polymorphic loci (PP).

Population	Country	Location	Lat	Lon	Species*	N	UHe	PP
BAR	Austria	Barwies	47°N	11°W	Oi	1	Na	Na
GOS	Austria	Gosau	48°N	14°W	Oi	2	0.107	19.4%
WEI	Austria	Weissenbach	48°N	14°W	Oi	1	Na	Na
YAG	Bulgaria	Yagodina	42°N	24°W	Oi	4	0.143	31.7%
DRA	Croatia	Drakulic Rijeka	45°N	16°W	Oi	4	0.165	38.5%
GOR	Croatia	Gornji Vhrovci	45°N	18°W	Oi	4	0.163	36.6%
OPO	Czech Republic	Oponelec	49°N	14°W	Oi	3	0.093	19.1%
PAR	Finland	Parainen/Pargas	60°N	22°W	Oi	1	Na	Na
BEL	France	Belfort	44°N	3°W	Oa	5	0.130	31.4%
COU	France	la Couvertoirade	44°N	3°W	Oa	4	0.135	28.0%
CUN	France	les Cuns	44°N	3°W	Oa	1	Na	Na
LAY	France	Layolle	44°N	3°W	Oa	6	0.158	37.2%
LIC	France	Licide	44°N	3°W	Oa	7	0.159	39.1%
LIQ	France	Les Liquisses	44°N	3°W	Oa	8	0.159	40.6%
MAS	France	Le Massegros	44°N	3°W	Oa	6	0.173	42.5%
MAZ	France	Les Mazes	44°N	3°W	Oa	7	0.174	46.2%
PEZ	France	La Pézade	44°N	3°W	Oa	5	0.162	36.6%
SAT	France	Saint Saturnin	44°N	3°W	Oa	5	0.147	36.6%
SER	France	Sériguas	44°N	3°W	Oa	5	0.158	38.5%
SGL	France	Saint-Georges de Lévéjac	44°N	3°W	Oa	5	0.173	40.0%
VES	France	Vessac	44°N	3°W	Oa	6	0.161	39.4%
LOO	France	Lôo	43°N	1°W	Oi	3	0.126	26.8%
CER	France	Cereste	44°N	6°W	Oi	3	0.124	25.8%
LAA	France	Lapan	47°N	2°W	Oi	3	0.146	31.1%
LAP	France	Lapanouse de Cernon	44°N	3°W	Oi	4	0.128	29.5%
NAN	France	Gorges du Nan	45°N	5°W	Oi	3	0.131	28.3%
QUI	France	Quilen	51°N	2°W	Oi	3	0.128	26.8%
RBO	France	Rambaud	45°N	6°W	Oi	2	0.104	18.8%
THH	France	Thorame-Haute	44°N	6°W	Oi	2	0.085	15.4%
THO	France	Thorame-Basse	44°N	7°W	Oi	5	0.155	36.3%
THU	France	La Thuile	46°N	6°W	Oi	5	0.187	43.4%
VAC	France	La Vacquerie	44°N	3°W	Oi	4	0.139	31.7%
FON	France	Fontaret	44°N	4°W	Oi, Oa	6	0.159	40.9%
TRI	France	La Trivalle	44°N	3°W	Oi, Oa	8	0.187	47.7%
VIL	France	le Villaret	45°N	4°W	Oi, Oa	10	0.174	48.9%
AUG	Germany	Augsburg	48°N	10°W	Oi	4	0.138	31.4%

WOR	Germany	Wormental	50°N	10°W	Oi	4	0.129	28.9%
CUD	Great Britain	Cudham	51°N	0°W	Oi	2	0.124	22.5%
HRC	Great Britain	Hutton Roof Crags	54°N	3°E	Oi	3	0.137	28.9%
DEL	Greece	Delvinaki	40°N	20°W	Oi	4	0.143	31.4%
GEA	Ireland	Lough Gealain	53°N	9°E	Oi	3	0.125	26.2%
MPI	Italy	Monte Pizi	42°N	14°W	Oi	1	Na	Na
PIA	Italy	Pian Crixia	45°N	8°W	Oi	3	0.126	26.5%
SEN	Italy	Senerchia	41°N	15°W	Oi	3	0.132	27.7%
SZC	Poland	Szczawnica	49°N	21°W	Oi	3	0.109	22.8%
KLA	Slovakia	Klastor	49°N	19°W	Oi	3	0.134	29.8%
KOR	Slovenia	Korada	46°N	14°W	Oi	4	0.148	35.7%
UNE	Slovenia	Unec	46°N	14°W	Oi	3	0.161	32.3%
EUG	Spain	Sta. Eugenia de Relat	42°N	2°W	Oi	1	Na	Na
RUE	Spain	Ruesta	43°N	1°E	Oi	1	Na	Na
TRO	Spain	Trokoniz	43°N	3°E	Oi	4	0.132	29.5%
BER	Spain	Bernues	42°N	1°E	Oi, Os	7	0.168	42.5%
ERR	Spain	Erro	43°N	1°E	Oi, Os	5	0.132	32.6%
OST	Spain	Ostiz	43°N	2°E	Oi, Os	4	0.137	31.1%
COM	Spain	Les Comes	42°N	2°W	Os	7	0.167	43.7%
CON	Spain	Contrast	42°N	1°W	Os	4	0.132	29.5%
SAV	Spain	Savalla del Comtat	42°N	1°W	Os	1	Na	Na
SSO	Spain	San Sadurni d'Osormort	42°N	2°W	Os	4	0.126	28.3%
BAK	Sweden	Baeck	57°N	17°W	Oi	3	0.126	26.2%
SJO	Sweden	Sjostorp	58°N	15°W	Oi	2	0.119	21.5%
BOU	Switzerland	Boudry	47°N	7°W	Oi	1	Na	Na
CRO	Switzerland	Crostand	47°N	7°W	Oi	3	0.129	26.5%
MER	Switzerland	Meride	46°N	9°W	Oi	1	Na	Na
PLA	Switzerland	Planige	46°N	8°W	Oi	2	0.110	20.0%
URS	Switzerland	Saint-Ursanne	47°N	7°W	Oi	3	0.122	26.8%

* Oi: *Ophrys* insectifera, Oa: *O. aymoninii* and Os: *O. subinsectifera*

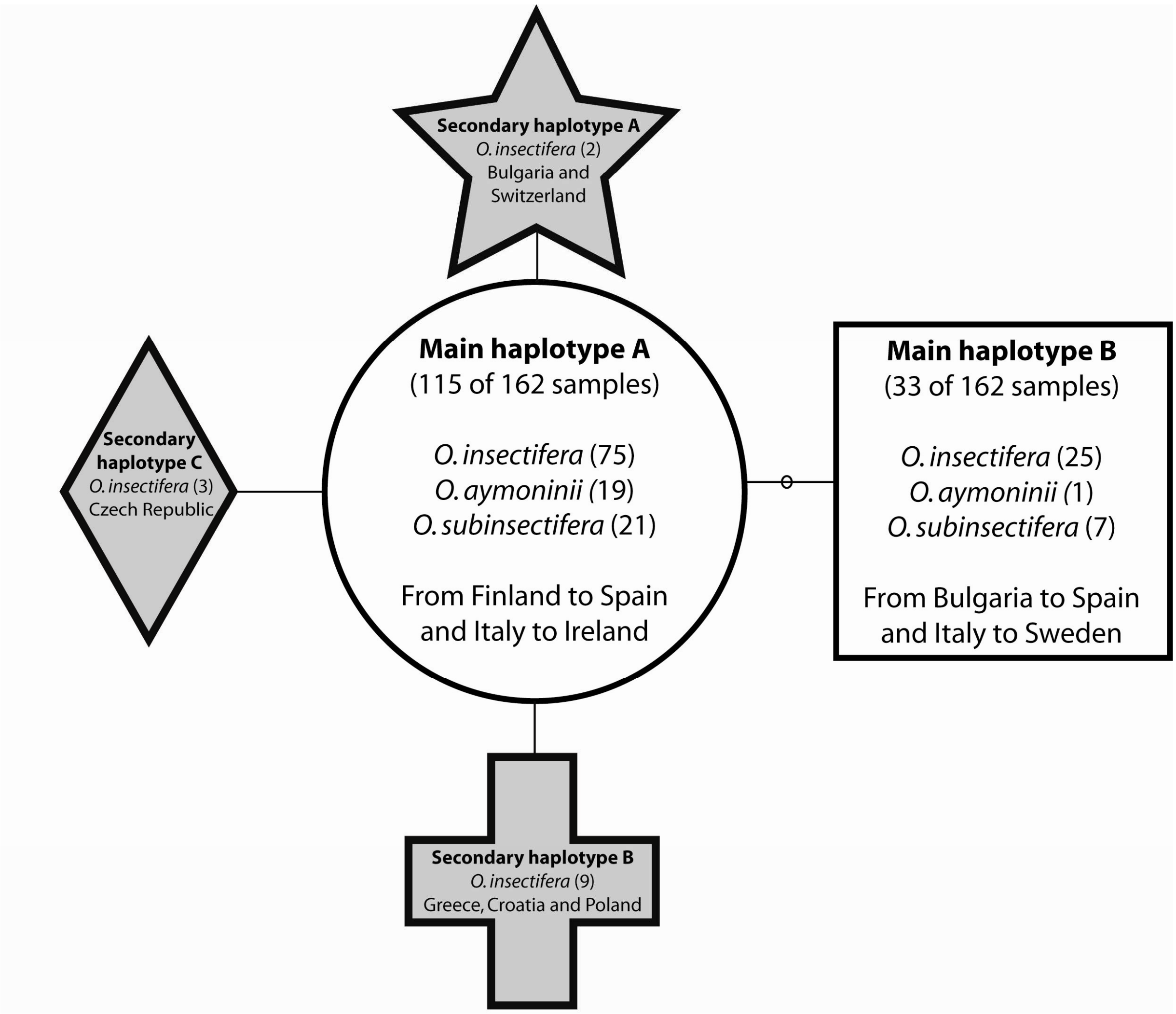


Figure 1: Schematised results of the statistical parsimony network analysis. Each haplotype is represented by a different shape (reported on Fig. 2). The total number of samples (detailed for each phenotypic species) and the general distribution of the corresponding haplotype are given.

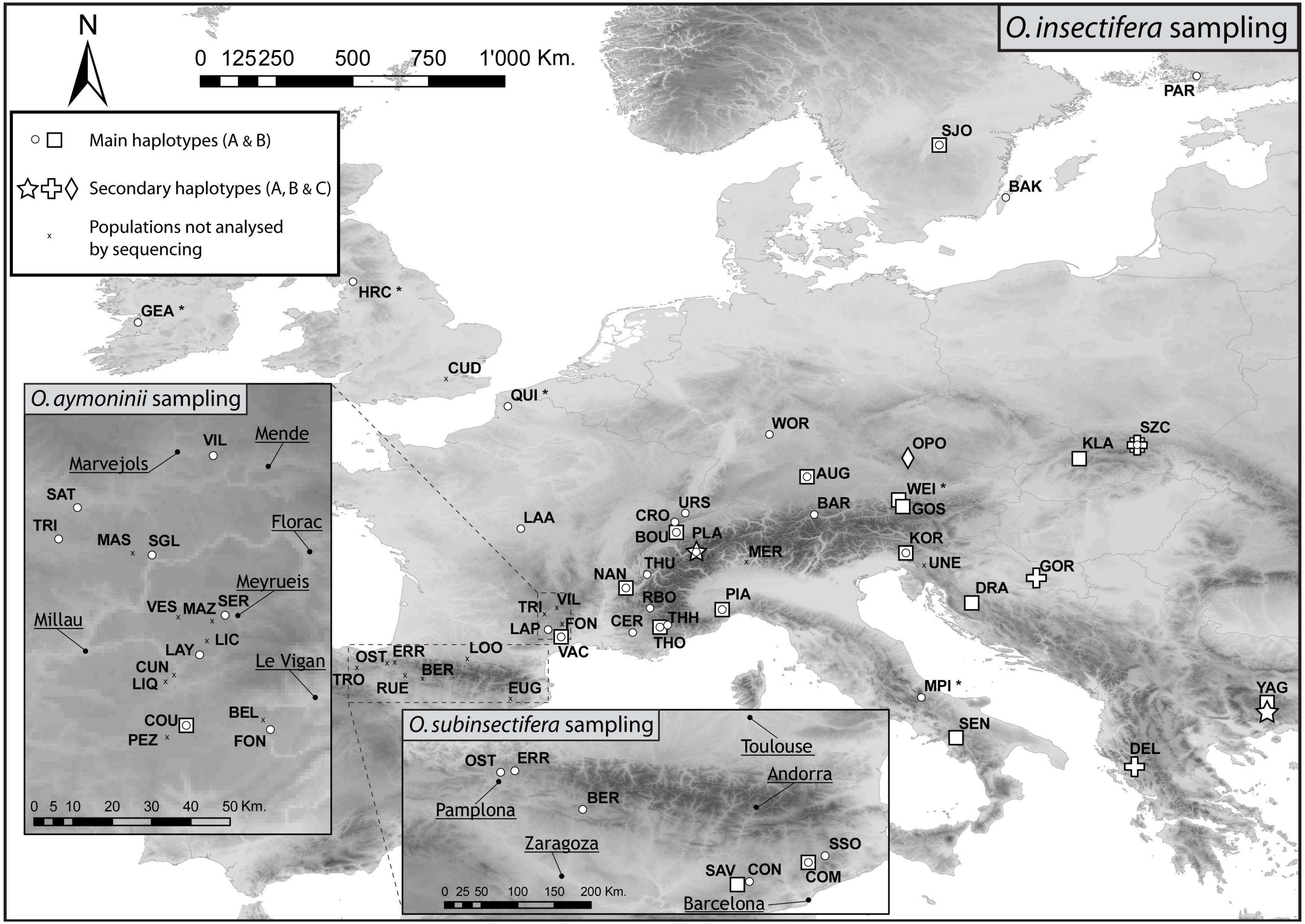


Figure 2: Map of sampled populations showing geographical localisation of haplotypes. Both main and secondary haplotypes are represented with similar symbols as in Fig. 1. Close-ups are provided for regions hosting endemic taxa. Populations represented by one single sequenced individual are marked with an asterisk.

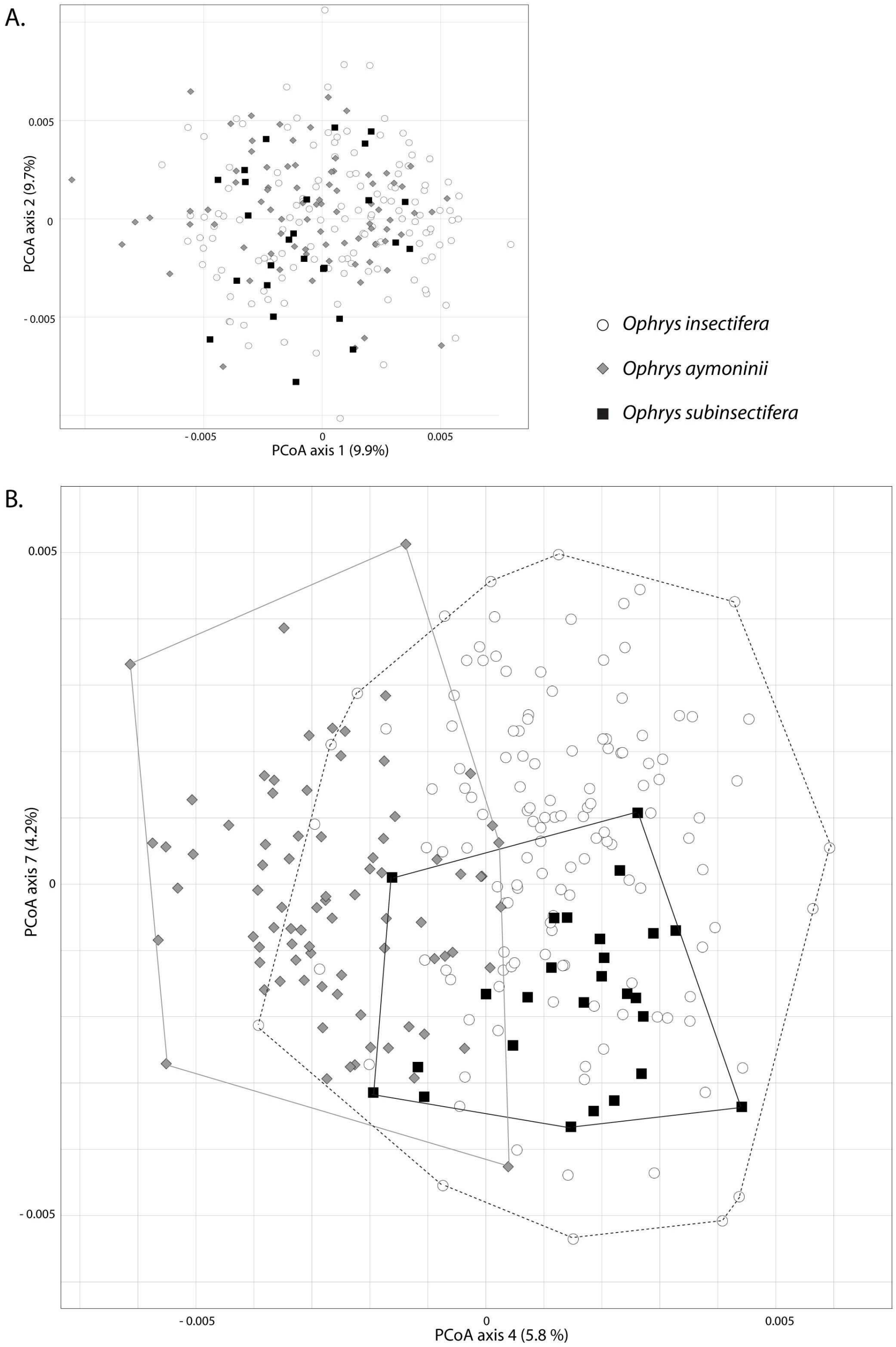


Figure 3: PCoA plot of all genotyped specimens within the *O. insectifera* group. Each of the three phenotypic species is displayed on the first two axes (A) as well as on axes four and seven (B).

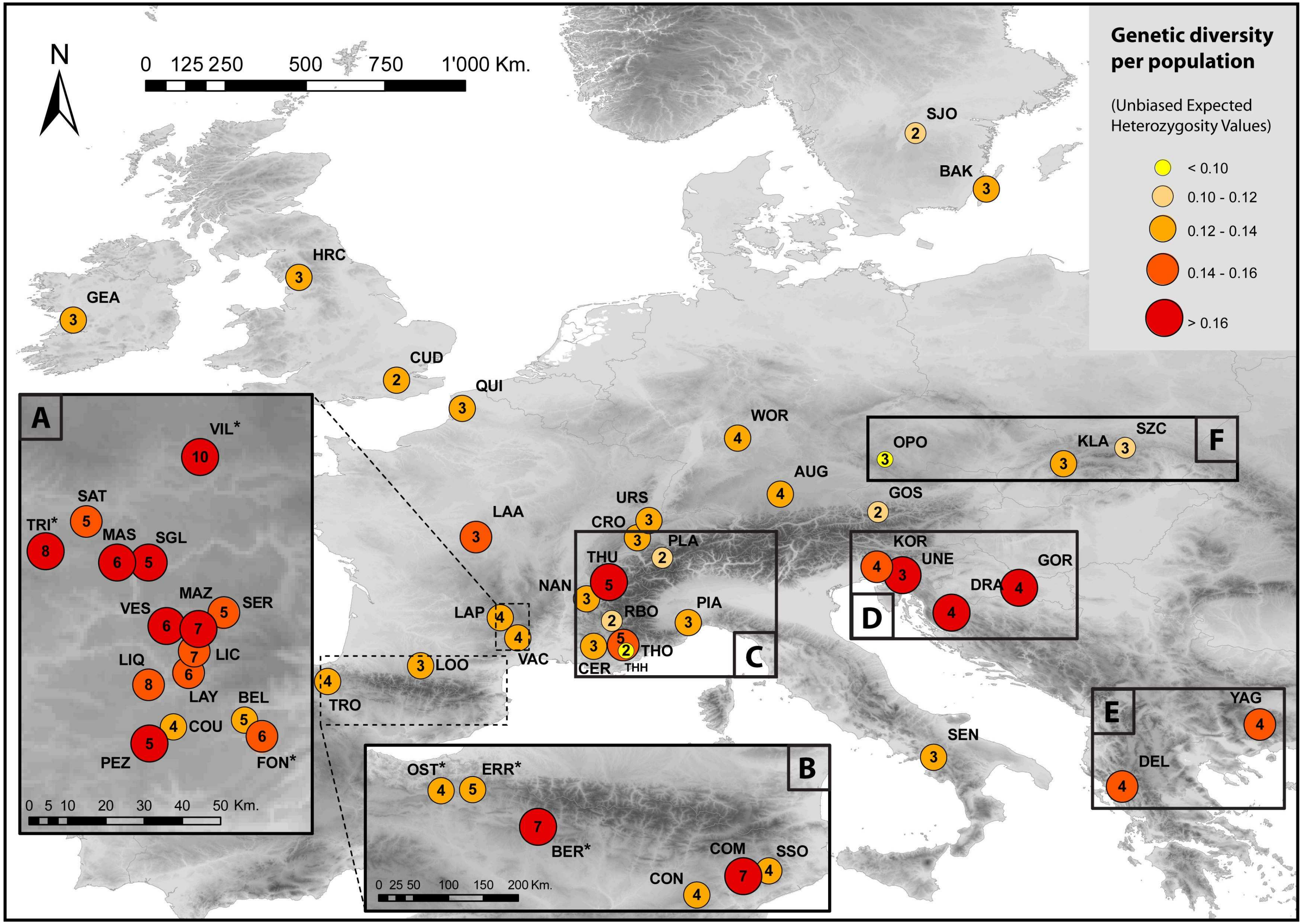


Figure 4: Pattern of genetic diversity for the *O. insectifera* group. Values of expected heterozygosities are showed according to five classes. The corresponding number of analysed individuals per population is also mentioned. Six regions (A-F) are framed for their importance in the diversity of the species' group, inferred both by AFLP and sequence analyses (see discussion). Populations of the two endemics (in frames A and B) are marked with an asterisk when they also contained *O. insectifera* samples.

CHAPTER III

Large-scale spatial genetic structure of the fly orchid (*Ophrys insectifera* group) reveals endemic lineages from different origins

Yann Triponez, Bertrand Schatz, Loïc Pellissier and Nadir Alvarez

(Version formatted according to *Conservation Genetics*)

Neuchâtel, August 2010

Large-scale spatial genetic structure of the fly orchid (*Ophrys insectifera* group) reveals endemic lineages from different origins

Triponez Y*, Schatz B[†], Pellissier L[§] and Alvarez N[§]

* corresponding author: yann.triponez@unine.ch
Laboratoire d'Entomologie Evolutive, Institut de Biologie
Université de Neuchâtel, Rue Emile-Argand 11, cp 158
CH-2009 Neuchâtel, Switzerland
Phone +41 32 718 31 62 - Fax +41 32 718 30 01

[†] Centre d'Ecologie Fonctionnelle et Evolutive, CNRS Montpellier
UMR 5175, 1919 Route de Mende
F-34293 Montpellier cedex 5, France
Phone +33 04 67 61 33 00 - Fax +33 04 67 41 21 38

[§] Departement d'Ecologie et Evolution, Faculté de biologie et de médecine
Université de Lausanne, Biophore
CH-1015 Lausanne. Switzerland
Phone +41 21 692 42 79 - Fax +41 21 692 41 65

Abstract

Quaternary climate oscillations strongly influenced the history and fate of most temperate organisms. Glaciations followed by warming periods shaped the distribution and the spatial genetic structure of species' lineages, generally leading to locally evolved genetic pools, sometimes even resulting in newly emerged endemic species. The European genus *Ophrys*

(Orchidaceae) is famous for its insects-like flower’s morphology, a direct consequence of its pseudocopulatory pollination strategy involving Hymenoptera males. A huge number of endemic *Ophrys* species was described recently, mostly on the basis of subtle morphological variations and unique pollinator dependence. The fly orchid group (*O. insectifera* sensu lato) is composed of three distinct morphological species, one widespread and two endemics, all of them genetically very similar. In this study we used a phylogeographic approach based on AFLP data combined to ecological niche modelling at the last glacial maximum (LGM), in order to discuss the evolution of endemism in the fly orchid group.

If the phylogeography of the whole group generally resulted in a strongly admixed genetic structure, two genetic lineages were geographically more restricted. As one endemic lineage could be attributed to the biogeography and history of the fly orchid, the second showed an interesting affinity with the French endemic *O. aymoninii*. This result encouraged the supplementary description of its genetic diversity pattern for further application in local conservation actions. In conclusion, despite strong genetic similarity within *Ophrys* species’ groups, further similar phylogeographical studies may be useful, notably in reinforcing the validation of conservation priorities for other endemic *Ophrys* species.

Keywords

Ophrys insectifera, endemics, spatial genetic structure, *O. aymoninii*, conservation genetics, AFLP

Introduction

Past climatic changes are among the most important historical factors that molded species richness and level of endemism within a given area, both proved to be inversely proportional to the variation in temperature since the Last Glacial Maximum (Jansson 2003). The spatial genetic structure of many organisms was shaped by Quaternary climatic oscillations, especially in temperate regions strongly subject to dramatic changes in habitats (Jansson 2003; Thompson 2005). For instance in Europe (Tribsch and Schönswetter 2003; Médail and Diadema 2009), the temperate tree flora was depauperated during Plio-Pleistocene climate changes (Svenning 2003). However, among the roughly estimated 12000 species for the whole European Flora, still many endemics can be found in most areas of Europe (Walters 1976). These endemic species might be the result of processes acting at different evolutionary scales. First, as a consequence of a wide extinction of lineages during harsh climatic periods, several species of Tertiary origin eventually did not get extinct, but managed to survive in refuges constituting currently relictual regional distributions; those can be seen as palaeo-endemics (Thompson 2005). Second, glaciations also permitted recent diversification of neo-endemic taxa, for example by isolating after polyploidization events either the diploid precursors, called patro-endemics, or the polyploid products called apo-endemics (Comes and Kadereit 1998; Thompson 2005). More commonly, range fragmentation of a widespread ancestral taxon into limited refugia during ice ages resulted in disjunctive distributions of closely related species, designated as schizo-endemics. This last category of endemics can lead to taxa potentially displaying any degree of divergence compared to the parental pool, depending on the age of disjunction (Thompson 2005). Even if regional differentiation did not lead to taxonomical entities recognized as endemics, numerous phylogeographic studies show that most species kept at least a trace of such genetically diverging lineages (e.g. Hewitt 1999; Schmitt 2009).

Ophrys orchids provide a nice case-study to investigate the evolution of endemism in Europe since it raises not only question regarding evolutionary processes, but also taxonomic issues (Bateman et al. 2009). Pillon and Chase (2006) noted that in Europe, a bias due to strong interests for this emblematic genus led to numerous and sometimes confusing descriptions of endemic taxa. The genus *Ophrys* comprises for some authors as far as ca. 250 described microspecies and it provides remarkable cases of species' groups enclosing both widespread species and local endemics that are very little, or even not, genetically differentiated (Devey et al. 2008). Several phylogenetic studies reported the recent (and sometimes ongoing) diversification of species groups within the genus (e.g. Schlüter et al. 2007; Devey et al .2009) and showed that endemic species within *Ophrys* do not seem to correspond to palaeo-endemics. Karyological studies demonstrated that, except for a few members within the sect. *Pseudophrys* or for some inter-specific hybrids (Delforge 2001, Bernardos et al. 2003), most *Ophrys* species have a diploid chromosomal number (2n=36; Greilhuber & Ehrendorfer 1975; Bianco et al. 1991; D'Emerico et al. 2005). Therefore occurrences of neo-endemics related to polyploidization events might be relatively rare and schizo-endemic species should be regarded as the main rule regarding patterns of endemism in *Ophrys*. Particularly, since *Ophrys* reproduction relies on sexually deceptive pollination systems, populations isolated from a widespread parent might get locally quickly adapted to the available pollinators. Large variation in current endemic phenotypes could then reflect repeated changes in morphologies occurring during isolation events as ancient as the Quaternary glaciations. Even so, in the case of sexually deceptive orchids, one should not neglect that neo-endemics can easily evolve quickly, through ecological sympatric speciation due to shifts in pollinator's attraction (e.g. Schiestl & Ayasse 2002; Schlüter et al. 2009). With or without geographical isolation, sources of neo-endemism in *Ophrys* could be multiple, as attested by the current abundance of species limited to very restricted areas.

In order to address patterns of endemism in *Ophrys*, we explored evolutionary history of the phylogenetically most basal species assembly within the genus (Devey et al. 2008): the fly orchid group. This aggregate of three species is represented by the emblematic and widespread *O. insectifera* L., together with two endemics from south-western Europe: *Ophrys aymoninii* BREISTROFFER, an endemic species from the karstic region of southern French Massif Central (known as "Grands Causses" region), and *O. subinsectifera* HERMOSILLA & SABANDO, a Spanish vicariant growing in the southern foothills of the Pyrenees. All three species have distinctive morphologies, emit unique scent and are associated with specific pollinator dependence (Borg-Karlson et al. 1993; Hermosilla et al. 1999; Vereecken 2009; Vereecken et al. in prep.). However in view of their confounding strong genetic similarity, they have been treated more as Evolutionary Significant Units, all deserving care (for several reasons) in conservation actions (Triponez et al., submitted). In order to describe phylogeographic patterns, we used fine-scale Amplified Fragment Length Polymorphism (AFLP) in combination with modelling of past species distribution at Last Glacial Maximum (LGM; -21'000 years), and investigated the following questions. 1) How is the genetic variation of the *O. insectifera* group structured in Europe and can it be explained by its past distribution during LGM when its range was more restricted than today? 2) Does the spatial genetic structure traduce endemism? In other words, are genetic lineages among the fly orchid group compatible with the occurrences of endemic species? 3) If endemic gene lineages do exist, to what extent do they contribute to the overall genetic diversity of the group?

The present study also underlines further questions concerning conservation priorities of endemic taxa. Endemics might be locally abundant but they generally reveal to be rare and thus more prone to be threatened. This statement is particularly true for narrow-endemics, taxa that are restricted to one or a few small populations (Kruckeberg & Rabinowitz 1985). Exceptional concerns are given to preserve such endemic organisms since they account for a large proportion in defining priority settings in species conservation (Schmeller et al. 2008). We previously mentioned that schizo-endemics can be of any level of divergence compared to parental or related taxa. However, shall we apply the same conservation rules regardless of the genetic and/or phenotypic level of differentiation of a given taxon? For instance, how conservation priorities should deal with cryptic endemics, i.e. endemic genetic lineages that are morphologically undifferentiated from other widespread lineages? After addressing these questions, also in the light of the strong taxonomical bias within this group, we will discuss the extent to which patterns of genetic diversity are informative when building up conservation priorities in endemic *Ophrys* taxa, illustrated by the particular case of the French endemic *O. aymoninii*

Methods

Sampling

The sampling of the three species among the *O. insectifera* group was carried out during the flowering periods in spring 2007 and 2008 (from April to June depending on the regions), in order to cover most of their respective distribution ranges. In each population visited, from three to ten leaves were collected and this fresh material was directly desiccated in silica gel (Chase and Hills 1991). Within zones of endemism, we collected only plants showing morphological features typical of a defined species and discarded putative hybrids.

AFLP analysis and procedure

Total genomic DNA was extracted from 10 mg of silica-dried leaf fragments using the DNeasy Plant Kit (Qiagen, Hilden, Germany). All following reactions were conducted in 96-well plates, in which samples were randomly distributed. In each plate, ten to 15 samples were duplicated for further reproducibility tests. AFLP analysis was slightly modified from the classical method described by Vos et al. (1995). The detailed protocols of digestion, ligation, pre-amplification and selective amplification are available by the first author. Final selective PCR products were analysed using the GeneScan technology with an automated capillary-sequencer (ABI 3730XL, Applied Biosystems, Foster City, CA; service provided by Macrogen Inc., Seoul, South Korea). Resulting fluorescent AFLP patterns were scored using GeneMapper 3.7 (Applied Biosystems) with some changes in default parameters (maximum bin width=1 bp, light peak smoothing, peak threshold=200 RFU). The allelic pattern obtained automatically was checked visually and followed by minor manual corrections.

A primer trial on a small number of samples (from different species and origin) was conducted using 16 primer combinations to identify pairs of selective primers that were repeatable and polymorphic enough within the fly orchid group. The two primer combinations *Eco*RI-ACAG/*Mse*I-CAA and *Eco*RI-ACAG/*Mse*I-CTG yielded suitable numbers of bands, levels of variation among loci and optimized reproducibility, and were chosen for fingerprinting application to the full sampling. Repeatability was tested by comparing AFLP patterns between samples and their replicates, with a final satisfying overall reproducibility rate of >90%. We scored the presence or absence of each AFLP marker in each individual plant. The matrices of the two scored primer pairs were concatenated into one binary matrix where individuals and bands were stored as lines and columns, respectively. A first scoring included all samples of the three species among the *O. insectifera* group and was used for further spatial genetic structure analysis. In order to avoid genetic structuring artefacts due to a too low or high final specimens harbouring extreme number of bands, we discarded all individuals standing out of a ten percent interval from the mean number of fragments for the total sampling. A second scoring including only samples of *O. aymoninii* was completed, and used as a new basis for subsequent genetic diversity analysis of this endemic taxon. Here, we discarded only individuals showing less than half of the mean number of fragments for the complete *O. aymoninii* sampling.

Spatial genetic structure analysis of the fly orchid group

We investigated the spatial genetic structure of the fly orchid group based on the AFLP matrix, through a Bayesian inference model-based algorithm using STRUCTURE 2.2 (Pritchard et al. 2000; Falush et al. 2007) to assign individuals into a user-defined number of genetic groups (hereafter K). We used the “no admixture” model and set the MCMC algorithm to 200,000 generations of burnin and 800,000 generations for data acquisition. Here, K values ranged from one to ten, with 15 replicates for each tested value. For each K value, replicates that 1) yielded the highest likelihood value and 2) did not present any genetic group with a value lower than 0.01 were considered as best runs within a given K, and considered for further analyses. For the best run among all, individuals were assigned to a group if their respective assignment probability was higher than 0.95. The result of the best run was displayed on a geographical map using ARCGIS 9.1 (ESRI), representing each population as a pie-chart showing the number of individuals assigned to each genetic group as well as non-assigned samples. Considering the high level of admixture, we explored a second approach based on the average assignments per population among the three best runs. Here, average percentages of assignment were calculated at the population level, by calculating the mean assignment to each cluster for all individuals within a population. Each population was displayed on geographical maps, as a pie-chart showing the average percentage of attribution to each group, using ARCGIS 9.1 (ESRI).

Fly orchid group distribution modelling for LGM

We modelled the distribution of the species *O. insectifera* (three taxa pooled) using 7 modelling techniques implemented in the BIOMOD (Thuiller et al. 2009) R package (see detailed protocol as supp. mat. 2). Model were calibrated from presences collected during the field sampling pooled with occurrences from the GBIF database (www.gbif.org) group and using 7 climatic layers from Worldclim (Hijmans et al. 2005). An Ensemble forecasting approach (Marmion et al. 2009) was then used to project a central tendency from a combination of the models obtained by the different niche-based modelling techniques into both current climate and past climate of the LGM computed by the CCSM circulation model. The resulting predicted distribution was illustrated on a map of Europe using ARCGIS 9.1 (ESRI), with most probable occurrences of the fly orchid during LGM displayed as strong dark intensity.

Genetic diversity pattern of endemics: the case of *O. aymoninii*

Using the AFLP scoring of samples belonging to *O. aymoninii*, we used the software AFLP-SURV (Vekemans 2002) to calculate the percentage of polymorphic loci (at the 5% level) and the expected heterozygosity (also called Nei’s gene diversity), for each population including a minimum of three individuals, and at different geographical levels (French administrative regions and causses, see results for exact definition). We obtained in addition the corresponding permutation tests for genetic differentiation among populations and among geographical regions. Absolute values of Nei’s gene diversity per population were categorized into five classes and displayed on geographical maps using ARCGIS 9.1 (ESRI). A centre point of distribution was calculated using the average latitude and longitude of all populations collected. The genetic diversity was correlated to both latitudinal and longitudinal distance (of a particular population) from this distribution centre and statistically tested using SigmaStat version 2.03 (SPSS Inc., Chicago, USA).

