UNIVERSITÀ DEGLI STUDI DELL’AQUILA

DIPARTIMENTO DI MEDICINA CLINICA,
SANITÀ PUBBLICA, SCIENZE DELLA
VITA E DELL’AMBIENTE

DOTTORATO DI RICERCA IN SCIENZE DELLA SALUTE E DELL’AMBIENTE

XXXIV ciclo

Curriculum

Scienze Ambientali

Titolo della tesi

Phylogeography of Italian endemic Flea Beetles (Chrysomelidae, Galerucinae, Alticini): biogeographic, systematic, and conservation implications

SSD BIO/05 – Zoologia

Dottorando	Tutor
Dr. Emanuele Berrilli	**Prof. Daniele Salvi**

Coordinatore del corso

Chiar.ma Prof.ssa M. Grazia Cifone

A.A. 2021/2022

SUMMARY

RIASSUNTO 7

ABSTRACT 10

GENERAL INTRODUCTION 13

- **Quaternary climatic oscillation in Europe: effect on biodiversity** 13
- **Glacial/interglacial cycles as triggers for isolation, diversification and build-up of genetic diversity patterns** 14
- **Different effect in different ecosystem: temperate vs alpine fauna** 14
- **Importance of Mediterranean peninsulas as glacial refugia** 15
- **The Italian Peninsula, a Mediterranean hot spot** 16
- **Gap in biodiversity knowledge: invertebrates and their importance for the study of biogeographic and evolutionary processes** 16
- **Alticini as a study model** 17
- **Alticini biodiversity in Europe and Italy** 18
- **Thesis aims and objectives** 19
- **References** 20

Part I 26

SPECIES SAMPLING, MORPHOLOGICAL AND MOLECULAR IDENTIFICATION 26

CHAPTER 1 27

Sampling 27

- **Distribution of Italian Alticini and identification of target species** 27
- ***Longitarsus laureolae*** 29
- ***Longitarsus zangherii*** 31
- ***Psylliodes biondii*** 32
- ***Psylliodes ruffoi*** 33
- ***Psylliodes springeri*** 34
- **Technique of sampling** 35
- **Problems of species identification, phenology, and rarity: the needs of a widespread and repeated sampling** 35
- **Sampling design and sampling results** 36
- **References** 39

CHAPTER 2 40

Integrative taxonomy and systematics of target species 40

- **Morphological techniques** 40
- **Molecular techniques and assessment** 40
- **Molecular and morphological mismatch in *Longitarsus* and *Psylliodes*: improving molecular and morphological tool for the systematics of target species** 41

References **42**
CHAPTER 3 .... **43**
Case study I **43**
Sharpening the DNA barcoding tool through a posteriori taxonomic validation: the case of *Longitarsus* flea beetles (Coleoptera: Chrysomelidae) **43**
  Introduction **43**
  Materials and methods **45**
    Sample collection and morphological identification 45
    DNA extraction, amplification, and sequencing 48
    Reference sequence dataset building 48
    Sequences' taxonomy assessment analyses 48
  Result **50**
    Original dataset 50
    Optimal Threshold Dataset 53
    Final dataset 54
  Discussion **56**
  Reference **59**
  Supporting information **64**
CHAPTER 4 .... **66**
Case study II **66**
Cryptic, sibling, or neither of the two? Systematics and taxonomic identification of *Psylliodes ruffoi and P. kiesenwetteri* (Chrysomelidae, Galerucinae, Alticini) based on a combined morphological and phylogeographic assessment **66**
  Introduction **66**
  Materials and methods **68**
    Study System 68
    Sampling 69
    DNA Extraction, Amplification, and Sequencing 69
    Molecular Species Delimitation 70
    Dataset and Sample preparation for morphological analysis 71
    Morphological assessment 71
    Phylogenetic analyses 71
  Results **72**
    Molecular species delimitation and preliminary morphological identification 72
    Morphometric results and new morphological diagnostic characters 73
    Phylogenetic relationship 77
  Discussion **77**

References **80**

CHAPTER 5 .... **84**

Case study III **84**

A phylogeographic and morphological assessment of the *Psylliodes springeri* species complex (Chrysomelidae, Alticini) uncovers new endemic diversity from the Majella mountains (southern Apennine) **84**

Introduction **84**

Materials and Methods **85**

Sample collection and morphological identification 85

DNA extraction 85

Species Delimitation 86

Phylogenetic analyses 86

Results **86**

Phylogenetic and species delimitation analysis 86

Discussion **90**

References **92**

Part II **96**

PHYLOGEOGRAPHY OF ITALIAN ENDEMIC ALTICINI **96**

CHAPTER 6 .... **97**

Refugia within refugia – assessing the link between Late Quaternary climate stability and spatial patterns in endemism and genetic divergence in the Iberian **97**

Introduction **97**

Materials and Methods **98**

Sample collection and morphological identification 98

DNA extraction and amplification 102

DNA sequence data analysis 102

Results **103**

Longitarsus laureolae 103

Longitarsus zangherii 105

Psylliodes ruffoi 107

Longitarsus springeri 108

Psylliodes biondii 111

Discussion **113**

References **116**

Supplementary **120**

CHAPTER 7 .... **125**

Effect of Quaternary climatic oscillations on temperate and alpine biotas: insights from comparative phylogeography and historical demography of Italian endemic Alticini along the Apennine mountains **125**

Introduction **125**

Materials and Methods **126**

Sample collection and morphological identification 126

DNA extraction 126

Results **128**

Longitarsus laureolae 128

Longitarsus zangherii 129

Psylliodes ruffoi 130

Longitarsus springeri 131

Psylliodes biondii 132

Discussion **132**

References **135**

CONCLUSION **138**

References **140**

ACKNOWLEDGEMENTS **141**

RIASSUNTO

Il clima è tra i principali fattori che influenza la distribuzione degli ecosistemi sul pianeta. Le variabili climatiche (es. temperatura, umidità, vento, precipitazioni) di una data regione, per lunghi periodi di tempo, rappresentano importanti pressioni selettive per l'evoluzione delle specie vegetali e animali. Negli ultimi due milioni di anni il clima sulla Terra ha subito fluttuazioni periodiche, risultate nelle ere glaciali. Le oscillazioni climatiche attraverso molteplici cicli glaciale/interglaciale ha avuto un ruolo chiave nella strutturazione della biodiversità attuale e nel plasmare la distribuzione delle specie e la loro diversità genetica, sia su scala regionale che globale. Le penisole mediterranee, balcanica, iberica e italiana, hanno agito da rifugi durante le glaciazioni pleistoceniche, fornendo habitat idonei e stabili per la sopravvivenza delle specie nei periodi climaticamente sfavorevoli. Diversi studi hanno documentato la presenza di molteplici rifugi all'interno della penisola italiana durante i periodi climatici del Pleistocene, che hanno reso questa penisola un hotspot regionale di biodiversità all'interno dell'hotspot del Mediterraneo. Tuttavia, la maggior parte delle conoscenze che abbiamo sui modelli filogeografici delle specie endemiche italiane provengono da studi che hanno utilizzato vertebrati come specie modello, sebbene rappresentino solo una piccola parte della biodiversità globale della specie (<1,5%). Ciò rende la nostra conoscenza della risposta ecologica delle specie ai cambiamenti climatici del passato tutt'altro che completa.

Nel mio progetto di dottorato, ho utilizzato il gruppo molto diversificato degli Alticini (Coleoptera Chrysomelidae, Galerucinae) endemici dell'Appennino centrale per studiare i processi ecologici, evolutivi e biogeografici innescati dai cicli delle glaciazioni quaternarie sui biota temperati e d'alta quota. Gli Alticini mostrano range di distribuzione da micro a macro scala, essendo presenti nella maggior parte degli ecosistemi montani, e hanno esigenze ecologiche molto selettive e ben note, rendendoli così ideali modelli per gli studi filogeografici. L'obiettivo principale del progetto di tesi è studiare i processi storici che hanno determinato la distribuzione della diversità genetica intraspecifica in specie endemiche dell'Appennino centrale e confrontare le risposte ecologiche di specie associate ad ambienti temperati e di alta quota ai cambiamento climatici del Quaternario. Per raggiungere questi obiettivi principali, ho adottato un approccio multi-taxa comprendente (i) un approccio tassonomico integrativo per delimitare le unità filogenetiche e tassonomiche e (ii) un approccio filogeografico multilocus per identificare i principali rifugi pleistocenici all'interno del sistema appenninico ed inferire i processi demografici e biogeografici che hanno determinato la struttura di popolazione e la distribuzione della diversità genetica nelle specie studiate. Sono state selezionate sei specie target in base alla loro distribuzione ed ecologia: *Longitarsus springeri*, *Psylliodes biondii* e *Psylliodes springeri* legate ad ambienti di alta montagna; *Longitarsus laureolae* e *Longitarsus zangherii* legati ad ambienti temperati e *Psylliodes ruffoi* legata ad ambienti xerofili. Sono state eseguite tre principali campagne di raccolta per campionare le popolazioni delle specie target.

Nella prima parte della tesi vengono presentati i risultati di tre studi di delimitazione morfologica e molecolare integrativa delle specie target, che sono stati necessari per delimitare le unità evolutive e

tassonomiche per lo studio filogeografico. In questi studi sono state evidenziate e risolte incongruenze tra identificazione molecolare e morfologica in tre principali taxa. In primo luogo, all'interno del genere *Longitarsus*, abbiamo osservato diversi problemi di identificazione errata delle specie e incongruenze tassonomiche all'interno dei database pubblici di sequenze di DNA (GenBank e BOLD). Ciò rende difficile implementare un'identificazione molecolare delle specie confrontando la sequenza di DNA di un gene standard (*cox1*) di un individuo con le sequenze depositate in un database pubblico attraverso un approccio di DNA barcoding. Per ovviare a questo problema, abbiamo implementato una *pipeline* metodologica con revisione tassonomica a posteriori che consente di migliorare la qualità del database di sequenze di riferimento (da GenBank e BOLD) e l'efficacia dello strumento di DNA barcoding. Nel secondo caso di studio, abbiamo osservato che alcuni campioni, morfologicamente identificati come *P. ruffoi*, avevano caratteri molecolari (cioè sequenze di DNA) tipici della specie affine *P. kiesenwetteri*. L'utilizzo di un approccio tassonomico integrativo e filogeografico ha permesso di delimitare le due specie *P. ruffoi* e *P. kiesenwetteri* da un punto di vista molecolare e morfologico, chiarendone la relazioni filogenetiche e tassonomiche. Nel terzo caso di studio, attraverso un approccio tassonomico integrativo, abbiamo identificato le popolazioni della Maiella di *P. biondii* come unità evolutiva distinta dalle restanti popolazioni appenniniche, verosimilmente una nuova specie criptica.