Results

AFLP fringerprinting

When scoring together all individuals from the three species of the *O. insectifera* group, the two primer combinations generated a total of 325 AFLP markers, each specimen (240 in total from 59 populations, see Table 1) yielding between 155 and 194 peaks. The scoring limited to the *O. aymoninii* individuals generated a total of 271 AFLP markers, and each of the 126 samples (from 16 populations) showed between 68 and 163 fragments.

Spatial genetic structure analysis of the *O. insectifera* group

Following Evanno et al. (2005), we obtained the best results in the STRUCTURE analysis when six groups (K=6) were considered, with log likelihood values of the three best runs standing above -38000. The individual-based approach of the best K=6 resulted in 206 out of 240 individuals (86%) attributed to one of the six genetic groups. The geographic projection of the genetic lineages for this best run is presented in Fig. 1, showing as well (for some populations) non-attributed individuals as transparent parts of the pie-charts. The approach based on the average assignments among the three best runs is presented in Fig. 2. This second approach was useful as we remarked that the three best runs did not exactly result in identical patterns, probably due to high rate of admixture.

As a first result visible in both Figs. 1 and 2, a large majority of populations showed high levels of admixture between the different lineages, with 54 among 59 populations assigned to more than one single STRUCTURE group. Four out of six lineages (in blue, green, pink and white on Figs. 1 and 2) are widely dispersed across Europe. Even if they show no clear geographic structure, some of these groups were still geographically patterned (based on the best-three-runs-average results showed in Fig. 2, considering the respective proportions of each genetic group within a population). In particular, the white genetic line showed greater proportion in most western (Atlantic and alpine) populations, but did not appear (or only at very low percentages) in southern Scandinavia and in many populations from the *O. aymoninii* endemic area. The blue group demonstrated highest rates along a large stripe running from Southern Alps to the Baltic Sea, but was missing in the British Isles and south-eastern Balkans. On the contrary, the green and pink genetic groups did not show any trends and appeared almost everywhere in Europe.

Two genetic lineages were however much more restricted to particular regions of Europe. The black group was exclusively present in south-western Balkans, with occurrences in Greek and Bulgarian populations. The red group showed very high proportions in many populations from the south-eastern distribution (France and Spain), with however rare occurrences elsewhere in Europe (these rare occurrences even disappeared in the non-illustrated third best run). Considering a 95% threshold of assignment percentage, 36 samples out of 52 (69.2%) attributed to this red group were members of the endemic species *O. aymoninii*, while the other samples came at a great majority (14 out of 16) from nearby French or Spanish populations of *O. insectifera* and *O. subinsectifera*.

When looking at the representation of the different genetic groups per species (see Fig. 3), *O. insectifera* showed quite homogenous proportions for each gene lineage (except for the rarer Balkan black group exclusive to this species). The Spanish endemic *O. subinsectifera* denoted some similarities with *O. insectifera*, except for the white genetic group’s dominance. However as noted above, this white group was also present at high proportions in many western populations of *O. insectifera*, but interestingly not in *O. aymoninii*. Finally, *O. aymoninii* was composed of two dominant lineages, red and pink, present in respectively 50% and 34% of the specimens. We remarked here once more the good match between this French endemic and the red genetic group, half of the defined individuals (36 out of 73) being classified within this group.

LGM distribution of the fly orchid

The result of the species distribution modelling at LGM is displayed in Fig. 4, as well as ice covering limit at the same period. The projection of the species distribution model of *O. insectifera* in the past climate of the LGM showed high probabilities of suitable climatic conditions, in different parts of Europe. In particular, the three classically paradigm refugia for temperate species (i.e. Iberian Peninsula, Italy and Balkans; see Hewitt 2000) were also found here. However, optimal conditions for *O. insectifera* seemed not to occur in the most southern parts of these regions. The Iberian refugia might then correspond to the northern half of Spain but did not extend southwards. The Italian refugia zone should have covered most of the Italian Peninsula (except the extreme south), the northern part of the current Adriatic Sea (which has remained dry at that period) and Slovenian plains. Finally, the Balkanic refugia might be represented by a main area covering southern Serbia, Macedonia, eastern Albania and northern Greece, with some extensions into southern Bulgaria. In addition, since currently *O. insectifera* shows a quite northern distribution and occupies a wide climatic niche (because this species can tolerate quite well colder conditions compared to most other *Ophrys* species; see supp. mat. 3), the model predicted also large surfaces occupied by the species during the LGM. The fly orchid might also have survived quite extensively out of the classical southern refugia, especially in a wide Atlantic zone western from a line running from the French Southern Alps to Brittany. Adequate climatic conditions modelled for southern UK, northern France and north from the Alps should be considered with care. Indeed, all these regions stood very close to the southern border of LGM icecap, and might rather represent an analytical bias due to oversampled areas in northern Europe.

Genetic diversity of *O. aymoninii*

Table 2 resumes genetic diversity indices for each *O. aymoninii* population, for causses (i.e. calcareous plateaus separated by deep gorges typical from French southern Massif Central) and for French administrative regions. The percentage of polymorphic loci and Nei's gene diversity per population ranged from 62% to 76.8% and from 0.209 to 0.322, respectively. Results obtained for two of them (CUN and TRI) should be considered with caution, considering their limited sampling (n < 5). The average Nei's gene diversity within populations was 0.252. The genetic diversity pattern shown in Fig. 5 indicates that populations from the centre of the distribution (e.g. LIC, MAZ, SER, SGL) are in general more diverse than peripheral populations (such as FON, PEZ, SAT or VIL). The previous observation was confirmed by correlations between genetic diversity and distribution centre distance (Fig. 6). Either performed on the north-south axis (Fig. 6A) as well as on the east-west axis (Fig. 6B), correlations revealed negative slopes (as supposed by the pattern displayed on Fig. 5). However, only the trend from the latitudinal correlation was significant through a Spearman rank order test (p=0.033). The 16 populations studied that covered a large part of the species' distribution, were genetically differentiated as confirmed by the permutation test for genetic differentiation among populations (F_{ST} = 0.073, P < 0.0001). However, STRUCTURE analyses resulted in no spatial genetic structure at the level of the *O. aymoninii* distribution (data not shown; best K = 3, all pops admixed). The permutation test for genetic differentiation showed the same absence of geographic differences, either among causses (F_{ST} = 0.01, P = 0.178) or among administrative regions (F_{ST} = 0.03, P = 0.267). At the level of the causses, Sauveterre was genetically less diverse than both Larzac and Noir. The genetic diversity level of administrative regions revealed that Aveyron was more diverse genetically than Lozère and Gard, and at a larger scale that region Midi-Pyrénées was slightly more diverse than Languedoc-Roussillon. As causses are spread over different administrative regions, we provide in addition a supplementary material including values for each detailed area (Table 3 in supp. mat. 1), as well as recommendations for local conservation actions.

Discussion

Heterogeneous levels of admixture at the European scale for the *O. insectifera* group

One of the main results of this study is the lack of marked spatial genetic structure at the European level for the *O. insectifera* species group. Almost all genetic lineage defined were widespread across Europe (as shown in Figs. 1 and 2) and more than 90% of the populations comprised specimens issued from different lineages. This high level of admixture might be a consequence of the biology of the species, which demonstrate long-distance dispersion potential due to dusk-like seeds, as for many other orchid species (Squirrell et al. 2001). This biological feature tends to homogenize the genetic identity of populations through long distances and maintain their low genetic differentiation (e.g. Alexandersson and Ågren 2000). From a historical point of view, distinct groups of lineages might potentially have an origin in Pleistocene refugia (as suggested by the LGM species modelling). We could for instance consider a French or Spanish origin of the white lineage, and a blue group coming from the Italian refugia, two areas of Europe that might have been partially isolated from each other during the LGM (and two paradigm sources of isolated lineages). These lineages, when re-colonizing, might have extended mostly western for the white and eastern for the blue. However, concomitant long distance dispersals might have progressively rendered the pattern unclear. Nowadays such lineages are no longer geographically delimited, at least in western and central Europe, as a result of rapid postglacial admixture tentatively related to the dispersal capacities of *Ophrys* minute seeds.

In addition, some of these lineages might have been connected during Pleistocene times. For instance, the large distribution area covering France at LGM might have allowed a continuous genetic flow between Iberian and Italian refugia, despite the barriers of Alpine and Pyrenean icecaps. In such a scenario only most isolated populations, for instance at the southern tip of LGM potential refugia, would have developed particular genetic identities. We hypothesise that, even during LGM, local selection and drift could have been regularly challenged by recurrent dispersal events, notably in this western area, that might have acted as a melting-pot. As a result, the pattern of historical lineages in central Europe is currently completely scrambled.

LGM distribution of the species could tentatively explain the presence of a geographically restricted lineage located in south-eastern Balkans (in black on Figs. 1 and 2). Indeed, although this region was a good candidate for hosting fly orchids during glaciations, it seemed to lack connexions to other European refugia, as most surroundings of central Dinaric Mountains (in Bosnia and Croatia) were inadequate for the species (see Fig. 4). This isolation might have subsisted until nowadays, as rare Greek and Bulgarian sampled populations still stand at the south-eastern tip of the species' distribution (see supp. mat. 3). But an even more plausible explanation for this Balkan endemic lineage, although maybe complementary to past distribution models, is described in Devey et al. (2009) where a comparable pattern was observed among species of the *O. fuciflora* species group. Here the genetic discontinuity running north–south through the Adriatic was associated to the effects of prevailing winds that would have impeded long-distance dispersal events from the south-eastern part of the distribution to the rest of Europe. Despite being only hypothetical especially concerning Devey's critical question "how long these prevailing winds are likely to have existed", this scenario is interestingly compatible with our current results on the fly orchid group.

Towards a characteristic genetic lineage for *O. aymoninii* ?

Whereas both climatic (prevailing winds direction) and historical (LGM isolated refugia) causes might explain the isolation of the Balkan group and partially the white and blue groups, the history and fate of the genetic lineage centred and strongly restricted to central France (in red on Figs. 1 and 2) is harder to explain in a biogeographic context. However, this lineage strongly overlaps with the distribution of *O. aymoninii*, in which it was strongly represented (i.e. half of all *O. aymoninii* specimens were attributed to this western lineage; Fig. 3). Despite strong genetic similarity among members of the fly orchid group, a trend towards a weak genetic differentiation between *O. aymoninii* and the two other members of the group was established (Triponez et al., submitted). As a consequence, the red lineage could be associated with phenotypic and/or ecological characteristics found in *O. aymoninii*, notably considering its strong affinity for predominant calcareous bedrocks in the Grand Causses region. This pattern might have arisen in the context of sympatric ecological speciation within this group (notably here through a change in the pollinator species, from the wasp *Argogorytes mystaceus* to the bee *Andrena combinata* as reported by Borg-Karlson et al. 1993; Delforge 2001; Vereecken 2009; Vereecken et al., in prep.), despite this scenario remains difficult to test accurately. In *O. aymoninii*, as in other deceptive orchid species, spatial limitation in the dispersal abilities is strongly related with the presence of highly specialised pollinator species (Cozzolino & Widmer 2005). However *A. combinata* is widely distributed across Europe; hence we could envisage why this endemic orchid and its associated genetic lineage are restricted to such a narrow region in France only if we assume that the ecological niches of *O. aymoninii* and of its pollinator exclusively overlap in this area. Long-distance dispersal of the red group via seeds was nevertheless observed (e.g., in Scandinavia, Italy; see Fig. 2). In the case of red lineage long-distance dispersal via *O. aymoninii* dispersers, these occurrences would be followed by extinction because of non-overlapping niches with *A. combinata*. But as this genetic lineage was also detected in *O. insectifera* samples (e.g. in populations LAP or LAA in France, or TRO in Spain, see Fig. 2), it might also disperse out of the range of the endemic via the *O. insectifera* form. In the latter case, the red genetic group would endure introgression, with other groups more common in the rest of European *O. insectifera* populations. These reasons might explain why occurrences of red lineage are relatively rare out of the endemic zone of *O. aymoninii*.

Since about three decades, various clues reported the individuality of *O. aymoninii* in view of its morphology, geographical repartition (Breistroffer 1981; Delforge 1983) and functional characteristics linked to pollination biology (Borg-Karlson et al. 1993; Delforge 2001; Vereecken 2009; Vereecken et al., in prep.). In addition to previous phenotypic considerations, the present study also revealed an “aymoninii genetic lineage”, attesting for the probable genetic individuality of the French endemic. In contrast, *O. subinsectifera* was not supported by the spatial genetic structure and thus not associated with a particular genetic lineage. We might here face a case of incomplete lineage sorting, especially concerning this Spanish endemic. In the same way than proposed in a complementary article on the fly orchid group (Triponez et al., submitted), we supposed that if *O. subinsectifera* was to be younger than *O. aymoninii*, it might also not stand as further as the French endemic in its genetic circumscription. It is finally possible that the AFLP primer pairs chosen were not sensitive enough to circumscribe this Iberian endemic taxon properly.

Diversity clues for future conservation of *O. aymoninii*

Because it is associated both to a characteristic gene lineage (as shown in the present study) and to specific pollination biology, the French endemic deserves particular conservation interest. Recent identification of threats foresaw a potential future inscription of *O. aymoninii* on the IUCN red list of threatened species (Schatz & Jacob 2008). Key zones to consider for conservation actions encompassing high levels of diversity are presented as supplementary material (supp. mat. 1). In summary we showed here that populations at the centre of the distribution were more diverse genetically compared to those in the periphery, an expected pattern predicted by the controversial central-marginal concept, also called the “abundant centre” hypothesis (Sagarin & Gaines 2002, Pfeifer et al 2009). This pattern might suggest a much localised outbreak of the *O. aymoninii* form associated with particular ecological (e.g. climate, type of bedrock) and physiological (flower shape and odours) properties of the species, and also including an overlapping niche with a specialized pollinator. From this central optimal distribution area, decreasingly suitable conditions bring the species to become less present, until too extreme environment for the species’ survival. Genetic diversity would then follow the same fate, showing highest levels at the centre (i.e. optimal) part of the distribution. We noted in addition that the plants belonging to central populations are sensibly bigger and more vigorous than specimens from populations situated at the margins of the *O. aymoninii* distribution (B. Schatz unpubl. data). As populations at the centre of the distribution were the most diverse genetically and that they most likely stand here at their ecological optimum, particular attention should be paid to preserve the habitats of *O. aymoninii* within this area (for more precisions see supp. mat. 1).

Conclusion

In the present study, we addressed a weak spatial genetic structure in the fly orchid group, with strong levels of admixture over its whole distribution area. However, our results also demonstrated the existence of two endemic genetic lineages. Although associated with different evolutionary processes, both *O. insectifera* specimens from the Balkans area and *O. aymoninii* have to be taken into account for the conservation priorities of the whole species group. Despite we could not address a proper genetic circumscription for the phenotypically well-distinguished *O. subinsectifera*, this taxon might also deserve preservation efforts since the lack of genetic identity might be explained by a recent evolutionary history. Similar studies on other *Ophrys* species groups including both endemics and widespread taxa might help in describing general trends in biogeographical patterns and addressing conservation priorities of this emblematic orchid genus**.**

Acknowledgments

We thank all scientists and keen amateurs who permitted to locate *Ophrys* populations across Europe: J.E. Arnold, A. Croce, C. Bernard, F. Dabonneville, B. Dolinar, P. Geniez, C.E. Hermosilla, L. Jacob, A. Jacquet, N. Margraf, J. L. Menos, N. Petrou, D. Průša, G. Scopece, R. Souche, A. Soulié and numerous others for punctual information. The authors are especially grateful to A. Espíndola for continuous help during field sessions and laboratory work, and to A. Sarr for support with AFLP analysis. This research was supported by the Swiss National Science Foundation (grant n° 3100A0-116778/1), the Parc naturel regional des Grands Causses (grants n° 019792 and n° 040763), the French Society of Orchidophily (SFO) and the Swiss Orchid Foundation at the herbarium Jeny Renz (Botanical Institute, Basel University).

Literature Cited

Alexandersson R, Ågren J (2000) Genetic structure in the nonrewarding, bumblebee-pollinated orchid *Calypso bulbosa*. Heredity 85: 401-409.

Bateman RM, Devey DS, Malmgren S, Bradshaw E, Rudall PJ (2009) Conflicting species concepts underlie perennial taxonomic controversies in *Ophrys*. In: Actes of the SFO colloquium, Montpellier.

Bernardos S, Garcia-Barriuso M, Sanchez-Anta MA, Amich F (2007) Composition, geographical affinities and endemism of the Iberian Peninsula orchid flora. Nord. J. Bot. 25: 227-237.

Bianco P, Demerico S, Medagli P, Ruggiero L (1991) Polyploidy and aneuploidy in *Ophrys*, *Orchis*, and *Anacamptis* (Orchidaceae). Plant Syst.Evol. 178: 235-245.

Borg-Karlson AK, Groth I, Agren L, Kullenberg B (1993) Form-specific fragances from *Ophrys insectifera* L. (Orchidaceae) attract species of different pollinator genera. Evidence of sympatric speciation? Chemoecology 4: 39-45.

Breistroffer M (1981) Notes succinctes sur quelques équivalences nomenclaturales d'espèces céveno-caussenardes et description d'une sous-espèce nouvelle d'Orchidée. B Soc. Bot. Fr.-Lett. 128: 69-72.

Chase MW, Hills HH (1991) Silica-gel: an ideal material for field preservation of leaf samples for DNA studies. Taxon 40: 215-220.

Comes HP, Kadereit JW (1998) The effect of quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 3: 432-438.

Cozzolino S, Widmer A (2005) Orchid diversity: an evolutionary consequence of deception? Trends Ecol. Evol. 20: 487-494.

Delforge P (1983) Remarques sur *Ophrys insectifera* subsp. *aymoninii* Breistroffer et description d'un hybride nouveau de cette sous-espèce: *Ophrys insectifera* nsubsp. *tytecaeana* Delforge. L’Orchidophile 55: 307-312.

Delforge P (2001) *Guide des orchidées d'Europe, d'Afrique du Nord et du Proche-Orient, 2nd ed.* Delachaux et Niestlé, Paris.

D'Emerico S, Pignone D, Bartolo G, Pulvirenti S, Terrasi C, Stuto S, Scrugli A (2005) Karyomorphology, heterochromatin patterns and evolution in the genus *Ophrys* (Orchidaceae). Bot. J. Linnean Soc. 148: 87-99.

Devey DS, Bateman RM, Fay MF, Hawkins JA (2008) Friends or relatives? Phylogenetics and species delimitation in the controversial European orchid Genus *Ophrys*. Ann. Bot.-London 101: 385-402.

Devey DS, Bateman RM, Fay MF, Hawkins JA (2009) Genetic structure and systematic relationships within the *Ophrys fuciflora* aggregate (Orchidaceae: Orchidinae): high diversity in Kent and a wind-induced discontinuity bisecting the Adriatic. Ann. Bot.-London 104: 483-495.

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

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes 7: 574-578.

Greilhuber J, Ehrendorfer F (1975) Chromosome-numbers and evolution in *Ophrys* (Orchidaceae). Plant Syst. Evol. 124: 125-138.

Hermosilla CE, Amardeilh JP, Soca R (1999) *Sterictiphora furcata* Villers, pollinisateur d'*Ophrys subinsectifera* Hermosilla & Sabando. L'Orchidophile 139: 247-254.

Hewitt GM (1999) Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. 68: 87-112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. Nature 405: 907-913.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25: 1965-1978.

Jansson R (2003) Global patterns in endemism explained by past climatic change. P. R. Soc. B 270: 583-590.

Kruckeberg AR, Rabinowitz D (1985) Biological aspects of endemism in higher-plants. Annu. Rev. Ecol. Syst. 16: 447-479.

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. Divers. Distrib. 15: 59-69.

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

Pfeifer M, Schatz B, Pico FX, Passalacqua NG, Fay MF, Carey PD, Jeltsch F (2009) Phylogeography and genetic structure of the orchid *Himantoglossum hircinum* (L.) Spreng. across its European central-marginal gradient. J. Biogeogr. 36: 2353-2365.

Pillon Y, Chase MW (2007) Taxonomic exaggeration and its effects on orchid conservation. Conserv. Biol. 21: 263-265.

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

Sagarin RD, Gaines SD (2002) The 'abundant centre' distribution: to what extent is it a biogeographical rule? Ecol. Lett. 5: 137-147.

Schatz B, Jacob L (2007) Enjeux pour la conservation des orchidées méditerranéennes. Journée scientifique au Parc naturel régional des Grands Causses, Millau, France.

Schiestl FP, Ayasse M (2002) Do changes in floral odor cause speciation in sexually deceptive orchids? Plant Syst. Evol. 234: 111-119.

Schlüter PM, Ruas PM, Kohl G, Ruas CF, Stuessy TF, Paulus HF (2007) Reproductive isolation in the Aegean *Ophrys omegaifera* complex (Orchidaceae). Plant Syst. Evol. 267: 105-119.

Schlüter PM, Ruas PM, Kohl G, Ruas CF, Stuessy TF, Paulus HF (2009) Genetic patterns and pollination in *Ophrys iricolor* and *O. mesaritica* (Orchidaceae): sympatric evolution by pollinator shift. Bot. J. Linnean Soc. 159: 583-598.

Schmeller DS, Gruber B, Budrys E, Framsted E, Lengyel S, Henle K (2008) National responsibilities in European species conservation: A methodological review. Conserv. Biol. 22: 593-601.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. Front. Zool. 6:9.

Squirrell J, Hollingsworth PM, Bateman RM, Dickson JH, Light MHS, MacConaill M, Tebbitt MC (2001) Partitioning and diversity of nuclear and organelle markers in native and introduced populations of *Epipactis helleborine* (Orchidaceae). Am. J. Bot. 88: 1409-1418.

Svenning JC (2003) Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. Ecol. Lett. 6: 646-653.

Thompson JD (2005) Plant Evolution in the Mediterranean. Oxford University Press, New York.

Thuiller W, Lafourcade B, Engler R, Araujo MB (2009) BIOMOD - a platform for ensemble forecasting of species distributions. Ecography 32: 369-373.

Tribsch A, Schonswetter P (2003) Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the Eastern Alps. Taxon 52: 477-497.

Vekemans X (2002) AFLP-SURV version 1.0. Distributed by the author, Free University of Brussels.

Vereecken NJ (2009) Pollinator-mediated selection, reproductive isolation and the evolution of floral traits in the genus *Ophrys* (Orchidaceae). Dissertation, Free University of Brussels.

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA-fingerprinting. Nucleic Acids Res. 23: 4407-4414.

Walters SM (1976) The conservation of threatened vascular plants in Europe. Biol. Conserv. 10: 31-41.

Tables

Table 1. Geographic indications for the investigated populations of the *O. insectifera* species group (listed by country), showing for each population which species were collected and the number of genotyped individuals (N).

Population	Country	Location	Lat	Lon	Species*	N
BAR	Austria	Barwies	47°N	11°W	Oi	1
GOS	Austria	Gosau	48°N	14°W	Oi	2
WEI	Austria	Weissenbach	48°N	14°W	Oi	1
YAG	Bulgaria	Yagodina	42°N	24°W	Oi	4
DRA	Croatia	Drakulic rijeka	45°N	16°W	Oi	4
GOR	Croatia	Gornji Vhrovci	45°N	18°W	Oi	4
OPO	Czech Republic	Oponelec	49°N	14°W	Oi	3
PAR	Finland	Parainen/Pargas	60°N	22°W	Oi	1
BEL	France	Belfort	44°N	3°W	Oa	5
COU	France	la Couvertoirade	44°N	3°W	Oa	4
CUN	France	les Cuns	44°N	3°W	Oa	1
LAY	France	Layolle	44°N	3°W	Oa	6
LIC	France	Licide	44°N	3°W	Oa	7
LIQ	France	Les Liquisses	44°N	3°W	Oa	8
MAS	France	Le Massegros	44°N	3°W	Oa	6
MAZ	France	Les Mazes	44°N	3°W	Oa	7
PEZ	France	La Pézade	44°N	3°W	Oa	5
SAT	France	Saint Saturnin	44°N	3°W	Oa	5
SER	France	Sérigas	44°N	3°W	Oa	5
SGL	France	Saint-Georges de Lévéjac	44°N	3°W	Oa	5
VES	France	Vessac	44°N	3°W	Oa	6
LOO	France	Lôo	43°N	1°W	Oi	3
CER	France	Cereste	44°N	6°W	Oi	3
LAA	France	Lapan	47°N	2°W	Oi	3
LAP	France	Lapanouse de Cernon	44°N	3°W	Oi	8
NAN	France	Gorges du Nan	45°N	5°W	Oi	3

QUI	France	Quilen	51°N	2°W	Oi	3
RBO	France	Rambaud	45°N	6°W	Oi	2
THH	France	Thorame-Haute	44°N	6°W	Oi	2
THO	France	Thorame-Basse	44°N	7°W	Oi	5
THU	France	La Thuile	46°N	6°W	Oi	5
VAC	France	La Vacquerie	44°N	3°W	Oi	4
FON	France	Fontaret	44°N	4°W	Oi, Oa	6
TRI	France	La Trivalle	44°N	3°W	Oi, Oa	4
VIL	France	le Villaret	45°N	4°W	Oi, Oa	10
AUG	Germany	Augsburg	48°N	10°W	Oi	4
WOR	Germany	Wormental	50°N	10°W	Oi	4
CUD	Great Britain	Cudham	51°N	0°W	Oi	2
HRC	Great Britain	Hutton Roof Crags	54°N	3°E	Oi	3
DEL	Greece	Delvinaki	40°N	20°W	Oi	4
GEA	Ireland	Lough Gealain	53°N	9°E	Oi	3
MPI	Italy	Monte Pizi	42°N	14°W	Oi	1
PIA	Italy	Pian Crixia	45°N	8°W	Oi	3
SEN	Italy	Senerchia	41°N	15°W	Oi	3
SZC	Poland	Szczawnica	49°N	21°W	Oi	3
KLA	Slovakia	Klastor	49°N	19°W	Oi	3
KOR	Slovenia	Korada	46°N	14°W	Oi	4
UNE	Slovenia	Unec	46°N	14°W	Oi	3
EUG	Spain	Sta. Eugenia de Relat	42°N	2°W	Oi	1
RUE	Spain	Ruesta	43°N	1°E	Oi	1
TRO	Spain	Trokoniz	43°N	3°E	Oi	4
BER	Spain	Bernues	42°N	1°E	Oi, Os	7
ERR	Spain	Erro	43°N	1°E	Oi, Os	5
OST	Spain	Ostiz	43°N	2°E	Oi, Os	4
COM	Spain	Les Comes	42°N	2°W	Os	7
CON	Spain	Contrast	42°N	1°W	Os	4
SAV	Spain	Savalla del Comtat	42°N	1°W	Os	1
SSO	Spain	San Sadurni d'Osormort	42°N	2°W	Os	4
BAK	Sweden	Baeck	57°N	17°W	Oi	3
SJO	Sweden	Sjostorp	58°N	15°W	Oi	2
BOU	Switzerland	Boudry	47°N	7°W	Oi	1
CRO	Switzerland	Crostand	47°N	7°W	Oi	3
MER	Switzerland	Meride	46°N	9°W	Oi	1
PLA	Switzerland	Planige	46°N	8°W	Oi	2
URS	Switzerland	Saint-Ursanne	47°N	7°W	Oi	3

* Oi: *Ophrys* insectifera, Oa: *O. aymoninii* and Os: *O. subinsectifera*

Table 2. Genetic diversity results for the endemic species *O. aymoninii*. The number of individuals analysed (N), the percentage of polymorphic loci (PLP) and the Nei's diversity value (Hj) are given for each population, as well as for different geographic levels.

		N	PLP	Hj
Populations				
BEL	Belfort	10	70.5	0.25592
COU	la Couvertoirade	5	65.3	0.20935
CUN	les Cuns	3	62	0.32222
FON	Fontaret	6	65.3	0.24697
LAP	Lapanouse de Cernon	3	63.5	0.21734
LAY	Layolle	10	67.9	0.23422
LIC	Licide	10	72.3	0.28029
LIQ	les Liquisses	10	76.8	0.23322
MAS	le Massegros	9	71.2	0.24773
MAZ	les Mazes	10	74.2	0.26856
PEZ	la Pézade	10	73.1	0.2384
SAT	Saint-Saturnin	5	67.5	0.2245
SER	Sérigas	10	69.7	0.27414
SGL	Saint-Georges de Lévéjac	6	65.7	0.2671
VES	Vessac	9	67.5	0.26695
VIL	le Villaret	10	71.2	0.24551
Causses				
	Larzac	44	72.3	0.25718
	Noir	49	67.9	0.25961
	Sauveterre	33	71.2	0.23646
French "Departement"				
	Aveyron	45	71.2	0.25786
	Gard	46	66.4	0.24658
	Lozere	35	66.8	0.25869
French administrative region				
	Languedoc-Roussillon	81	66.8	0.2491
	Midi-Pyrenees	45	71.2	0.25786

Figures and figure legends

Figure 1: Spatial genetic structure of the *O. insectifera* group for the best run of the STRUCTURE analysis (K=6). The six genetic groups obtained are represented by different colours. Each population, represented as a pie-chart, shows the proportion of samples attributed to a particular genetic lineage. Pie-charts are proportional to the number of analysed individuals. Non-assigned samples at the 0.95 probability level are also shown as a transparent part of pie-charts.

Figure 2: Spatial genetic structure of the *O. insectifera* group as an average of the three best runs of the STRUCTURE analysis (K=6). The six genetic groups obtained are represented by different colours. Each population, represented as a pie-chart, shows the average proportion of assignment (among the three best runs) for each genetic lineage. Populations followed by an asterisk are represented by one single individual analysed.

Figure 3: Respective proportion of the six genetic lineages obtained for the best run of the STRUCTURE analysis (K=6), for each of the three species of the *O. insectifera* group.

Figure 4: Modelling of the potential European distribution of the *O. insectifera* group at last glacial maximum (-21000 ybp), also showing main icecaps. Darker is the area, higher was the probability for this area to provide optimal ecological conditions.

Figure 5: Pattern of genetic diversity for the endemic *O. aymoninii*. Values of Nei's diversity are showed per population according to five classes. The corresponding number of analysed individuals per population is also mentioned.

Figure 6: Geographical trends of genetic diversity for sampled *O. aymoninii* populations. The value of Nei's genetic diversity per population decreases with increasing latitudinal (A) and longitudinal distance (B) from the distribution centre.

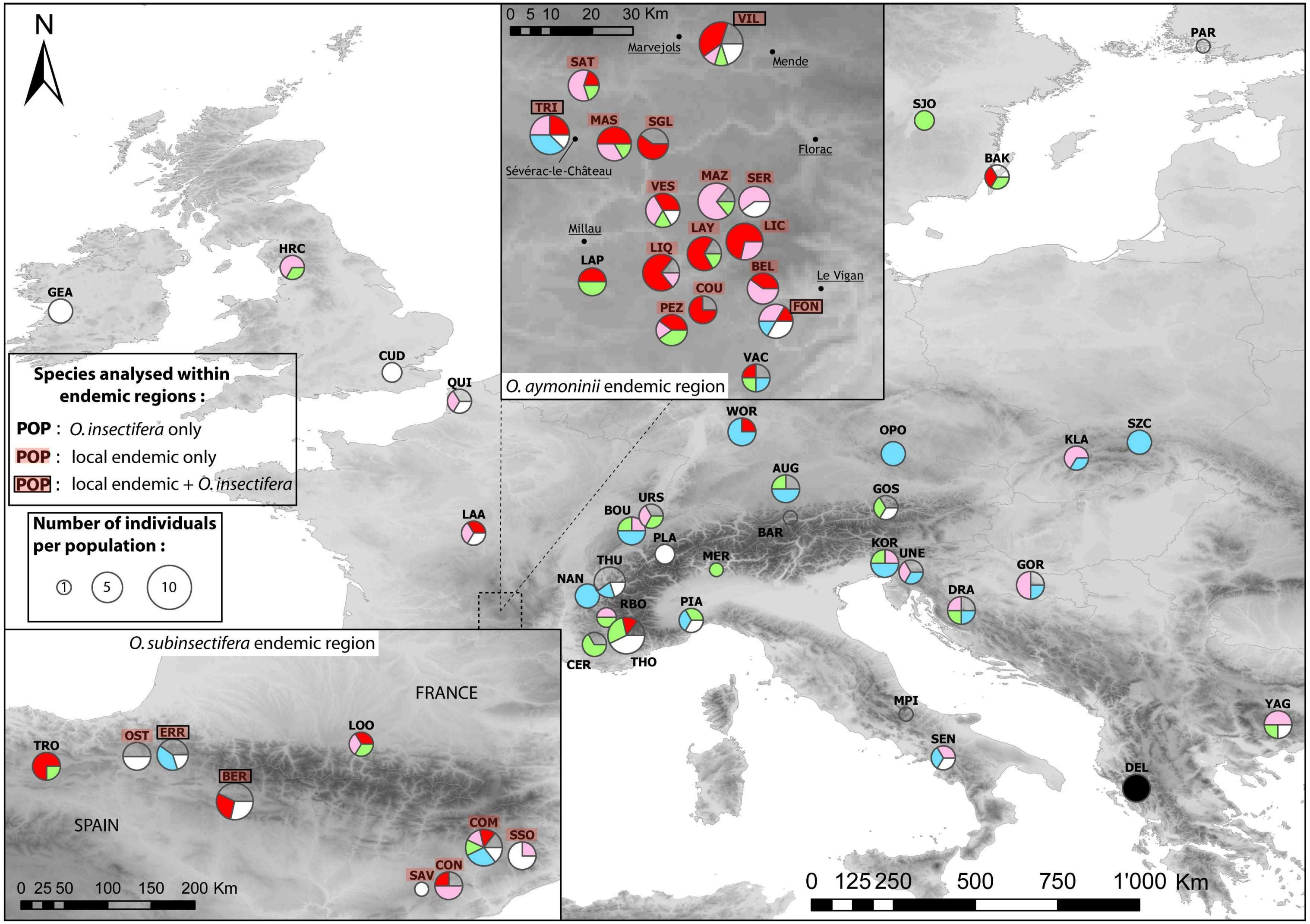


Figure 1: Spatial genetic structure of the *O. insectifera* group for the best run of the STRUCTURE analysis (K=6). The six genetic groups obtained are represented by different colours. Each population, represented as a pie-chart, shows the proportion of samples attributed to a particular genetic lineage. Pie-charts are proportional to the number of analysed individuals. Non-assigned samples at the 0.95 probability level are also shown as a transparent part of pie-charts.

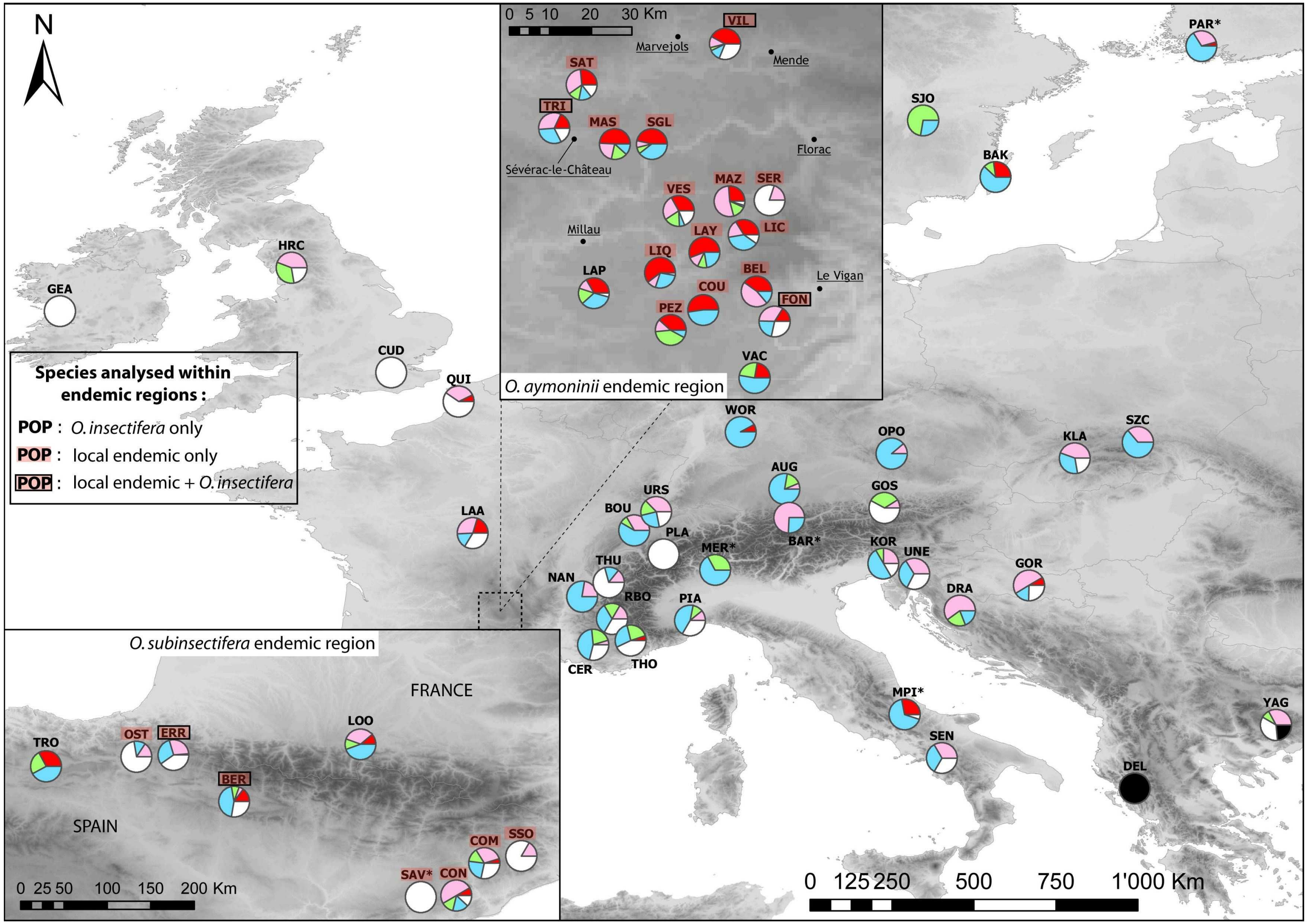


Figure 2: Spatial genetic structure of the *O. insectifera* group as an average of the three best runs of the STRUCTURE analysis (K=6). The six genetic groups obtained are represented by different colours. Each population, represented as a pie-chart, shows the average proportion of assignment (among the three best runs) for each genetic lineage. Populations followed by an asterisk are represented by one single individual analysed.

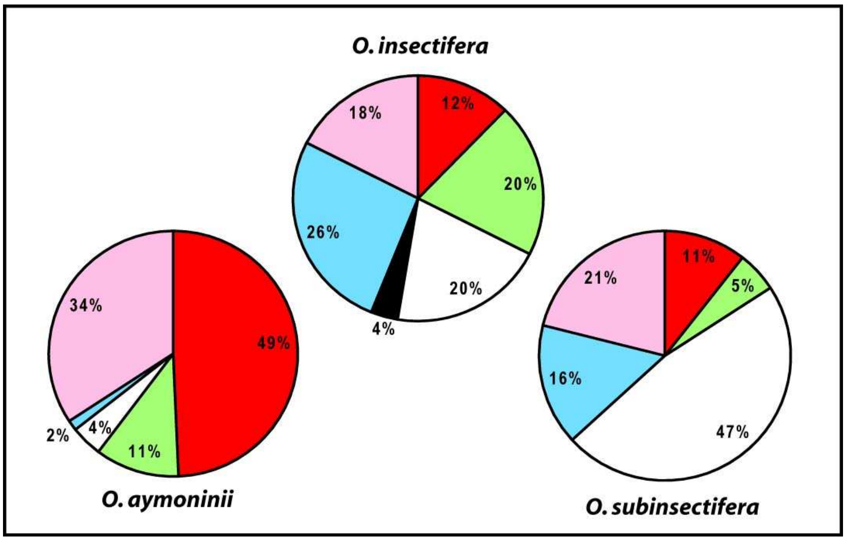


Figure 3: Respective proportion of the six genetic lineages obtained for the best run of the STRUCTURE analysis (K=6), for each of the three species of the *O. insectifera* group.

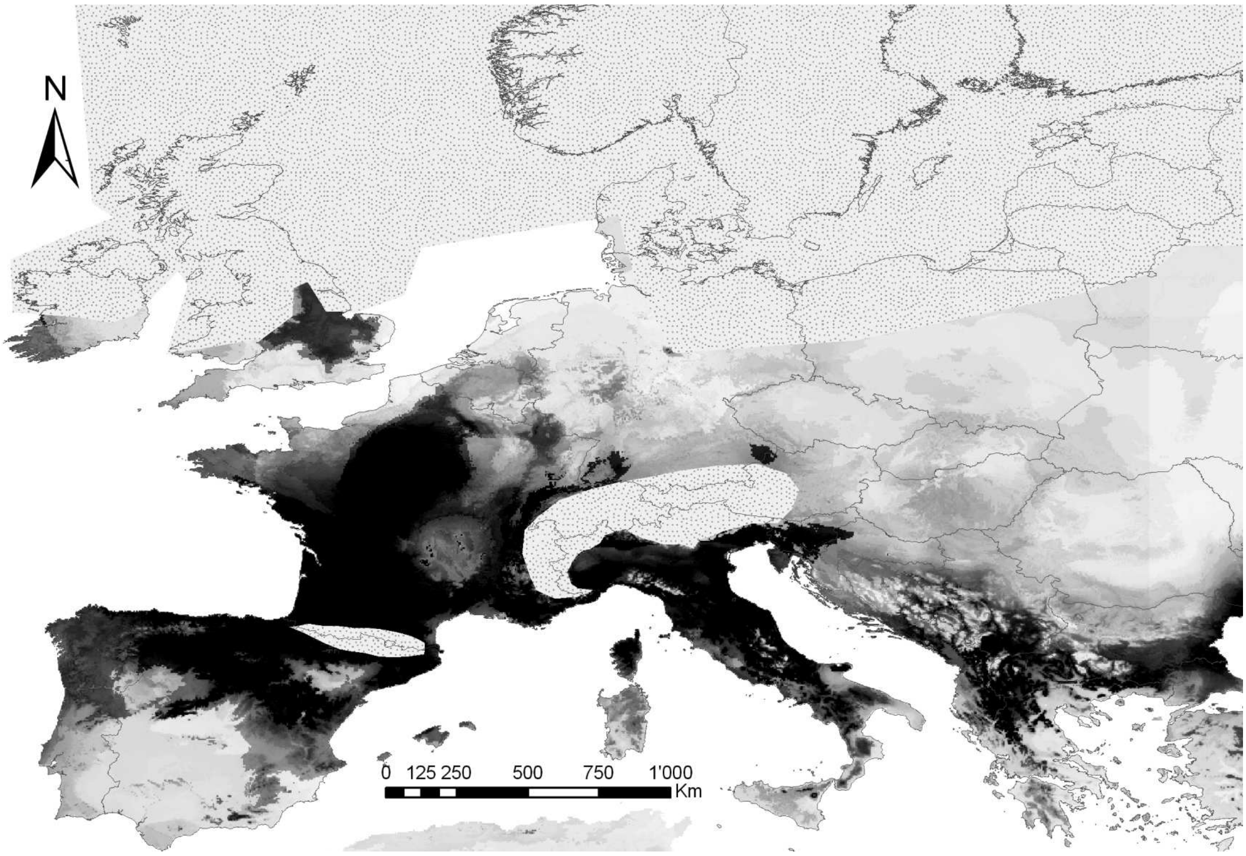


Figure 4: Modelling of the potential European distribution of the *O. insectifera* group at last glacial maximum (-21000 ybp), also showing main icecaps. Darker is the area, higher was the probability for this area to provide optimal ecological conditions.

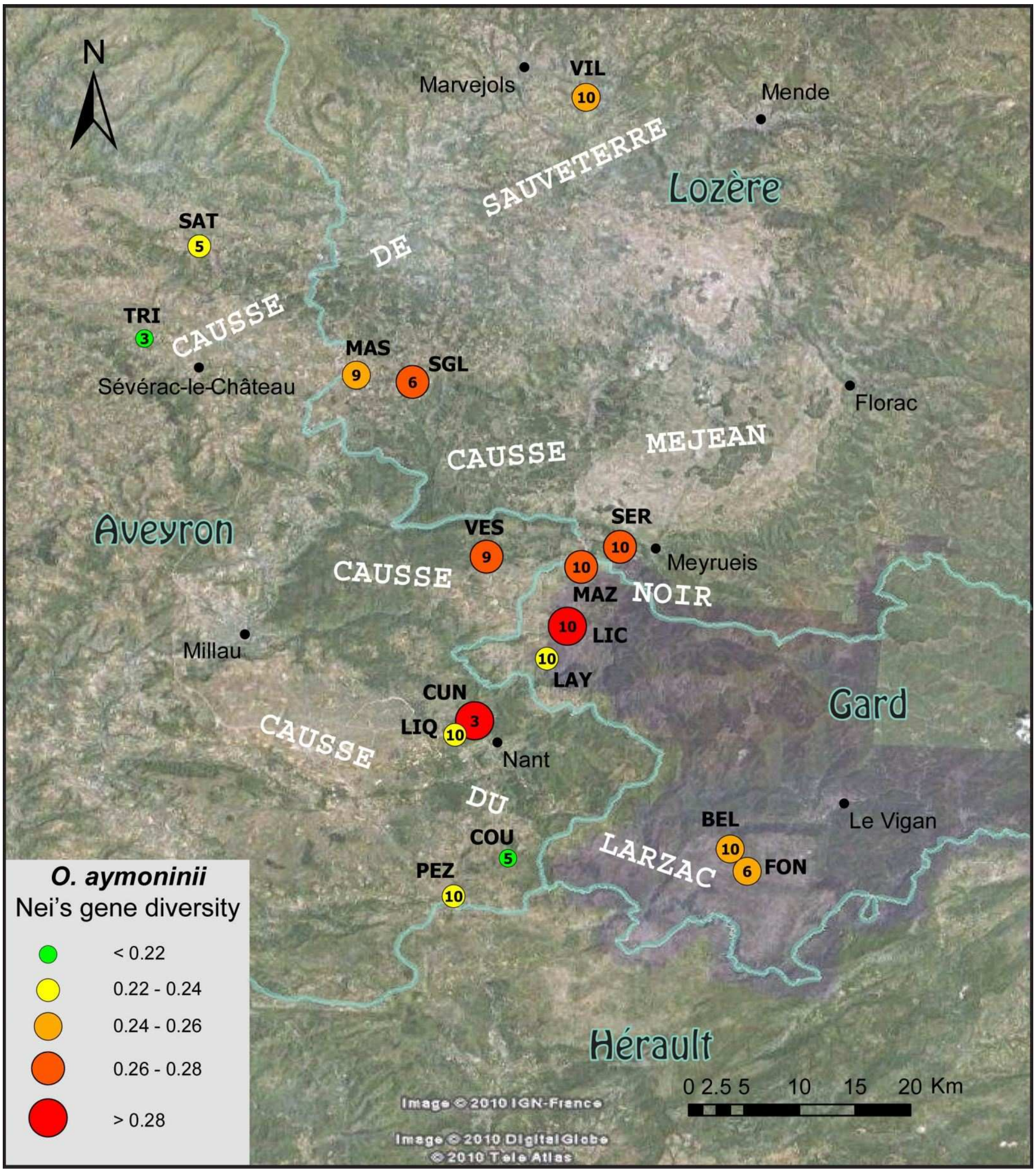


Figure 5: Pattern of genetic diversity for the endemic *O. aymoninii*. Values of Nei's diversity are showed per population according to five classes. The corresponding number of analysed individuals per population is also mentioned.

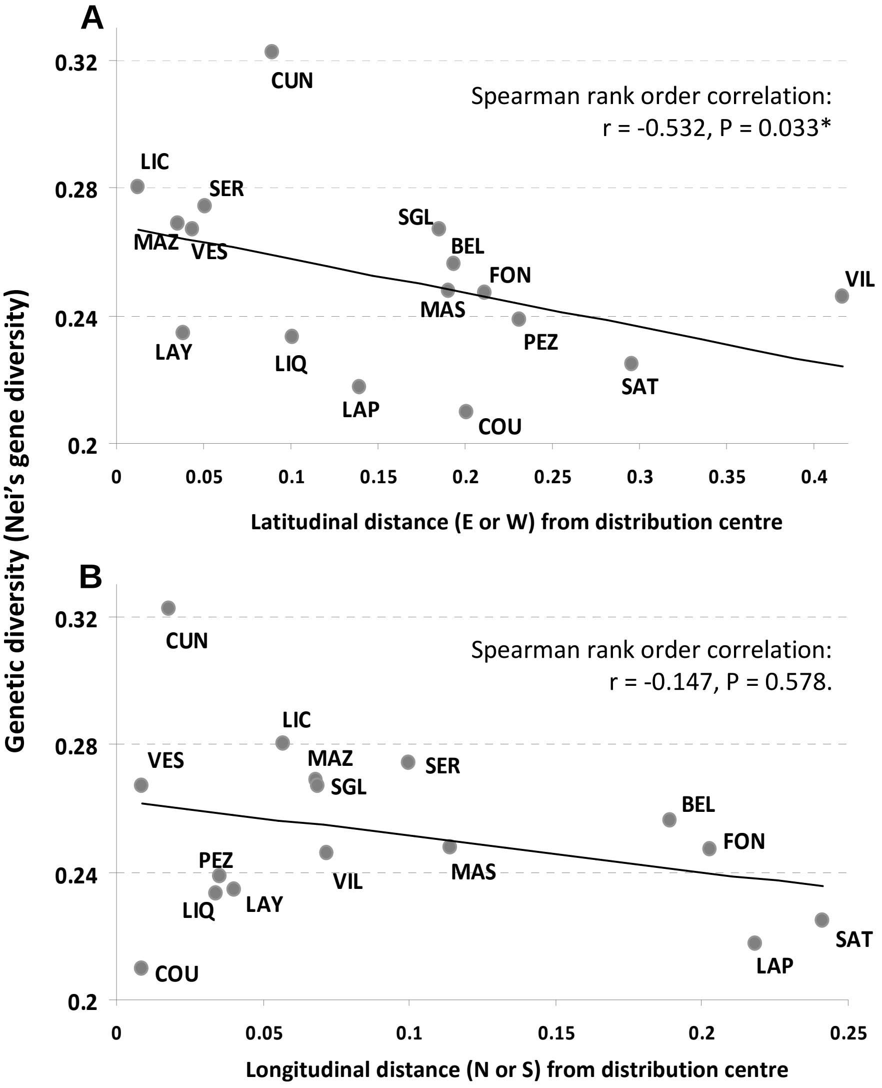


Figure 6: Geographical trends of genetic diversity for sampled *O. aymoninii* populations. The value of Nei's genetic diversity per population decreases with increasing latitudinal (A) and longitudinal distance (B) from the distribution centre.

Supplementary material 1

Diversity clues for future conservation of *O. aymoninii*

Recent identification of threats for *O. aymoninii* (populations monitoring combined to habitat's characterisation) foresaw a potential future inscription of the species on the IUCN red list of threatened species (Schatz & Jacob 2008). The present study, covering a large zone of the species' distribution, gave an opportunity to describe the pattern of genetic diversity for *O. aymoninii.* Such results would ideally complement the circumscription of important zones for the protection of the endemic species.

Table 3. Detailed results of Nei's gene diversity per causse, per administrative regions and for combined data. Cross results also shown the concerned populations used for calculating diversity indices.

		REGION / DEPARTMENT		
		Midi-Pyrénées 0.258	**Languedoc-Roussillon** 0.249	
		Aveyron 0.258	**Gard** 0.247	**Lozère** 0.259
CAUSSE	**Larzac** 0.257	(COU, CUN, LIQ, PEZ) 0.268	(BEL, FON) 0.246	**-**
	Noir 0.260	(VES) 0.267	(LAY, LIC, MAZ) 0.254	(SER) 0.274
	Sauveterre 0.236	(TRI, SAT) 0.223	**-**	(MAS, SGL, VIL) 0.244

As an important result of the genetic diversity study on *O. aymoninii*, we showed that populations at the centre of the distribution were more diverse genetically compared to those in the periphery (see Fig. 5 of main article). As reported here in details (Table 3), we could identify key zones to consider for conservation actions concerning this endemic species, with some particular areas encompassing high levels of genetic diversity.

At the level of French administrative areas (or “department”), Lozère in region Languedoc-Roussillon and Aveyron in region Midi-Pyrénées (where regional protection is already effective) gave the best results. For both departments, their southern parts appeared to be more diverse genetically, with highest levels either on the causse Noir part of Lozère and on both causse Noir and causse du Larzac areas of Aveyron.

At the causse level, causse Noir showed the highest Nei’s diversity. This causse, extending on three different departments, showed less diversity in its central part (situated on Gard). However, one should remark that higher values obtained at both ends of causse Noir (located on Lozère in the east and Aveyron in the west) represent results from one single population per region. As a final contrasting note considering general diversity of causse Noir, we remind that it harboured the lowest percentage of polymorphic loci (see Table 2 in main article). Causse du Larzac was somewhat similar to causse Noir considering genetic diversity, with slightly inferior value for Nei’s diversity, but harbouring the highest percentage of polymorphic loci. Western part of Larzac situated on Aveyron seemed to be more diverse than its eastern part on Gard. Finally, the lowest diversity values were obtained for the northernmost causse de Sauveterre. It showed the opposite trend of Larzac, i.e. highest values of diversity in the east (on Lozère) than in the west (on Aveyron). Finally, complementary populations from causse Méjean should be added in the future to unravel the complete diversity pattern of *O. aymoninii*. But for this last causse, regarding its central location we also would expect it to hold above-than-average diversity values (similar to causse Noir or Larzac), especially in its western half.

As recommendations concerning the conservation of the endemic, adding regional protection in the Languedoc-Roussillon to currently existing protection in Midi-Pyrénées would of course ideally preserve the full species distribution. Protection at the department level in Lozère and Gard would perhaps be more focused, in view of *O. aymoninii* populations’ occurrences within the region, even if one should not forget that additional populations are found in Hérault as well (Schatz &

Jacob 2008). But more than large-scale efforts, optimal habitats should be preserved firstly so that currently known populations could maintain. As we showed that populations at the centre of the distribution were mostly diverse genetically, particular attention should be paid to the habitats of *O. aymoninii* within this area.

Litterature cited :

Schatz B, Jacob L (2007) Enjeux pour la conservation des orchidées méditerranéennes. Journée scientifique au Parc naturel régional des Grands Causses. p. 88, Millau, France.

Supplementary material 2

We extracted occurrences of the species in Europe from the GBIF database (www.gbif.org). We kept only the data with a minimal spatial accuracy of 5km. These data were pooled with the occurrences collected during the fieldwork. Since the occurrences were highly aggregated in some part of Europe, we randomly selected a subset of occurrences with a minimal distance of 50km. Most modeling techniques require not solely information about presences but also absences to determine the suitable conditions for a given species, thus we selected 10'000 pseudo-absences randomly. The modeling techniques will then discriminate the conditions where there are presences from the background environment (Witz and Guisan, 2009). GBIF is known to be highly biased for western European countries (Yesson et al., 2007). Therefore, we did not select pseudo-absences in Eastern Europe countries but limited them to the European Union with Switzerland, Norway and Balkan countries also included. Selecting pseudo-absences in unsampled countries can significantly bias the response curves of the models (Phillips et al., 2009), even if the risk is limited with the studied species as it is known to occur mostly in Western Europe. The resulting presences and pseudo-absences were used in the following species distribution modeling.

We run single-models with 7 climatic variables originated from Worldclim (Hijmans et al., 2005) at a resolution of 2.5 arc-minute (around 5 km): maximum temperature of warmest month (bio5), minimum temperature of coldest month (bio6), mean temperature of warmest quarter (bio10), mean temperature of coldest quarter (bio11), precipitation seasonality (bio15), precipitation of the warmest quarter (bio18) and precipitation of the coldest quarter (bio19).

We modeled the distribution of the species using the BIOMOD tool (Thuiller et al., 2009), as implemented for R software. We used 8 different niche-based modeling techniques : (1) generalized linear model (GLM), (2) generalized additive model (GAM), (3) classification tree analysis (CTA), a classification method which select the best trade-off between the number of leaves of the tree and the explained deviance, (4) artificial neural networks (ANN), a machine learning method, with the mean of three runs used to provide predictions and projections, as each simulation gives slightly different results, (5) mixture discriminant analysis (MDA), a classification method that uses MARS function for the regression part of the model, (6) multivariate adaptive regression splines (MARS), (7) generalized boosting model (GBM), a machine learning method which combines a boosting algorithm and a regression tree algorithm to construct an 'ensemble' of trees, and (8) Random Forest (RF), a machine learning method which is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. These 8 methods are implemented in the BIOMOD package.

In order to evaluate the predictive performance of the species distribution model, we used a random subset of 70% of the data to calibrate every model, and used the remaining 30% for the evaluation. Models were evaluated using a relative operating characteristic (ROC) curve and the Area Under the Curve (AUC) (Fielding and Bell, 1997). We replicated the data splitting 10 times and calculated the average AUC of the repeated split-sample, which gives a more robust estimate of the predictive performance of each model.

Finally, each model was projected into both current and past climate conditions with the Worldclim data of the CCSM circulation model for the last glacial maximum (LGM; -21’000 years) downscaled at a resolution of 2.5 arc-minute. In order to obtain the central tendency of these distributions, accounting for variations among modeling techniques, we applied a weighted average of the 8 modeling techniques based on the predictive power (AUC). Ensemble forecasting approaches have been shown to significantly improve the accuracy of species distribution models (Marmion et al., 2009).

Literature cited:

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24, 38-49.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978.

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modeling. Diversity and Distributions 15, 59-69.

Phillips SJ, Dudik M, Elith J, Graham C, Lehmann A, Leathwick J, & Ferrier S (2009) Sample selection bias and presence-only models of species distributions: implications for selection of background and pseudo-absences. Ecological Applications 19, 181-197.

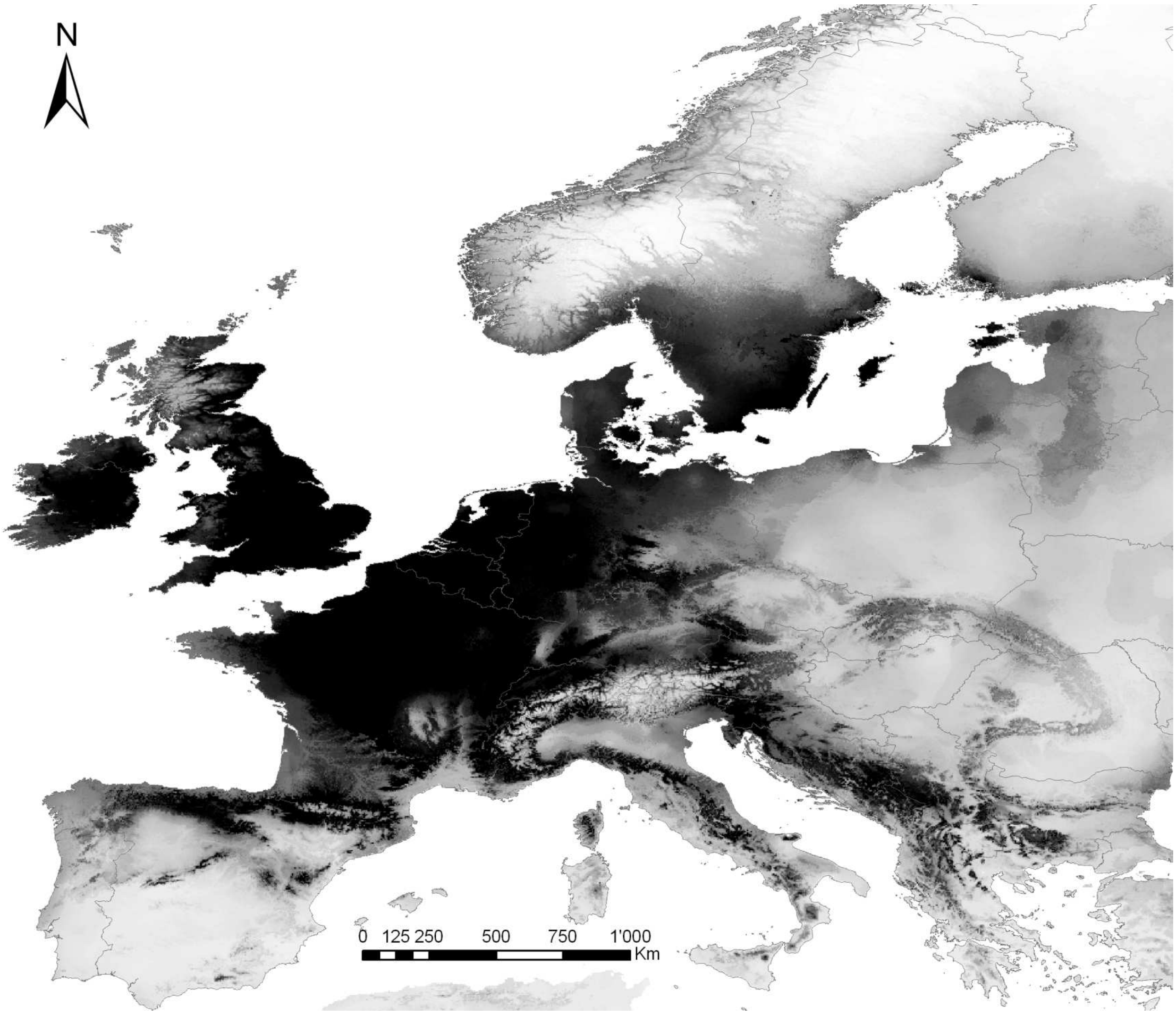
Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD - a platform for ensemble forecasting of species distributions. Ecography 32, 369-373.

Wisz MS, Guisan A (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecology 9:8.

Yesson C, Brewer PW, Sutton T, Caithness N, Pahwa JS, Burgess M, Gray WA, White RJ, Jones AC, Bisby FA, Culham A (2007) How Global Is the Global Biodiversity Information Facility? PLoS ONE 2, e1124.

Supplementary material 3

This map represents the current potential distribution modelled for *O. insectifera*. The areas with a high probability of hosting appropriate ecological conditions for the fly orchid are displayed in black.



CHAPTER IV

Biogeography of an oil-collecting mutualism: tight interactions involve shared refugia but independent dispersal of partner species

Yann Triponez, Anahí Espíndola, Lucas Bassin, Loïc Pellissier and Nadir Alvarez

(Version formatted according to *Ecology Letters*)

Neuchâtel, August 2010

Biogeography of an oil-collecting mutualism: tight interactions involve shared refugia but independent dispersal of partner species

Yann TRIPONEZ*, Anahí ESPÍNDOLA, Lucas BASSIN, Loïc PELLISSIER and Nadir ALVAREZ

* Corresponding author

Yann TRIPONEZ, yann.triponez@unine.ch, Laboratoire d'Entomologie Evolutive, Institut de Biologie, Université de Neuchâtel, Rue Emile-Argand 11, cp 158, CH-2009 Neuchâtel, Switzerland. Phone +41 32 718 31 62 - Fax +41 32 718 30 01

Anahí ESPÍNDOLA, maria.espindola@unine.ch, Laboratoire d'Entomologie Evolutive, Institut de Biologie, Université de Neuchâtel, Switzerland

Lucas BASSIN, lucas.bassin@unil.ch, Laboratoire d'Entomologie Evolutive, Institut de Biologie, Université de Neuchâtel, Switzerland

Loïc PELLISSIER, loic.pellissier@unil.ch, Département d'Ecologie et Evolution, Université de Lausanne, Switzerland

Nadir ALVAREZ, nadir.alvarez@unil.ch, Département d'Ecologie et Evolution, Université de Lausanne, Switzerland

Running title: Biogeography of a plant-insect mutualism

Type of article: *Letter*

Number of words in the abstract: 150

Number of words in the manuscript as a whole: 7043

Number of words in the main text: 5594

Number of references: 38

Number of figures: 6

Number of tables: 1

Abstract

Mutualistic interactions between plants and insects are suitable case-studies for testing simple hypotheses in comparative phylogeography. Because of the dependence among mutualists, one could expect to see partners dispersing in parallel, showing a global trend towards congruence in their respective spatial genetic structures. Here we compare the phylogeographic patterns of three species involved in a specialized pollination system, i.e., the oil-producing plant *Lysimachia vulgaris* and two oil-collecting bees within the genus *Macropis*. Our analysis combines fine-scale spatial genetic structuring and past ecological niches modeling. Results report an expected wide overlap in potential refugia during last glacial maximum between the plant and its pollinators. However the current spatial genetic structures reveal that re-colonization pathways of the three species were not similar. These results are discussed in view of the different dispersal abilities and ecological preferences of the actors, considering as well additional non-obligate phylogenetically-related partners, which punctually allow maintaining this mutualism.

Key words

Comparative phylogeography, oil-collection mutualism, *Lysimachia*, *Macropis*, spatial genetic structuring, western Palearctic

INTRODUCTION

The spatial distribution of interacting organisms is driven by a wide diversity of processes, ranging from coevolution and local adaptation, to historical processes (Thompson & Cunningham 2002). Such processes should, however, also influence the fate of lineages within species (Thiel-Egenter *et al.* submitted) and more particularly, the intra-specific spatial genetic structure (SGS) of interacting species as well as the level of congruence in the spatial arrangement of their respective gene pools (Alvarez *et al.* 2010). Whereas in parasitic interactions, organisms are affected by antagonistic evolutionary forces (e.g., regarding their dispersal abilities and patterns of gene flow) and are likely to demonstrate incongruent patterns of genetic structure (e.g. Althoff & Thompson 1999), in mutualisms, one would expect to see partners dispersing in parallel, showing a global trend towards congruence in their respective SGS (e.g., Anderson *et al.* 2004). However, whether or not the kind of interaction (i.e., parasitic vs. mutualistic) is correlated with the level of congruence among interacting species is still a debated question, and other factors (e.g., differences in generation times and population sizes) seem to interfere with this prediction (see Alvarez *et al.* 2010). Mutualisms between insects and plants are among the most thoroughly studied biotic interactions (Bronstein 1994) and could represent nice case-studies for testing the congruence level in the respective SGS of mutualistic species. Because multi-partner systems seem much too complex to be apprehended concretely, it is preferable to first consider narrow mutualisms (i.e. including as few partners as possible) which allow tests of straightforward hypotheses in the simplest comparative phylogeography framework.

Cooperative relations between plants and insects exist under many diversified forms, among which pollination interactions, ant-mediated seed dispersal (myrmecochory) and plant protection, as well as more subtle digestive mutualism (Anderson & Midgley 2003; Bronstein *et al.* 2006). Specialised pollination systems encompass a large number of tight interactions and are therefore well suited to test for congruence in the respective SGS of interacting species. For instance, the co-evolutionary processes taking place between plants and insects involved in nursery pollination systems such as the yucca/yucca moth (Pellmyr 2003) or the fig/fig wasp systems (Cook & Rasplus 2003; Machado *et al.* 2005) have shown trends towards co-cladogenesis of both partners at a wide scale of time (e.g., Machado *et al.* 2001). In such systems, the specificity of the interaction has driven extensive diversification processes both in plants and insects, as attested by the large number of species found in these groups. However, defining species boundaries in such ancient and highly-diversified groups might be difficult and testing hypotheses in the field of comparative phylogeography might not be evident. Moreover, as a common feature to all nursery pollination systems (and more generally to most plant-insect mutualisms), the one-to-one rule of a single pollinating insect per plant species is generally not respected and is rather rare among the previous examples (Bronstein *et al.* 2006). Cheater species parasiting the interaction often accompany the mutualistic pollinator(s), with sometimes unclear limits between the two categories. For instance, in another, less diversified, nursery pollination system composed of the European globeflower (*Trollius europaeus*) and its associated flies within the genus *Chiastocheta*, the different pollinators differ in the timing of their life cycle and in the number of eggs laid: whereas some species visit early-days flowers, lay very few eggs and are therefore efficient mutualistic partners, other species lay many eggs on old flowers (or even on fruits), thus do not necessary participate to pollination and should be more considered as parasitic. In between, several other *Chiastochaeta* species that successively visit the flower all along its blooming period show intermediate strategies, potentially contributing to pollination (Despres & Jaeger 1999).

Less documented than nursery pollination, oil-producing plants and oil-collecting bees represent another specialised pollination system (Vogel 1988). Oil-producing flowers, which can be found worldwide in 11 different plant families, are pollinated by a few bee genera within Melittidae and Apidae (Hymenoptera). The oil-collecting bees use oil combined to pollen as larval food, or for water-resistant lining of larval cells (Renner & Schaefer 2010). All oil-collecting bee species have developed special branched hairs on legs or abdomen to facilitate oil transportation (Rasmussen & Olesen 2000), an ancient adaptation already observed in the 53 Mya old fossil species *Palaeomacropis eocenicus* (Michez *et al.* 2007). Strong constraints such as the specific nutritional requirements and the temporal match with the timing of plant flowering have driven insects to a strong level of specificity to a given oil-producing plant. This is notably the case of the oil-collecting bees associated with several temperate yellow-blossoming *Lysimachia* (Myrsinaceae; Vogel 1976). A large proportion of species within this genus, native either of temperate Eurasia (subgenus *Lysimachia* s. str.) or northern America (subgenus *Seleucia*; Hao *et al.* 2004), are fertilised by oil-collecting bees from the genus *Macropis* (Hymenoptera, Melittidae). Here as well, strict one-to-one relationships between oil-collecting bees and *Lysimachia* species do not represent the general rule. However, even if several oil-producing *Lysimachia* species are visited by more than one single species of oil-collecting bees (Cane *et al.* 1983; Simpson *et al.*1983), most plant species demonstrate a trend towards a high specificity to a few effective visiting oil-collecting bees (Pekkarinen *et al.* 2003; Celary 2004; Bassin *et al.* in prep.). Oil flowers mutualisms may therefore represent suitable case studies for testing simple hypotheses in comparative phylogeography.

Here, we investigate the comparative phylogeography of the yellow loosestrife *L. vulgaris* L. (i.e., the most widespread oil-producing species of *Lysimachia* in Eurasia) and its two associated *Macropis* oil-collecting bees in the western Palearctic, *M. europaea* Warncke 1973 and *M. fulvipes* Fabricius 1805. These two species present large distributions spreading from Atlantic coasts of Europe to central Russia and even as far as eastern Russia for the latter (Michez & Patiny 2005). Despite a switch in their altitudinal distribution (Bassin *et al.* in prep), these two widespread *Macropis* species can be found together in sympatric populations (Celary 2004; Pekkarinen *et al.* 2003), except in Anatolia and Caucasus where *M. europaea* is absent, or in the British Islands and in the Scandinavian Peninsula where *M. fulvipes* is missing. These two European *Macropis* species have been reported visiting *L.*

vulgaris as source for oil collection (Simpson *et al.* 1983; Celary 2004; Michez & Patiny 2005*i*). However, whereas *M. europaea* seems strongly monolectic on *L. vulgaris*, *M. fulvipes* can rely on other native *Lysimachia* species as optional sources of oil, for instance *L. punctata* (Dötterl & Schäffler 2007, Y. Triponez pers. obs) and *L. nummularia* (Celary 2004; Bassin *et al.* in prep). A third European species, *M. frivaldszkyi* Mocsary 1878, also associated with *L. vulgaris* (Y. Triponez pers. obs.) shows a more restricted distribution (i.e., Balkans, Anatolia and northern Syria) and is not analyzed in the present study (but see Bassin et al. *in prep*).