Nella seconda parte della tesi è stato condotto uno studio filogeografico comparativo sulle specie target di Alticini, rappresentative di specie temperate e criofile, endemiche dell'Appennino, utilizzando un set di dati multilocus (marcatori genetici nucleari e mitocondriali). Le significative differenze in termini di struttura genetica di popolazione osservate nelle specie target, suggeriscono una risposta individuale di ciascuna di queste ai cambiamenti climatici pleistocenici, con marcate differenze in termini di distribuzione geografica e tempo di divergenza dei lignaggi intra-specifici. I risultati indicano inoltre un effetto maggiore dell'ultimo periodo glaciale sulle specie temperate rispetto alle specie legate agli ambienti mesofili. Al contrario, le specie criofile hanno subito un processo di isolamento in aree montane idonee che corrisponde al cosiddetto modello delle *sky island*. In generale, per tutte le specie si può osservare che, indipendentemente dal tempo di origine dei principali lignaggi intraspecifici, in ogni complesso montuoso è presente un distinto set di aplogruppi, indicando che ogni catena montuosa sia servita da rifugio indipendente durante i cambiamenti climatici pleistocenici. Nell’ultimo capitolo della tesi, i risultati delle ricostruzioni della demografia storica attraverso EBSP corroborano il modello di risposta individualistica delle specie alle glaciazioni passate. Alcune specie, come *L. laureolae* e *L. springeri*, mostrano trend demografici storici spiegabili con processi di contrazione glaciale ed espansione postglaciale tipici delle specie temperate. Al contrario, altre specie hanno mostrato andamenti diversi, che non possono essere spiegati dai classici modelli di espansione/contrazione, e che sono probabilmente derivati da processi specie-specifici che si sono verificati a scala regionale.

I risultati di questa tesi hanno permesso di documentare i modelli filogeografici ed i processi evolutivi che hanno determinato la marcata struttura genetica e l’elevata diversità della fauna appenninica. È stato mostrato come la sinergia tra eterogeneità geografica ed ecologica di questa regione e le risposte ecologiche individuali delle specie ai cambiamenti climatici pleistocenici, abbia determinato distinti processi evolutivi e demografici negli endemiti appenninici, anche in specie con simili esigenze ecologiche e che attualmente vivono in stretta sintopia.

I risultati di questa tesi contribuiscono alla conoscenza delle dinamiche evolutive e demografici di una componente sotto studiata della fauna endemica appenninica, permettendo di comprendere meglio i processi che hanno generato l'hotspot di diversità e il biota di questa regione. Questi risultati sono cruciali per la pianificazione di strategie per la conservazione e gestione a lungo termine della biodiversità appenninica, specialmente nel contesto dei cambiamenti climatici in atto.

ABSTRACT

The climate is one of the main factors that influence the distribution of ecosystems on the planet. Climatic variables (e.g., temperature, humidity, wind, precipitation) of a given region over extended periods of time represent important selective agents for the evolution of plants and animal species. Over the past two million years the climate on Earth has undergone periodic fluctuations, resulting most notably in the Ice Ages. The climate oscillation through multiple glacial/interglacial cycles have played a key role in structuring current biodiversity patterns and shaping the distribution of species and their genetic diversity both at the regional and global scale. The Mediterranean peninsulas, the Balkan, the Iberian and the Italian peninsula, acted as refugia during Pleistocene glaciations, providing suitable and stable habitats for the survival of species during climatically unfavorable periods. Within the Italian Peninsula several studies documented the establishment of multiple refugia during Pleistocene climatic periods, making Italy a regional biodiversity hotspot within the Mediterranean global biodiversity hotspot. However, most of the knowledge we have on phylogeographic patterns of Italian endemic species come from studies that have used vertebrates as model species, although they represent only a small part of the global biodiversity of the species (<1.5%). This makes our appraisal of the ecological response of species to past climatic changes far from complete.

In my PhD project, I used the highly diverse group of flea beetles endemic of the central Apennines (Coleoptera Chrysomelidae, Galerucinae, Alticini), to investigate the ecological, evolutionary and biogeographic processes primed by Quaternary glaciations cycles on temperate and high-altitude biota. These insects show distribution patterns from micro to macro scale, being present in most mountain ecosystems, and have restrictive and well-known ecological requirements, thus making them ideal phylogeographic models. The main aim of the thesis project is to study the historical processes that have shaped current patterns of genetic diversity within different endemic Alticni species of Apennines and to compare how species with different ecological requirements have responded to Quaternary climate change. To achieve these main aim, I have adopted a multi-taxa approach including (i) an integrative taxonomic approach to delimit the phylogenetic and taxonomic units, and (ii) a multilocus phylogeographic approach to identify the main Pleistocene refugia within the Apennine system, and to infer demographic and biogeographic processes underlying the formation of the observed intraspecific diversity patterns. Six target species were selected based on their distribution and ecology: *Longitarsus springeri*, *Psylliodes biondii* and *Psylliodes springeri* linked to high mountain environments; *Longitarsus laureolae* and *Longitarsus zangherii* linked to temperate environments and *Psylliodes ruffoi* linked to xerophilous environments. Three large sampling campaigns were performed to sample the populations of target species.

A first integrative morphological and molecular delimitation of the target species was done as a fundamental and preliminary step to the phylogeographic study. In this framework, three main issues related to the mismatch between molecular and morphological identification were pointed out and resolved. First, within the genus *Longitarsus*, we observed several instances of species misidentification and taxonomic

incongruences within the public databases of DNA sequences (GenBank and BOLD). This makes difficult to implement a molecular identification of species by comparing the DNA sequence of a target gene (*cox1*) of a sampled individual with deposited sequences in public database through a DNA barcoding approach. To overcome this drawback, we implemented a methodical pipeline with *a posteriori* taxonomic revision that allow improving the quality of the reference-sequence library (from GenBank and BOLD) and the effectiveness of the DNA barcoding tool. Second, we observed that some samples, morphological identified as *P. ruffoi*, had molecular characters (i.e. DNA sequences) typical of the related species *P. kiesenwetteri*. The use of an integrative taxonomy approach within a phylogeographic framework allowed to delimit the two species *P. ruffoi* and *P. kiesenwetteri* from a molecular and morphological point of view, clarifying their phylogenetic and taxonomic relationship. Third, following the molecular assessment of the Maiella population of *P. biondii*, we observed that this population show a remarkable genetic distance relative to other populations of *P. biondii*, suggesting they are possibly two distinct species. Thanks to an integrative taxonomy approach, we were able to solve these taxonomic and systematic problems, to identify a new lineage (possibly a new species) endemic to the Maiella mountains, and thus to identify the evolutionary units on which to deepen the phylogeographical study.

In the second part of the thesis, a comparative phylogeographic study was carried on the target species representing temperate and non-temperate species of Alticini endemic to the Apennine, using a multilocus dataset (nuclear and mitochondrial genetic markers). The observed remarkable differences in population structure between the target species support an idiosyncratic pattern underlying individualistic response of the species to past climatic changes, with marked differences in terms of geographic distribution and time of divergence of intra-specific lineages. We also observed a greater effect of the last glacial period on temperate species compared to species related to mesophilic environments. On the contrary, cryophilic species have undergone a process of isolation in suitable mountain areas that closely correspond to the so-called sky islands model. In general, for all studied species it can be observed that, regardless of the time of origin of the main intraspecific lineages, each mountain chain has its own private haplogroups, suggesting that each of these mountain system acted as a separate refugium during Pleistocene climatic oscillations. In the last chapter of the thesis, I showed that also the reconstruction of historical demographics through EBSP corroborates the model of individualistic response to past glaciations. Some species, such as *L. laureolae* and *L. springeri*, seem to present demographic histories that can be explained by the glacial contraction and postglacial expansion typical of temperate species. Conversely, other species have shown different trends, which cannot be explained by the classical expansion and contraction processes, and which are probably derived from species-specific processes occurring at a smaller (regional) scale.

The results of this thesis have allowed us to document the phylogeographic patterns and the underlying evolutionary processes that have determined the high genetic structure and the diversity of the Apennine fauna. It has been shown how the interplay between the geographical and ecological heterogeneity of this region and the species-specific ecological response to past climatic changes has determined distinct evolutionary and demographic processes in Apennine endemics, even in species with similar ecological requirements and that currently live in close syntopy.

This thesis adds a significant piece of knowledge on the evolutionary and demographic dynamics of understudied fauna endemic to the Italian peninsula. This allowed us to improve our understanding of the processes that generated a biodiversity hotspot within this region and that determined the assembly of its biota. These results are crucial for planning informed strategies for the long term conservation and management of the Apennine biodiversity, especially within the context of ongoing climate change.

GENERAL INTRODUCTION

Quaternary climatic oscillation in Europe: effect on biodiversity

The current species distribution is the combined result of present and past processes. During the Quaternary, the current and most recent of the three periods of the Cenozoic era (from 2.588 ± 0.005 million years ago to today) (Cohen et al., 2013), in particular, major climatic events had a greatest effect on the distribution of species and ecosystems on Earth and therefore on the establishment of current geographic patterns of biodiversity. The Arctic ice cover stabilized about 2.4 million years ago, during the early Quaternary. From then until 0.9 million years ago, the polar ice caps had cycles of approximately 41,000 years of expansion and contraction. From 0.9 million years ago to today, cyclical periods of about 100,000 years have followed one another with increasingly dramatic events. The Croll-Milankovitch theory proposes that the regular periodicity of the ice age cycles is governed by changes in the Earth’s orbit around the Sun (Bennett and Bennett, 1997; Williams et al., 1998). The principal orbital eccentricity, the variation of the axial inclination, and the precession due to the Earth’s axial oscillation, modify the insolation and consequently the energy it receives. The synergistic effect of orbital changes and ocean current flows has led to significant climate changes (Webb et al., 1997; Williams et al., 1998). The Quaternary period was the stage of these major climatic oscillations, with longer and colder glacial periods alternating with shorter and warmer interglacial periods. During glacial periods, the polar ice caps expanded and compressed warmer environments towards the Equator (Williams et al., 1998) and large mountain blocks around the world undergone a considerable glaciation. In addition to this ‘direct’ climatic effect on the displacement of temperate and alpine habitats, glaciations also caused an indirect effect on the geography of emerged landforms and therefore of terrestrial and coastal habitats. Indeed, the large volume of accumulated ice reduced global sea levels by up to 120 m during glacial maxima. This has led to the creation of land bridges between islands and continental landmasses in different parts of the world. Although these major climatic changes have not had the same effect across the planet, they have certainly strongly influenced the response of ecosystems and individual species (Hewitt, 2000).