In this study, we aim at comparing the SGS of the main partners involved in this mutualism (i.e., *L. vulgaris*, *M. europaea* and *M. fulvipes*) to understand the role of historical changes in the respective distribution of intra-specific lineages in mutualistic species. All three species present large distribution ranges ideal to infer wide-scale phylogeographical patterns. In order to describe these patterns, we use fine-scale Amplified Fragment Length Polymorphism (AFLP) in combination with modelling of past species distribution at the Last Glacial Maximum (LGM; -21'000 years), and investigate how is the genetic variation of *L. vulgaris*, *M. europaea* and *M. fulvipes* structured in Europe, and whether or not each one can be explained by its respective past distribution during LGM. We further test the extent to which the phylogeographic patterns (i.e. LGM potential refugia and post-glacial re-colonization pathways) found in oil-collecting bees are compatible with those unravelled in *L. vulgaris*, as hypothesised given the reciprocal dependence of mutualistic species.

METHODS

Sampling

The sampling of *L. vulgaris* and of the two *Macropis* species was carried out during flowering periods in summer 2007 and 2008 (from June to September depending on the regions). In each plant population visited, five leaves were collected randomly on different plants across the population. This fresh material was desiccated in silica gel (Chase and Hills 1991). Additionally, when blossoming and weather conditions were favourable for pollinators' activity, we also sampled the oil-collecting bees. Females were collected directly on the flowers when pollinating, as patrolling males were caught in flight with the help of an insect net. All pollinators were stored in 70% ethanol. We used the determination key by Michez and Patiny (2005) and a binocular for recognising between the different *Macropis* species. Some extra-material for genetic analysis of *M. europaea* and *M. fulvipes* was also obtained thanks to personal entomological collections of Denis Michez and Lucas Bassin.

AFLP analysis and procedure

For plants, total genomic DNA was extracted from 10 mg of silica-dried leaf fragments using the DNeasy Plant Kit (Qiagen, Hilden, Germany). For insects, in order to avoid plant DNA contamination via pollen grains stuck to the bees' exoskeleton, total genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) from the wing muscles of the bees carefully removed and rinsed. All following reactions were conducted in 96-well plates, in which samples were randomly distributed. In each plate, ten to 15 samples were duplicated for further reproducibility tests. AFLP analysis was slightly modified from the classical method described by Vos *et al.* (1995). The detailed protocols of digestion, ligation, pre-amplification and selective amplification are available by the first author. Final selective PCR products were analysed using the GeneScan technology with an automated capillary-sequencer (ABI 3730XL, Applied Biosystems, Foster City, CA; the service was provided by Macrogen Inc., Seoul, South Korea). Resulting fluorescent AFLP patterns were scored using GeneMapper 3.7 (Applied Biosystems) with some changes in default parameters (maximum bin width=1 bp, light peak smoothing, peak threshold=200 RFU). The allelic pattern obtained automatically was checked visually and followed by minor manual corrections.

For both *L. vulgaris* and *Macropis* spp., a primer trial on a small number of samples from different origin was conducted using 12 (for plants) and six (for bees) different primer combinations to identify pairs of selective primers that were repeatable and polymorphic enough. For *L. vulgaris*, the two primer combinations *Eco*RI-ACA/*Mse*I-CTA and *Eco*RI-ATA/*Mse*I-CAC yielded suitable numbers of bands, levels of variation among loci and optimized reproducibility, and were chosen for fingerprinting application to the full sampling. For *M. europaea* and *M. fulvipes*, the final primer combinations chosen were *Eco*RI-ACA/*Mse*I-CAA and *Eco*RI-ACG/*Mse*I-CAA. Reproducibility was tested by comparing AFLP patterns between a sample and its given replicate, and the threshold for an acceptable reproducibility rate of a primer combination was set to 90% overall. We scored the presence or absence of each AFLP marker in each individual plant and bee species. The matrices of the two scored primer pairs were concatenated into one binary matrix where individuals and bands were stored as lines and columns, respectively. In order to avoid genetic structuring artefacts due to a too low or too high final number of bands, we discarded all individuals standing out of a ten percent interval from the mean number of fragments for the total sampling.

Gene pool structure analyses

We investigated the spatial genetic structure of the plant and of the two bees based on the AFLP matrices, through a Bayesian inference model-based algorithm using STRUCTURE 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2007) to assign individuals into a user-defined number of genetic groups (hereafter K). For the diploid plants, we doubled each sample's haplotype. For haplodiploid bees, following the general instructions for dominant marker inputs, after doubling each sample's haplotype, we analyzed males and females all together by entering specifically in the haploid males one haplotype as missing data (D. Falush pers. comm.). We used the "admixture" model (except for the *M. europaea* data set in which we used the "no admixture" algorithm because the "admixture" model failed in converging properly despite a large number of trials) and set the MCMC algorithm to 200,000 generations of burnin and 800,000 generations for data acquisition. Here, K values ranged from one to 15, with five replicates for each tested value. The replicate that yielded the highest likelihood value was considered as the best run (i.e. the best K value), and was chosen for further analyses. However, for the *L. vulgaris* data set, in which the likelihood values still did not reach a maximum after K=15, we considered as best K value the replicate that preceded the smallest increase between likelihoods of two consecutive K values. For the best K values considered in each species, average percentages of assignment per genetic group were calculated at the population level (deduced from the assignments at the individual level within a population). The results for each population were displayed on geographical maps, with pie-charts showing the average percentage of assignment to each group, using ARCGIS 9.1 (ESRI).

Distributions and niches modelling

We modelled the distribution of each individual insect and plant species using seven modelling techniques implemented in the BIOMOD (Thuiller *et al.* 2009) R package (see detailed protocol as Appendix 1, also containing the detailed methods for ecological niches comparisons, see below). Models were calibrated with presences corresponding to our sampling points together with occurrences from the GBIF database (www.gbif.org) and using seven climatic layers from Worldclim as in Hijmans *et al.* (2005). An Ensemble forecasting approach (Marmion *et al.* 2009) was then used to project for each species a central tendency from a combination of the models obtained by the different niche-based modelling techniques into 1) the current climate and 2) the past climate at the LGM computed by the CCSM circulation model. The resulting predicted distributions were illustrated for each species, in the present and at LGM, on maps of Europe using ARCGIS 9.1 (ESRI). Here most probable occurrences were represented by corresponding increased dark intensity.

In order to better distinguish the most probable refugia zones for each species, we selected from past predictions only the most reliable probabilities, deduced from present observed values. In order to do so, we extracted from models obtained for the present the probability values at our sampled locations only, using the Spatial Analyst Tools from the ArcToolbox (ESRI). Then we calculated the average ($\overline{X}$) probability and standard deviation (SD) of the extracted values for present. The selection on the LGM models consisted in keeping only the probability values > $\overline{X}$ - SD. Finally, as we supposed that the most probable refugia corresponded to regions where both plant and insect occurred together (either *L. vulgaris* and *M. europaea*, or *L. vulgaris* and *M. fulvipes*), we projected the shared areas for each plant-insect couple, with the “Intersect” option of the Analysis Tools of the ArcToolbox (ESRI). The resulting selected most likely refugia (for each species), as well as the areas shared between plant and insects were illustrated at the European scale using ARCGIS 9.1 (ESRI).

RESULTS

Data collection and AFLP

In total we analysed 95 populations of *L. vulgaris*, all of them represented by two to five individual plants. We could also collect *Macropis* in 53 of them (56%), containing sometimes only *M. europaea* (36 populations) or only *M. fulvipes* (nine populations), and occasionally the two bee species in sympatry (in eight *L. vulgaris* population). Nine supplementary populations of bees (three of *M. europaea* and six of *M. fulvipes*, without a corresponding *L. vulgaris* population) were also added to the sampling: additional populations of *M. fulvipes* were collected on other *Lysimachia* species, either on *L. punctata* in Turkey and Georgia, or on *L. nummularia* in one Swiss site; supplementary *M. europaea* populations were obtained from the personal collections of D. Michez and L. Bassin. Altogether we analysed 23 populations of *M. fulvipes* and 47 of *M. europaea*. All details concerning the number of individuals analysed per population of *L. vulgaris*, *M. europaea* and *M. fulvipes* as well as geographical information are reported in Table 1.

The two primer couples for AFLP genotyping resulted in the following outcomes. Among 339 samples analyzed for *L. vulgaris*, we obtained 282 different alleles, each individual yielding a minimum of 153 and a maximum of 188 fragments. In *M. europaea*, we genotyped a total of 106 fragments among 179 samples, with each individual presenting at least 32 bands, up to a maximum of 61. Finally, 79 individuals of *M. fulvipes* were genotyped, with a total of 128 alleles and each sample encompassing between 42 and 70 fragments.

Spatial genetic structure

In *L. vulgaris*, the optimal number of genetic clusters was six (i.e., the smallest increase in the likelihoods of two consecutive runs occurred between six and seven; Δ log likelihood [K=6; K=7] = 14.5; best log likelihood [K=6] among five replicates = -56334.8). A clear longitudinal structure (from east to west) was observed for the different genetic lineages (Fig. 1). Spatially-restricted lineages were detected for: eastern Anatolia and Georgia (in yellow on Fig. 1), western Anatolia, southern Balkans and Carpathians (in red), Corsica (in white) and the Atlantic zone (in blue). Two lineages (in orange and green) were less restricted to a particular region, showing large distributions spreading all across Europe. However one among the two (in green) showed larger proportions in the southern half of Europe (France, Italy, Spain and Switzerland).

In *M. europaea*, the best K value was obtained when four genetic groups were considered (best log likelihood [K=4] among five replicates = -5203). When displayed on the map of Europe (Fig. 2), the geographic structure of these four genetic lineages showed an overall north-south pattern. This trend was especially obvious considering a group shared by Spanish and Italian populations and rarely extending to the North (in blue). The genetic structure of northern Europe was more scrambled with three lineages widely dispersed from France to southern Balkans. One (in orange) was largely dominant in most populations. Another (in green) was mainly present in populations from the north-eastern quarter of Europe, and was even dominant in a few ones (e.g. in Germany or Hungary). Finally the last (in red) was irregularly widespread (e.g. absent from France, Germany and Italy) and always represented in minority.

In *M. fulvipes,* the optimal number of genetic clusters was three (best log likelihood [K=3] among five replicates = -4685.2). These were geographically organized from east to west as follows (see Fig. 3): the first group (in red) was strongly prevailing in Anatolia and southern Caucasus, although detected in small proportions in Carpathians; the second genetic lineage (displayed in orange) dominated all populations from central and Eastern Europe as far as the Baltic area although it was also detected in minor proportions in central Anatolia; finally, the third lineage (in blue) was dominant for all Spanish populations.

Potential distribution modelling

The models of potential distributions obtained for each of the three species are represented both for current and LGM ecological conditions. The potential distributions for the present are shown as Supporting Information. These mostly matched with the ranges of distribution reported for each species (i.e., for bees, see Michez and Patiny 2005, and for *L. vulgaris*, see Hulten and Fries 1986). Some inaccuracies were however recorded, for instance the predicted presence of *M. fulvipes* in the British Isles or Scandinavian Peninsula, or the potential occurrence of *M. europaea* in Turkey or Georgia. Main flaws of these models resided in the fact that they represent the appropriate ecological conditions for a given organism, but do not take into consideration its biogeographical history. More details regarding the current distributions and evolutionary histories of *M. europaea* and *M. fulvipes* are available in Bassin *et al.* (in prep).

The raw past models for LGM are shown, for each species, as a background layer in Figs. 1-3 (together with the corresponding spatial genetic structure results). The selected past distributions showing only the most probable refugia are presented in Figs. 4-6, comparing patterns between *L. vulgaris* and *M. europaea* (Fig. 4), between *L. vulgaris* and *M. fulvipes* (Fig. 5) and finally between the two bee species (Fig. 6).

For *Lysimachia vulgaris,* several southern regions in the western Palearctic were likely to offer suitable conditions for the yellow loosestrife during LGM (see background of Fig. 1 for raw model, or Figs. 4A and 5A for the selected most probable refugia showed in grey and black). Many regions around the Black Sea were favourable, particularly (i) the area comprised between the Azov Sea and western Caucasus, (ii) a zone comprising northern Syria and Irak as well as most of Anatolia (although with a clear central disjunction), and (iii) a large part of south-eastern Balkans especially southern Bulgaria, European Turkey, Macedonia and central Greece. A wide portion of the Italian Peninsula (plus Corsica) was also suitable except for the extreme south in Puglia and Sicily. In France, the area comprised between the Alps and the Pyrenees as well as the French Atlantic coast, extending along northern Spanish coast, were potentially suitable for *L. vulgaris*. Finally, the central plateaus of Spain could also have hosted the plant during LGM.

For *M. europaea*, the raw LGM potential distribution (in background of Fig. 2) showed two large areas with suitable ecological conditions (in north-eastern Spain and south-western France, as well as in Italy). Southern Balkans presented a third, although more limited, potential survival area for the bee during LGM. Even if many areas fitted both for *M. europaea* and *L. vulgaris* (in black in Fig. 4A and B), the range of the bee was smaller than that of the plant. This trend is especially true in the eastern half of Europe, where the plant should have found optimal conditions over a much larger area (in grey in Fig. 4A). There, only a thin belt along the eastern and southern coasts of Black sea, as well as few areas around the current Albanian-Macedonian-Greek borders could have hosted *M. europaea* as pollinator of *L. vulgaris* during the last ice age. The climatic conditions of Western Europe might have been more convenient for this oil-collecting bee, including a good match with the plant, and even showing areas appropriate for *M. europaea* only (visible in grey on Fig. 4B).

For *M. fulvipes*, the LGM raw prediction for *M. fulvipes* is once more presented as background of Fig. 3. Optimal zones during LGM in the west of Europe spanned almost uninterruptedly northern and eastern Spain, France (except the northern third) and Italy. The Balkans region (from current central Croatia to eastern Bulgaria) and surrounding areas of the Black Sea were also largely suitable. The congruence between potential refugia for *M. fulvipes* and *L. vulgaris* (Fig. 5A and B) was not as good as the one observed previously for *M. europaea* and the plant. Interestingly, whereas large areas of Western Europe and the Balkans were optimal for *M. fulvipes* but not for the plant (in grey in Fig. 5B), a contrasting trend (i.e. suitable for *L. vulgaris* but not for the bee) was observed in Anatolia (in grey in Fig. 5A). Comparing most probable refugia for the two bee species (in black and grey of Figs. 4B and 5B), showed that *M. fulvipes* could potentially be much more widespread than *M. europaea* during glaciations. Finally we observed that co-occurrences between *M. europaea* and *L. vulgaris* were a subset of those between *M. fulvipes* and the plant (Fig. 6), i.e. all potential areas with both *M. europaea* and *L. vulgaris* at LGM were also potential areas for *M. fulvipes*. Regions where *M. fulvipes* should have occurred alone with *L. vulgaris* (showed in grey in Fig. 6) were mostly located in Eastern Europe and Anatolia.

DISCUSSION

Inferring SGS and modelling past distribution are complementary powerful analytical tools to reveal patterns in comparative phylogeography

As presented in the backgrounds of Figs. 1-3, the correspondence between LGM distributions of *L. vulgaris* and of the two *Macropis* species was rather good, although showing regional differences. Because the respective presence/absence source datasets for modelling distributions of the plant and the bees are mostly independent (i.e., geographical presences and absences for each studied species were obtained from several unrelated sources), this overlap is not a bias of the method: it reflects shared ecological optimums among species rather than a tautological correlation caused by putatively co-sampled points.

Past distribution models suggest that both locally and regionally, each species might be able to live in areas where the corresponding partner is absent. This pattern is likely to be explained by the ecology of this pollination system, in which species can also rely on other additional partners (see below). Moreover, *L. vulgaris* might even survive in the complete absence of any insect pollinator (via autogamy or vegetative reproduction; Vogel 1976, 1986) as it might be the case today in far-northern Europe. This trend becomes obvious when displaying the current potential distribution of the yellow loosestrife in comparison to the oil-collecting bees (see Supporting Information), for instance in Scotland, Finland or Norway. As shown by the ecological niches overlays (presented as Supporting Information), the plant demonstrates a larger ecological niche, supporting more humid ecological conditions than *Macropis* spp, as well as a wider range of suitable temperatures (either colder or warmer).

SGS patterns brought complementary information and largely confirmed the nature of areas showing high presence probabilities during LGM. For example in *L. vulgaris*, the two distinct genetic lineages in Anatolia (see Fig. 1) were compatible with the two corresponding optimal areas predicted for the plant in the region. The lineage endemic from Corsica proved that not only did the island provide optimal conditions for the plant, but also that *L. vulgaris* should have been strongly isolated there once. In *M. europaea*, additional information was supplied by the geographic distribution of the genetic lineage limited to Iberian and Italian Peninsulas (in blue on Fig. 2). Circumscription of this lineage to two regions showing optimal conditions during LGM suggested substantial genetic connexions on the eastern-western axis between populations from Spain and Italy. In *M. fulvipes*, the continuous distribution of an Eastern genetic lineage at the southern shore of Black Sea (Fig. 3) revealed that the potential disjunctive distribution of *L. vulgaris* (discussed above) did not influence the dispersal routes of this bee: here again strong eastern-western connexions between populations should have existed during LGM. The Iberian-Italian and Black Sea lineages found in *M. europaea* and *M. fulvipes*, respectively, attest for the capacity of the bees to disperse longitudinally. In contrast, both species seem much more constrained while dispersing towards the North: natural barriers tend to impede the postglacial northern movements of bees (see below).

Globally, a simple visual comparison between the SGS of the plant and each of the two bees shows an overall low level of congruence. The only parallel that can be drawn lies in the overall genetic homogeneity in a wide zone enclosing middle-range to high latitudes, from the Atlantic coast to the Carpathians (in orange in Fig. 3).

Overlapping areas of high-probability presence in the modelled LGM distributions of *L. vulgaris* and of the two *Macopis* bees suggest most probable refugia and the role played by additional partners

Refugia zones predicted for *M. europaea* were largely included in those expected for *L. vulgaris* (black areas in Fig. 4A and B). In contrast, large areas with suitable ecological conditions for *M. fulvipes* but not for the plant might have existed during LGM in France and Balkans (Fig. 5B; see below). This result is in agreement with ecological data provided by Bassin *et al.* (in prep), in which the authors provided evidence for a stronger trend to monolectism and a higher level of ecological stenoecy in *M. europaea* than in *M. fulvipes*. As a consequence, it can be predicted that *M. europaea* should be the most frequent pollinator of *L. vulgaris* in Western Europe. Even if ecological conditions proved to be suitable for both bees in the western half of the continent, *M. europaea* might be more competitive on *L. vulgaris* (considering its stenoecic status and its tight association with the plant.). The reported absence of *M. fulvipes* in Great Britain and western Scandinavia nowadays might also confirm that this species was less present in the western half of the continent. In contrast, most regions of south-eastern Balkans, Anatolia and the surroundings of Black Sea were likely to host *M. fulvipes* as the main pollinator of *L. vulgaris* during LGM, since *M. europaea* is more scarcely distributed in this area according to the past modelled distribution (and to its current absence from Anatolia and Caucasus)

“Lonely” refugia (i.e., a term that we define as an area with high probability presence during LGM in one but not the other species of the mutualism) are mainly observed in *M. fulvipes* (grey areas in Fig. 5B) and *L. vulgaris* (grey areas in Figs. 4A and 5A). If such predicted areas corresponded to realized refugia during LGM, the two species should have locally relied on other partners: (i) *M. fulvipes* could have found another oil-providing plant, such as *L. nummularia* that currently shows a distribution similar to *L. vulgaris* in the western Palearctic but that is not dependant on *Macropis* species for pollination (Bittrich & Kadereit 1988; Bassin *et al.* in prep), or *L. punctata* with which the bee is today strongly associated around the Black Sea ; (ii) *Lysimachia vulgaris* is likely to have survived during LGM in several parts of Anatolia that were not suitable for *M. europaea* and *M. fulvipes*, by relying on a third *Macropis* species, *M. frivaldszkyi*, whose distribution nowadays covers most parts of South-Eastern Europe and the Middle-East. Punctual or long-lasting associations with additional partners not only support the presence of such “lonely” refugia during LGM, but also might have been significant for the survival of the mutualistic system in providing supplementary suitable areas for the three main species to overcome harsh climatic periods.

Natural barriers impede the northern re-colonization of bees, not of plants

Despite *L. vulgaris* harbours two genetic lineages with a relatively narrow geographical distribution in Corsica (as a consequence of a long-lasting isolation on the island) and in eastern Anatolia and Georgia (as the Caucasus Mountains might represent a strong barrier to northern colonization; see Seddon *et al.* 2002), the largest part of the European distribution of the plant is covered by genetic lineages with a wide latitudinal range: e.g., from southern Spain to Northern Ireland (in blue in Fig. 1) or from southern Turkey to Latvia (in red in Fig. 1). Therefore, the plant does not seem particularly limited in its northern/southern dispersal and the northern re-colonization of Europe from southern refugia might have progressively followed the retreat of inhospitable habitats. Consequently the actual genetic variation of *L. vulgaris* is mostly structured longitudinally and shows interesting suture zones, for instance in the Baltic region, where different lineages meet far from their predicted origin. Most populations in central and Western Europe show, however, a strong level of admixture resulting from two lineages with a weak spatial structure. The wide distribution of these lineages could be caused by the strong ecological dependence on water of the yellow loosestrife and subsequent frequent water-mediated longitudinal dispersal of seeds via main European rivers.

In contrast to the pattern illustrated in the plant, the SGS of oil-collecting bees revealed a stronger importance of longitudinal natural barriers during northern postglacial extension: the lineages characteristic of the Iberian Peninsula (for both species), Italy (for *M. europaea* only) and Anatolia (for *M. fulvipes* only) do not expand further North. Barriers represented by the Alps and the Pyrenees might therefore have acted as an important obstacle to the northern re-colonization of genetic lineages from southern refugia from Iberia and Italy (see Taberlet *et al.* 1998 for other illustrations of this paradigm). In Anatolia, dispersal to northern areas was impeded by a water natural barrier: during LGM in this region stood a large freshwater lake with a location and surface equivalent to the current Black Sea (although with a lower water level; Ryan *et al.* 1997).

Evaluating the biogeographical history of other lineages in bees (i.e., those more largely spread in central Europe) is difficult with our current results. South-western France or southern Balkans refugia both represent putative refugia areas from which recolonization might have taken place for *M. europaea* (currently absent from Anatolia and Caucasus) and *M. fulvipes* (today absent from most of north-western Europe), respectively. Obstacles such as mountain ranges (or large water basins) might therefore have impacted more intensively on insects' than on plants' latitudinal dispersal, despite presumable faster dispersal capacity of bees by flight. In contrast, bee lineages restricted to southern areas that were not delimited by barriers such as mountains ranges or large water basins during LGM, might have rapidly spread towards northern areas.

Conclusion

Despite the strong reciprocal dependence of the three species studied in this specialised pollination system, our results show that the phylogeographic patterns found in pollinators are not congruent with those revealed in the host plant. Re-colonization pathways tend to be different, notably because natural barriers impede the displacement into northern territories of some species (the bees), but not of others (the plant). In contrast, putative refugia in the western Palearctic mainly overlap. Our working hypothesis forecasting comparable SGS among mutualistic partners should therefore be balanced by the respective biological and ecological characteristics of each species involved (e.g., dispersal abilities, population sizes; for a discussion, see Alvarez *et al.* 2010). Another factor driving incongruence in the respective biogeographic histories of these species is the non-exclusive nature of obligate reciprocity in this specialised pollination system (which punctually incorporates other potential partners). The phylogeographic framework addressed in the current study shows that processes driving historical changes in co-evolutionary processes are subtle. Considering a whole network of interacting species in non-exclusive systems therefore represents the next step for an exhaustive biological understanding of mutualistic interactions.

ACKNOWLEDGEMENTS

The authors are sincerely grateful to all people that brought their precious help during fieldwork sampling: R. Arnoux, A. Favre, I. Hiltpold, D. Kharazishvili and B. Prudat. A special thank to A. Sarr for support with AFLP analysis. We wanted to more particularly thank D. Michez for providing supplementary samples for DNA analysis, geographical databases for niches modeling and for providing helpful remarks. This research was supported by the Swiss National Science Foundation (grant n° 3100A0-116778/1). The fieldwork was additionally granted by the travel grants commission of the Swiss Academy of Sciences

(SCNAT), and by the University of Neuchâtel (grant “Marguerite Wütrich and A. Mathey-Dupraz”).

REFERENCES

Althoff, D.M. & Thompson, J.N. (1999). Comparative geographic structures of two parasitoid-host interactions. *Evolution*, 53, 818-825.

Alvarez, N., Kjellberg, F., McKey, D. & Hossaert-McKey, M. (2010). Phylogeography and historical biogeography of obligate specific mutualisms. In: *The Geography of Host-Parasite Interactions* (eds. Morand, S. & Krasnov, B.). Oxford University Press, UK, pp. 31-39.

Anderson, B. & Midgley, J.J. (2003). Digestive mutualism, an alternate pathway in plant carnivory. *Oikos*, 102, 221-224.

Anderson, B., Olivieri, I., Lourmas, M. & Stewart, B.A. (2004). Comparative population genetic structures and local adaptation of two mutualists. *Evolution*, 58, 1730-1747.

Bittrich, V. & Kadereit, J. (1988). Cytogenetical and geographical aspects of sterility in *Lysimachia nummularia*. *Nord. J. Bot.*, 8, 325-328.

Britton, N.L. & Brown, A. (1913). *An illustrated flora of the northern United States, Canada and the British Possessions*. [WWW document]. URL http://www.plants.usda.gov/java/profile?symbol=LYVU

Bronstein, J.L. (1994). Our current understanding of mutualism. *Q. Rev. Biol.*, 69, 31-51.

Bronstein, J.L., Alarcon, R. & Geber, M. (2006). The evolution of plant-insect mutualisms. *New Phytol.*, 172, 412-428.

Cane, J.H., Eickwort, G.C., Wesley, F.R. & Spielholz, J. (1983). Foraging, grooming and mate-seeking behaviors of *Macropis nuda* (Hymenoptera, Melittidae) and use of *Lysimachia ciliata* (Primulaceae) oils in larval provisions and cell linings. *Am. Midl. Nat.*, 110, 257-264.

Celary, W. (2004). A comparative study on the biology of *Macropis fulvipes* (Fabricius, 1804) and *Macropis europaea* Warncke, 1973 (Hymenoptera : Apoidea : Melittidae). *Folia Biol.-Krakow*, 52, 81-85.

Chase, M.W. & Hills, H.H. (1991). Silica-gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon*, 40, 215-220.

Cook, J.M. & Rasplus, J.Y. (2003). Mutualists with attitude: coevolving fig wasps and figs. *Trends Ecol. Evol.*, 18, 241-248.

Despres, L. & Jaeger, N. (1999). Evolution of oviposition strategies and speciation in the globeflower flies *Chiastocheta* spp. (Anthomyiidae). *J. Evol. Biol.*, 12, 822-831.

Dötterl, S. & Schäffler, I. (2007). Flower scent of floral oil-producing *Lysimachia punctata* as attractant for the oil-bee *Macropis fulvipes*. *J. Chem. Ecol.*, 33, 441-445.

Falush, D., Stephens, M. & Pritchard, J.K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes*, 7, 574-578.

Hao, G., Yuan, Y.M., Hu, C.M., Ge, X.J. & Zhao, N.X. (2004). Molecular phylogeny of *Lysimachia* (Myrsinaceae) based on chloroplast trnL-F and nuclear ribosomal ITS sequences. *Mol. Phylogenet. Evol.*, 31, 323-339.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, 25, 1965-1978.

Hulten, E. & Fries, M. (1986). *Atlas of North European vascular plants: north of the Tropic of Cancer I-III*. Koeltz Scientific Books, Königstein.

Machado, C.A., Jousselin, E., Kjellberg, F., Compton, S.G. & Herre, E.A. (2001). Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. *Proc. R. Soc. Lond. Ser. B-Biol. Sci.*, 268, 685-694.

Machado, C.A., Robbins, N., Gilbert, M.T.P. & Herre, E.A. (2005). Critical review of host specificity and its coevolutionary implications in the fig/fig-wasp mutualism. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 6558-6565.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.*, 15, 59-69.

Michez, D., Nel, A., Menier, J.J. & Rasmont, P. (2007). The oldest fossil of a melittid bee (Hymenoptera : Apiformes) from the early Eocene of Oise (France). *Zool. J. Linn. Soc.*, 150, 701-709.

Michez, D. & Patiny, S. (2005). World revision of the oil-collecting bee genus *Macropis* Panzer 1809 (Hymenoptera : Apoidea : Melittidae) with a description of a new species from Laos. *Ann. Soc. Entomol. Fr.*, 41, 15-28.

Pekkarinen, A., Berg, O., Calabuig, I., Janzon, L.A. & Luig, J. (2003). Distribution and co-existence of the *Macropis* species and their cleptoparasite *Epeoloides coecutiens* (Fabr.) in NW Europe (Hymenoptera : Apoidea, Melittidae and Apidae). *Entomol. Fenn.*, 14, 53-59.

Pellmyr, O. (2003). Yuccas, yucca moths, and coevolution: A review. *Ann. Mo. Bot. Gard.*, 90, 35-55.

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

Rasmussen, C. & Olesen, J.M. (2000). Oil flowers and oil-collecting bees. *Scandinavian Association for Pollination Ecology honours Knut Fegri*, 39, 23-31.

Renner, S.S. & Schaefer, H. (2010). The evolution and loss of oil-offering flowers: new insights from dated phylogenies for angiosperms and bees. *Philos. Trans. R. Soc. B-Biol. Sci.*, 365, 423-435.

Ryan, W.B.F., Pitman, W.C., Major, C.O., Shimkus, K., Moskalenko, V., Jones, G.A.*, et al.* (1997). An abrupt drowning of the Black Sea shelf. *Mar. Geol.*, 138, 119-126.

Seddon, J.M., Santucci, F., Reeve, N. & Hewitt, G.M. (2002). Caucasus Mountains divide postulated postglacial colonization routes in the white-breasted hedgehog, *Erinaceus concolor*. *J. Evol. Biol.*, 15, 463-467.

Simpson, B.B., Neff, J.L. & Seigler, D.S. (1983). Floral biology and floral rewards of *Lysimachia* (Primulaceae). *Am. Midl. Nat.*, 110, 249-256.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.*, 7, 453-464.

Thompson, J.N. & Cunningham, B.M. (2002). Geographic structure and dynamics of coevolutionary selection. *Nature*, 417, 735-738.

Thuiller, W., Lafourcade, B., Engler, R. & Araujo, M.B. (2009). BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373.

Vogel, S. (1976). *Lysimachia* : Ölblumen des Holartiks. *Naturwissenschaften*, 63, 44-45.

Vogel, S. (1986). Ölblumen and Ölsammelnde Bienen: Zweite Folge *Lysimachia* und *Macropis*. In: *Akademie der Wissenschaftenund der Literatur Mainz* (ed. Verl. FS) Stuttgart, Germany.

Vogel, S. (1988). The oil-bee oil-flower relationships: parallelism and other aspects of their evolution in space and time. *Z. Zool. Syst. Evol.*, 26, 341-362.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M.*, et al.* (1995). AFLP- a new technique for DNA-fingerprinting. *Nucleic Acids Res.*, 23, 4407-441.

Table 1: List of sampled populations for *L. vulgaris*, *M. europaea* and *M. fulvipes*, also presenting the geographic coordinates and the number of individuals genetically analyzed. The first column corresponds to numbers displayed on the populations from Figs 1-3.

N° on map	Population	Latitude (N)	Longitude	Country	Place	*L. vulgaris*	*M. europaea*	*M. fulvipes*
1	BAK	41.765	43.484	Georgia	Bakuriani	-	-	5
2	KHA	41.999	43.656	Georgia	Khashuri	5	-	3
3	KOB	41.858	41.786	Georgia	Kobuleti	3	-	-
4	TKI	42.377	43.037	Georgia	Tkibuli	-	-	1
5	ABA	40.601	31.282	Turkey	Abant Gölü	3	-	-
6	BEL	38.264	34.291	Turkey	Belısarma	5	-	-
7	DAD	41.481	33.603	Turkey	Daday	2	-	4
8	DAG	39.574	39.864	Turkey	Dağyolu	2	-	-
9	ERF	41.859	34.744	Turkey	Erfelek	-	-	4
10	ILG	41.140	34.064	Turkey	Ilgaz Geçidi	-	-	5
11	YAY	38.061	28.773	Turkey	Yayla Gölü	4	-	-
12	RIL	42.102	23.091	Bulgaria	Rila	2	4	-
13	TUL	42.574	25.576	Bulgaria	Tulovo	3	3	-
14	KOZ	41.056	21.036	Macedonia	Kozjak	4	5	-
15	DES	44.042	21.537	Serbia	Despotovac	3	5	2
16	KRU	43.105	22.688	Serbia	Krupačko jezero	5	3	-
17	FEL	46.701	23.590	Romania	Feleacu	3	-	3
18	FRA	47.548	25.765	Romania	Frasin	3	-	-
19	OIT	46.067	26.372	Romania	Col de Oituz	2	-	-
20	PAS	47.712	23.777	Romania	Pasul Gutai	2	-	1
21	DIO	47.932	19.062	Hungary	Diosjeno	4	3	-
22	HAJ	47.695	21.655	Hungary	Hajduhadhaz	5	3	-
23	HOD	46.356	20.209	Hungary	Hodmezovasarhely	2	4	-
24	KBA	46.660	17.126	Hungary	Kis-Balaton	4	5	-
25	DOM	49.016	21.673	Slovakia	Domasa Dobra	5	1	5
26	SLO	48.746	19.241	Slovakia	Slovenska Lupča	4	-	-
27	HOS	49.687	12.582	Czech Republic	Hostka	5	5	-
28	SUC	49.486	16.762	Czech Republic	Suchy	4	-	5
29	JOH	47.500	14.549	Austria	Johnsbach	3	-	-
30	RUC	46.591	14.567	Austria	Ruckersdorf	2	5	-
31	GJE	45.727	14.408	Slovenia	Gorni Jezero	3	4	-
32	LOK	45.395	14.691	Croatia	Lokve	3	-	-
33	VRH	45.334	13.920	Croatia	Vrh	3	4	-
34	AND	46.504	11.238	Italy	Andriano	4	-	-
35	BAG	44.312	8.048	Italy	Bagnasco	4	-	1
36	CAS	43.777	10.628	Italy	Castelvecchio di Compito	4	5	-
37	MAT	41.406	14.406	Italy	Lago Matese	4	-	-
38	RIE	42.510	12.753	Italy	Rieti	4	4	-
39	SOA	45.209	10.734	Italy	Soave	2	-	-
40	VIV	45.441	8.006	Italy	Viverone	4	5	-
41	BAL	47.402	9.619	Switzerland	Balgach	4	4	-
42	BEG	46.427	6.242	Switzerland	Begnins	5	5	-
43	BOL	46.161	8.863	Switzerland	Bolle di Magadino	4	3	-
44	MON	46.234	7.338	Switzerland	Lac de Montorge	3	5	-
45	PRE	47.285	7.429	Switzerland	Les Préhaies	4	-	-
46	VAU	47.190	7.318	Switzerland	Vauffelin	-	-	5
47	BAZ	47.680	-0.206	France	Bazouges	4	5	-
48	BDL	45.771	6.242	France	Bout du Lac d'Annecy	4	4	-
49	BIG	42.640	9.459	France	Etang de Biguglia	5	-	-
50	BOR	45.440	2.439	France	Bort-les-Orgues	4	-	-
51	BOV	49.855	2.378	France	Boves	4	4	-
52	CAP	44.301	-0.255	France	Captieux	5	4	-
53	COX	47.973	2.316	France	Combreux	4	-	-
54	JUG	48.405	-2.325	France	Jugon-les-Lacs	4	1	-
55	MAV	49.421	0.533	France	Marais Vernier	3	1	-
56	ROL	47.951	5.257	France	Rolampont	4	-	-
57	SCA	43.606	4.336	France	Etang de Scamandre	3	4	-

Table 1 continued.

N° on map	Population	Latitude (N)	Longitude	Country	Place	*L. vulgaris*	*M. europaea*	*M. fulvipes*
58	ARI	42.991	-3.975	Spain	Arija	4	-	-
59	ASP	42.715	-1.158	Spain	Aspurz	3	4	-
60	BEC	40.403	-5.627	Spain	Becedas	3	-	4
61	CAE	42.221	3.105	Spain	Castello d'Empuries	2	-	-
62	CAM	42.226	-2.627	Spain	Sierra de Camero Nuevo	3	-	-
63	CER	42.947	-4.492	Spain	Cervera de Pisuerga	1	-	-
64	CUE	40.192	-2.113	Spain	Cuenca	4	4	-
65	HOR	39.361	-4.615	Spain	Horcajo de los Montes	4	-	-
66	LOZ	40.902	-3.863	Spain	Rio Lozoya	2	3	2
67	MES	38.487	-2.358	Spain	Mesones	4	-	-
68	MIN	41.974	-8.737	Spain	Rio Minho	2	2	-
69	PIE	39.048	-4.239	Spain	Piedrabuena	2	-	5
70	REN	42.504	-4.729	Spain	Renodo	-	4	-
71	RIO	40.324	-6.646	Spain	Rio Riofrio	4	-	-
72	SAL	40.957	-5.671	Spain	Salamanca	-	4	-
73	SAN	42.115	-6.734	Spain	Lago di Sanabria	3	-	5
74	SON	42.974	-3.805	Spain	Soncillo	-	-	5
75	ZAO	40.798	-2.155	Spain	Zaorejas	5	-	-
76	AGH	54.521	-6.314	Ireland	Aghalee	5	-	-
77	CLO	53.332	-7.980	Ireland	Clonmacnoise	5	-	-
78	BEE	55.001	-3.721	Great Britain	Beeswing	5	-	-
79	BIL	51.022	-0.478	Great Britain	Billingshurst	4	-	-
80	CLE	51.429	-2.832	Great Britain	Clevedon	4	-	-
81	HEL	53.457	-2.693	Great Britain	St. Helens	2	-	-
82	BAU	50.489	3.822	Belgium	Baudour	-	3	-
83	WAM	50.110	5.435	Belgium	La Wamme	3	-	-
84	LOO	51.611	5.076	Netherlands	Loon op Zand	5	-	-
85	TER	53.394	5.313	Netherlands	Terschelling	3	-	-
86	DIP	50.919	13.681	Germany	Dippoldiswalde	4	5	1
87	HOL	53.368	9.642	Germany	Holvede	5	3	-
88	LEM	51.723	6.995	Germany	Lembeck	4	3	-
89	MUM	48.601	8.192	Germany	Mummelsee	3	-	-
90	SCH	53.711	10.838	Germany	Schaalsee	4	-	-
91	TRO	51.683	11.041	Germany	Trollblumenwiese	4	2	4
92	WOL	48.421	10.599	Germany	Wolbach	4	5	-
93	LIS	56.234	10.167	Denmark	Lisbjerg Skov	4	-	-
94	MAR	54.769	11.506	Denmark	Maribo	4	5	-
95	UGG	57.587	10.142	Denmark	Uggerby A	3	1	-
96	HAL	59.130	11.489	Norway	Halden	3	-	-
97	LAS	59.170	10.178	Norway	Lastelanded	3	-	-
98	NOR	61.395	7.284	Norway	Gaupne	3	-	-
99	BOK	56.403	13.600	Sweden	Boekholm	5	2	-
100	LJU	58.802	15.452	Sweden	Ljusfallshammar	2	5	-
101	BOB	53.947	16.599	Poland	Bobolice	4	-	-
102	KET	49.844	19.214	Poland	Kety	4	5	1
103	DZE	56.655	24.933	Latvia	Dzelmes	3	5	4
104	LAU	57.178	22.692	Latvia	Lauciene	3	3	4

FIGURES

Figure 1: Spatial genetic structure of *L. vulgaris* for the the best STRUCTURE run (K=6). The six genetic groups are represented by different colours. Each population, represented as a pie-chart, shows the average percentages of assignment per genetic lineage, calculated at the population level. The background represents the modelling of the potential distribution of the plant at last glacial maximum (LGM), also showing main icecaps. Different levels of grey shading represent different probabilities of past presence (i.e., the darker is the area, the higher was its probability to provide optimal LGM ecological conditions for a given species).

Figure 2: Similar to Fig. 1 but for the bee *M. europaea*, with the spatial genetic structure corresponding to the best STRUCTURE run at K=4.

Figure 3: Similar to Figs. 1 and 2 but for the bee *M. fulvipes* (with the best STRUCTURE run at K=3).

Figure 4: Comparison of most probable refugia zones between *L. vulgaris* and *M. europaea.* The area in black represents refugia shared by both species. The “lonely” refugia (areas enclosing one single member of the mutualism) are shown in grey on both maps respectively for the plant (A) or the bee (B).

Figure 5: Similar to Fig. 4 but comparing most probable refugia between *L. vulgaris* and *M. fulvipes*.

Figure 6: Map showing the refugia zones shared by *L. vulgaris* and *M. europaea* (in black) and by *L. vulgaris* and *M. fulvipes* (in grey). The black area is a subset of the grey one.

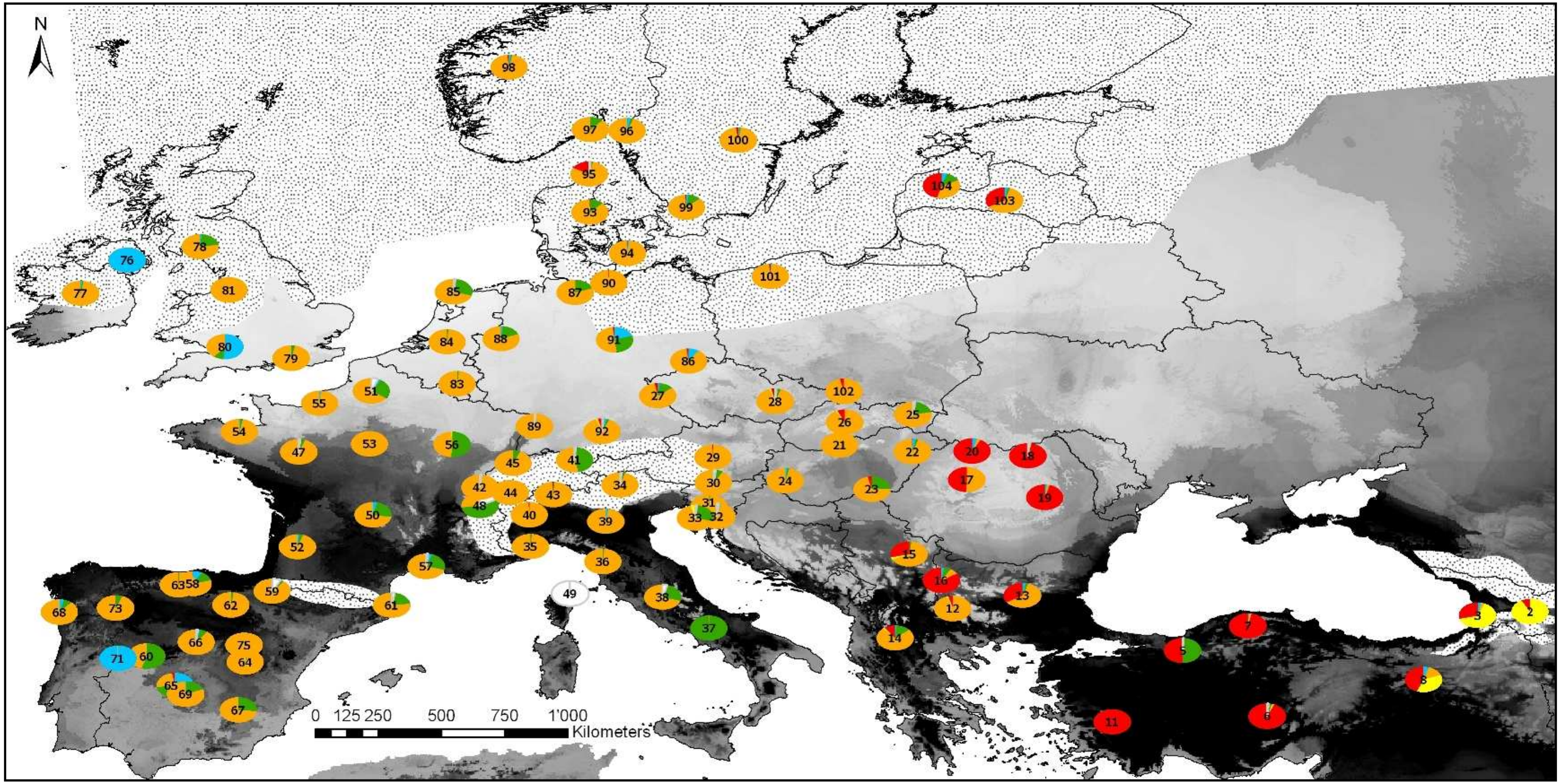


Figure 1: Spatial genetic structure of *L. vulgaris* for the the best STRUCTURE run (K=6). The six genetic groups are represented by different colours. Each population, represented as a pie-chart, shows the average percentages of assignment per genetic lineage, calculated at the population level. The background represents the modelling of the potential distribution of the plant at last glacial maximum (LGM), also showing main icecaps. Different levels of grey shading represent different probabilities of past presence (i.e., the darker is the area, the higher was its probability to provide optimal LGM ecological conditions for a given species).

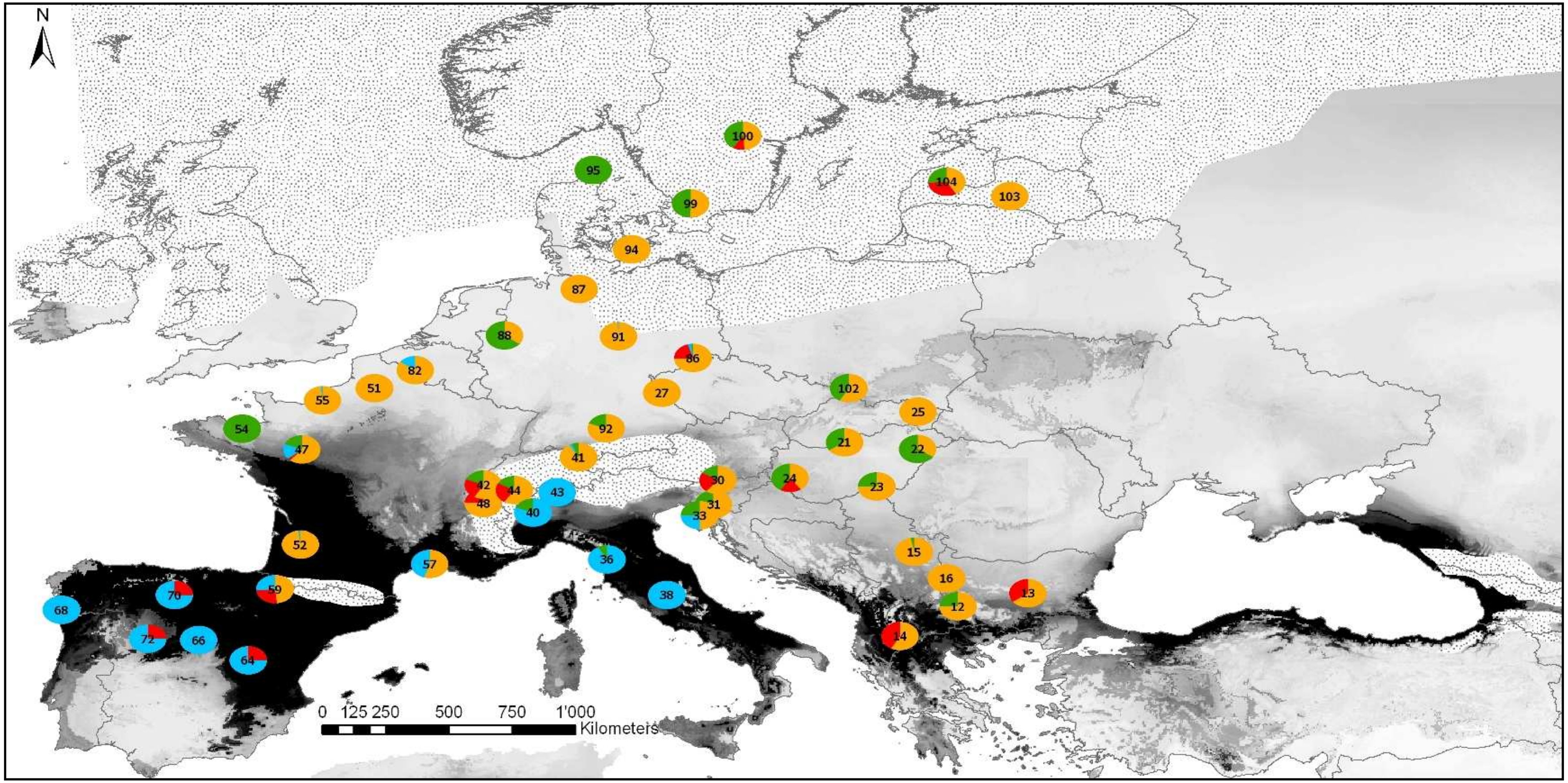


Figure 2: Similar to Fig. 1 but for the bee *M. europaea*, with the spatial genetic structure corresponding to the best STRUCTURE run at K=4.

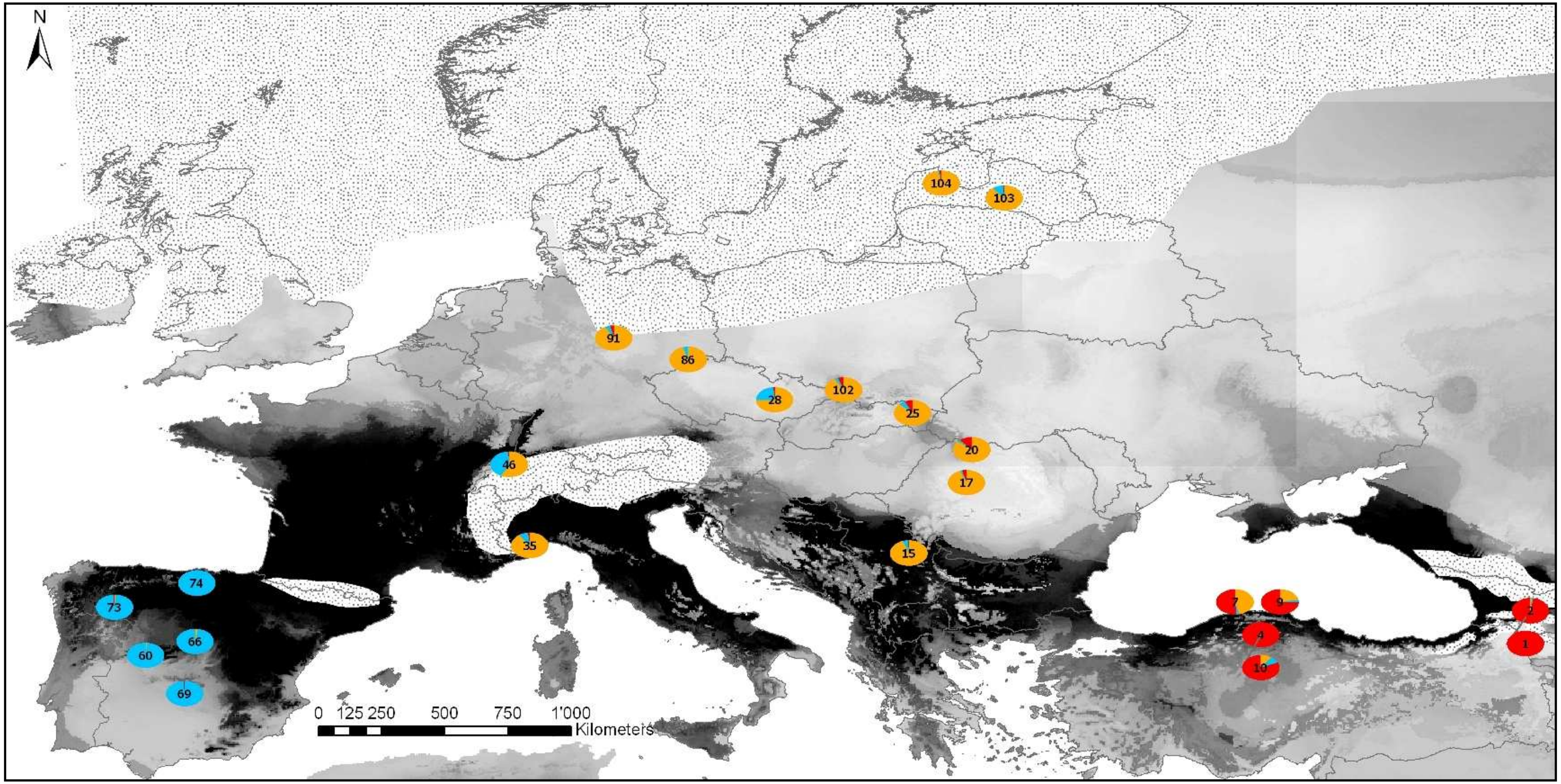


Figure 3: Similar to Figs. 1 and 2 but for the bee *M. fulvipes* (with the best STRUCTURE run at K=3).

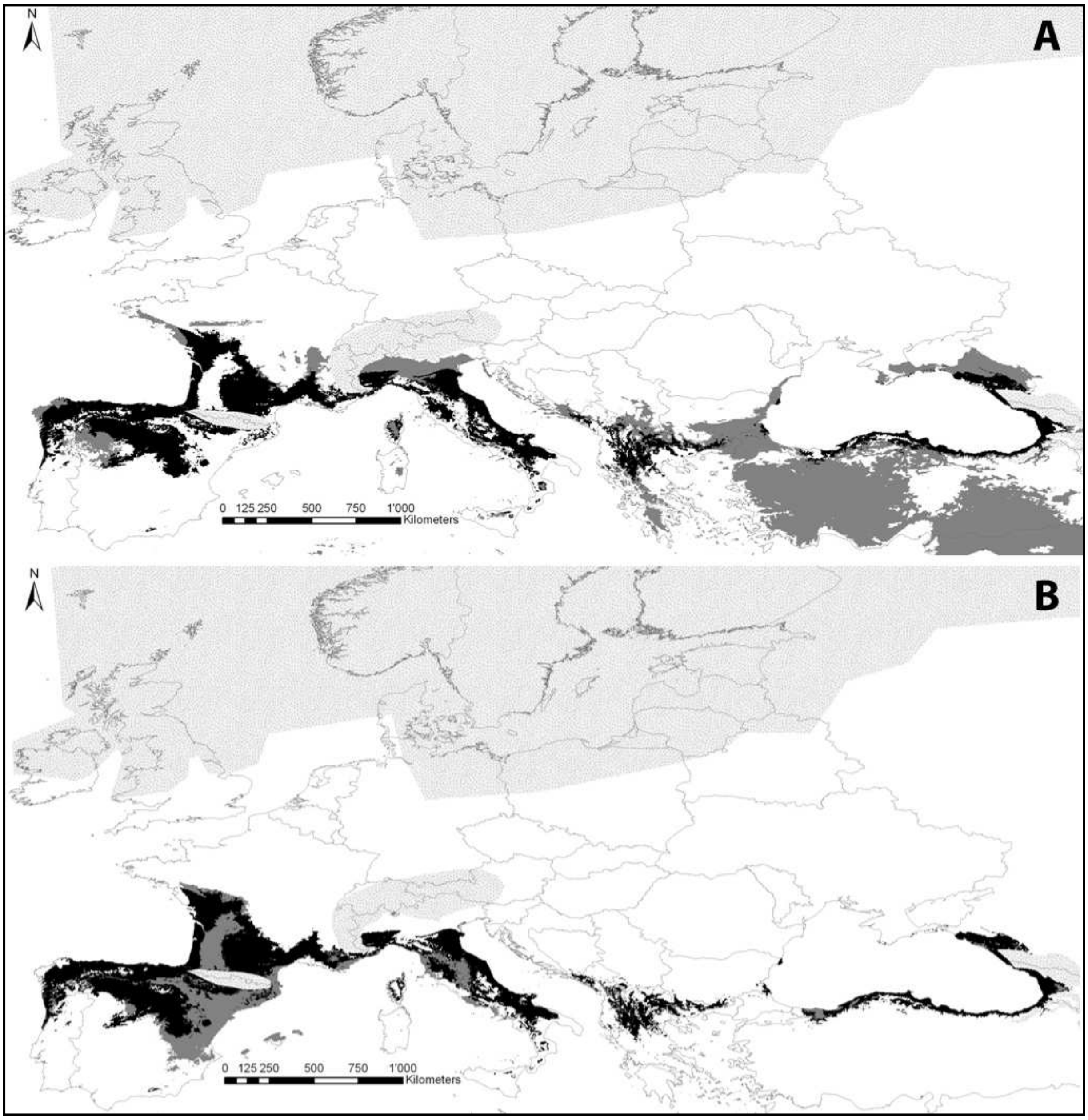


Figure 4: Comparison of most probable refugia zones between *L. vulgaris* and *M. europaea.* The area in black represents refugia shared by both species. The "lonely" refugia (areas enclosing one single member of the mutualism) are shown in grey on both maps respectively for the plant (A) or the bee (B).

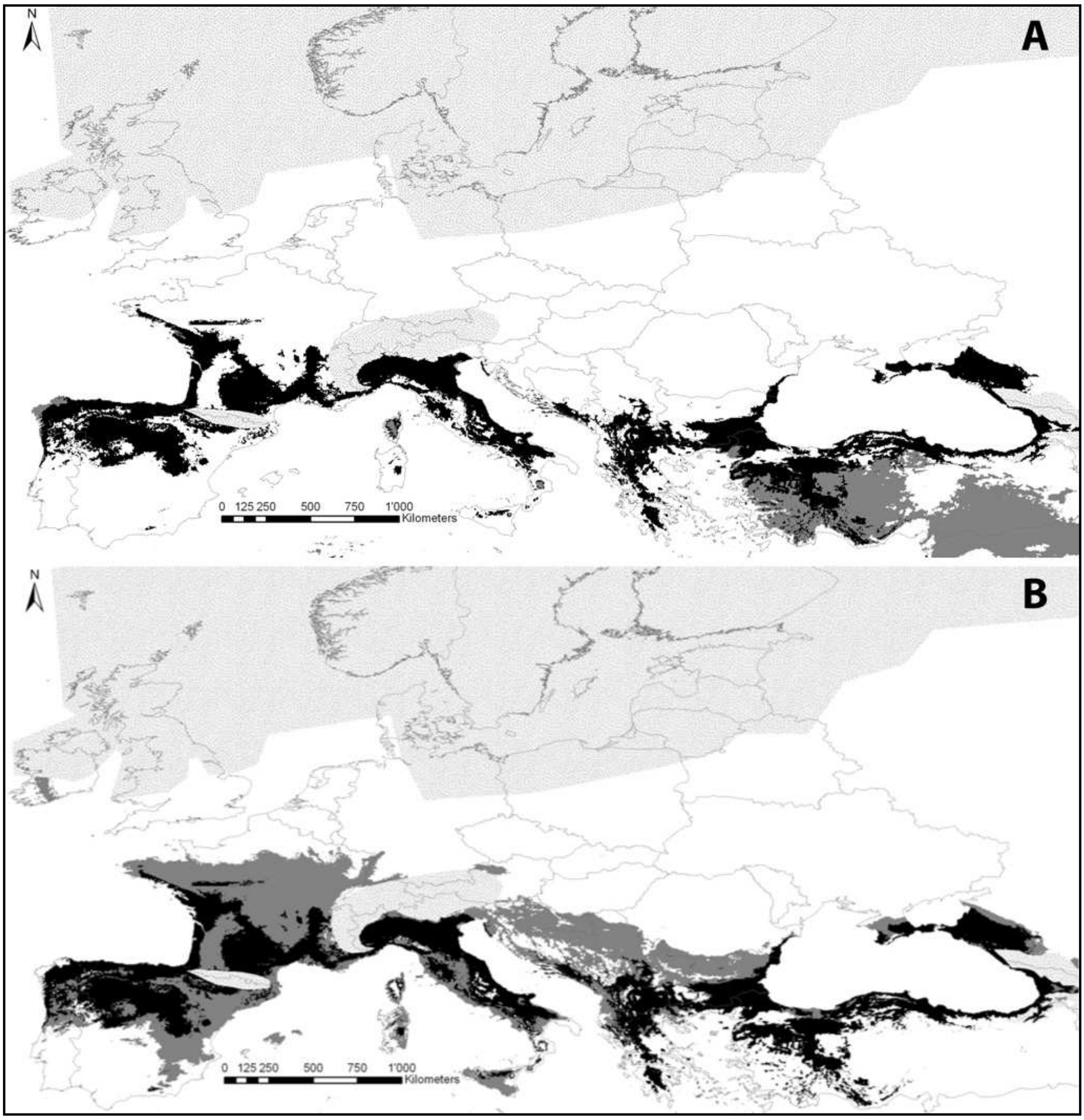


Figure 5: Similar to Fig. 4 but comparing most probable refugia between *L. vulgaris* and *M. fulvipes*.

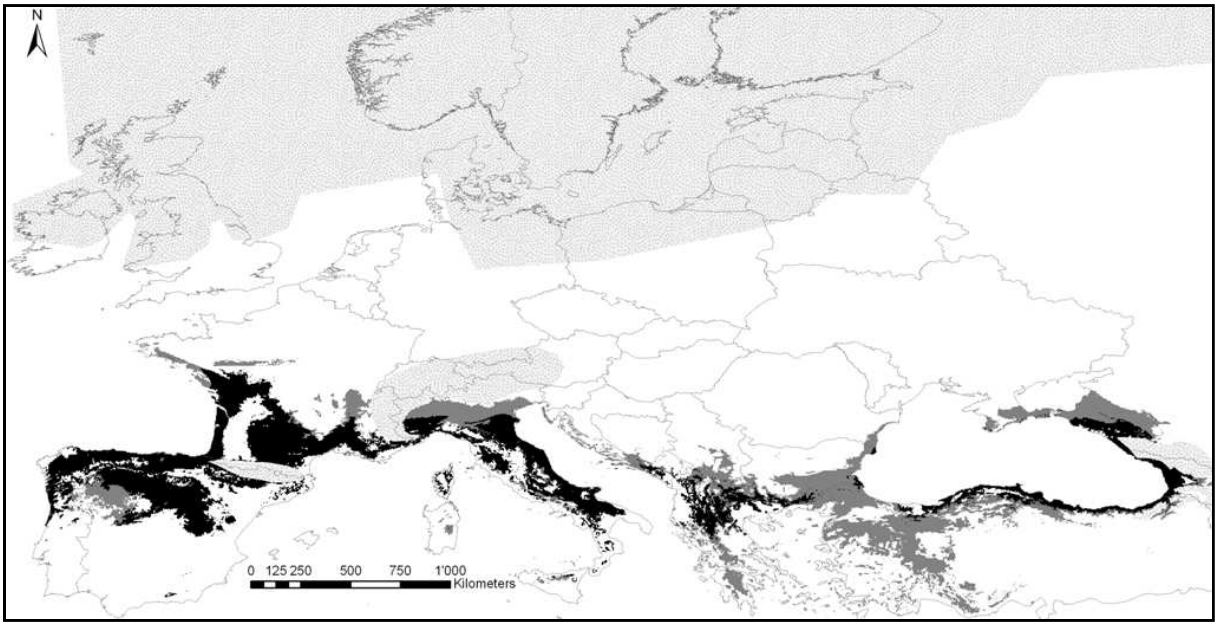
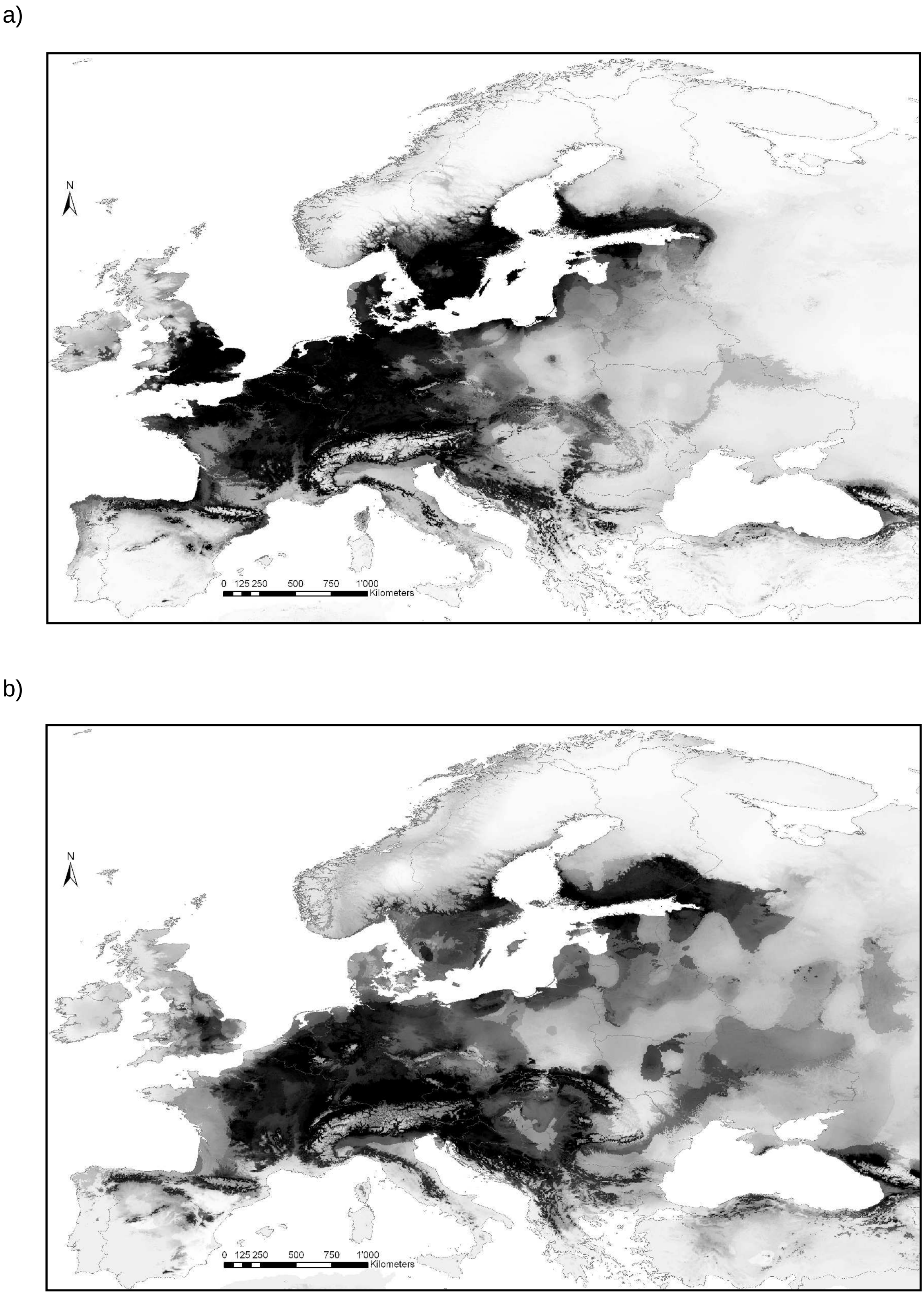


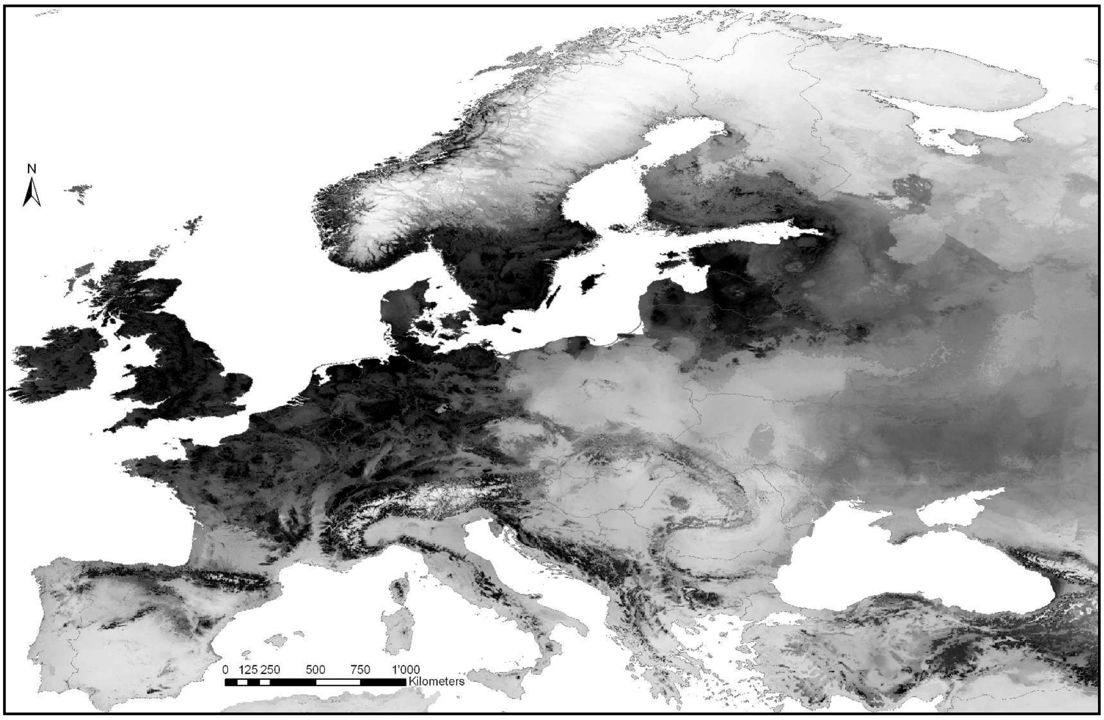
Figure 6: Map showing the refugia zones shared by *L. vulgaris* and *M. europaea* (in black) and by *L. vulgaris* and *M. fulvipes* (in black plus grey). The black area is a subset of the grey one.

SUPPORTING INFORMATION

Current potential distributions modeled for *Macropis europaea* (a), *M. fulvipes* (b) and *Lysimachia vulgaris* (c).