The orographic complexity of Europe has resulted in a complex and eventful biogeographical history during the Quaternary. First of all, three large Mediterranean peninsulas in the southern Europe, namely the Iberian, Italian and Balkan peninsulas played a key role as refugia of temperate biotas during glacial phases (Taberlet et al., 1998; Hewitt, 1999; Habel et al., 2005) allowing the persistence of populations of temperate species that become extinct at higher latitudes as a consequence of the glaciations. These climate changes have imposed a cyclical contraction-expansion of land and habitats, resulting in a parallel contraction and expansion of the species' ranges in a pattern known as the "expansion-contraction" (EC) pattern (Provan & Bennet 2008). During these range changes, Mediterranean Sea forms an important southern barrier, limiting the southward movement of the biota during cold periods. Furthermore, the east-west orientation of the main mountain ranges of the Alps and the Pyrenees has acted as a more or less important barrier to species expansion (Taberlet et al., 1998). To get an idea of the effect of glaciations on this continent during the peak of the last glacial maximum, around 20,000 years ago, the Scandinavian ice sheet covered much of Great Britain and northern Europe. Permafrost, tundra and steppe have spread as far as southern France, from the Alps to the Black Sea (Frenzel, 1992; Hewitt, 2003). The main mountain ranges of the continent, such as Alps, Pyrenees, Balkans and Apennines, have had significant glaciations. Moreover, the accumulation of terrestrial ice has reduced the Mediterranean sea level by up to 150 m (Denton and Hughes, 1981; Tooley, 1993). This has caused several islands or archipelagos in the Mediterranean to expand their size or connect with the mainland.

Glacial/interglacial cycles as triggers for isolation, diversification and build-up of genetic diversity patterns

Knowledge of the biogeographical patterns of the Quaternary period in Europe was initially based on fossil data, in particular on the extensive network of pollen nuclei, which allowed to describe the changes in the species distribution during the ice ages (Huntley and Birks, 1983; Bennett and Bennett, 1997; Hewitt, 1999). The introduction of molecular techniques and the birth of phylogeography, have enabled an ever greater understanding of the evolutionary processes and patterns that occurred at a specific and intraspecific level during the Quaternary (Avise, 2000; Hewitt, 2000, 2004). Specifically, during this period, each species has undergone cycles of range expansion–contractions following the sudden and cyclic climatic changes. These strong climatic oscillations between glacial and interglacial phases have shaped the distribution of many species and consequently have been an important mechanism behind the processes of divergence and speciation (Hewitt, 1996; Roy, 1997; Hewitt, 2000, 2003, 2004). Habitat fragmentation and displacement (Richard, 1991; Elias, 1996; Davis and Shaw, 2001) have led to multiple opportunities for isolation and divergence of populations, with intense interruptions of gene flow based on environmental heterogeneity and the presence of geographical barriers. This has resulted in improved formation of intraspecific genetic diversity driven by origination of lineage in isolated refugia and their subsequent mixing due to secondary contact (Canestrelli et al., 2010; Feliner, 2011; Vega et al., 2010) and, in some cases, to species divergence during the Pleistocene (Hewitt, 2000; Carstens and Knowles, 2007).

Different effect in different ecosystem: temperate vs alpine fauna

Glacial cycles had affected range and demography of species in different ways, depending on their ecological requirements. Cold-adapted species tend to contract their range in the northern and mountainous areas during warm interglacials, while they experience range expansions towards the south and into the lowlands during glacial periods. This spatial contraction-expansion were also accompanied by demographic (i.e. population size) contraction-expansion. In contrast, species adapted to warm climates will exhibit range and demographic expansion during interglacial periods and underwent to demographic contraction of populations that will retreat southward and into lowlands during glacial phases (Hewitt, 1996, 2004; Provan and Bennett, 2008; Schmitt, 2017). However, although these patterns have been found on several species (Taberlet et al., 1998), there are further evidence for more complex scenarios. As for the alpine taxa, linked to the alpine climate that occurs at high altitudes and above the tree line, many are endemic and have a low dispersal capacity. For example, many high mountain insects have lost their ability to fly (Hodkinson, 2005; Margraf et al., 2007; Schmitt, 2009) and many plants present limited seed dispersion (Morgan and Venn, 2017), which greatly reduces their spread ability. This made implausible a pattern of large areal changes during the different phases of the glacial cycles. Furthermore, it is not certain that the conditions now present in the high altitude environments are the same as those found in the lowlands during the glacial maximums; indeed, many species now living in the high mountains have not undergone significant range expansions at low altitude in the cold periods precisely due to several adverse environmental factors in these environments (Schmitt et al., 2006). Finally, although the glacial periods lasted very long, the diversity of species and geographic structure do not seem to show a pattern that reflects an extensive mixture of Alpine organisms during this period (Lohse et al., 2011). Regarding the temperate species, adapted to warm climates, mounting evidences show that the classical scenarios of glacial contraction and postglacial expansion does not always explain the observed pattern and empirical data. In particular, some temperate species of Mediterranean coastal areas have maintained stable populations or have even seen to expand their range during the glacial periods, because of the synergic effect of greater environmental availability due to glacial marine regression (i.e. lowering of sea levels) and the formation of mild coastal microclimates due to the mitigating action of the sea (Canestrelli et al., 2007; Canestrelli and Nascetti, 2008; Salvi et al., 2014). Thus, depending on the species and regional context, the expansion-contraction model can be more complex than previously thought because of additional climate-linked environmental changes due to sea-level (and shorelines) changes.

Importance of Mediterranean peninsulas as glacial refugia

The Mediterranean peninsulas, the Iberian, the Italian and the Balkan peninsula, thanks to their arrangement and the heterogeneity of their territories, served as refugia for most of the European flora and fauna (Hewitt, 1996, 1999, 2000). Even if the term refugium has been used in different ways, creating some confusion (Provan and Bennett, 2008), the glacial refugium is an area where environmental suitability for temperate species was maintained through quaternary climatic cycles. Therefore, within these refugia ancient genetic diversity and distinct genetic lineages have persisted during unfavorable periods of the Pleistocene (Médail and Diadema, 2009; Feliner, 2011). Although it has been established that each species have had a distinct ecological response to Pleistocene climate changes, the importance of the Mediterranean peninsulas as reservoirs of biodiversity is undoubted (Weiss and Ferrand, 2007; Hewitt, 2011). Indeed, northern Europe hosts a lower genetic variety than southern Europe in terms of number of species and genetic diversity, and this has been called the “southern richness and northern purity” pattern by Hewitt (Hewitt, 1996, 1999, 2000).

Moreover, given the complex topography of southern European peninsulas (with the presence of large mountain groups alternating with valleys and plains, and the long coastline often close to archipelagos), multiple processes of allopatric divergence between populations isolated took place. A scenario called “refugia within refugia” i.e. the persistence of smaller areas within the peninsular refugia where a large number of lineages remained during the subsequent Pleistocene climatic oscillation (Hewitt, 1996; Gómez and Lunt, 2007; Provan and Bennett, 2008). This explains the great intraspecific genetic diversity and richness of endemic species that have been observed in Mediterranean refugia (Canestrelli et al., 2010; Canestrelli et al., 2012; Cresti et al., 2019; Gonçalves et al., 2009; Maura et al., 2014; Médail and Diadema, 2009; Razgour et al., 2015; Salvi et al., 2013; Vega et al., 2010).

The Italian Peninsula, a Mediterranean hot spot

Italy represents one of the most important hotspots of plant and animal biodiversity within the Mediterranean global biodiversity hotspot (Myers et al., 2000; Zachos and Habel, 2011). Its geographical position in the center of the Mediterranean, its close connection to the two Iberian and Balkan peninsulas, the simultaneous presence of considerable mountain ranges and latitudinal extent, has determined a wide variety of environments and climatic contexts. About half of the plant species and about a third of all animal species currently present in Europe are found in Italy. According to the Italian Check List, drawn up by the Ministry of the Environment, Italy hosts 57,468 animal species, of which 4,777 (8.6%) are endemic and nearly 12,000 plant species of which about 6,700 (13%) are endemic (Minelli et al., 1993; Minelli, 1996). In addition to the number of species, the importance of the Italian peninsula has been extensively investigated at the interspecific and genetic diversity levels. Numerous studies have shown that both at the peninsular and at the insular level, Italy has been the scene of intra- and interspecific evolutionary processes (Canestrelli et al., 2007; Canestrelli and Nascetti, 2008; Vega et al., 2010; Senczuk et al., 2019; Menchetti et al., 2021; Schmitt et al., 2021).

Gap in biodiversity knowledge: invertebrates and their importance for the study of biogeographic and evolutionary processes

Most of the knowledge we have on phylogeographic patterns on an Italian or European scale is due to studies that have, almost exclusively, used vertebrates as model species, despite vertebrates only account for a tiny portion of global species biodiversity (<1.5% of species diversity; Cardoso et al., 2011). On the other hand, invertebrates are the most representative organisms of world biodiversity in terms of richness, abundance, and biomass. Just to give some examples, more than half of the known species are insects and only the beetles (order Coleoptera) comprise at least 10 times the number of species of all vertebrates and over 25% of all the species described (Cardoso et al., 2011). Of all known animal species in Italy, about 95% are invertebrates (Minelli et al., 1993). Invertebrates embrace an impressive morphological and ecological variety that allow them to live in all terrestrial environments on Earth. They can be scavengers, phytophagous, symbionts, parasites, endo- and ectoparasites, hyperparasitoids, or predators. From a biogeographic point of view, invertebrates can have very large (cosmopolitan) distributions or can be endemic of narrow regions, some also showing punctiform distributions. Given the variety of species, sizes and functional roles, the often short generation times and the rapid evolutionary rates, many invertebrate taxa are excellent evolutionary and ecological models and can play a key role as indicators of habitat change and thus of fundamental importance in the perspective of conservation actions (Kremen et al., 1993; Diniz-Filho et al., 2010).

Alticini as a study model

Alticini are a monophyletic tribe related to the tribe Galerucini, with which they form the subfamily Galerucinae in the family Chrysomelidae (Bouchard et al., 2011). Alticini are commonly referred to as flea beetles because of their characteristic enlarged hind femur due to the presence of the Maulik's organ, which confers the ability to jump as a defense mechanism against predation (Nadein and Betz, 2016). Alticini have small to moderate body size (usually 0.5–17.8 mm in length (Lee and Yu, 2021)) and, like other Chrysomelids, have a high range of colorations. The tribe Alticini comprises about 534 genera and more than 8,000 species worldwide, making it the largest Chrysomelid tribe (Iannella, 2021). In addition, we expect that a large portion of the diversity of this tribe is still unknown to science.

The distribution patterns of flea beetles are very diverse, ranging from cosmopolitan or sub-cosmopolitan species, to strictly and locally distributed species. The areas with the richest diversity are the tropical regions of South America, Africa and Asia (Biondi and D'alessandro, 2013). Like other representatives of the Chrysomelid family, flea beetles feed exclusively on plants and thus, they are present in any habitat with vegetation. This means that they are present from the coastal areas (e.g. behind the dunes), up to the high altitude environments. Trophic specializations at both larval and adult stage vary, but in general these insects make small holes on the external surface of leaves, stems and, more rarely, flowers of a wide range of higher plants (Feeny et al., 1970; Salvi et al., 2019). Regarding the trophic range, this also varies greatly from species to species and is possible to find: (i) monophagous species with adults feeding on one or two systematically closely related plant genera; (ii) oligophagous species with adults feeding on plant genera from one or two systematically closely related families; and (iii) polyphagous species with adults feeding on many botanical species that are not systematically closely related (Biondi et al., 2013).