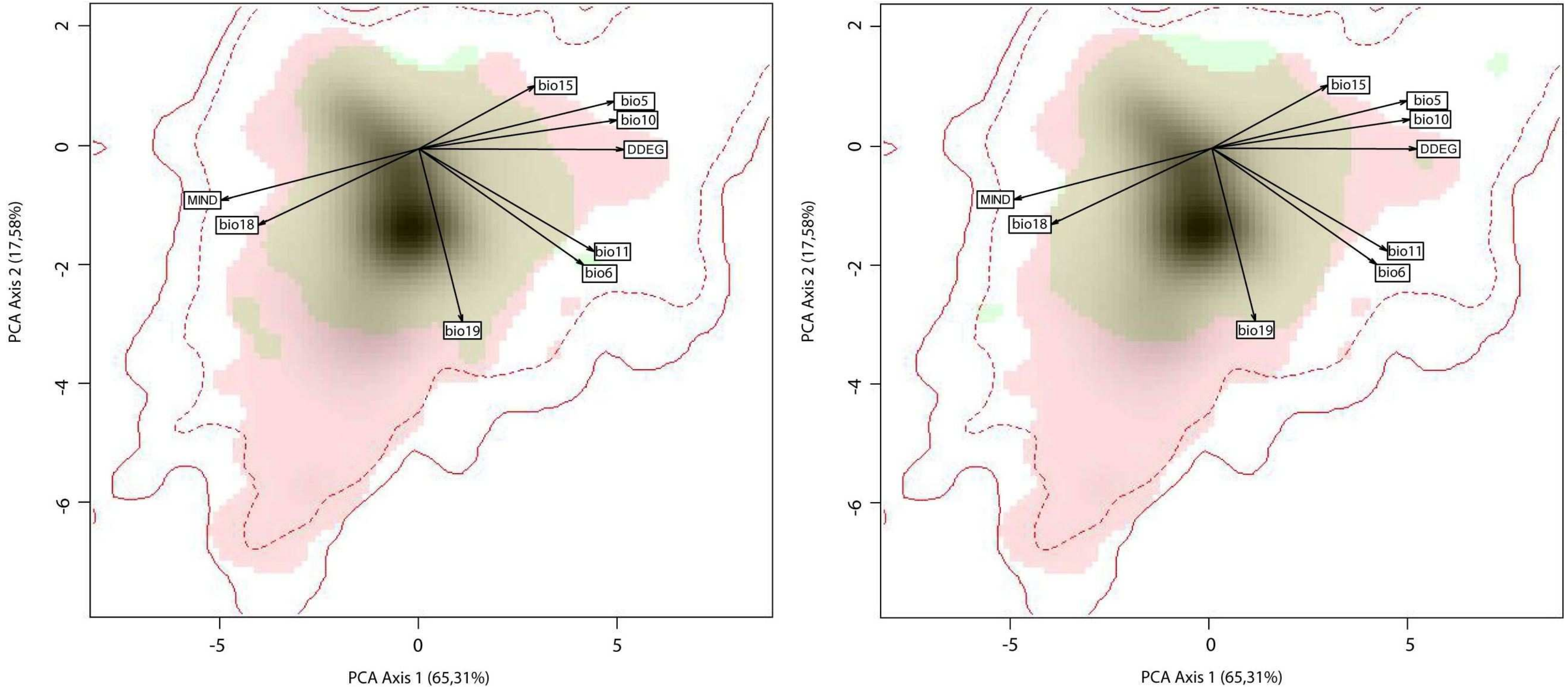
c)



SUPPORTING INFORMATION

Ecological niches comparison

The niches comparisons between *L. vulgaris* (in red) and *M. europaea* (left, in blue), or *M. fulvipes* (right, in blue) display the larger ecological niche of the plant compared to both *Macropis* species. Following the Worldclim variables, the yellow loosestrife is expected to support more humid ecological conditions (bio18, bio19) and shows a wider range of suitable temperatures, either colder (bio6, bio11) or warmer (bio5, bio10). The respective niches of bees, almost entirely comprised within the one of the plant (grey areas), illustrate their dependence on this particular oil-providing plant. On the opposite *L. vulgaris* may cope without these two main partners, more particularly in the previously described ecological conditions, by relying on other pollinators, on vegetative propagation or on selfing.



APPENDIX 1

Detailed method for the past and present distribution modelling of *Lysimachia* and *Macropis* spp.

We extracted occurrences of each species in Europe from the GBIF database (www.gbif.org). We kept only the data with a minimal spatial accuracy of 5km. These data were pooled with the occurrences collected during the fieldwork and, for *Macropis* spp., with the database published by Michez & Patiny (2005). Since the occurrences were highly aggregated in some part of Europe, we randomly selected a subset of occurrences with a minimal distance of 50km. Most modeling techniques require not solely information about presences but also absences to determine the suitable conditions for a given species, thus we selected 10’000 pseudo-absences randomly. The modeling techniques will then discriminate the conditions where there are presences from the background environment (Witz and Guisan, 2009). GBIF is known to be highly biased for western European countries (Yesson et al., 2007). Therefore, we did not select pseudo-absences in Eastern Europe countries but limited them to the the Western and Balkanic countries. Selecting pseudo-absences in unsampled countries can significantly bias the response curves of the models (Phillips et al., 2009). The resulting presences and pseudo-absences were used in the further species distribution modeling.

We ran single-models with seven climatic variables from Worldclim (Hijmans et al., 2005) at a resolution of 2.5 arc-minute (around 5 km): maximum temperature of warmest month (bio5), minimum temperature of coldest month (bio6), mean temperature of warmest quarter (bio10), mean temperature of coldest quarter (bio11), precipitation seasonality (bio15), precipitation of the warmest quarter (bio18) and precipitation of the coldest quarter (bio19).

We modeled the distribution of each species using the BIOMOD package (Thuiller et al., 2009), as implemented in R CRAN. We used eight different niche-based modeling techniques : (1) generalized linear model (GLM), (2) generalized additive model (GAM), (3) classification tree analysis (CTA), a classification method which select the best trade-off between the number of leaves of the tree and the explained deviance, (4) artificial neural networks (ANN), a machine learning method, with the mean of three runs used to provide predictions and projections, as each simulation gives slightly different results, (5) mixture discriminant analysis (MDA), a classification method that uses MARS function for the regression part of the model, (6) multivariate adaptive regression splines (MARS), (7) generalized boosting model (GBM), a machine learning method which combines a boosting algorithm and a regression tree algorithm to construct an 'ensemble' of trees, and (8) Random Forest (RF), a machine learning method which is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. These eight methods are implemented in the BIOMOD package.

In order to evaluate the predictive performance of the species distribution models, we used a random subset of 70% of the data to calibrate every model, and used the remaining 30% for the evaluation. Models were evaluated using a relative operating characteristic (ROC) curve and the Area Under the Curve (AUC) (Fielding and Bell, 1997). We replicated the data splitting 10 times and calculated the average AUC of the repeated split-sample, which gives a more robust estimate of the predictive performance of each model.

Finally, each model was projected into both current and past climate conditions with the Worldclim data of the CCSM circulation model for the last glacial maximum (LGM; -21’000 years) downscaled at a resolution of 2.5 arc-minute. In order to obtain the central trend of these distributions, accounting for variations among modeling techniques, we applied a weighted average of the eight modeling techniques based on the predictive power (AUC). Ensemble forecasting approaches have been shown to significantly improve the accuracy of species distribution models (Marmion et al., 2009).

Detailed method for the niche comparison between *Lysimachia* and *Macropis* spp.

In order to compare the niches of *L. vulgaris* and of the two *Macropis* species, we extracted the climatic information for the occurrences from the Worldclim bioclimatic variables (Hijmans *et al.* 2005). We selected several variables important for the ecology of species studied, in particular reflecting conditions of temperature and precipitation: bio5, bio6, bio10, bio11, bio18, bio19. Additionally, we calculated two additional layers from Wordclim data: degree-day (DDEG) and averaged moisture index during the year (MIND). Recently, tools have been developed to compare the niche of a species in different part of its range or to compare closely related species to investigate niche conservatism (Broennimann *et al.* submitted). The climatic niche of *L. vulgaris* was compared to the respective niches of both *M. europaea* and *M. fulvipes* in a gridded climatic space using a principal component analysis (PCA). First, we ran the PCA on the climatic environment available represented by 30’000 pixels randomly selected throughout Europe. Each pixel in the environmental space corresponds to a unique set of environmental conditions present at one or more sites in the geographical space. The occurrences of the groups of populations were then projected in the climatic space available. Finally, we applied a kernel density function to determine the "smoothed" density of occurrences of each pixel in the environmental space for each studied species. We divided the density of occurrence by the density of the environment in each focal pixel to obtain a measure of the density of the species relative to the availability of climate. This ratio was finally rescaled between 0 and 1. This approach has been shown with virtual species to be very robust to compare climatic niche of species (Broennimann *et al.* submitted). The similarity between niches was afterwards measured using the Schoener’s *D* metric, ranging from 0 (no niche overlap) to 1 (complete overlap), and its associated statistical tests (tests of niche conservatism), proposed originally by Warren *et al.* (2008). In order to test if the niches of the groups were significantly different, occurrences were pooled and randomly split, maintaining the number of occurrences as in the original datasets. The niche overlap statistics (D, see Warren *et al.* 2008) was then recalculated. This process was repeated 100 times and a histogram of simulated values was constructed. The comparison of observed and simulated values of the test statistic allows significance testing for niche equivalency. If the observed value falls within the density of 95% of the simulated values, the null hypothesis of niche equivalency cannot be rejected. We also applied a niche similarity test *sensu* Warren *et al.* (2008). Rejection of the null hypothesis indicates that the niche models of two species are more similar (or different) than would be expected by chance. Rejection of the null hypothesis also indicates that the observed niche differentiation between species is a function of habitat selection and/or suitability rather than simply an artefact of the underlying environmental differences between the suite of habitats available to the two species (Peterson *et al.* 1999; Warren *et al.* 2008).

References

Fielding, A.H. & Bell J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.*, 24, 38-49.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, 25, 1965-1978.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modeling, *Divers. Distrib.*, 15, 59-69.

Michez, D. & Patiny, S. (2005). World revision of the oil-collecting bee genus *Macropis* Panzer 1809 (Hymenoptera : Apoidea : Melittidae) with a description of a new species from Laos. *Ann. Soc. Entomol. Fr.*, 41, 15-28.

Peterson, A.T., Soberon, J. & Sanchez-Cordero, V. (1999). Conservatism of ecological niches in evolutionary time. *Science*, 285,1265-1267.

Phillips, S.J., Dudik, M., Elith, J., Graham, C., Lehmann, A., Leathwick, J., *et al.* (2009). Sample selection bias and presence-only models of species distributions: implications for selection of background and pseudo-absences. *Ecol. Appl.*, 19, 181-197.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009). BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373.

Warren, D.L., Glor, R.E. & Turelli, M. (2008). Environmental niche equivalency versus conservatism : quantitative approaches to niche evolution. *Evolution*, 62, 2868–2883.

Wisz, M.S. & Guisan, A. (2009). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data, BMC Ecology, 9:8.

Yesson, C., Brewer, P.W., Sutton, T., Caithness, N., Pahwa, J.S., Burgess, M., *et al.* (2007). How global is the Global Biodiversity Information Facility? *PLoS ONE*, 2, e1124.

GENERAL DISCUSSION

1. Contribution of the current thesis to European phylogeography: consistent patterns and original results

General phylo- and biogeographic results presented in this thesis are in agreement with paradigm patterns for the Western Palearctic, established progressively since the end of the 20th century. In each of the plant and insect systems investigated, we could retrieve classical potential refugia previously described for both temperate and cold-adapted species (e.g. Schönswetter et al. 2005, Médail & Diadema 2009, Schmitt 2009, Stewart et al. 2010). Re-colonization pathways roughly fitted to one or another of the main paradigm patterns established up to now (Hewitt 2000, 2001, Taberlet & Cheddadi 2002). However, complementary analyses of the plants and insects studied here revealed other attractive, new or unexpected outcomes, especially concerning the genetic connection between some European regions. For instance, the *Oreina speciosa / alpestris* superspecies showed strong genetic links between mountains, notably between the Balkans and the Alps, or between the Black Forest and the Carpathians. The phylogeographic pattern of *Macropis europaea* revealed genetic similarity between Italian and Iberian peninsulas. These are some unusual results when comparing with previous European phylogeographic studies. Discussing wide-scale similarities for all systems analyzed in this study however remains difficult, considering the large range of organisms investigated in this thesis. Nevertheless, despite regions of Europe played a different role in each organism/system, three general considerations are common to all plants and insects investigated here.

Firstly Eastern Europe, and more particularly the Balkans, stands for a key area for understanding the geographic structure of genetic variation in European widespread species. In all the systems addressed, this region is indeed situated at the interface between distinct lineages. It also notably represents the contact zone of genetically well-established European and Anatolian gene pools. Finally, it corresponds to a source of historically, and currently, isolated genetic lineages (e.g. Naciri & Gaudeul 2007). Particular attention should be paid to this region in future research, as in addition, the Balkanic area was neglected in most previous phylogeographic works at the European scale.

Second, in the systems studied here, central Europe usually displays intricate spatial genetic structures compared to most southern peninsulas (especially in temperate widespread species). This pattern might be due to the biology of species (minute seeds, relation to waterways, easy dispersal by flight) and to the absence of important barriers in central and northern Europe. Such a result is surprising in view of previously established paradigms for western Palearctic (Hewitt 2000, 2001, Taberlet et al. 1998), and we would have expected a more clear-cut lineages structure in central Europe. Different approaches (e.g., using additional genetic markers) as well as a larger sampling of individuals per population might have allowed more precise trends for central and northern Europe. Overall, it is nevertheless primordial to first consider the type of organism studied and its associated biological properties, in order to understand processes behind genetic patterns.

Finally, the combination of different methods allows outlining a complete picture of the phylogeographic patterns and underlying processes in a given species. This implies associating different genetic methods (proved to perform well when dealing with genetically little-variable organisms [e.g. Devey et al. 2008]) with other approaches such as spatial genetic structure analysis and ecological distribution modelling (e.g. Pease et al. 2009), an optimal method to test and discuss hypotheses on both genetic and ecological bases.

2. Phylogeography and species concept: two highly related topics

Combining spatial genetic structure with past & present ecological niche modelling allows a wider understanding of the recent evolutionary history in both single taxa and groups of related species (see Chapters I, II and III). In this thesis, the latter was associated with antithetic patterns: from (i) species classically (i.e., morphologically) considered as distinct that, however, were not (or weakly) genetically differentiated (in the Fly Orchid group), to (ii) similar morphological species that only partially corresponded to distinct phylogenetic clades, and for which an intricate history was revealed (in the *Oreina speciosa* / *alpestris* complex); but also (iii) several taxa classically considered as synonymous descriptions of the same species that finally resulted in two genetically well-isolated entities (in *Lysimachia punctata* / *verticillaris*). For all these closely-related species systems, previous phylogenies including only a few specimens of the species group did not reveal enough details concerning the species groups' histories and evolution (Anderberg et al. 2007, Devey et al. 2008, Dobler et al. 1996). In contrast, here, our large-scale phylogeographic approach based on a wide sampling (covering most of the species' distribution ranges available throughout Europe) has allowed defining with a high level of accuracy the biogeographic history of the studied taxa. It seems now obvious that further studies dealing with species delimitations among closely-related organisms should as much as possible consider phylogeographic approaches with a large-scale sampling in order to optimally resolve taxonomical and evolutionary questions.

This is especially true when considering the range of molecular techniques currently available, very innovative and generally affordable.

Intra-specific genetic structure and species concept are tightly related topics: phylogeographic patterns could help in delimiting species or at least distinct genetic entities, and species should be correctly defined in order to perform accurate phylo- and biogeographic analyses.

3. Comparative phylogeography in mutualists: patterns independently driven by the particular biogeographic history of each species

When applied to the particular case of ecological interactions (i.e., in the pollination mutualism studied in Chapter IV and Appendices 1 and 2), comparative phylogeography permits to conclude that mutualists might have shared a large part of their refugia, but also that their respective spatial genetic structures are generally independent. Whereas we expect most refugia to be shared among interacting species (because of the mutualistic nature of involved partners [Alvarez et al. 2010]), addressing independent spatial genetic structures in such partners was a much more unforeseen result. However, it seems manifest that dispersal abilities as well as other life-history traits (e.g., generation time) characteristic of each type of organism (rather than ecological interactions) are strongly prevailing to explain the history of genetic lineages. More unexpected is the conclusion that quickly-dispersing organisms (e.g., flying insects such as *Macropis* bees) are more limited in their dispersal by natural barriers than are slowly-dispersing species (e.g., *Lysimachia* plants). To understand such constrains, it is necessary to take into account the species' ecological niches: in species with a narrow niche, even if dispersal abilities are high, the “realized” dispersion will be strongly tempered by a low probability of finding a suitable habitat; in species with a wider niche, even if associated with low dispersal abilities, step by step dispersal could take place relatively rapidly *via* a wide range of suitable habitats. The fate of genetic lineages in related species could then be drastically different depending, for instance, on the stenoecic / euryoecic status of a species (as in *M. europaea* vs. *M. fulvipes*). Last but not least, considering spatial genetic patterns only might fail in explaining correctly the past histories of organisms (e.g., by missing the necessary ecological clues brought out by ecological niche modelling or by not considering the history of a species into a wider context involving specific but also additional non-obligate partners). Further investigations including coalescence analyses based on past ecological niche modelling (Carstens & Richards 2007) might allow investigating one step further the link between patterns and processes.

4. Conclusion

After many years of single-species phylogeographic studies performed on numerous species of plants, animals, fungi, insects and other invertebrates, the time has come to perform comparative research on associated organisms. Ideally, future phylogeographic studies should combine genetic structure analyses with ecological descriptions and niche modelling. Such phylogeographic approaches would also probably be useful in the perpetual attempts to delimit and describe species, as well as in trying to define a hypothetical unified species concept. This thesis points out which aspects could be combined in order to unravel phylogeographic patterns among closely-related species and interacting organisms. Hopefully some of the approaches performed here might be constructive, for instance in the elaboration of future comparative European biogeographic studies involving large numbers of species, e.g., from a same ecological community. Understanding the history of communities and the processes involved in shaping and maintaining species' compositions through space and time might represent the next step for further phylogeographic research.

References

Alvarez N, Kjellberg F, McKey D, Hossaert-McKey M. 2010. Phylogeography and historical biogeography of obligate specific mutualisms. In: *The Geography of Host-Parasite Interactions* (eds. Morand, S. & Krasnov, B.). Oxford University Press, UK, pp. 31-39.

Anderberg AA, Manns U, Källersjö M. 2007. Phylogeny and floral evolution of the *Lysimachieae* (Ericales, Myrsinaceae): evidence from ndhF sequence data. *Willdenowia* **37:** 407-421.

Carstens BC, Richards CL. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* **61:** 1439-1454

Devey DS, Bateman RM, Fay MF, Hawkins JA. 2008. Friends or relatives? Phylogenetics and species delimitation in the controversial European orchid Genus *Ophrys*. *Annals of Botany* **101:** 385-402.

Dobler S, Mardulyn P, Pasteels JM, Rowell-Rahier M. 1996. Host-plant switches and the evolution of chemical defense and life history in the leaf beetle genus *Oreina*. *Evolution* **50:** 2373-2386.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907-913.

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology* **10:** 537-549.

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

Naciri Y, Gaudeul M. 2007. Phylogeography of the endangered *Eryngium alpinum* L. (Apiaceae) in the European Alps. *Molecular Ecology* **16:** 2721-2733.

Pease KM, Freedman AH, Pollinger JP, et al. 2009. Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology* **18:** 1848-1862

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.

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

Stewart JR, Lister AM, Barnes I, Dalen L. 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B-Biological Sciences* **277:** 661-671.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7:** 453-464.

Taberlet P, Cheddadi R. 2002. Quaternary refugia and persistence of biodiversity. *Science* **297:** 2009-2010.

APPENDIX 1

Ecological niche overlap in sister species: how do oil-collecting bees *Macropis europaea* and *M. fulvipes* (Hymenoptera: Melittidae) avoid competition and hybridization?

Lucas Bassin, Nadir Alvarez, Loïc Pellissier and Yann Triponez

(Version formatted according to *Apidologie*)

Neuchâtel, August 2010

Ecological niche overlap in sister species: how do oil-collecting bees *Macropis europaea* and *M. fulvipes* (Hymenoptera: Melittidae) avoid competition and hybridization?

Lucas BASSIN, Nadir ALVAREZ, Loïc PELLISSIER and Yann TRIPONEZ*

Lucas BASSIN
Laboratoire d'Entomologie Evolutive, Institut de Biologie, Université de Neuchâtel, Switzerland
lucas.bassin@unil.ch

Nadir ALVAREZ
Département d'Ecologie et Evolution, Université de Lausanne, Switzerland
nadir.alvarez@unil.ch

Loïc PELLISSIER
Département d'Ecologie et Evolution, Université de Lausanne, Switzerland
loic.pellissier@unil.ch

*Yann TRIPONEZ,
Laboratoire d'Entomologie Evolutive, Institut de Biologie
Université de Neuchâtel, Rue Emile-Argand 11, cp 158
CH-2009 Neuchâtel, Switzerland
Phone +41 32 718 31 62 - Fax +41 32 718 30 01
yann.triponez@unine.ch

Abstract

Oil-collecting bees are found worldwide and always in association with mutualistic oil flowers. In the Western Palearctic, three oil-collecting bee species within the genus *Macropis* (Hymenoptera, Melittidae) interact in a tight pollination mutualism with species of the only European oil-producing plant genus *Lysimachia* L. (Myrsinaceae). Two of these oil-collecting bees (*M. europaea* and *M. fulvipes*) show overlapping geographic distributions, comparable morphologies and similar ecological characteristics (e.g. habitat's type, floral preferences). In view of these similarities we presume that hybridization should occur between the two species, unless potential variation among the species' ecological niches prevents it, decreasing at the same time competition for resources. Using modern genetic analyses and ecological niche modelling on a large bee sampling throughout Europe, we discuss new perspectives on the ecology and evolutionary history of this mutualism.

Keywords

Macropis / mutualism / *Lysimachia* / ecological niche / hybridization

1. Introduction

Mutualisms between insects and plants are widely recognized as indispensable components in ecosystems functioning (Bronstein, 2006). Among the most fascinating plant-insect mutualisms, nursery pollination systems have been largely documented, notably in the fig/fig wasps and the yucca/yucca moths interactions (e.g. Pellmyr, 2003; Cook and Rasplus, 2003; Machado et al., 2005). In contrast, specific pollination systems involving oil-producing plants and oil-collecting bees have been much less studied (Vogel, 1988). In such systems, oil-offering flowers, found worldwide in 11 different plant families, are pollinated by specifically adapted oil-collecting bees, distributed in a few genera within Melittidae and Apidae families. Such bees use oil combined to pollen as larval food, or/and for water-resistant lining of larval cells (Renner and Schaefer, 2010). Oil-collecting bees have developed special branched hairs on legs or abdomen to facilitate oil transportation (Rasmussen and Olesen, 2000), an adaptation already observed on the 53 Mya old fossil species *Palaeomacropis eocenicus* (Michez et al., 2007), which attests for the ancient nature of coevolutionary processes within this mutualism.

In temperate habitats, only a few species have developed such adaptations. Among them are found species within the genus *Macropis* Panzer 1809 (Hymenoptera, Melittidae), one of the 14 genera included in Melittidae (Michener 2000). The 16 species of the genus are subdivided into three subgenera (*Macropis* s. str., *Paramacropis* Popov and Guiglia 1936 and *Sinomacropis* Michener 1981), spanning the Holarctic (Michez and Patiny 2005). All *Macropis* taxa visit, collect oil and pollinate the yellow flowers of *Lysimachia* species (Myrsinaceae; Vogel, 1976). Evidence of floral oil was found in at least 75 among the 191 species of *Lysimachia* (Vogel, 1986, 1988; Hao et al., 2004). Here the oil is produced by elaiophores (oil secreting trichomes) located at the basal part of petals and anther's filaments (Simpson and Neff, 1983). These plants are native either of temperate Eurasia (subgenus *Lysimachia* s. str.) or northern America (subgenus *Seleucia*; Hao et al., 2004). As in the great majority of mutualistic interactions, strict one-to-one relationships between oil-collecting bees and *Lysimachia* species do not represent the general rule; hence each plant species is generally visited by more than one species of oil-collecting bees, depending on localisation and period of flowering (Cane et al., 1983; Simpson and Neff, 1983; Pekkarinen et al., 2003; Celary, 2004).

Three species of *Macropis* are found in the western Palearctic ecozone: *Macropis europaea* Warncke 1973, *M. frivaldszkyi* Mocsary 1878 and *M. fulvipes* Fabricius 1805 (Michez and Patiny, 2005). All three species can be found together in sympatric populations (Celary, 2004; Pekkarinen et al., 2003; L. Bassin and Y. Triponez pers. obs.). Among them, *M. frivaldszkyi* (for which very little documentation relative to its ecological characteristics is currently available) shows the narrowest distribution (Balkans, Anatolia, Syria). In contrast, *M. europaea* and *M. fulvipes* are widely distributed and hold rather similar ranges across Europe, although the former species is more restricted eastwards (in Russia) and even absent from Anatolia or Caucasus, and the latter does not reach the Scandinavian Peninsula (i.e. Sweden and Norway) and the British Islands (Michez and Patiny, 2005). These two species display a high level of similarity in their morphology, behavior, habitat and floral preferences. For instance, a comparative study established in southern Poland (Celary, 2004) showed that all features concerning the nest's architecture and the larval stages differ almost only in size (always slightly bigger for *M. fulvipes*). The most obvious difference between the two species seems to stand in their respective phenologies, with a flight period delayed of about ten days (independently observed in different regions of Europe; Michez, 2002; Pekkarinen et al., 2003; Celary, 2004). Such a delay in flight periods might be meaningful for resources partitioning in sympatric *Macropis* populations, especially in places where only one *Lysimachia* species is available (Celary, 2004). The floral preferences of bees (reviewed by Michez and Patiny, 2005) reveal that *L. vulgaris* seems to be the main resource of oil for both *M. europaea* and *M. fulvipes*. It might even be the only one for the strictly monoloectic *M. europaea*. *Macropis fulvipes* was additionally observed on *L. nummularia* (sometimes presumed as its preferred plant species [Michez, 2002; Pekkarinen et al., 2003]) and on *L. punctata* (especially on its native area in Anatolia and Caucasus; Y. Triponez pers. obs.).

In the current study, our aim is to address whether or not *M. europaea* and *M. fulvipes* are able to efficiently avoid too strong competition and counter-selected hybridization, despite their similar ecologies and frequent syntopy on *L. vulgaris*. Based on a large sampling all over Europe, we modeled the bees' ecological niches and analyzed the sampled specimens using both gene sequencing and AFLP genotyping, in order to answer the following questions. 1) How distantly related are these two species from a genetic point of view? 2) Can hybrids between *M. europaea* and *M. fulvipes* be detected and how frequently? 3) On which ecological properties do the ecological niches of the two species differ? 4) How these potential ecological differences affect their co-existence and allow them to avoid competition for resources? So far, no study has provided objective elements to answer these questions.

2. Material and Method

2.1. Field work

As a first step to identify potential populations of *Macropis*, we used international and national freely accessible databases to locate sites of *L. vulgaris*, *L. nummularia* and *L. punctata* (in its native area) across Europe. The sampling and the observations of *Macropis* in Europe were performed during summers 2006 to 2008, each year between June and August depending on the region. Because of the tight interaction between the plant and the bee, as soon as a *Lysimachia* population was found, *Macropis* females could be catch easily while pollinating the flowers. From time to time, patrolling males were also collected in flight with an insect net. We aimed at capturing between five and ten bees per population. All insects were conserved in Ethanol 70%. The three *Macropis* species were determined following Michez and Patiny (2005). A complementary study describing the pollinators of *L. nummularia* was carried out in parallel in June and July 2008 (see Electronic-only material).

2.2. Genetic distance between *M. europaea* and *M. fulvipes* and detection of putative hybrids

Total genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) from the wing muscles of the bees carefully removed and rinsed, in order to avoid plant DNA contamination via pollen grains stuck to the bees' exoskeleton. To estimate the genetic distance between *M. europaea* and *M. fulvipes*, we chose four individuals (from distant origins) for each bee species and sequenced three mtDNA regions classically used for barcoding: 1) 16s ribosomal RNA (*16s rRNA*, using the primers LR-N-13398 and LR-J-12883 from Simon et al., 1994), 2) partial cytochrome oxidase I (*COI*, with the primers C1-J-1751 and C1-N-2191 from Simon et al., 1994) and 3) partial cytochrome B (*CytB*, using primers designed by Belshaw and Quicke, 1997). Fragments were amplified using a standard 30 µl PCR mix. The PCR were run in a TGradient thermocycler (Biometra, Goettingen, Germany) using the same program for all mtDNA regions (the detailed composition of PCR mix and conditions of PCR amplification steps are available by the last author). The PCR product purification and sequencing was carried out by Macrogen Inc. (Seoul, South Korea). Sequencing was performed using forward primers for all mtDNA regions under BigDye TM terminator cycling conditions, purifying the products using ethanol precipitation and running them into an Automatic Sequencer 3730xl (Applied Biosystems, Foster City, USA). Sequences were manually corrected using the software Chromas Pro 1.34 (Technelysium, Helensvale, Australia). Alignment was carried out for each mitochondrial region using ClustalW Multiple Alignment (Thompson et al., 1997) as implemented in the software BioEdit 7.0.5.3 (Hall, 1999), followed by minor manual corrections. Genetic distances between species and general genetic statistics were calculated using MEGA 4 (Kumar et al., 2008).

In order to detect potential hybridization among *Macropis* species, we used Amplified Fragment Length Polymorphism (AFLP) genotyping. AFLP analysis followed the classical method described by Vos et al. (1995), with slight modifications (detailed protocol available upon request to the last author). Reactions were conducted in 96-well plates, in which samples were randomly distributed. We used restriction enzymes *EcoRI* and *MseI* for DNA digestion. A primer trial on a small number of samples was conducted using six different primer combinations to identify pairs of selective primers that were repeatable and polymorphic enough. Each individual sample was fingerprinted with the two primer combinations *EcoRI*-ACA/*MseI*-CAA and *EcoRI*-ACG/*MseI*-CAA. Final selective PCR products were analyzed using the GeneScan technology with an automated capillary-sequencer (ABI 3730XL, Applied Biosystems, Foster City, CA; service provided by Macrogen Inc., Seoul, South Korea). Resulting fluorescent AFLP patterns were scored using GeneMapper 3.7 (Applied Biosystems) with some changes in default parameters (maximum bin width=1 bp, light peak smoothing, peak threshold=200 RFU). We scored the presence or absence of each AFLP marker in each specimen. The matrices of the two scored primer pairs were concatenated into one single binary matrix where individuals and bands were stored as lines and columns, respectively. Multivariate ordinations using the principal coordinate analysis (PCoA) approach were produced on this distance matrix to investigate genetic relationships among specimens and taxa using the software Ginkgo 1.5.8 (Bouxin, 2005). The PCoA relied on the Jaccard similarity coefficient (Jaccard, 1908) to avoid grouping terminals on the basis of shared zeros.

2.3. Ecological niche modeling in *M. europaea* and *M. fulvipes*

We modeled the current distribution of *M. europaea* and *M. fulvipes* using seven modeling techniques implemented in the BIOMOD (Thuiller et al., 2009) R package (see detailed protocol provided as Electronic-only material). Models were calibrated from presences collected during the field sampling pooled with occurrences from the GBIF database group (www.gbif.org) and using seven climatic layers from Worldclim (Hijmans et al., 2005). An Ensemble forecasting approach (Marmion et al., 2009) was then used to project a central tendency from a combination of the models obtained by the different niche-based modeling techniques into current climate conditions in Europe. Ecological niches of both species were displayed on the same referential, a multi-dimensional scale represented by the two first axes of a PCA (see detailed method provided as Electronic-only material). We finally used the geographical data recorded during field work to calculate the correlation between the presence of each bee species and both altitude and latitude, and tested if the two species showed significant differences regarding both factors by performing a t-test and a one-way ANOVA (for altitude) and a Mann-Whitney Rank Sum test (for latitude). All statistical analyses were performed with the program SigmaStat version 2.03 (SPSS Inc., Chicago, USA).

3. Results

3.1. *Macropis* sampling

In total, 94 *Macropis* populations (composed of one or two species) were collected in Europe, Turkey and Georgia (Table 1; Fig. 1). *Macropis europaea* was most frequently collected, with a total sampling of 419 individuals. It was followed by *M. fulvipes* (198 individuals) and *M.* frivaldszkyi (8 individuals). We observed 58 populations comprising *M. europaea* only (in white on Fig. 1), 18 with only *M. fulvipes* (in grey) and a single with only *M. frivaldszkyi* (in black). Sympatric populations were also found, mainly composed of *M. europaea* and *M. fulvipes* (15 sites in Western and central Europe); *Macropis frivaldszkyi* was observed in sympatry with *M. europaea* in a single location in Serbia (population KRU) and with *M. fulvipes* in a single location in Turkey (population ERF). Regarding floral preferences, *M. europaea* was only sampled on *L. vulgaris*, whereas *M. fulvipes* and *M. frivaldszkyi* were caught on both *L. vulgaris* and *L. punctata*. Furthermore two individuals of *M. fulvipes* were also found on *L. nummularia* in one single site in Switzerland (visited during the pollination study presented as Electronic-only material).

3.2 Genetic distance between *M. europaea* and *M. fulvipes*

Amplification of mitochondrial genes *COI*, *16s rRNA* and *CytB* resulted in alignments of 423 bp, 485bp and 150 bp respectively. The largest genetic distance between *M. europaea* and *M. fulvipes* was obtained with *CytB* (0.15±0.03), for which 24 nucleotides were variable (22 parsimony-informative). *COI* was the second most variable gene (distance of 0.09±0.01), with 35 polymorphic sites (34 parsimony-informative). Finally *16s rRNA* resulted in the smallest distance among species (0.04±0.01) and showed 22 variable nucleotides (21 parsimony-informative).

3.3. Hybridization in *Macropis*

In total, 291 individuals (198 *M. europaea*, 87 *M. fulvipes* and six *M. frivaldszkyi*) were analyzed by AFLP genotyping in order to detect hybridization between *Macropis* species. The automatic scoring of the two primer pairs resulted in a total of 471 alleles (out of which 447 were parsimony-informative), each individual yielding a minimum of 120 and a maximum of 219 fragments. The PCoA analysis (Fig. 2) showed a strong partitioning of species when considering the first three axes (explaining respectively 37.9%, 7.9 % and 5.9% of total variance, i.e., 51.7% of cumulative variance). Hybridization was rare, and only one single individual (a female, sampled in TUL, holding *M. europaea* morphology) showed an intermediate position in the PCoA, typical of hybrids, situated halfway between the scatter-plots representative of *M. europaea* and *M. fulvipes*. In addition, three other samples, all from LOZ and showing a *M. fulvipes* morphology, clustered together with *M. europaea* samples. This pattern is hard to explain unless we invoke labeling error during DNA extraction in samples from the LOZ population.

3.4 Mean altitude and latitude comparison between *M. europaea* and *M. fulvipes*

Overall, *M. europaea* was collected at a significantly lower mean altitude (405 m. a.s.l.) than was *M. fulvipes* (732 m. a.s.l.) (t-test, $t = 4.904$, $P < 0.001$). The ANOVA also showed that the mean altitudes of collection were significantly different ($df = 2$, $P < 0.001$; see Fig. 3) between allopatric populations of *M. europaea* (360 m. a.s.l), of *M. fulvipes* (867 m. a.s.l) and sympatric sites (580 m. a.s.l.). As expected, sampling altitudes significantly decreased as latitude increased, with a strong negative correlation between altitude and latitude for both species (Pearson Product Moment Correlation, $r =-0.54$, $P < 0.0001$ for *M. europaea* and $r = -0.66$, $P < 0.0001$ for *M. fulvipes*). A Mann-Whitney Rank Sum test showed that there was no significant difference in the sampling latitude between the two species ($P = 0.065$).

3.5 Ecological niches modeling comparison between *M. europaea* and *M. fulvipes*

The overlapping representation of the respective ecological niches of the two *Macropis* species on the two first PCA axes based on the Worldclim variables shows that the niche of *M. fulvipes* is slightly larger than that of *M. europaea* (presented as Electronic-only material). The Schoener's D metric with its associated statistical tests (Warren et al., 2008) revealed that the ecological niches modeled for each *Macropis* species were significantly different (D = 0.722, P = 0.019), although in relation to the total available environment, the niches remained quite comparable, with the two bee species globally using analogous environmental space. Generally *M. fulvipes* occupies a slightly larger and more diffuse ecological niche, tolerating colder, more humid but also drier conditions than *M. europaea*. Consequently, differences in the patterns of potential European distribution could be observed for the two species. Figure 4 shows the current modeled distributions of both species. In general *M. europaea* (Fig. 4A) showed a more "temperate oceanic" pattern, with wide optimal areas in central Europe (north from the Alps), as well as in land zones adjacent to Atlantic, Baltic and North Seas. The species also did not extend much eastern than the Carpathians. In contrast, the potential distribution of *M. fulvipes* (Fig. 4B) could rather be qualified of "temperate continental", showing a clear Eastern shift of its optimum, especially in the Balkans and the Carpathians, into Russia and around the Black Sea. Globally, suitable areas (in different grey shadings) are more extended in *M. fulvipes* than in *M. europaea*, although optimal environments (in black) are rarer or, in any case, more diffuse. A close-up on the Alpine region confirmed the previous results forecasting *M. fulvipes* to be present at higher elevations (see Fig. 4).

4. Discussion

4.1 Important genetic distance impedes post-zygotic hybridization between *Macropis* species

Even if hybridization between species might be counter-selected because hybrids tend to fall into adaptive valleys (Coyne, 1998), different cases of hybridization were described in bees (e.g. Ribble, 1973; Hall, 1990; Rinderer et al., 1991; Nascimento et al., 2000). Based on the assumptions that *M. europaea* and *M. fulvipes* demonstrate (i) similar behaviors and ecologies (e.g., for pollination and reproduction), (ii) comparable morphologies and (iii) frequent syntopy (i.e., they are often found in sympatry, collecting oil on the same plant), it can be expected that hybridization might occur between both species despite their high-level of genetic divergence (15% on *CytB*). However, because the slight delay in their respective phenologies can be considered as a pre-zygotic barrier to their reproduction, the frequency of such events remained difficult to predict. In the current study, hybridization is shown to be rare, since among 291 individuals analyzed only one single confirmed hybrid was identified (0.3% of the total sampling).

The high level of divergence found between these two taxa, which however branch together as two sister terminal clades in an ongoing phylogenetic study on Mellittidae including six *Macropis* taxa (D. Michez, pers. comm.), is at the upper limit values observed for sister species within a same genus (Sheffield et al., 2009). Such a high genetic distance cannot be due to recent events spanning Pleistocene's climatic oscillations (in contrast to the intra-specific phylogeographic history of lineages within each species; see Triponez et al. in prep.) and the divergence between *M. europaea* and *M. fulvipes* is probably relatively ancient, taking place in the Pliocene or even the Miocene. As a consequence, reproductive incompatibility between these species is probably well established, as attested by the rare evidence of hybridization addressed in the present study. However, more intense screening of the populations from Balkans (we should keep in mind that the single hybrid detected was collected in the Bulgarian population TUL) could reveal further hybridization events, which could also imply *M. frivaldszkyi*, whose level of divergence with the two other bees has not been evaluated yet.

4.2 New insights on the ecological preferences and spatial distributions of *M. europaea* and *M. fulvipes*

Additional pre-zygotic barriers preventing hybridization between *M. europaea* and *M. fulvipes* might reside in several ecological properties that reduce the probability of co-occurrence between the two species. In our study, at least three observations confirm the shift in ecological niches of the two species: (i) despite model predictions forecasted rather similar optimal areas for the two species in a large part of central Europe (Fig. 4), we have identified only scarce sympatric populations (15 sites on our whole European sampling, see Fig. 1); (ii) occurrences in our sampling were spatially more continuous for *M. europaea* than for *M. fulvipes*, the latter being most frequent in mountainous areas; (iii) niches were significantly different as attested by the significant test for the Schoener's metric (i.e., the one of *M. fulvipes* being slightly larger than that of *M. europaea*). When confronting (i) and (ii) to the paradigm of ecological valence *sensu* Dajoz (1978), we can consider *M. fulvipes* as an euryoecic species tolerating a larger ecological range (e.g. potential habitats, flower's preference) than *M. europaea*, which is more stenoecic. Indeed, *M. fulvipes* is not only present in more climatically harsh regions (e.g. in continental Europe, in the Alps; see Fig. 4), but it is also able to collect oil from other *Lysimachia* species (Fig. 1, also see the *L. nummularia* pollination study in the Electronic-only material), even if *L. vulgaris* remains its more frequent oil-providing plant species in Europe. In contrast, *M. europaea* shows higher probabilities of occurrence in more temperate climatic zones and there are no convincing evidences yet of *M. europaea* feeding on other oil-providing plants (data compiled until here seem much anecdotic). Consequently, *M. europaea* is also expected to be more competitive than *M. fulvipes* when both species are found in sympatry on *L. vulgaris*. The less competitive *M. fulvipes* would then replace *M. europaea* in habitats with harsher conditions, as shown by the significantly higher mean altitude where *M. fulvipes* was collected.

Sympatric sites are therefore not as frequent as expected by the ecological niche's models, because they could be located only at transition areas, where *M. europaea* reaches its ecological limits and becomes less competitive. As a matter of fact, the 15 sympatric sites interestingly stand at a mean altitude of 580 m. a.s.l., the almost exact halfway point between mean altitudes of each bee species.

4.3 Other factors driving the *Lysimachia – Macropis* mutualism

In addition to the varying ecological properties of each bee species, other forces inherent to the mutualistic interaction with the oil-providing plant could additionally influence the respective distribution of bees. The scents emitted by *Lysimachia* flowers could for instance play a role in differential attraction of the two species. In a preliminary GC-MS chemical analysis performed on the *L. vulgaris* flower's bouquet, one particular substance (junipene [C15H24], a tricyclic sesquiterpene) was found in quite high amounts in four *L. vulgaris* populations pollinated by *M. fulvipes*, but was absent from two populations pollinated by *M. europaea* (*L.* Bassin unpubl. data). Could this substance induce a species-specific attraction and consequently represent a supplementary factor influencing the bees' distribution? Further chemical analysis in a larger proportion of the populations sampled in this study as well as biotests performed on bees would allow testing this hypothesis.

Acknowledgments

The authors want to warmly thank all people that bring their precious help during fieldwork sampling: R. Arnoux, P. Bassin, A. Espíndola, A. Favre, I. Hiltpold, D. Kharazishvili and B. Prudat. A special thank to A. Sarr for support with AFLP analysis and to B. Schatz for collaboration with the chemical analyses. We especially want to thank D. Michez for providing supplementary samples for DNA analysis, geographical databases for niches modeling and for providing helpful comments. We are grateful to F. Amiet and J.-P. Haenni for the precise determinations of pollinators. We also thank the CRSF (Centre du Réseau Suisse de Floristique) for providing geographic coordinates of Swiss *Lysimachia* populations. This research was supported by the Swiss National Science Foundation (grant n° 3100A0-116778/1). The fieldwork was additionally granted by the travel grants commission of the Swiss Academy of Sciences (SCNAT), and by the M. Wütrich and A. Mathey-Dupraz Foundation (University of Neuchâtel).

References

Belshaw R., Quicke D.L.J. (1997) A molecular phylogeny of the Aphidiinae (Hymenoptera: Braconidae), Mol. Phylogenet. Evol. 73, 281-293.

Bouxin G. (2005) Ginkgo, a multivariate analysis package, J. Veg. Sci. 163, 355-359.

Bronstein J.L., Alarcon R., et al. (2006) The evolution of plant-insect mutualisms, New Phytol. 1723, 412-428.

Cane J.H., Eickwort G.C., et al. (1983) Foraging, grooming and mate-seeking behaviors of *Macropis nuda* (Hymenoptera, Melittidae) and use of *Lysimachia ciliata* (Primulaceae) oils in larval provisions and cell linings, Am. Midl. Nat. 1102, 257-264.

Celary W. (2004) A comparative study on the biology of *Macropis fulvipes* (Fabricius, 1804) and *Macropis europaea* Warncke, 1973 (Hymenoptera : Apoidea : Melittidae), Folia Biol.-Krakow 521-2, 81-85.

Cook J.M., Rasplus J.Y. (2003) Mutualists with attitude: coevolving fig wasps and figs, Trends Ecol. Evol. 185, 241-248.

Coyne J.A., Orr H.A. (1998) The evolutionary genetics of speciation, Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. 353, 287-305.

Dajoz R. (1978) Précis d'écologie, Dunod, Paris.

Hall H.G. (1990) Parental analysis of introgressive hybridization between African and European honeybees using nuclear-DNA RFLPs, Genetics 1253, 611-621.

Hall T.A. (1999) BioEdit:a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp 41, 95-98.

Hao G., Yuan Y.M., et al. (2004) Molecular phylogeny of *Lysimachia* (Myrsinaceae) based on chloroplast trnL-F and nuclear ribosomal ITS sequences, Mol. Phylogenet. Evol. 311, 323-339.

Hijmans R.J., Cameron S.E., et al. (2005) Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol. 2515, 1965-1978.

Jaccard P. (1908) Nouvelles recherches sur la distribution florale, Bull. Soc. Vaud. Sci. Nat. 22, 223-270.

Kumar S., Nei M., et al. (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences, Brief. Bioinform. 94, 299-306.

Machado C.A., Robbins N., et al. (2005) Critical review of host specificity and its coevolutionary implications in the fig/fig-wasp mutualism, Proc. Natl. Acad. Sci. U. S. A. 102, 6558-6565.

Marmion M., Parviainen M., et al. (2009) Evaluation of consensus methods in predictive species distribution modelling, Divers. Distrib. 151, 59-69.

Michener C.D. (2000) The bees of the world, The Johns Hopkinds University Press, Baltimore.

Michez D., 2002. Monographie systématique, biogéographique et écologique des Melittidae (Hymenoptera, Apoidea) de l’Ancien Monde – Premières données et premières analysesFaculté Universitaire des Sciences Agronomiques de Gembloux, Gembloux.

Michez D., Nel A., et al. (2007) The oldest fossil of a melittid bee (Hymenoptera : Apiformes) from the early Eocene of Oise (France), Zool. J. Linn. Soc. 1504, 701-709.

Michez D., Patiny S. (2005) World revision of the oil-collecting bee genus *Macropis* Panzer 1809 (Hymenoptera : Apoidea : Melittidae) with a description of a new species from Laos, Ann. Soc. Entomol. Fr. 411, 15-28.

Nascimento V.A., Matusita S.H., et al. (2000) Evidence of hybridization between two species of *Melipona* bees, Genet. Mol. Biol. 231, 79-81.

Pekkarinen A., Berg O., et al. (2003) Distribution and co-existence of the *Macropis* species and their cleptoparasite *Epeoloides coecutiens* (Fabr.) in NW Europe (Hymenoptera : Apoidea, Melittidae and Apidae), Entomol. Fenn. 141, 53-59.

Pellmyr O. (2003) Yuccas, yucca moths, and coevolution: A review, Ann. Mo. Bot. Gard. 901, 35-55.

Rasmussen C., Olesen J.M. (2000) Oil flowers and oil-collecting bees, Scandinavian Association for Pollination Ecology honours Knut Fegri 39, 23-31.

Renner S.S., Schaefer H. (2010) The evolution and loss of oil-offering flowers: new insights from dated phylogenies for angiosperms and bees, Philos. Trans. R. Soc. B-Biol. Sci. 3651539, 423-435.

Ribble D.W. (1973) Hybridization between three species of andrenid bees, Ann. Entomol. Soc. Am. 666, 1281-1286.

Rinderer T.E., Stelzer J.A., et al. (1991) Hybridization between European and Africanized honey-bees in the neotropical Yucatan peninsula, Science 253, 309-311.

Sheffield C.S., Hebert P.D.N., et al. (2009) DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies, Mol. Ecol. Resour. 9, 196-207.

Simon C., Frati F., et al. (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers, Ann. Entomol. Soc. Am. 876, 651-701.

Simpson B.B., Neff J.L., et al. (1983) Floral biology and floral rewards of *Lysimachia* (Primulaceae), Am. Midl. Nat. 1102, 249-256.

Thompson J.D., Gibson T.J., et al. (1997) The clustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 24, 4876-4882.

Thuiller W., Lafourcade B., et al. (2009) BIOMOD - a platform for ensemble forecasting of species distributions, Ecography 323, 369-373.

Vogel S. (1976) *Lysimachia* : Ölblumen des Holartiks, Naturwissenschaften 63, 44-45.

Vogel S. (1986) Ölblumen and Ölsammelnde Bienen: Zweite Folge *Lysimachia* und *Macropis*, in: Verl. F.S. (Ed.), Akademie der Wissenschaftenund der Literatur Mainz, Stuttgart, Germany.

Vogel S. (1988) The oil-bee oil-flower relationships: parallelism and other aspects of their evolution in space and time, Z. Zool. Syst. Evol. 265, 341-362.

Vos P., Hogers R., et al. (1995) AFLP- a new technique for DNA-fingerprinting, Nucleic Acids Res. 2321, 4407-4414.

Warren D.L., Glor R.E., et al. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution, Evolution 6211, 2868-2883.

Tables

Table I: *Macropis* populations sampled in Europe, Anatolia and Caucasus, showing the detailed number of individuals collected per species, as well as the corresponding geographical data.

	***Macropis* species collected** (number of individuals)			**Geographical information** (Latitude, longitude: decimal degrees / Elevation: meters above see level)				
Population	***M. europaea***	***M. fulvipes***	***M. frivaldszkyi***	**Country**	**Place**	**Latitude**	**Longitude**	**Elevation**
ARI	2	7	-	Spain	Arija-Llano	42.991	-3.975	835
ASP	6	-	-	Spain	Aspurz	42.715	-1.158	560
BAG	-	2	-	Italy	Bagnasco	44.312	8.048	481
BAK	-	15	-	Georgia	Bakuriani	41.765	43.484	1552
BAL	8	-	-	Switzerland	Balgach	47.402	9.619	317
BAU	3	-	-	Belgium	Baudour	50.489	3.822	65
BAZ	7	-	-	France	Bazouges	47.680	-0.206	34
BDL	6	-	-	France	Bout du Lac (d'Annecy)	45.771	6.242	464
BEC	-	9	-	Spain	Becedas	40.403	-5.627	1094
BEG	10	-	-	Switzerland	Begnins	46.427	6.242	477
BOK	4	-	-	Sweden	Boekholm	56.403	13.600	113
BOL	7	-	-	Switzerland	Bolle di Magadino	46.161	8.863	200
BON	5	-	-	Switzerland	Bonfol	47.469	7.174	441
BOV	11	-	-	France	Boves	49.855	2.378	405
CAM	13	-	-	Spain	Sierra de Camero Nuevo	42.226	-2.627	793
CAP	12	-	-	France	Captieux	44.301	-0.255	95
CAR	7	-	-	Spain	Carrion de Los Condes	42.335	-4.617	820
CAS	7	-	-	Italy	Castelvecchio di Compito	43.777	10.628	2
CER	1	13	-	Spain	Cervera de Pisuerga	42.947	-4.492	1090
CHA	4	1	-	Switzerland	Chavanne-des-bois	46.313	6.138	459
CML	4	-	-	France	Camélas	42.629	2.684	310
CRU	7	-	-	Spain	Santa Cruz de Campezo	42.675	-2.361	558
CUE	7	-	-	Spain	Cuenca	40.192	-2.113	960
DAD	-	6	-	Turkey	Daday	41.481	33.603	1015
DAG	-	-	5	Turkey	Dağyolu	39.574	39.864	1296
DAM	3	-	-	Switzerland	Damphreux	47.472	7.112	425
DES	5	2	-	Serbia	Despotovac	44.042	21.537	311
DIO	5	-	-	Hungary	Diosjeno	47.932	19.062	207
DIP	6	1	-	Germany	Dippoldiswalde	50.919	13.681	432
DOM	1	5	-	Slovakia	Domasa Dobra	49.016	21.673	240
DZE	6	4	-	Latvia	Dzelmes	56.655	24.933	48
ECO	3	1	-	Switzerland	Ecogia	46.291	6.137	440
ERF	-	10	2	Turkey	Erfelek	41.859	34.744	682
FEL	-	5	-	Romania	Feleacu	46.701	23.590	670
FUN	7	-	-	Switzerland	Poutafontana-Uvrier	46.246	7.410	498
GJE	12	-	-	Slovenia	Gorni Jezero	45.727	14.408	566
GRA	8	-	-	Switzerland	Les Grangettes	46.395	6.901	376
GRAN	3	1	-	Switzerland	Granois	46.254	7.337	972
GRI	5	-	-	Switzerland	Grimisuat	46.256	7.392	795
HAJ	3	-	-	Hungary	Hajduhadhaz	47.695	21.655	184
HOD	7	-	-	Hungary	Hodmezovasarhely	46.356	20.209	64
HOL	3	-	-	Germany	Holvede	53.368	9.642	39
HOS	7	-	-	Czech Republic	Hostka	49.687	12.582	515
ILG	-	5	-	Turkey	Ilgaz Geçidi	41.140	34.064	1471
JON	3	-	-	Switzerland	Jongny	46.500	6.881	663
JUG	2	-	-	France	Jugon-les-Lacs	48.405	-2.325	26
KBA	9	-	-	Hungary	Kis-Balaton	46.660	17.126	123
KET	5	1	-	Poland	Kety	49.844	19.214	398
KHA	-	7	-	Georgia	Khashuri	41.999	43.656	689
KOZ	9	-	-	Macedonia	Kozjak	41.056	21.036	877
KRU	3	-	1	Serbia	Krupačko jezero	43.105	22.688	410
LAU	3	13	-	Latvia	Lauciene	57.178	22.692	119
LEM	3	-	-	Germany	Lembeck	51.723	6.995	46
LIP	1	-	-	Romania	Lipova	46.716	27.234	230
LJU	1	-	-	Sweden	Ljusfallshammar	58.802	15.452	54
LOC	6	-	-	Switzerland	Etang du Loclat	47.018	6.998	437
LOD	2	-	-	Switzerland	Loderio	46.379	8.973	348
LOZ	2	4	-	Spain	Rio Lozoya	40.902	-3.863	1125
LUG	2	-	-	Switzerland	Lugnez	47.486	7.115	438
LUN	-	11	-	Spain	Los Barrios de Luna	42.844	-5.862	1030

Table I: continued

	***Macropis* species collected** (number of individuals)			**Geographical information** (Latitude, longitude: decimal degrees / Elevation: meters above see level)				
Population	***M. europaea***	***M. fulvipes***	***M. frivaldszkyi***	**Country**	**Place**	**Latitude**	**Longitude**	**Elevation**
MAR	8	-	-	Denmark	Maribo	54.769	11.506	1
MARA	5	-	-	Switzerland	Marais de la Brocassière	46.225	7.323	495
MAV	2	-	-	France	Marais Vernier	49.421	0.533	0
MDV	5	-	-	Switzerland	Moulin de Vert	46.180	6.027	356
MIN	3	-	-	Spain	Rio Miño	41.974	-6.646	2
MOE	4	-	-	Switzerland	Prévessins-Moëns	46.254	6.057	464
MOM	-	1	-	Spain	Mombeltran	40.247	-5.018	530
MON	7	-	-	Switzerland	Lac de Montorge	46.234	7.338	645
PAS	-	1	-	Romania	Pasul Gutai	47.712	23.777	838
PIE	-	5	-	Spain	Piedrabuena	39.048	-4.239	553
PRE	2	2	-	Switzerland	Les Préhaies	47.285	7.429	620
PRE2	-	2	-	Switzerland	Les Préhaies	47.287	7.431	637
PRT	3	-	-	France	Portiragnes	43.302	3.332	3
PRW	4	-	-	Belgium	Péruwelz	50.508	3.589	30
PUE	14	-	-	Spain	La Puebla de Labarca	42.473	-2.593	399
RAS	2	3	-	Spain	Rascafria-Lozoya	40.910	-3.870	1133
REN	6	-	-	Spain	Renodo	42.504	-4.729	900
RIE	5	-	-	Italy	Rieti	42.510	12.753	373
RIL	6	-	-	Bulgaria	Rila	42.102	23.091	470
RIO	-	1	-	Spain	Rio Frio	40.324	-6.646	824
RUC	11	-	-	Austria	Ruckersdorf	46.591	14.567	493
SAL	5	-	-	Spain	Salamanque	40.957	-5.671	782
SAN	-	20	-	Spain	Lago di Sanabria	42.115	-6.734	1000
SCA	9	-	-	France	Etang de Scamandre	43.606	4.336	-3
SON	-	5	-	Spain	Soncillo-Herbosa	42.974	-3.805	825
SUC	-	17	-	Czech Republic	Suchy	49.486	16.762	664
TKI	-	1	-	Georgia	Tkibuli	42.377	43.037	1229
TRO	2	11	-	Germany	Trollblumenwiese	51.683	11.041	488
TUL	4	-	-	Bulgaria	Tulovo	42.574	25.576	318
VAU	-	6	-	Switzerland	Vauffelin	47.190	7.317	706
VIV	14	-	-	Italy	Viverone	45.441	8.006	239
VRH	9	-	-	Croatia	Vrh	45.334	13.920	1
WOL	5	-	-	Germany	Wolbach/Woerbeschwang	48.421	10.599	437
ZAO	19	-	-	Spain	Zaorejas	40.798	-2.155	943

Figures

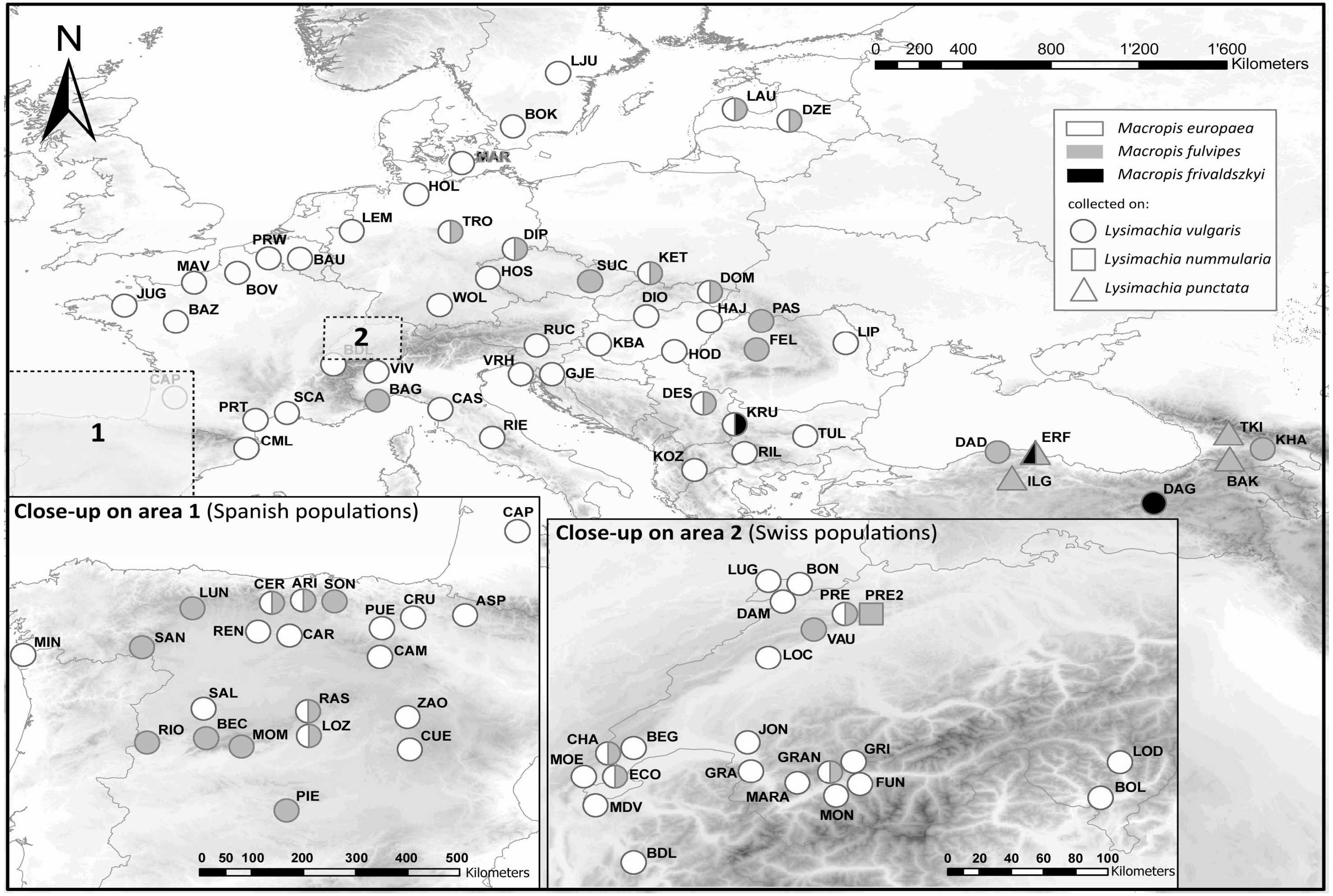


Figure 1: Map showing all *Macropis* populations sampled. Populations are displayed using different colors representing the three *Macropis* species, as well as different forms corresponding to three *Lysimachia* species on which the bees were collected.

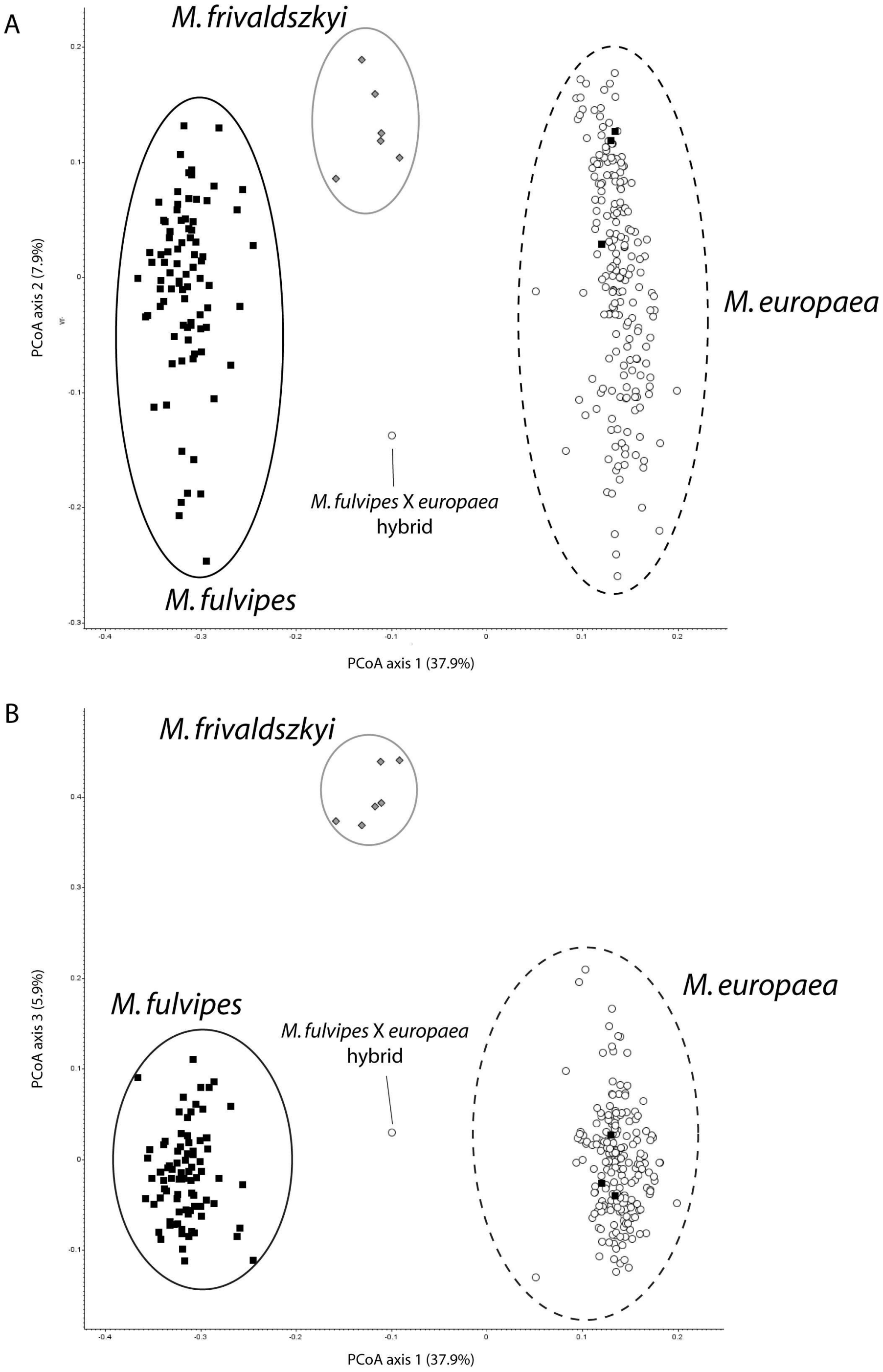


Figure 2: PCoA plot of all Macropis specimens genotyped, showing the potential hybrids detected. Each of the three phenotypic species is displayed on the first two axes (A) as well as on axes one and three (B).

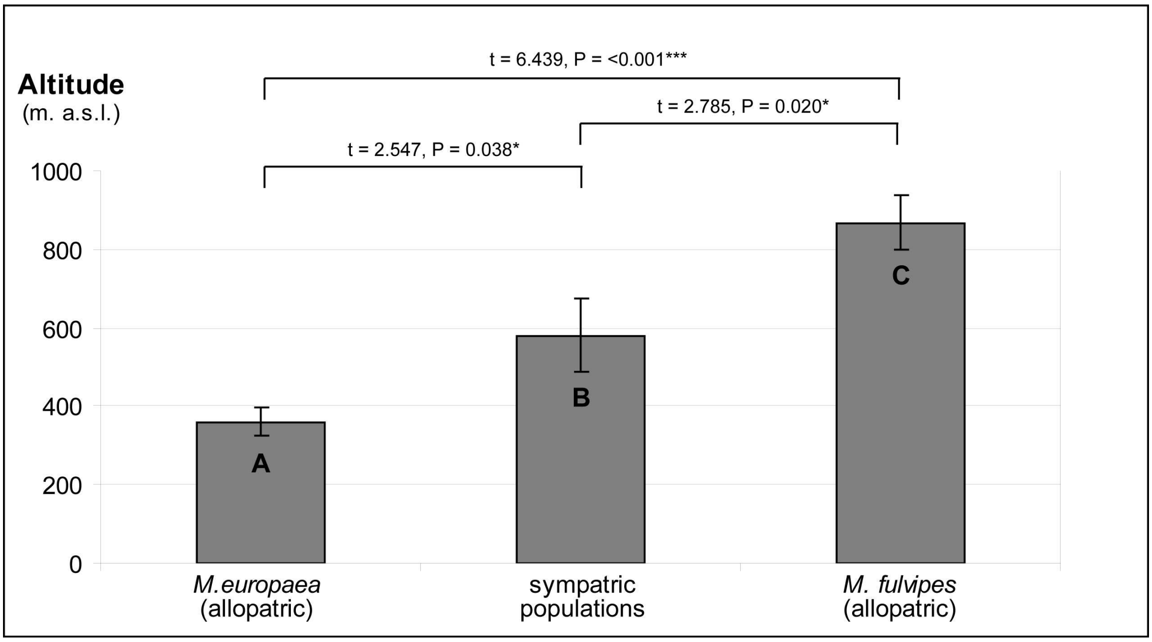


Figure 3: Mean sampling altitudes for the two bee species, either considering allopatric populations (left and right) or sympatric populations (centre). The results of the pairwise comparison (post-hoc Bonferroni t-test) are shown above for the corresponding pair of means compared.

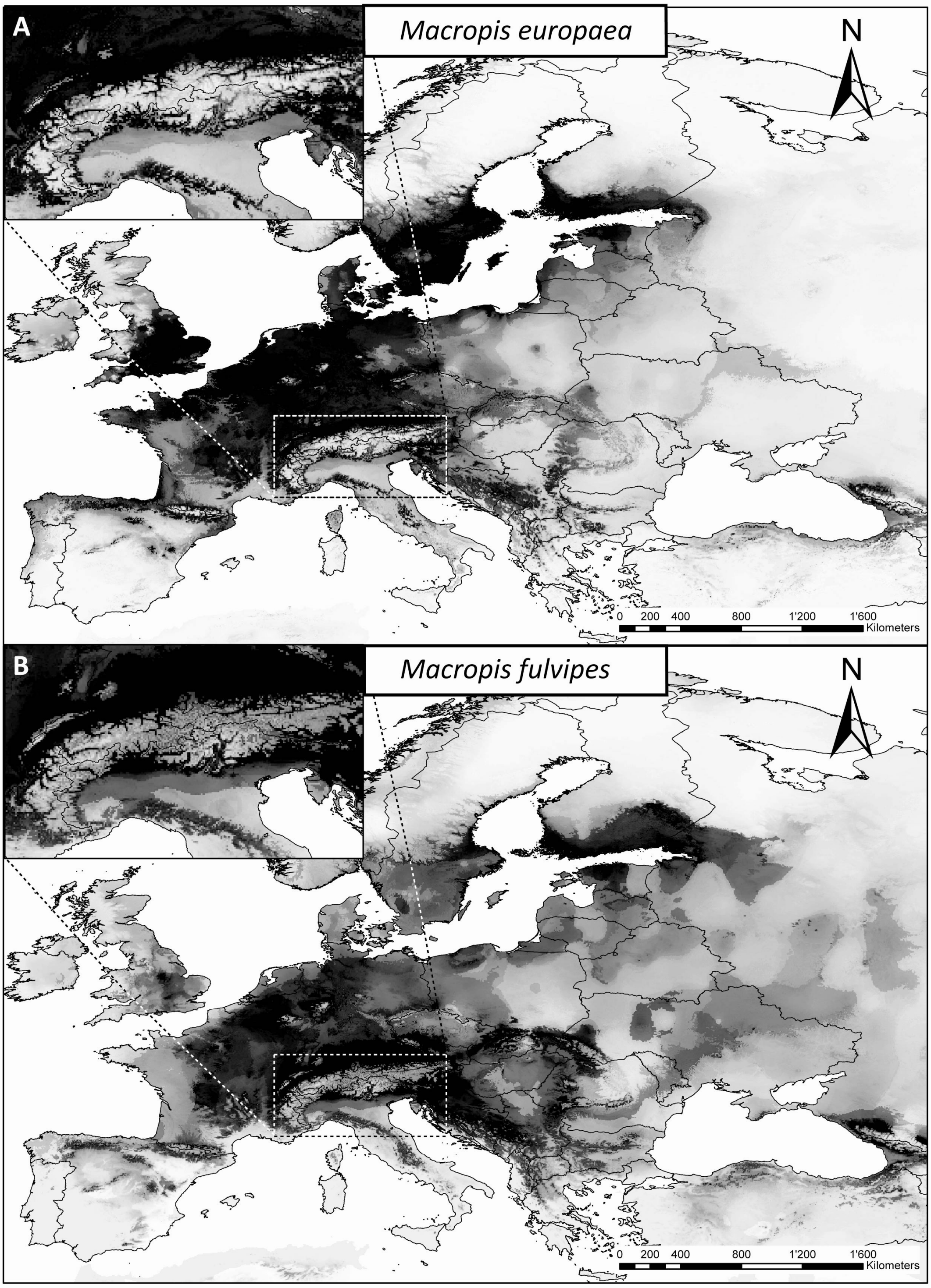


Figure 4: Modeling of the current potential European distributions of *M. europaea* (A) compared to *M. fulvipes* (B). A close-up on the Alpine region is provided for each species, in order to compare the potential distributions in altitude.

Electronic-only material

1. A descriptive pollination study of *Lysimachia nummularia* in western Switzerland

Method

The sampling of pollinators of *L. nummularia* was carried out during the months of June and July 2008 in ten different Swiss populations (see Table II and Fig. 5). In order to optimally describe associated pollinators, each population was visited at least twice during that period (a few sites were even visited three times), and on-site observations lasted between one and two hours. Observations were carried out by sunny weather at the hottest hours of the day (from 11h to 16h). If a pollination behavior was presumably observed, the insect was caught with a net and stocked in EtOH 70%. Determination of pollinators was done with the help of J.-P. Haenni for Diptera (Natural History Museum of Neuchâtel, Switzerland) and F. Amiet for Hymenoptera (Solothurn, Switzerland).

Table II: List of the ten Swiss *L. nummularia* populations studied, including the respective geographical and co-occurring *Lysimachia* species information.

Population	Localisation	Latitude (dec.degrees)	Longitude (dec.degrees)	Elevation (m. a.s.l.)	Sympatry
BRO	Broc	46.614	7.096	689	*L. vulgaris*
CHA	Champoz	47.248	7.287	814	*L. nemorum*
GRA	Les Grangettes	46.395	6.905	375	*L. vulgaris*
JON	Jongny	46.506	6.89	669	*L. vulgaris*
LOC	Loclat	47.017	6.996	437	*L. vulgaris*
MDV	Moulin de Vert	46.18	6.026	356	*L. vulgaris*
POR	Porrentruy	47.417	7.062	459	*L. vulgaris*
PRE2	Les Préhaies	47.287	7.43	637	-
VAU	Vauffelin	47.19	7.318	703	*L. vulgaris*
VIL	Villars	47.029	6.915	691	-

Results and discussion

Our observations resulted in pollination records only in three sites (VIL, CHA and PRE2) as no pollinators were observed in the other populations. The insects observed can be classified into two main categories: hoverflies (Diptera, Syrphidae) and wild bees. As the sampled hoverflies belonged to generalist and common species (*Melanostoma scalare* Fabricius 1794, *Platycheirus albimanus* Fabricius 1781, *Episyrphus balteatus* De Geer 1776 and *S. vitripennis* Meigen 1822) they probably did not account for very effective pollinators, performing only short and superficial visits on the flowers. Wild bees pollinating *L. nummularia* in the present study belonged to two different genera: *Lasioglossum* Curtis 1833 (Hymenoptera, Halictidae) and *Macropis* Panzer 1809 (Hymenoptera, Melittidae).