This strong association with plants means that, among Alticini there are species considered as crop pests worldwide (Metspalu et al., 2014). For example, among the species of the genus *Epitrix* Foundras, four are listed as quarantine species by the European and Mediterranean Plant Protection Organization (EPPO), due to damage caused to potatoes (Germain et al., 2013). Another pest species is *Psylliodes chrysocephala* Linneaus, native to the Palearctic but introduced in North America, which is known to cause severe damage to cabbage or other Brassicaceae crops. In general, there are several species belonging to this tribe whose crop damage has a significant economic impact (Coral Şahin et al., 2019). In some cases, however, the effectiveness in damaging plants makes some flea beetles exceptional biocontrol agents of weeds, among them we find species belonging to the genera *Longitarsus* Berthold, *Altica* Müller, *Aphthona* Chevrolat and *Chaetocnema* Stephens (Jonsen et al., 2001; Konstantinov et al., 2001; Cagáň et al., 2006; Coral Şahin et al., 2019). One of the most noteworthy applications is probably that of *Chaetocnema tibialis* (Illiger), adopted for the control of the invasive wild amaranth (Cagáň et al., 2006).

Furthermore, the IUCN has included a number of Alticini species in its red lists, underlining their importance also from a conservation point of view. Some examples are the *Longitarsus mellissi* Wollaston, 1871 from the island of Sant Elena, listed in 2019 as Critically Endangered due to the small size of its range and the fragmentation of the environment in which it lives, and the species *Mniophilosoma obscurum* Gillerfors, listed in 2016 as Critically Endangered, endemic to Flores (Azores, Portugal) present only in the small fragments of native forest on the island.

Alticini biodiversity in Europe and Italy

The biodiversity of Alticini in Europe and in Italy is due the different climatic and geographical conditions (Konstantinov et al., 2009) and, as mentioned above, by the biogeographical processes related to the Quaternary glacial cycles. Regarding the European fauna, about 400 species and 32 genera of Alticini are known, with about 20% of species being endemic to this region (De Jong et al., 2014). However, this estimate needs to be updated as, in Italy alone, about 340 species and 25 genera are reported (Biondi, 1990; Biondi et al., 2013). Alticini have the highest number of endemic species among the Italian Chrysomelids. In fact, there are 38 endemic species in the tribe, which represent 11.18% of the entire Italian flea beetle fauna. The western Alps region is the one with the greatest presence of endemic species, followed by the central Apennine region and the Sicilian-Southern Apennine region.

The high diversity of this group and its distribution from macro to micro scale make Alticini excellent models for the study of ecological, evolutionary and biogeographic processes (Baselga et al., 2015). The well documented chorological and ecological information of this group makes it highly representative of the overall biodiversity of a given ecosystem. Moreover, their very close association with plants makes these insects true proxies of the status of the plant component of the environments in which they live (Biondi et al., 2013). This makes Alticini an ideal model for ecological and evolutionary studies, and to answer the questions of this thesis project.

Thesis aims and objectives

The aims of my PhD project are to investigate the phylogeographic patterns of endemic Alticini species of Apennines, and to compare the effect of Quaternary glaciations cycles on temperate and high-altitude biotas. To achieve these aims I adopted a multi-taxa approach that allows making robust inferences on the processes that have shaped current diversity patterns. The following objectives were set out to fulfill the main aims of the thesis:

1. Selection of target species based on their ecology and geographic distribution.
2. Assessment of the systematics of target species using an integrative (morphological and molecular) taxonomic approach.
3. Identification of the main genetic lineages and of the refugial areas for temperate and high-mountain species by means of a phylogeographic approach.
4. Multi-locus reconstruction of the historical demography of each species to compare the ecological response of temperate and high mountain species to Pleistocene glacial cycles.

References

Avise, J.C., 2000. Phylogeography: the history and formation of species. Harvard university press.

Baselga, A., Gómez-Rodríguez, C., Vogler, A.P., 2015. Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. Glob. Ecol. Biogeogr. 24, 873–882. https://doi.org/10.1111/geb.12322

Bennett, K.D., Bennett, K., 1997. Evolution and ecology: the pace of life. Cambridge University Press.

Biondi, M., 1990. Elenco commentato dei crisomelidi alticini della fauna italiana. Fragm. Entomol. Romana 22, 109–183.

Biondi, M., D'alessandro, P., 2013. Genus-group names of Afrotropical flea beetles (Coleoptera: Chrysomelidae: Alticinae): Annotated catalogue and biogeographical notes. EJE 107, 401–424.

Biondi, M., Urbani, F., D'Alessandro, P., 2013. Endemism patterns in the Italian leaf beetle fauna (Coleoptera, Chrysomelidae). ZooKeys 177.

Bouchard, P., Bousquet, Y., Davies, A.E., Alonso-Zarazaga, M.A., Lawrence, J.F., Lyal, C.H., Newton, A.F., Reid, C.A., Schmitt, M., Ślipiński, S.A., 2011. Family-group names in Coleoptera (Insecta). ZooKeys 1.

Cagáň, Ľ., Tóth, P., Tóthová, M., 2006. Population dynamics of *Chaetocnema tibialis* Illiger and *Phyllotreta vittula* (Redtenbacher) on the weed Amaranthus retroflexus L. and cultivated Amaranthus caudatus L. Plant Prot. Sci. 42, 73–80.

Canestrelli, D., Aloise, G., Cecchetti, S., Nascetti, G., 2010. Birth of a hotspot of intraspecific genetic diversity: notes from the underground. Mol. Ecol. 19, 5432–5451.

Canestrelli, D., Cimmaruta, R., Nascetti, G., 2007. Phylogeography and historical demography of the Italian treefrog, *Hyla intermedia*, reveals multiple refugia, population expansions and secondary contacts within peninsular Italy. Mol. Ecol. 16, 4808–4821.

Canestrelli, D., Nascetti, G., 2008. Phylogeography of the pool frog *Rana* (*Pelophylax*) *lessonae* in the Italian peninsula and Sicily: multiple refugia, glacial expansions and nuclear–mitochondrial discordance. J. Biogeogr. 35, 1923–1936.

Canestrelli, D., Salvi, D., Maura, M., Bologna, M.A., Nascetti, G., 2012. One Species, Three Pleistocene Evolutionary Histories: Phylogeography of the Italian Crested Newt, Triturus carnifex. PLoS ONE 7(7): e41754.

Cardoso, P., Erwin, T.L., Borges, P.A., New, T.R., 2011. The seven impediments in invertebrate conservation and how to overcome them. Biol. Conserv. 144, 2647–2655.

Carstens, B.C., Knowles, L.L., 2007. Shifting distributions and speciation: species divergence during rapid climate change. Mol. Ecol. 16, 619–627.

Cohen, K. M., Finney, S. C., Gibbard, P. L., & Fan, J. X., 2013. The ICS international chronostratigraphic chart. Episodes Journal of International Geoscience, 36(3), 199-204.

Coral Şahin, D., Magoga, G., Özdikmen, H., Montagna, M., 2019. DNA Barcoding as useful tool to identify crop pest flea beetles of Turkey. J. Appl. Entomol. 143, 105–117.

Cresti, L., Schönswetter, P., Peruzzi, L., Barfuss, M.H., Frajman, B., 2019. Pleistocene survival in three Mediterranean refugia: origin and diversification of the Italian endemic *Euphorbia gasparrinii* from the E. verrucosa alliance (Euphorbiaceae). Bot. J. Linn. Soc. 189, 262–280.

Davis, M.B., Shaw, R.G., 2001. Range shifts and adaptive responses to Quaternary climate change. Science 292, 673–679.

Denton, G.H., Hughes, T.J., 1981. The last great ice sheets. Wiley.

De Jong, Y. et al., 2014. Fauna Europaea - all European animal species on the web. Biodiversity Data Journal 2: e4034. doi: 10.3897/BDJ.2.e4034.

Diniz-Filho, J.A.F., De Marco Jr, P., Hawkins, B.A., 2010. Defying the curse of ignorance: perspectives in insect macroecology and conservation biogeography. Insect Conserv. Divers. 3, 172–179.

Elias, S.A., 1996. The ice-age history of national parks in the Rocky Mountains. Smithsonian Inst Press.

Feeny, P., Paauwe, K.L., Demong, N.J., 1970. Flea beetles and mustard oils: host plant specificity of *Phyllotreta cruciferae* and *P. striolata* adults (Coleoptera: Chrysomelidae). Ann. Entomol. Soc. Am. 63, 832–841.

Feliner, G.N., 2011. Southern European glacial refugia: a tale of tales. Taxon 60, 365–372.

Frenzel, B., 1992. Atlas of paleoclimates and paleoenvironments of the Northern Hemisphere. Geogr. Res. Inst. Hung. Acad. Sci. Bp. Gustav Fisch. Verl. Stuttg. Jena N. Y.

Germain, J.-F., Chatot, C., Meusnier, I., Artige, E., Rasplus, J.-Y., Cruaud, A., 2013. Molecular identification of *Epitrix* potato flea beetles (Coleoptera: Chrysomelidae) in Europe and North America. Bull. Entomol. Res. 103, 354–362.

Gómez, A., Lunt, D.H., 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula, in: Phylogeography of Southern European Refugia. Springer, pp. 155–188.

Gonçalves, H., Martinez-Solano, I., Pereira, R., Carvalho, B., García-París, M., Ferrand, N., 2009. High levels of population subdivision in a morphologically conserved Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence from mtDNA, microsatellites and nuclear genealogies. Mol. Ecol. 18, 5143–5160.

Habel, J. C., Schmitt, T., & Müller, P., 2005. The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: the phylogeography of the Marbled White butterfly (Satyrinae, Lepidoptera). J. Biogeogr. 32(8), 1489-1497.

Hewitt, G., 2003. Ice ages: species distributions, and evolution, in: Evolution on Planet Earth. Elsevier, pp. 339–361.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.

Hewitt, G.M., 2011. Mediterranean peninsulas: the evolution of hotspots, in: Biodiversity Hotspots. Springer, pp. 123–147.

Hewitt, G.M., 2004. The structure of biodiversity–insights from molecular phylogeography. Front. Zool. 1, 1–16.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. 68, 87–112.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247–276.

Hilty, J., Merenlender, A., 2000. Faunal indicator taxa selection for monitoring ecosystem health. Biol. Conserv. 92, 185–197.

Hodkinson, I.D., 2005. Terrestrial insects along elevation gradients: species and community responses to altitude. Biol. Rev. 80, 489–513.

Huntley, B., Birks, H.J.B., 1983. Atlas of past and present pollen maps for Europe, 0-13,000 years ago. Cambridge University Press.

Iannella, M., D'alessandro, P., De Simone, W., & Biondi, M., 2021. Habitat specificity, host plants and areas of endemism for the genera-group *Blepharida* sl in the afrotropical region (Coleoptera, Chrysomelidae, Galerucinae, Alticini). Insects, 12(4), 299.

Jonsen, I.D., Bourchier, R.S., Roland, J., 2001. The influence of matrix habitat on *Aphthona* flea beetle immigration to leafy spurge patches. Oecologia 127, 287–294.

Konstantinov, A., Korotyaev, B., Volkovitsh, M., 2009. Insect Biodiversity in the Palearctic Region, in: Insect Biodiversity: Science and Society. pp. 107–162. https://doi.org/10.1002/9781444308211.ch7

Konstantinov, A., Volkovitsh, M., Cristofaro, M., 2001. New data on Palearctic *Aphthona* (Coleoptera: Chrysomelidae) with description of a new species: taxonomic and faunistic results of biological control exploration. Entomol. News 112, 31–41.

Kremen, C., Colwell, R., Erwin, T., Murphy, D., Noss, R. and, Sanjayan, M., 1993. Terrestrial arthropod assemblages: their use in conservation planning. Conserv. Biol. 796–808.

Lee, C. F., & Yu, S. F., 2021. Redescription of *Podontia lutea* (Olivier, 1790) and *P. dalmani* Baly, 1865 (Insecta: Coleoptera: Chrysomelidae: Galerucinae: Alticini), with Notes on the Biology of *P. lutea*. 台灣農業研究.

Lohse, K., Nicholls, J.A., Stone, G.N., 2011. Inferring the colonization of a mountain range—refugia vs. nunatak survival in high alpine ground beetles. Mol. Ecol. 20, 394–408.

Maura, M., Salvi, D., Bologna, M. A., Nascetti, G., & Canestrelli, D., 2014. Northern richness and cryptic refugia: phylogeography of the Italian smooth newt *Lissotriton vulgaris meridionalis*. Biol. J. Linn. Soc. 113(2), 590-603.

Margraf, N., Verdon, A., Rahier, M., Naisbit, R.E., 2007. Glacial survival and local adaptation in an alpine leaf beetle. Mol. Ecol. 16, 2333–2343.

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

Menchetti, M., Talavera, G., Cini, A., Salvati, V., Dincă, V., Platania, L., Bonelli, S., Balletto, E., Vila, R., Dapporto, L., 2021. Two ways to be endemic. Alps and Apennines are different functional refugia during climatic cycles. Mol. Ecol. 30, 1297–1310.

Metspalu, L., Kruus, E., Ploomi, A., Williams, I.H., Hiiesaar, K., Jõgar, K., Veromann, E., Mänd, M., 2014. Flea beetle (Chrysomelidae: Alticinae) species composition and abundance in different cruciferous oilseed crops and the potential for a trap crop system. Acta Agric. Scand. Sect. B—Soil Plant Sci. 64, 572–582.

Minelli, A., 1996. La checklist delle specie della fauna italiana. Un bilancio del progetto. Boll. Mus. Civ. Storia Nat. Verona 20, 249–261.

Minelli, A., Ruffo, S., La Posta, S., 1993. Checklist delle specie della fauna italiana. Calderini.

Morgan, J.W., Venn, S., 2017. Alpine plant species have limited capacity for long-distance seed dispersal. Plant Ecol. 218, 813–819.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature, 403(6772), 853-858.

Nadein, K., & Betz, O., 2016. Jumping mechanisms and performance in beetles. I. Flea beetles (Coleoptera: Chrysomelidae: Alticini). J. Exp. Biol. , 219(13), 2015-2027.

Provan, J., Bennett, K.D., 2008. Phylogeographic insights into cryptic glacial refugia. Trends Ecol. Evol. 23, 564–571.

Razgour, O., Salicini, I., Ibáñez, C., Randi, E., Juste, J., 2015. Unravelling the evolutionary history and future prospects of endemic species restricted to former glacial refugia. Mol. Ecol. 24, 5267–5283.

Richard, P., 1991. Pielou, EC (1991): After the Ice Age. The return of Life to Glaciated North America. The University of Chicago Press, Chicago, 366 p. ill., 16 x 23, 5 cm, $24.95 US. Géographie Phys. Quat. 45, 379–380.

Roy, M.S., 1997. Recent diversification in African greenbuls (Pycnonotidae: Andropadus) supports a montane speciation model. Proc. R. Soc. Lond. B Biol. Sci. 264, 1337–1344.

Salvi, D., D'alessandro, P., Biondi, M., 2019. Host plant associations in Western Palaearctic *Longitarsus* flea beetles (Chrysomelidae, Galerucinae, Alticini): a preliminary phylogenetic assessment. ZooKeys 856, 101.

Salvi, D., Harris, D. J., Kaliontzopoulou, A., Carretero, M. A., & Pinho, C., 2013. Persistence across Pleistocene ice ages in Mediterranean and extra-Mediterranean refugia: phylogeographic insights from the common wall lizard. BMC Evol Biol, 13(1), 1-18.

Salvi, D., Schembri, P.J., Sciberras, A., Harris, D.J., 2014. Evolutionary history of the Maltese wall lizard Podarcis filfolensis: insights on the 'Expansion–Contraction'model of Pleistocene biogeography. Mol. Ecol. 23, 1167–1187.

Schmitt, T., 2017. Molecular biogeography of the high mountain systems of Europe: an overview. High Mt. Conserv. Chang. World 63–74.

Schmitt, T., 2009. Biogeographical and evolutionary importance of the European high mountain systems. Front. Zool. 6, 1–10.

Schmitt, T., Fritz, U., Delfino, M., Ulrich, W., Habel, J.C., 2021. Biogeography of Italy revisited: genetic lineages confirm major phylogeographic patterns and a pre-Pleistocene origin of its biota. Front. Zool. 18, 1–13.

Schmitt, T., Hewitt, G., Müller, P., 2006. Disjunct distributions during glacial and interglacial periods in mountain butterflies: Erebia epiphron as an example. J. Evol. Biol. 19, 108–113.

Senczuk, G., Harris, D.J., Castiglia, R., Litsi Mizan, V., Colangelo, P., Canestrelli, D., Salvi, D., 2019. Evolutionary and demographic correlates of Pleistocene coastline changes in the Sicilian wall lizard *Podarcis wagleriana*. J. Biogeogr. 46, 224–237.

Taberlet, P., Fumagalli, L., Wust-Saucy, A., Cosson, J., 1998. Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. 7, 453–464.

Tooley, M., 1993. Long term changes in eustatic sea level. Clim. Sea Level Change Obs. Proj. Implic. Warrick R 81–107.

Vega, R., Amori, G., Aloise, G., Cellini, S., Loy, A., Searle, J.B., 2010. Genetic and morphological variation in a Mediterranean glacial refugium: evidence from Italian pygmy shrews, *Sorex minutus* (Mammalia: Soricomorpha). Biol. J. Linn. Soc. 100, 774–787.

Webb, R.S., Rind, D.H., Lehman, S.J., Healy, R.J., Sigman, D., 1997. Influence of ocean heat transport on the climate of the Last Glacial Maximum. Nature 385, 695–699.

Weiss, S., Ferrand, N., 2007. Phylogeography of southern European refugia. Springer.

Williams, M., Dunkerley, D., De Deckker, P., Kershaw, P., Chappell, J., 1998. Quaternary Environments., 2^{nd} edn.(Arnold: London.).

Zachos, F. E., & Habel, J. C., 2011. Biodiversity hotspots: distribution and protection of conservation priority areas. Springer Science & Business Media.

Part I

SPECIES SAMPLING, MORPHOLOGICAL AND MOLECULAR IDENTIFICATION

CHAPTER 1

Sampling

Distribution of Italian Alticini and identification of target species

Sampling activities were performed during three years (2019-2021) in the suitable season (from April to September). Target species were selected based among endemic elements of the Apennines fauna. An endemic species is defined as "a species showing a distribution restricted to a geographical area, delimited by natural elements, and independent of administrative borders" (Biondi, 2006; Biondi et al., 2013); whereas an "area of endemism" is defined as "a geographic region comprising the distributions of two or more monophyletic taxa, that exhibit a phylogenetic and distributional congruence, and have their respective relatives occurring in other such defined regions" (Harold and Mooi, 1994). Endemic alticini from the Apennines provide an ideal model for phylogeographic studies given their high ecological specialization and low dispersal abilities that make these species particularly sensitive to climate-linked habitat changes.

Most of the Apennine endemic species are restricted to specific sectors of the Apennine chain (D'Alessandro and Biondi, 2007). There are groups of taxa exclusive to the northern, central, and southern Apennines (Fig. 1). In the central and southern Apennines the distribution of endemic taxa appears wider. In fact, here we find the largest number of endemic Alticini species in the mountain range. In the central Apennines three endemic species are present in the high-altitude environments *Psylliodes springeri* Leonardi, *Psylliodes biondii* Leonardi and *Longitarsus springeri* Leonardi; and two endemics are present in beech wood environments: *Longitarsus zangherii* Warchałowskii and *Psylliodes urbaniae* Biondi & D'Alessandro. Endemic elements also occur in the southern Apennines and many of them are distributed from the northernmost part of Calabria to the mountainous regions of north Sicily. This is the case of *Longitarsus laureolae* Biondi, related to beech wood environments and *Psylliodes ruffoi* Leonardi, *Psylliodes feroniae* Leonardi and *Psylliodes leonhardi* Heikertinger, occurring in more xeric environments. However, *Psylliodes urbaniae*, *Psylliodes feroniae* and *Psylliodes leonhardi*, as well as three endemic taxa of the northern Apennines, *Psylliodes solarii* Leonardi, *Psylliodes parodii* Leonardi or *Psylliodes fiorellae* Leonardi tend to have very disjunct, extremely restricted or punctiform distributions, making them unsuitable for a phylogeographic study. Also the endemic species of the genera *Orestia* Germar and *Minota* Kutschera, were not considered due to their rarity and the difficulty of sampling, which can only be carried out by sieving the soil. In conclusion, the six target species selected for the thesis project were as follows: *Longitarsus laureolae, Longitarsus springeri, Longitarsus zangherii, Psylliodes biondii, Psylliodes ruffoi,* and *Psylliodes springeri*.

The information on the distribution of the above mentioned endemic species were collected from bibliographic material (Leonardi, 1975a, 1975b; Biondi, 1990, 1988; D'Alessandro and Biondi, 2007; Leonardi, 2007; Biondi et al., 2013), monographs (Warchałowski, 2013), public databases (Latella et al., 2005) and data from the entomological collection of Prof. Maurizio Biondi stored at the University of L’Aquila. A database was built including all the distribution occurrence record of the target species. For each record, the collection locality and, when present, the geographical coordinates were marked. In the first case the collection locality was used as a centroid for the coordinates, in the second case the coordinates were reported in decimal degrees (both in WGS84). Both data were managed in a GIS environment to allow the construction of distribution maps that were used for planning the sampling design.

Fig.1 Study region and geographic sectors of interest: Northern Apennines, Central Apennines, Southern Apennines, and Sicily.