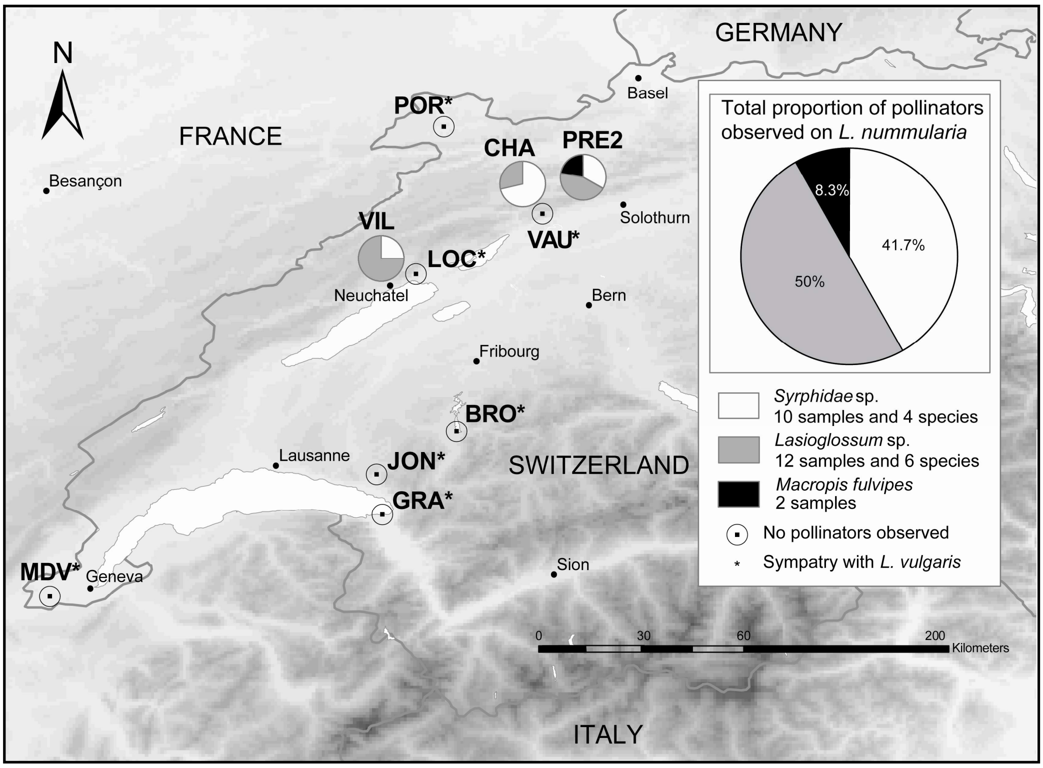


Figure 5: Map of the *L. nummularia* populations studied, with the relative proportion and number for each type of pollinators

Lasioglossum was dominantly observed, with 12 individuals belonging to the six following species: *L. morio* FABRICIUS 1793 (the most frequent with six individuals captured), *L. lineare* SCHENCK 1869 (two individuals), *L. calceatum* SCOPOLI 1763, *L. fulvicorne* KIRBY 1802, *L. puncticolle* MORAVITZ 1872 and finally *L. rufitarse* ZETTERSTEDT 1838 (one single individual per species). Typical pollination behavior for *Lasioglossum* consisted in taking some pollen on back legs with regular movements towards the bottom of the flower. These are not the first observations of *Lasioglossum* species on *L. nummularia* (e.g. Teppner, 2005), and some species observed here (e.g. *L. morio*, *L. calceatum* and *L. rufitarse*) apparently pollinate a wide range of plants (e.g. from other orchids such as *Cypripedium calceolus*, to the tomato *Solanum lycopersicum*; Teppner, 2005; Erneberg and Holm, 1999; Bittrich and Kaderheit, 1988). Consequently, *L. nummularia* should be unspecifically pollinated by species of *Lasioglossum* living near-at-hand. Finally we note that bees from this genus were also observed visiting other *Lysimachia* species, such as *L. vulgaris* (in numerous European populations, from Spain to Turkey (L. Bassin and Y. Triponez pers. obs.) and even *L. nemorum* (e.g. in population PRE; L. Bassin pers. obs.). The visits of these bees to *L. nummularia* are however not as frequent as in the relationship between *Macropis* and *L. vulgaris*.

Two individuals of the oil-collecting bee *Macropis fulvipes* FABRICIUS 1804 (previously reported as regular pollinator of *L. nummularia*; e.g Simpson and Neff, 1983; Vogel, 1986) were also caught, although in only one population (PRE2). Even if not strictly sympatric, we should note that this population stood less than 200 m. from a population of *L. vulgaris* (where *M. fulvipes* was also captured). In comparison, it is interesting to see that at the sympatric site VAU, *M. fulvipes* was caught only on *L. vulgaris*, neglecting nearby plants of *L. nummularia*. According to these observations *M. fulvipes* rarely visits *L. nummularia* and seems to give priority to *L. vulgaris* when available. Consequently, the assumed preference of *M. fulvipes* for *L. nummularia* (e.g. Michez, 2002) should be reconsidered, at least in Switzerland.

Our results add knowledge to the reproductive biology of *L. nummularia*. At first, the plant does not clearly depend on insects for its reproduction. Indeed, on most sites no pollinators were observed and, if pollinators were present, their visits seemed very sporadic. Asexual reproduction is therefore probably more frequent in *L. nummularia* than in other European *Lysimachia* species, although the seedset and quality of descendants is probably lower than with cross-pollinated sexual reproduction (Simpson & Neff, 1983; Bittrich & Kaderheit, 1988; Batygina, 2005; Hoffman, 2005). Moreover, according to our observations, pollinators could play a role in maintaining populations of *L. nummularia* fit, as populations without pollinators were often unhealthy and composed of a few individuals. Finally, our study confirms that, even if different native insects visit the flowers, small bees and more specifically specimens from the genus *Lasioglossum* should be the more effective vectors of *L. nummularia* pollen.

References

Batygina T. (2005) Sexual and asexual processes in reproductive systems, Acta Biol. Cracov. Bot. 47, 51-60.

Bittrich V., Kaderheit J. (1988) Cytogenetical and geographical aspects of sterility in *Lysimachia nummularia*, Nord. J. Bot. 4, 325-328.

Erneberg M, Holm B. (1999) Bee size and pollen transfer in *Cypripedium calceolus* (Orchidaceae), Nord. J. Bot. 19, 363-367.

Hoffmann F. (2005) Biodiversity and pollination. Flowering plants and flower-visiting insects in agricultural and semi-natural landscapes. University of Groningen.

Michez D. (2002) Monographie systématique, biogéographique et écologique des Melittidae (Hymenoptera, Apoidea) de l'Ancien Monde – Premières données et premières analyses. Faculté Universitaire des Sciences Agronomiques de Gembloux.

Simpson B.B., Neff J.L., et al. (1983) Floral biology and floral rewards of *Lysimachia* (Primulaceae), Am. Midl. Nat. 1102, 249-256.

Teppner H. (2005) Pollinators of tomato, *Solanum lycopersicum* (Solanaceae), in central Europe, Phyton-Ann. Rei Bot. A. 45, 217-235.

Vogel S. (1986) Ölblumen and Ölsammelnde Bienen: Zweite Folge *Lysimachia* und *Macropis*, in: Verl. F.S. (Ed.), Akademie der Wissenschaftenund der Literatur Mainz, Stuttgart, Germany.

2. Detailed methods for the present distribution modelling of *Macropis*

We extracted occurrences of the species in Europe from the GBIF database (www.gbif.org). We kept only the data with a minimal spatial accuracy of 5km. These data were pooled with the occurrences collected during fieldwork. Since the occurrences were highly aggregated in some parts of Europe, we randomly selected a subset of occurrences with a minimal distance of 50km. Most modeling techniques require not solely information about presences but also absences to determine the suitable conditions for a given species, thus we selected 10’000 pseudo-absences randomly. The modeling techniques discriminate the conditions for presences and absences, based on the background environment (Witz and Guisan, 2009). GBIF is known to be highly biased for western European countries (Yesson et al., 2007). Therefore, we did not select pseudo-absences in Eastern Europe countries but limited them to the Western and Balkanic countries. Selecting pseudo-absences in unsampled countries can significantly bias the response curves of the models (Phillips et al., 2009). The resulting presences and pseudo-absences were used in the further species distribution modeling.

We ran single-models with the seven following climatic variables from Worldclim (Hijmans et al., 2005) at a resolution of 2.5 arc-minute (around 5 km): maximum temperature of warmest month (bio5), minimum temperature of coldest month (bio6), mean temperature of warmest quarter (bio10), mean temperature of coldest quarter (bio11), precipitation seasonality (bio15), precipitation of the warmest quarter (bio18) and precipitation of the coldest quarter (bio19).

We modeled the distribution of the species using the BIOMOD package (Thuiller et al., 2009), as implemented in R CRAN (REFERENCE). We used eight different niche-based modeling techniques : (1) generalized linear model (GLM), (2) generalized additive model (GAM), (3) classification tree analysis (CTA), a classification method that selects the best trade-off between the number of leaves of the tree and the explained deviance, (4) artificial neural networks (ANN), a machine learning method, with the mean of three runs used to provide predictions and projections, as each simulation gives slightly different results, (5) mixture discriminant analysis (MDA), a classification method that uses MARS function for the regression part of the model, (6) multivariate adaptive regression splines (MARS), (7) generalized boosting model (GBM), a machine learning method which combines a boosting algorithm and a regression tree algorithm to construct an 'ensemble' of trees, and (8) Random Forest (RF), a machine learning method that is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest.

In order to evaluate the predictive performance of the species distribution model, we used a random subset of 70% of the data to calibrate every model, and used the remaining 30% for the evaluation. Models were evaluated using a relative operating characteristic (ROC) curve and the Area Under the Curve (AUC) (Fielding and Bell, 1997). We replicated the data splitting ten times and calculated the average AUC of the repeated split-sample, which gives a more robust estimate of the predictive performance of each model.

Finally, each model was projected into current climate conditions with the Worldclim data of the CCSM circulation model for the last glacial maximum (LGM; -21’000 years) downscaled at a resolution of 2.5 arc-minute. In order to obtain the central tendency of these distributions, accounting for variations among modeling techniques, we applied a weighted average of the eight modeling techniques based on the predictive power (AUC). Ensemble forecasting approaches have been shown to significantly improve the accuracy of species distribution models (Marmion et al., 2009).

References

Fielding A.H., Bell J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models, Environmental Conservation 24, 38-49.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A. (2005) Very high resolution interpolated climate surfaces for global land areas, International Journal of Climatology 25, 1965-1978.

Marmion M., Parviainen M., Luoto M., Heikkinen R.K., Thuiller W. (2009) Evaluation of consensus methods in predictive species distribution modeling, Diversity and Distributions 15, 59-69.

Phillips S.J., Dudik M., Elith J., Graham C., Lehmann A., Leathwick J., & Ferrier S. (2009) Sample selection bias and presence-only models of species distributions: implications for selection of background and pseudo-absences, Ecological Applications 19, 181-197.

Thuiller W., Lafourcade B., Engler R., Araújo M.B. (2009) BIOMOD - a platform for ensemble forecasting of species distributions, Ecography 32, 369-373.

Wisz M.S., Guisan A. (2009). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data, BMC Ecology 9:8.

Yesson C., Brewer P.W., Sutton T., Caithness N., Pahwa J.S., Burgess M., Gray W.A., White R.J., Jones A.C., Bisby F.A., Culham A. (2007) How Global Is the Global Biodiversity Information Facility?, PLoS ONE 2, e1124.

3. Niches’ overlap for *M. europaea* and *M. fulvipes*

Method

In order to compare the niche of the two *Macropis* species, we extracted the climatic information for the occurrences from the bioclimatic variables of Worldclim (Hijmans et al., 2005). We selected several variables important for the ecology of the bees, in particular reflecting conditions of temperature and precipitation: bio5, bio6, bio10, bio11, bio18, bio19. Also, we calculated two additional layers from Worldclim data: degree-day (DDEG) and averaged moisture index (MIND) during the year. Recently, tools have been developed to compare the niche of species in different part of its range or to compare closely related species to investigate niche conservatism (Broennimann et al., submitted). Climatic niches of the two species, *M. europaea* and *M. fulvipes* were compared in a gridded climatic space using a principal component analysis (PCA). First we run the PCA on the climatic environment available represented by 30’000 pixels randomly selected throughout Europe. Each pixel in the environmental space corresponds to a unique set of environmental conditions present at one or more sites in the geographical space. The occurrences of the groups of populations were then projected in the climatic space available. Finally, we applied a kernel density function to determine the "smoothed" density of occurrences of each pixel in the environmental space for each bee species. We divided the density of occurrence by the density of the environment in each focal pixel to obtain a measure of the density of the species relative to the availability of climate. This ratio is finally rescaled between 0 and 1. This approach has been shown with virtual species to be very robust to compare climatic niche of species (Broennimann et al., submitted). The similarity between niches was afterwards measured using the Schoener’s *D* metric and its associated statistical tests (tests of niche conservatism), proposed originally by Warren et al. (2008). This metric, which ranges from 0 (no niche overlap) to 1 (complete overlap). In order to test if the niches of the groups are significantly different, occurrences are pooled and randomly split, maintaining the number of occurrences as in the original datasets. The niche overlap statistics (D, see Warren et al. 2008) are then recalculated. This process is repeated 100 times and a histogram of simulated values is constructed. The comparison of observed and simulated values of the test statistic allows significance testing for niche equivalency. If the observed value falls within the density of 95% of the simulated values, the null hypothesis of niche equivalency cannot be rejected. We also applied a niche similarity test in the way of Warren et al. (2008). Rejection of the null hypothesis indicates that the niche models of two species are more similar (or different) than would be expected by chance. Rejection of the null also indicates that the observed niche differentiation between species is a function of habitat selection and/or suitability rather than simply an artifact of the underlying environmental differences between the suite of habitats available to the two species (Peterson et al. 1999; Warren et al. 2008).

Results and discussion

Here we can see that the niche of *M. fulvipes* is slightly larger than that of *M. europaea* (represented by a larger surface of non-overlapping blue than of non-overlapping red of Fig. 6). According to the Worldclim variables, it can be deduced that *M. fulvipes* is found within habitats with colder and drier winters, with hotter summers and with more precipitations (typical conditions of the continental area). On the opposite, *M. europaea* depends on higher winter temperatures as well as on more winter precipitations, typical conditions of the oceanic climate.

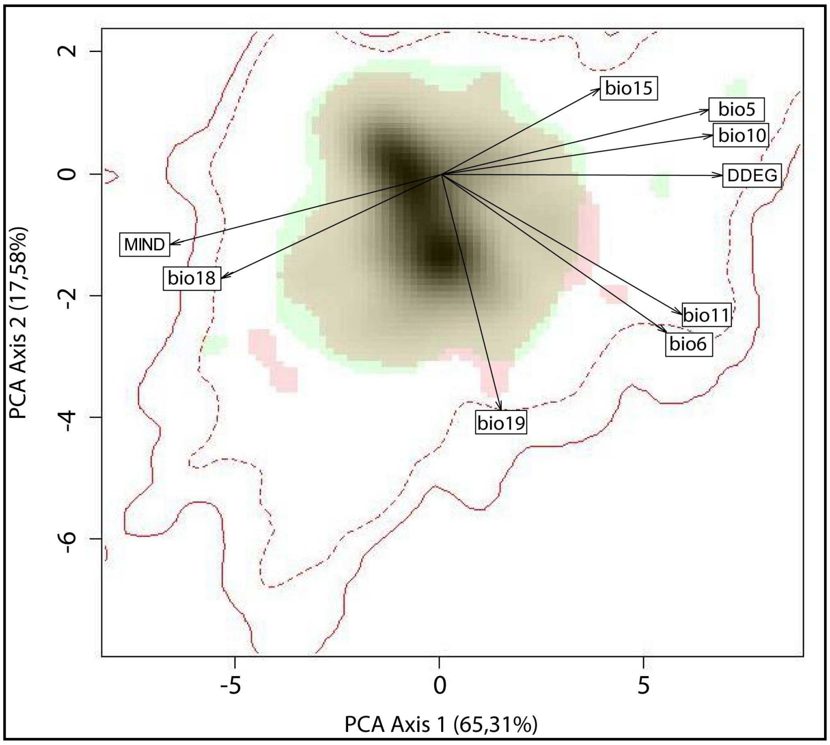


Figure 6: Respective ecological niches of *M. europaea* (in red) and *M. fulvipes* (in blue) displayed on the two first coordinates of the environmental space delimited by the climatic variables of Worldclim. Niche overlap is represented in brown.

References

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, 25, 1965-1978.

Peterson, A.T., Soberon, J. & Sanchez-Cordero, V. (1999). Conservatism of ecological niches in evolutionary time. *Science*, 285,1265-1267.

Warren, D.L., Glor, R.E. & Turelli, M. (2008). Environmental niche equivalency versus conservatism : quantitative approaches to niche evolution. *Evolution*, 62, 2868–2883.

APPENDIX 2

Genetic status of *Lysimachia punctata* and *L. verticillaris*

Yann Triponez and Nadir Alvarez

Neuchâtel, August 2010

Genetic status of *Lysimachia punctata* and *L. verticillaris*

Yann Triponez and Nadir Alvarez

Method

In order to define the genetic status of the two morphologically closely-related species, the Dotted Loosestrife *Lysimachia punctata* L. (Myrsinaceae) and the Whorled Loosestrife *L. verticillaris* SPRENG (also described as *L. verticillata* (GREENE) HAND.-MAZZ., *L. verticillata* M. BIEB or *L. punctata* var. *verticillata* (BIEB.) KLATT), we collected individuals across their native area in mountainous regions surrounding the Black Sea (Ferguson 1972), as well as from cultivated sites in northern Europe (Fig. 1). The 11 wild populations sampled (with population codes in parentheses) were found in the following countries: Bulgaria (VIT, BEK), Georgia (BAK, TKI), Greece (KON, KRA), Ukraine/Crimea (KRS, PER) and Turkey (ERF, ILG, YUC). In addition, the cultivated samples came from gardens situated in Norway (NOR) and Scotland (BAR). One leave per plant was cut and dried in silica-gel (Chase and Hills 1991).

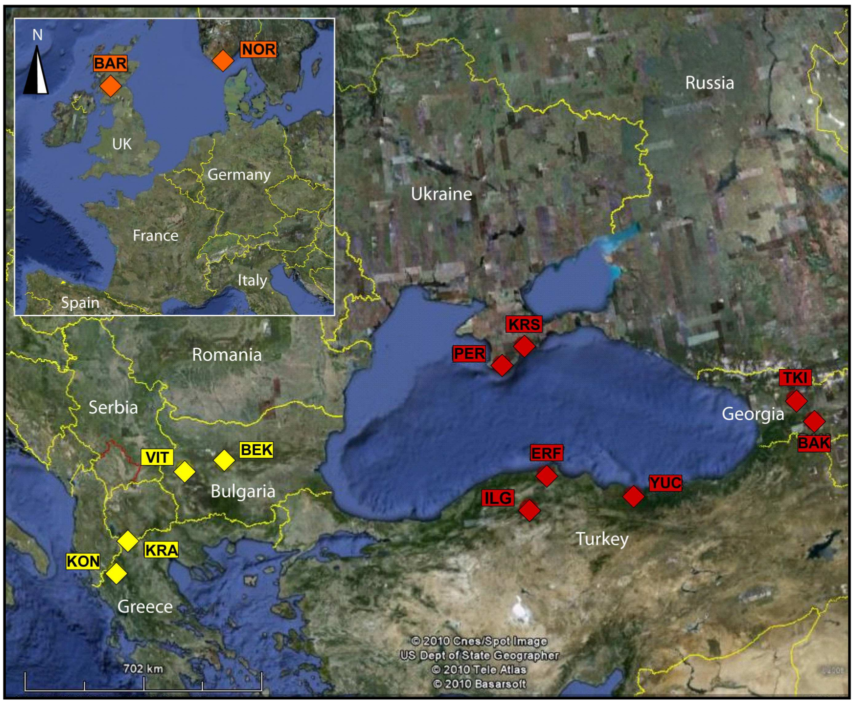


Figure 1: Sampling of *L. punctata* and *L. verticillaris* across its native area (main map), also showing the location of the cultivated samples collected (smaller map).

DNA was extracted with the DNeasy Plant Kit (Qiagen, Hilden, Germany). AFLP analysis was performed with the two primer combinations *Eco*RI-ACA/*Mse*I-CTA and *Eco*RI-ATA/*Mse*I-CAC (method modified from Vos et al. 1995). The detailed protocols of digestion, ligation, pre-amplification and selective amplification are available by the first author. Final selective PCR products were analysed using the GeneScan technology with an automated capillary-sequencer (ABI 3730XL, Applied Biosystems, Foster City, CA; the service was provided by Macrogen Inc., Seoul, South Korea). Resulting fluorescent AFLP patterns were scored using GeneMapper 3.7 (Applied Biosystems). The final binary matrix was analyzed by a principal coordinate analysis (PCoA) relying on the Jaccard similarity coefficient using the software Ginkgo 1.5.8 (Bouxin, 2005).

Results and discussion

This study was mainly based on confusion in the literature concerning *L. punctata* L., *L. verticillaris* SPRENGEL, *L. verticillata* (GREENE) HAND.-MAZZ.,and *L. verticillata* M. BIEB. For instance, Huxley (1992) considered *L. verticillata* as very similar to *L. punctata*. Currently all species seem to be accepted as synonymous as reported in the Flora Europaea (Ferguson 1972) or by the International Plant Names Index (IPNI), except *L. verticillata* (GREENE) HAND.-MAZZ. for the latter, considered as a correct species. Leblebici (1978) also described *L. punctata* and *L. verticillaris* as distinct (i) primarily by the difference in their petiole length, (ii) but also considering their non-overlapping ranges, *L. punctata* occurring in Europe and Western Turkey while *L. verticillaris* is found in northern and eastern Turkey, Caucasus, Crimea and Northern Iran. More recently, McAllister (1999) confirmed Leblebici's observations and completed the list of morphological differences between *L. punctata* and *L. verticillaris*. He also reported that both species could be found naturalized or cultivated in the British Isles.

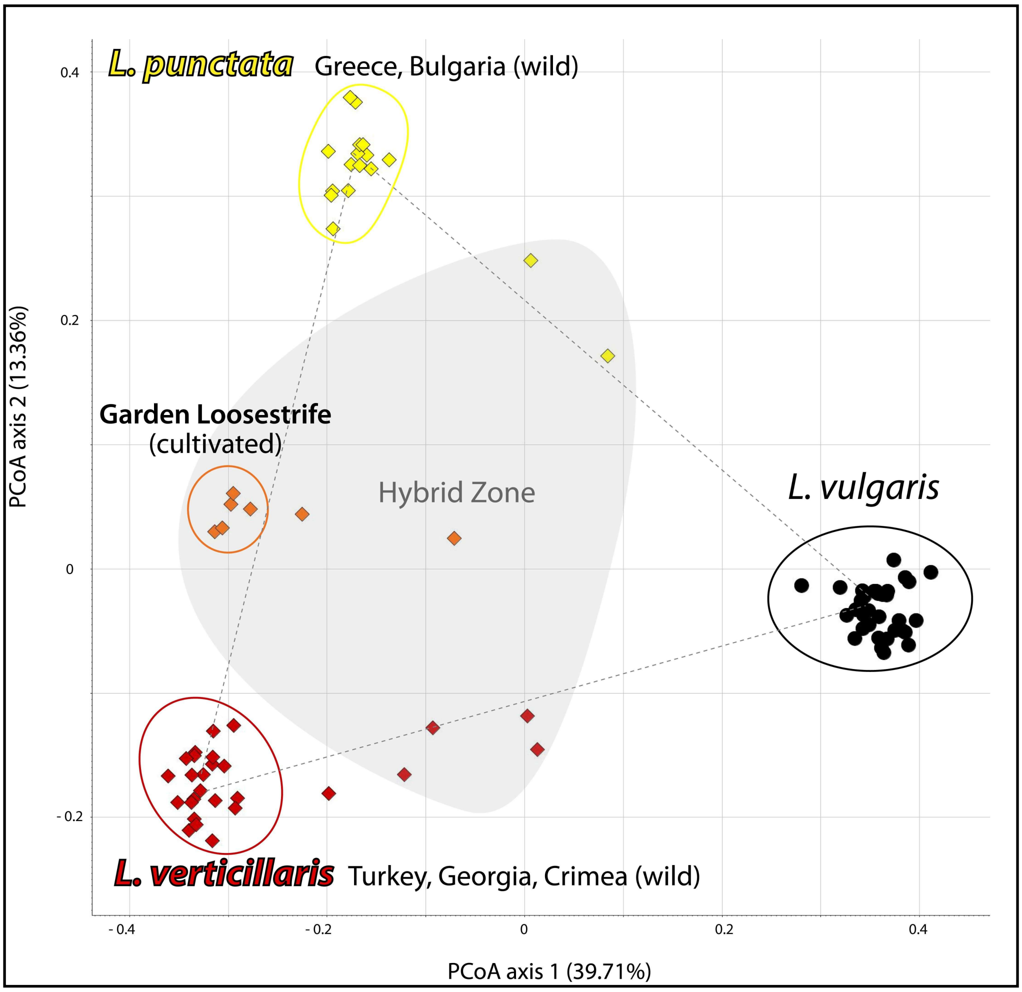


Figure 2: PCoA plot of all analyzed samples.

The current study brings the genetic confirmation to Leblebici's conclusion. Indeed, our AFLP results (Fig. 2) revealed a strong genetic divergence between *L. punctata* individuals from southern Balkans (Greek and Bulgarian populations, in yellow on Fig. 2) and *L. verticillaris* samples from northern, eastern and southern regions around the Black Sea (Crimean, Georgian and Turkish populations, in red on Fig. 2). When comparing these two scatterplots to the position of European samples from the widely distributed species *L. vulgaris* L. (considered as the sister species to *L. punctata* in a recent molecular phylogeny; Manns & Anderberg 2005) collected in Spain, Germany or Georgia (see Chapter IV), it becomes evident that such a divergence in *L. punctata* could not simply be due to spatial genetic variation at the intra-specific level: recent historical processes such as isolation in different refugia during Pleistocene glaciations are not likely to explain such a high level of divergence, and lineages divergence might be as ancient as the Pliocene.

Our study addresses two additional results regarding species hybridization. First, it is interesting to see that the *L. punctata* cultivated samples (in orange on Fig. 2) are located exactly halfway between the wild *L. punctata* and *L. verticillaris*. This typical hybrid position might reveal a recent history of artificial crossings in this plant widely used in horticulture and proves that Garden Loosestrife cultivars might generally represent hybrids between these two species. Secondly, some individuals showing intermediate positions may confirm that hybridization is possible between the widely spread *L. vulgaris* and *L. punctata*, *L. verticillaris* or even the cultivars. Here, natural hybridization could take place via cross-pollination mediated by shared oligolectic pollinators such as the mutualistic oil-collecting bees *Macropis* (Hymenoptera: Melittidae) or by more generalist pollinators such as bees from the genus *Lasioglossum* (Hymenoptera: Halictidae), both observed pollinating these *Lysimachia* (Triponez et al. in prep, Bassin et al. in prep).

In view of the current genetic results confirming Leblebici (1978) observations and conclusions, we propose that the name *L. punctata* L. should be kept only for populations within the European distribution area, as the species *L. verticillaris* SPRENGEL must be definitively used concerning the plants found in Anatolia, Caucasus or Crimea. The sampling of supplementary populations from western Anatolia would permit to define where the contact zone between the two species is situated and if natural hybrids could be found. Finally, a molecular dating based on chloroplastic sequences would be useful in order to define the spatio-temporal processes involved in the divergence between *L. punctata* and *L. verticillaris*.

References

Bouxin G. (2005) Ginkgo, a multivariate analysis package. J. Veg. Sci. 163, 355-359.

Chase M.W. and Hills H.H. (1991) Silica-gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* 40, 215-220.

Ferguson L.F. (1972). *Lysimachia* L. in Tutin T.G. et al. eds. Flora Europaea 3:26-21. Cambridge University Press, Cambridge.

Huxley A. et al. eds. (1992). The new Royal Horticultural Society Dictionary of gardening. Royal Horticultural Society, London.

Leblebici E. (1978) *Lysimachia* L. in Devis P.H. ed. Flora of Turkey 6:135-138. Edinburgh University Press, Edinburgh.

McAllister H.A. (1999) *Lysimachia punctata* L. and *L. verticillaris* Sprengel (Primulaceae) naturalised in the British Isles. Watsonia 22, 279-281.

Manns U. and Anderberg A.A. (2005) Molecular phylogeny of *Anagallis* (Myrsinaceae) based on ITS, *trn*L-F and *ndh*F sequence data. Int. J. Plant. Sci. 166, 1019-1028.

Vos P., Hogers R., Bleeker M., Reijans M., Vandelee T., Hornes M., et al. (1995) AFLP- a new technique for DNA-fingerprinting. Nucleic Acids Res. 23, 4407-441.

Yann TRIPONEZ, Biologist

Pierre-à-Mazel 54
2000 Neuchâtel
Switzerland

+41 32 718 31 62 (lab)
+41 78 825 02 25 (mobile)
+41 32 724 76 64 (home)
yann.triponez@unine.ch

Age: 29 (09.05.1981)
Nationality: Swiss
Marital status: single

LANGUAGES

French: mother tongue
English: very good skills
German: good skills
Italian: good skills
Serbo-Croatian: beginner

INTERESTS / LEISURE

Nature: biological conservation, general entomology and botany, ecology,
Sports: ski, trekking, volleyball, beach-volley, swimming
Other: Travel, landscapes and macro-photography, music

EDUCATION

April 2006 - May 2010:	**PhD thesis** *"Multi-species phylogeographic investigations in closely-related taxa and mutualistic plant-insect systems"*; University of Neuchâtel (Switzerland), Laboratory of Evolutionary Entomology (E-VOL) http://www2.unine.ch/entomo
October 2000 – March 2006:	**MSc in Behaviour, Ecology and Evolution** "*Evolutionary history of the Oreina speciosa / O. alpestris species complex: an approach based on phylogeography, chemotaxonomy and ecology*"; University of Neuchâtel, E-VOL laboratory.
1995-2000:	**High School Certificate** (Type B: Latin-English), Lycée-Collège de la Planta, Sion (Switzerland)
1987-1995:	Primary & Secondary school, Centre Scolaire de Crans-Montana, Crans-Montana (Switzerland)

COMPUTATIONAL SKILLS

- Microsoft Windows Office (2003, 2007)
- Adobe Illustrator, Photoshop, InDesign
- ArcView GIS (ESRI)
- Statistics (SigmaStat, SPSS, Gingko)
- Populations genetics (Structure, GenAlEx, AFLPsurv)
- Phylogenetics (BioEdit, MEGA, PAUP, MrBayes)
- Molecular markers analysis (PeakScanner, GeneMapper)

PROFESSIONAL EXPERIENCE

April 2006 – May 2010: **Assistant** at the E-VOL Laboratory, University of Neuchâtel

- Lectures and practical work about insect systematics and evolutionary ecology for 2nd and 3rd year students
- Teaching of Insects Biology class (Bachelor of Biology)
- Supervision of student's Master thesis works
- Planning and coordination of fieldwork missions in Europe
- Molecular biology laboratory work
- Writing and presenting research results in international journals and conference

2004 – 2006: **Volontary assistant,** University of Neuchâtel

- Help for practical work about insect systematics for 2nd year students

May – October 2004: **Internship** at the National Centre of Competence in Research (NCCR, University of Neuchâtel) with Jérôme Moreau and Betty Benrey

- Help for field work and laboratory experiments on the grapevine pest species *Lobesia botrana*

September – October 2003: **Internship** on tropical ecology and botany at the Masoala National Park (Madagascar), with Prof. Philippe Küpfer (University of Neuchâtel) and in collaboration with the Botanical and Zoological Park of Tsimbazaza (PBZT) in Antananarivo:

- Inventory and description of plant species
- Inventory and monitoring of butterfly species

Several high-season periods (2002-2005): **Ski instructor** at the Swiss Ski School, Crans-Montana, Switzerland

July – August 2002: **Work at the Alpine Garden** of "la Chanousia" in Italy:

- public reception and guided presentation in the garden
- collecting and determining plants for the garden
- gardening

REFERENCES

Prof. **Martine Rahier**
Rectorat
Faubourg du Lac 5a
CH - 2000 Neuchâtel
Tél. : +41 (0)32 718 10 25
Fax : +41 (0)32 718 10 21
martine.rahier@unine.ch

Dr. **Nadir Alvarez**
University of Lausanne
DEE, UNIL Sorge, Biophore
CH – 1015 Lausanne
Phone: +41 (0)21 692 42 47
Fax: +41 (0)21 692 42 65
nadir.alvarez@unil.ch

Dr. **Bertrand Schatz**
CNRS Montpellier / CEFE
UMR 5175
1919 Route de Mende
F-34293 Montpellier cedex 5
Phone : +33 (0)4 67 61 32 62
Fax: +33 (0)4 67 41 21 38
bertrand.schatz@cefe.cnrs.fr

CONFERENCES & SYMPOSIA

10 February 2010: **Plant and Animal Conservation Ecology (PACE10 Symposium)**,
Neuchâtel, SWITZERLAND
Oral presentation: *“Classical species delimitation among the Fly Orchid group does not allow predicting patterns of genetic diversity”*

13-16 October 2009: **DIVERSITAS Open Science Conference 2**,
Cape Town, SOUTH AFRICA
Poster presentation: “Lineage *boundaries among the Fly Orchid group (*Ophrys insectifera *s.l.): consequences of an intricate evolutionary history on conservation strategies*”

6-7 November 2008: **La Biodiversité en Valais – Bilan et Perspectives**,
Sion, SWITZERLAND

30 January –
1st February 2008: **Plant Species Concepts and Evolution**,
Neuchâtel, SWITZERLAND

15-16 February 2007: **Biology 07**,
Zürich, SWITZERLAND

13-14 September 2006: **Conflicts of Interest in Mutualistic Interactions**,
Neuchâtel, SWITZERLAND

PUBLICATIONS

Triponez Y., Espíndola A., Bassin L., Pellissier L., Alvarez N. Large Biogeography of an oil-collecting mutualism: tight interactions involve shared refugia but independent dispersal of partner species *(in prep)*

Triponez Y., Schatz B., Pellissier L., Alvarez N. Large-scale spatial genetic structure of the fly orchid (*Ophrys insectifera* group) reveals endemic lineages from different origins *(in prep)*

Bassin L., Alvarez N., Pellissier L., Triponez Y. Large Ecological niche overlap in sister species: how do oil-collecting bees *Macropis europaea* and *M. fulvipes* (Hymenoptera: Melittidae) avoid competition and hybridization? *(submitted to Apidologie)*

Triponez Y., Schatz B., Alvarez N. Hidden and visible diversity in the fly orchid (*Ophrys insectifera*): respective contributions in defining evolutionary significant units *(submitted to Conservation Biology)*

Triponez Y., Naisbit R.E., Rahier M., Alvarez N. Discordances between phylogenetic and morphological patterns of alpine leaf beetles attest for intricate history of lineages in postglacial Europe *(encouraged to resubmission in Molecular Ecology, resubmitted)*

Triponez Y., Naisbit R.E., Jean-Denis J.B., Rahier M., Alvarez N. 2007. Genetic and environmental sources of variation in the autogenous chemical defence of a leaf beetle. *Journal of Chemical Ecology* **33** (11): 2011-2027

D'Alessandro M., Held M., Triponez Y., Turlings T.C.J. 2006. The role of indole and other shikimic acid derived maize volatiles in the attraction of two parasitic wasps. *Journal of Chemical Ecology* **32** (12): 2733-2748