Longitarsus laureolae

Endemic species of the central southern Apennines, present from Abruzzo to Sicily. It shows the greatest morphological affinities with *L. candidulus* (Foudras), a W-Mediterranean species, and *L. leonardii* Doguet, a suborophilic Pyrenean species. It is present in localities between 1000 m and 1600 m in humid beech forest environments where it is found on its only currently known host plant, *Daphne laureola* (Thymelaeaceae) (Biondi, 1988) (Fig. 2).

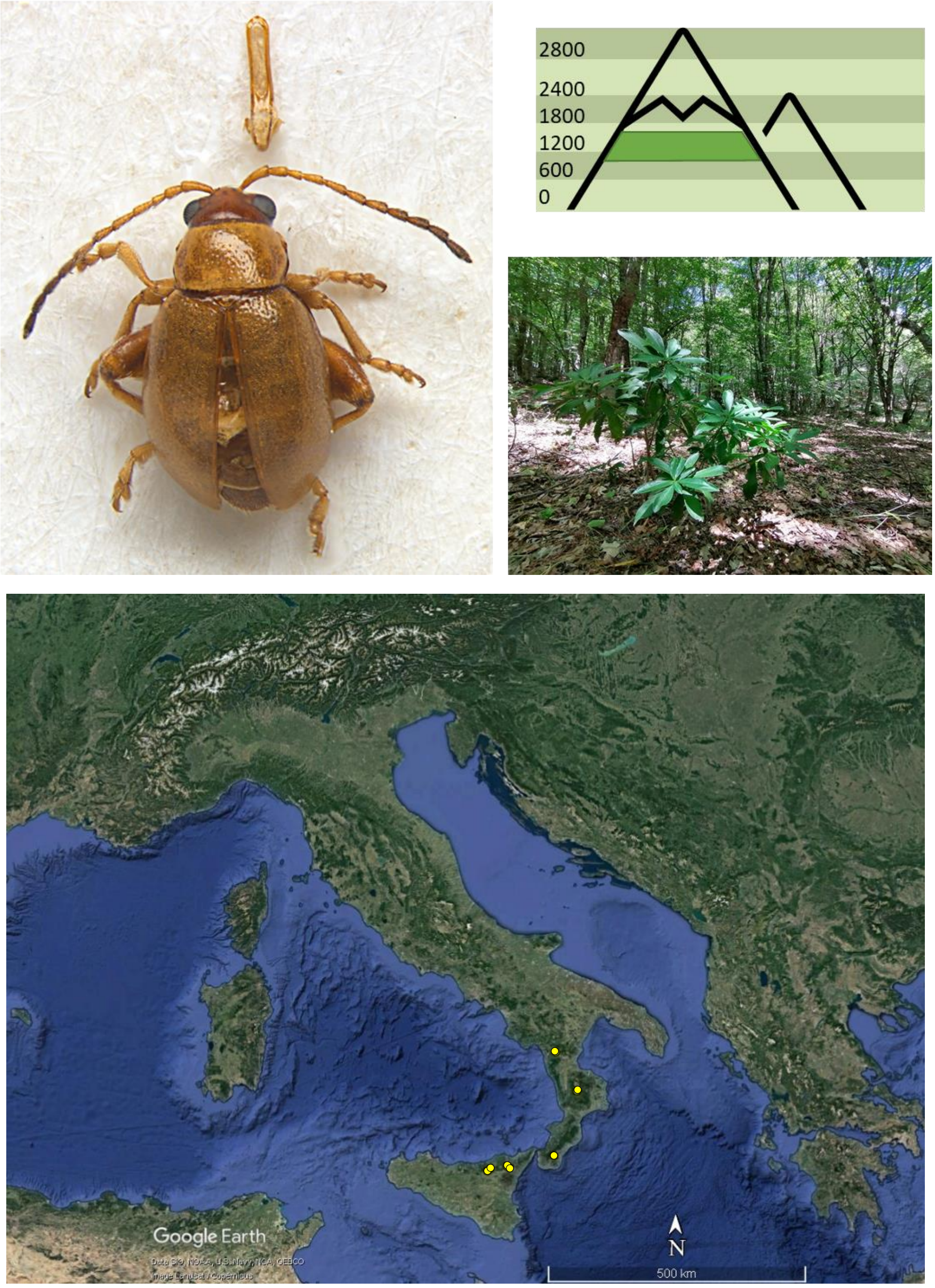


Fig.2 Habitus and aedeagus of *L. laureolae* specimen, altitudinal range of the species and a photo of the host plant *Daphne laureola* in its environment. Below, the distribution map based on historical records.

Longitarsus springeri

Endemic species of the central Apennines, present from the Sibillini Mountains to the Matese. From a systematic point of view, it appears to be well isolated from other European entities even if an affinity has been hypothesized for the species of the group of *L. succineus* (Foudras) (Leonardi, 1975b; D'Alessandro and Biondi, 2007; Farina, 2018). It has two distinct color forms, a light one known only in the valley of Lake Pilato and a dark one, present in the rest of the area (Farina, 2018). It lives in high altitude environments, with populations ranging from 1700 m up to 2700 m. The species feeds on Asteraceae, specifically it is known to live on *Senecio squalidus* L. *rupestris* (Waldst. & Kit) Greuter and on *Jacobaea erratica* (Bertol.) Fourr. (Biondi and Di Casoli, 1996) (Fig. 3).

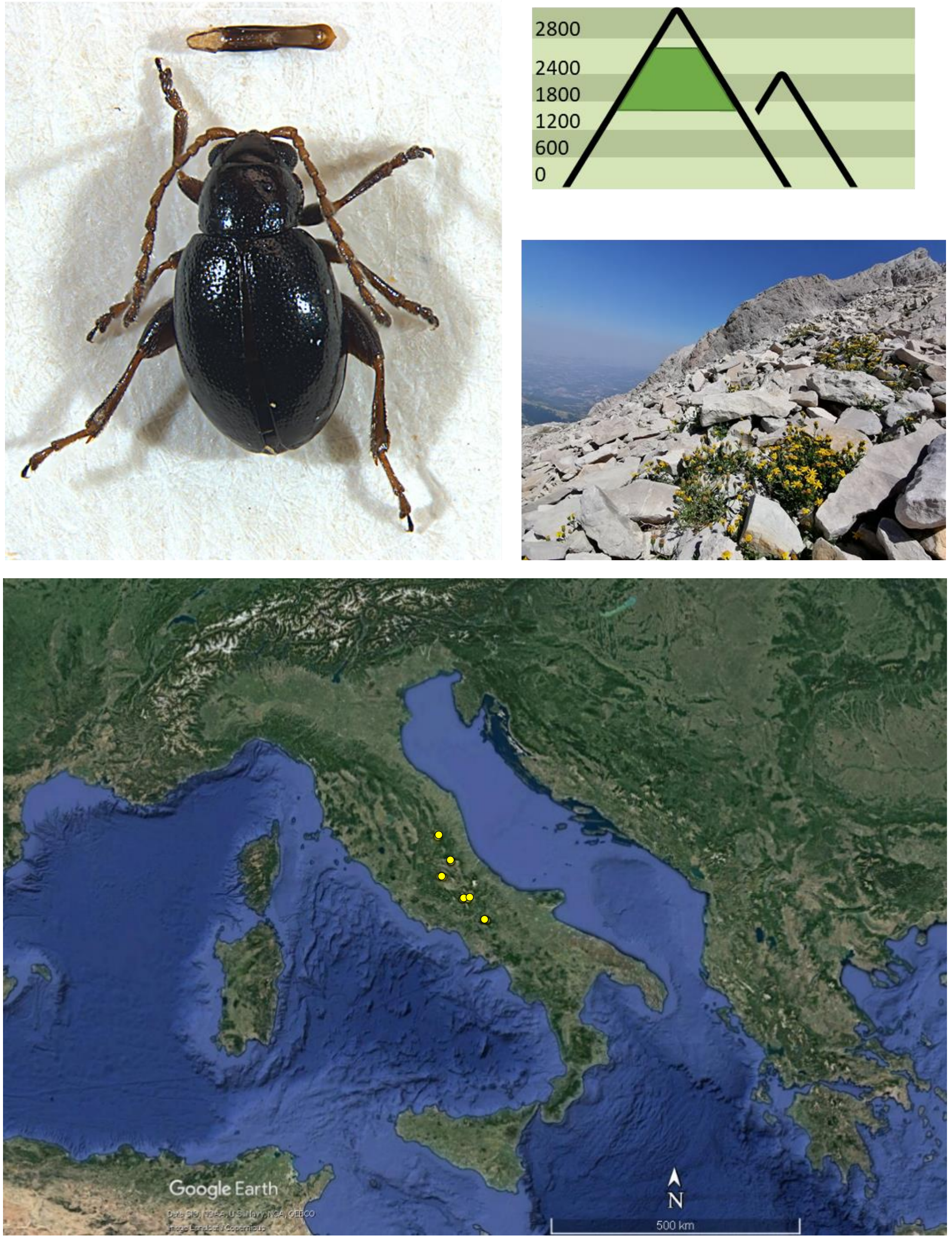


Fig.3 Habitus and aedeagus of *L. springeri* specimen, altitudinal rang of the species and a photo of the host plant *Senecio squalidus L. rupestris* in its environment. Below, the distribution map based on historical records.

Longitarsus zangherii

Species endemic to the central-northern Apennines, present from Emilia Romagna to Abruzzo. It is associated with beech forest environments between 770 m and 1600 m, and it is often found near water courses as it feeds on plants of the genus *Petasites* Miller that is linked to aquatic environments (Fig. 4).

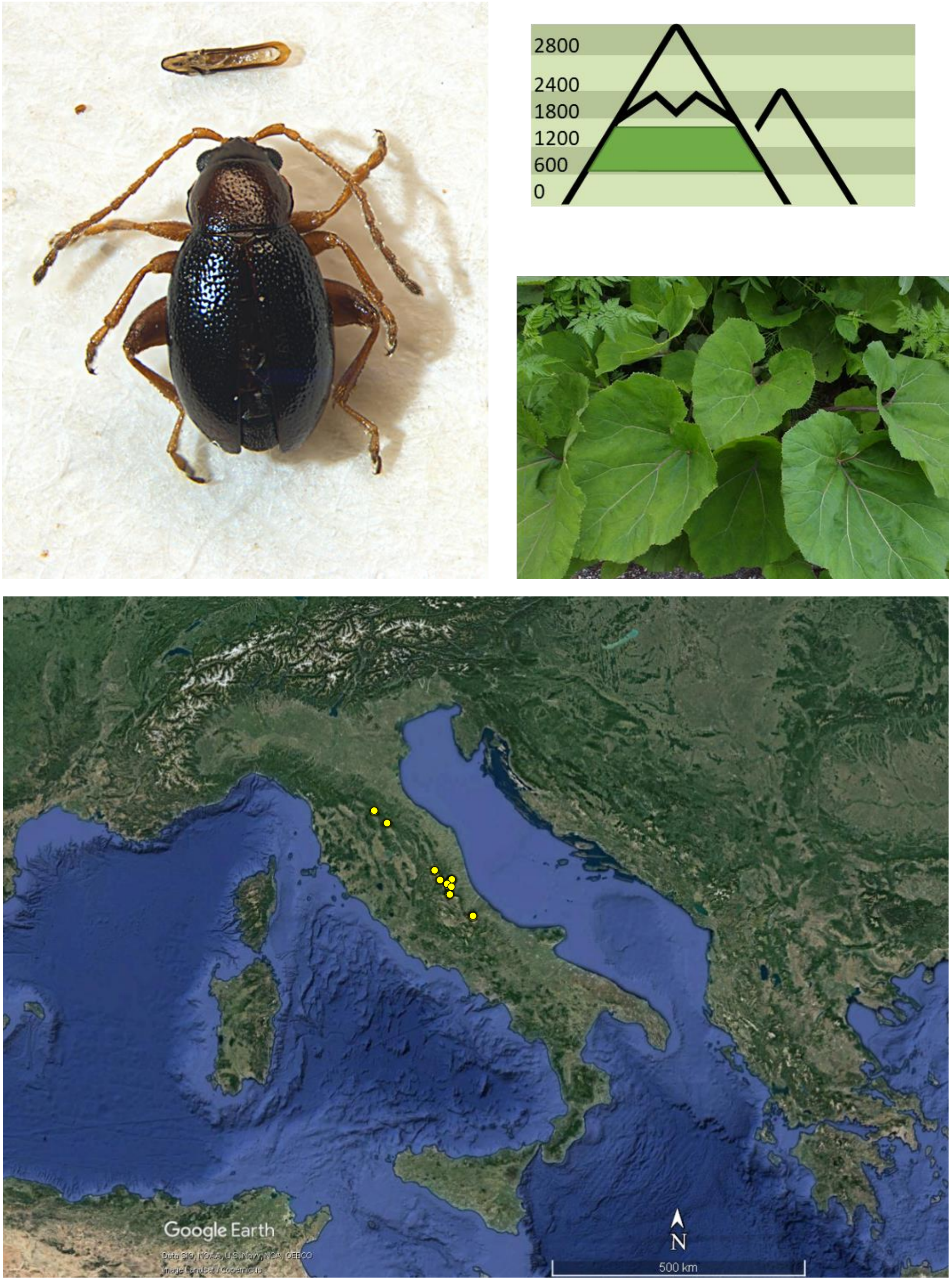


Fig.4 Habitus and aedeagus of *L. zangherii* specimen, altitudinal rang of the species and a photo of the host plant *Petasites* in its environment. Below, the distribution map based on historical records.

Psylliodes biondii

Endemic species to the central Apennines, with a distribution similar to that of *L. springeri*. From a taxonomic point of view, this species is very similar to *Psylliodes springeri* Leonardi, a micro endemism present in the valley of Lake Pilato, and to *Psylliodes picipes* Redtenbacher, a species with an alpine distribution. The Maiella population of *P. biondii*, due to the scarce material available and some slightly distinctive morphological traits, has not been included in the typical series by Leonardi and it requires further investigation (Leonardi, 2007). It lives in high altitude environments, with populations ranging from 1800 m up to 2700 m. *P. biondii* like other related species feeds on Brassicaceae, in particular the two known host plants of this species are *Isatis apennina* Ten. ex Grande and *Erysimum pseudorhaeticum* Polatschek. The populations of the northernmost part of the range feed on *I. apennina*, while, moving southwards, the species shifts to *E. pseudorhaeticum* (Fig. 5).

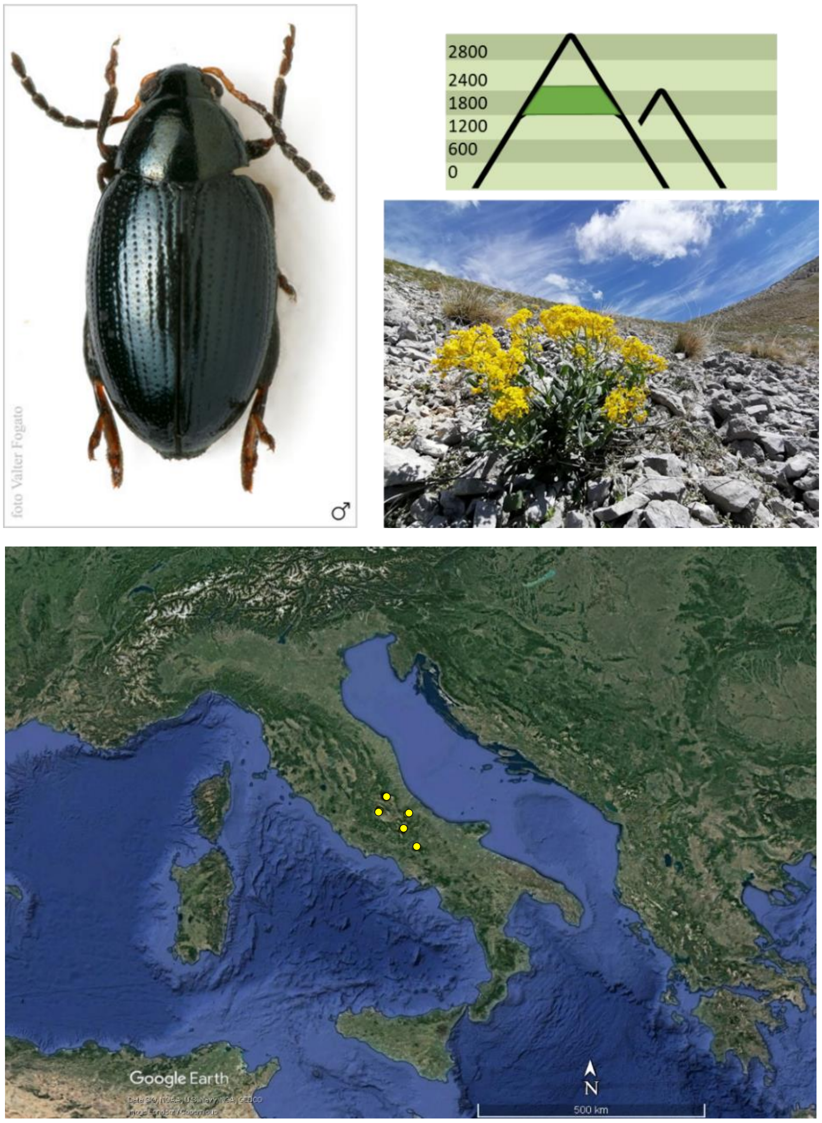


Fig.5 Habitus of *P. biondii* specimen (photo of Valter Fogato from www.Chrysomelidae.it), altitudinal rang of the species and a photo of the host plant *Isatis apennina* in its environment. Below, the distribution map based on historical records.

Psylliodes ruffoi

Endemic species to the southern Apennines, it is present from lower Campania to the northern part of Sicily. The species is part of the *Psylliodes gibbosa* group, presenting several morphological affinities with both *P. gibbosa* and *Psylliodes kiesenwetteri*, with which it shares a large part of the range (Leonardi, 1975a). It occurs over a wide altitude range, from 400 m up to 1600 m. It is found in ecotones between xeric woods and semi-arid clearings. There is still a lot of uncertainty as regards the host plant of *Psylliodes ruffoi* could be find, as this species has been collected on different types of plants (Baviera and Biondi, 2015). Generally it is associated with the Poaceae (Nadein, 2008) (Fig. 6).

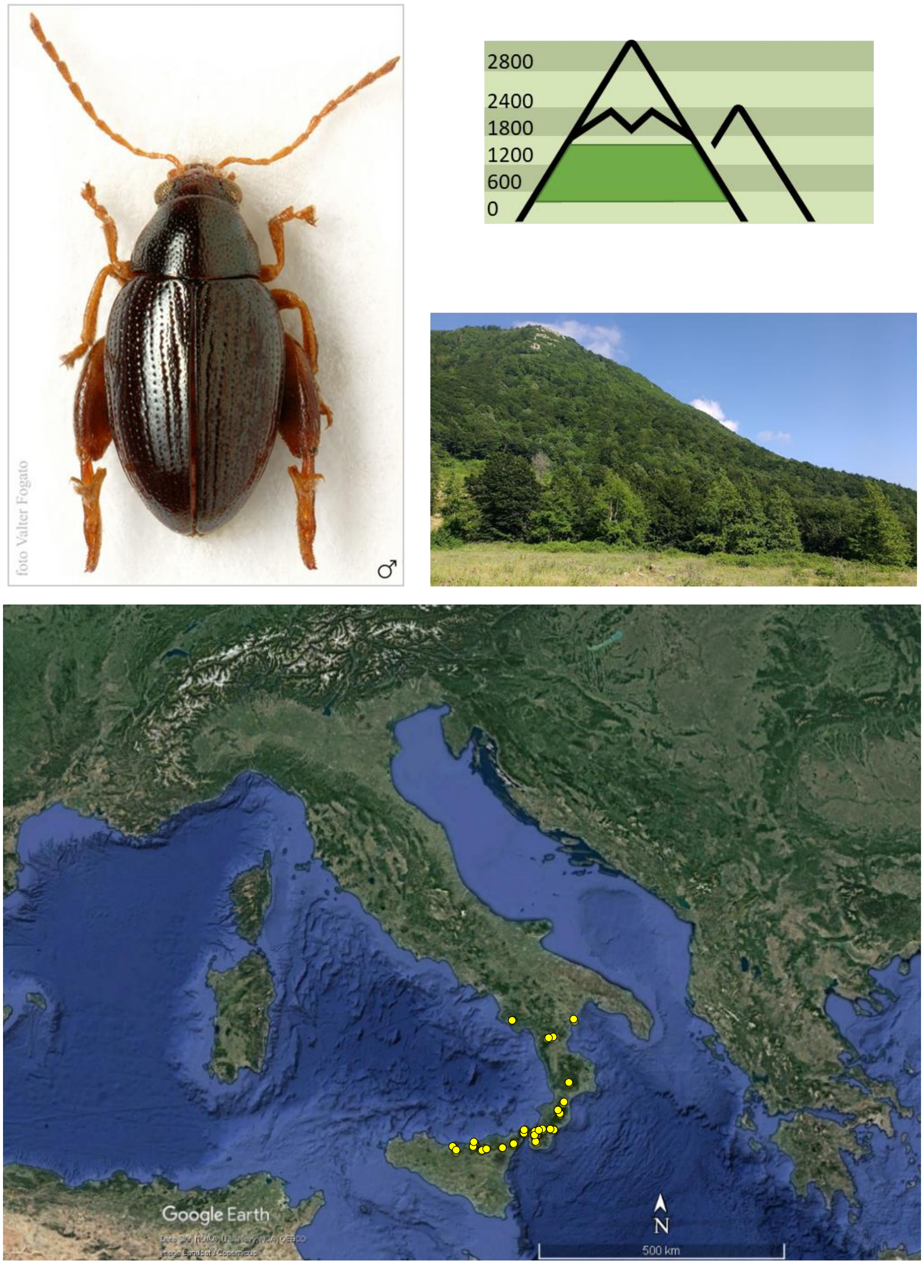


Fig.6 Habitus of *P. ruffoi* specimen (photo of Valter Fogato from www.Chrysomelidae.it), altitudinal rang of the species and a photo of its environment. Below, the distribution map based on historical records.

Psylliodes springeri

A microendemic species known only from the valley of Lake Pilato. As already mentioned, this species is morphologically similar to *P. biondii*, with which it also shares the ecology, as it lives in the same environments and feeds on the same host plant, *Isatis apennina* (Leonardi, 2007). It is precisely because of this affinity with *P. biondii* that it was selected as a target species, despite its punctiform range (Fig. 7).

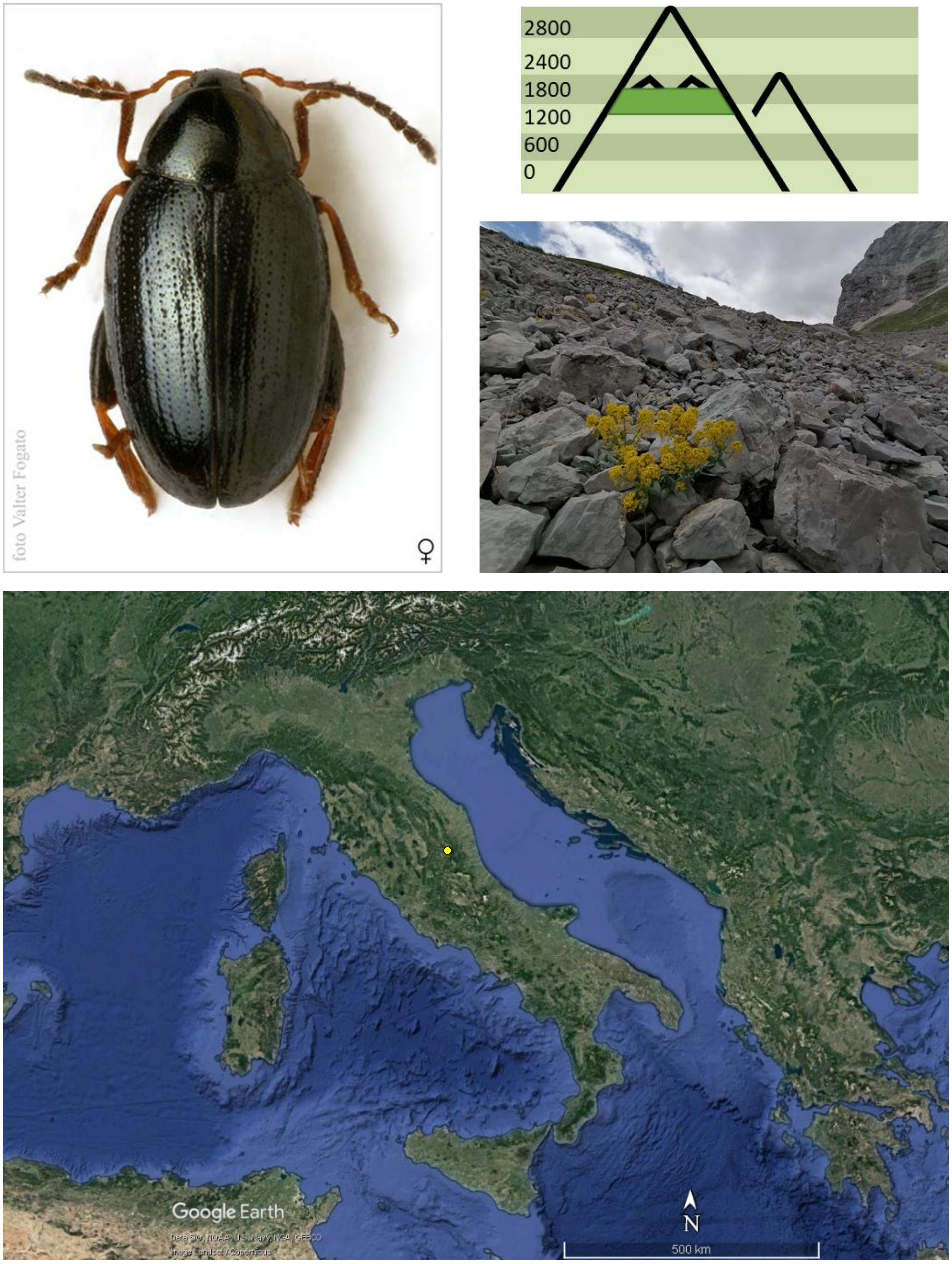


Fig.7 Habitus of *P. springeri* specimen (photo of Valter Fogato from www.Chrysomelidae.it), altitudinal range of the species and a photo of the host plant *Isatis apennina* in its environment. Below, the distribution map based on historical records.

Technique of sampling

Alticini beetles were sampled by beating on host plants (when known) or on vegetation with an entomological net or an entomological umbrella. Then, specimens were collected using an aspirator. Indeed, unlike other beetles, Alticini cannot be sampled by entomological traps, such as pitfalls, malaise or odorous traps. Collected specimens were stored in alcohol (EtOH) at 95° in 2 ml microtubes conical screw cap with O-ring. One or more tubes were assigned to each collection locality based on the number of specimens captured and the geographical coordinates of each locality were taken via GPS. Where possible, the plant on which the specimens were collected was marked.

Problems of species identification, phenology, and rarity: the need of a widespread and repeated sampling

The objective difficulties in capturing small-sized species that escape classical trapping systems such as the ones listed above, is not the only difficulty to consider when planning sampling of Alticini. One of the main problems of this tribe of beetles is that recognition at the species level is not possible in the field. This makes discrimination between the target species and similar species living in the same environments impractical in the field. The only possible sampling strategy is to focus on the environments or host plants on which the target species is present and to orientate at the genus level, which can often be identified in the field. In any case, even this taxonomic level requires subsequent confirmation by the taxonomist.

Another challenge for sampling Alticini is due to their phenology and rarity. In fact, many species linked to high altitude mountain environments have shorter phenology, and usually their collection is restricted to a short temporal range (one-two months). Sampling species with short phenology is particularly challenging because of the inter-annual seasonal variability, so that, depending on the annual weather conditions, the period of activity can vary from year to year even by several weeks. Species related to coastal environments or low and medium altitude have relatively longer phenology, and for them there is usually a good number of records. Moreover, most of the available data came from single sampling campaigns dating back to the 1980s and 1970s, or even to the end of the 19th century. This means that, even if the knowledge on the distribution is quite exhaustive, little is known on the species' phenology, and it is also possible that the species does not occur in those areas anymore. All this makes it very difficult to design a sample from a temporal point of view.

For all of these reasons it has been necessary to set a widespread and sampling design of Alticini repeated throughout the season.

Sampling design and sampling results

The occurrence of the target species were used to delineate sampling-areas in the northern, central and southern Apennines and Sicily based on the higher density of records as estimated in a GIS environment, using Qgis version 3.16.15. Sampling areas were selected by making a minimum convex polygon on the points and then applying a buffer of 50 km. The areas with the highest presence of records of target species were selected through the concentration map with kernel density estimate applying a 10 km radius.

In addition to the information on the distribution of the species, sampling design was based on as much information as possible on the biology of the target species, and in particular on:

- **Phenology**: in order to sample all the target species selected during the three years of PhD, the samplings were designed taking into account the known phenology of the target species to maximize sampling in each season (Tab. 1).

	January	February	March	April	May	June	July	August	September	October	November	December
Longitarsus laureolae					🟩	🟩	🟩	🟩	🟩	🟩		
Longitarsus springeri							🟩	🟩	🟩			
Longitarsus zangherii					🟩	🟩	🟩	🟩	🟩			
Psylliodes biondi					🟩	🟩	🟩					
Psylliodes ruffoi					🟩	🟩	🟩	🟩				
Psylliodes springeri					🟩	🟩	🟩					

Tab.1 The months of activity, known in literature, for the target species are shown in green.

- **Host Plant**: data on host plants of target species were collected through literature research. Distribution and ecological data of the host plants were considered to provide further data on the possible presence of target species.
- **Ecology**: environmental suitability data of the species were considered when present (Urbani et al., 2017).

Once the sampling areas were defined and biological information were collected, sampling paths were identified covering all recorded reporting locations. In addition, other intermediate locations falling within the sampling macro areas were selected to sample other areas of possible presence of the species.

Samplings was performed in three main field campaigns:

- Southern session I: from 17/06/2019 to 01/07/2019 along the southern Apennine chain and the northern part of Sicily;

- Southern session II: from 25/06/2020 to 04/07/2020 along the southern Apennine chain and the northern part of Sicily to complete the collections carried out in the first year;
- Northern session: from 17/7/2020 to 10/8/2020 along the central northern Apennine chain, the Tuscan-Emilian Apennines, and the entire Alpine arc.

In addition to these three main campaigns, several punctual campaigns (with one or a few populations as a target) were carried out to refine the sampling, particularly in the central Apennine area, the richest in target species. Preliminary sampling of the Alpine arc was also carried out for discover possible unknown locations of target species in this region and to collect material for possible future studies (Fig. 8).

Fig.8 Reported in red are the locality sampled during the three years of PhD.

Overall, about 370 sites between Apennines and the Alps were sampled and almost 2000 individuals belonging to 18 genera and about 150 species identified on a morphological basis were collected at the end of the three years. For the target species, almost all known sites were sampled. Furthermore, new occurrence sites were identified for *L. springeri*, *L. zangherii* and *L. laureolae*. In particular, for *L. laureolae*, which to date was known only up to the southern part of Pollino massif, a population was identified in Abruzzo, on the south-east part of Maiella massif, extending the range of this species by about 300 km. *Longitarsus springeri* was reported for the first time in Maiella, on Mount Viglio and on some secondary peaks of the central Apennines, such as Mount Camicia of the Gran Sasso massif and in Monti della Meta. Finally, *L. zangherii* was reported for the first time in the Aremogna plain, expanding its known range southwards, which until now stopped in the northern part of the Maiella massif, about 30 km further north (Tab. 2).

	Populations sampled/tot	**New populations**
Longitarsus laureolae	5/5	1
Longitarsus springeri	9/9	3
Longitarsus zangherii	8/8	6
Psylliodes biondi	6/6	0
Psylliodes ruffoi	11/13	0
Psylliodes springeri	1/1	0

Tab.2 The table shows, for each species, the number of populations sampled, the number of known populations and the number of new populations discovered during the sampling.

References

Baviera, C., Biondi, M., 2015. The Alticini (Coleoptera: Chrysomelidae, Galerucinae) of Sicily: recent records and updated checklist. Atti Della Accad. Peloritana Dei Pericolanti-Cl. Sci. Fis. Mat. E Nat. 93, 2.

Biondi, M., 2006. Recenti aspetti metodologici in biogeografia storica. Atti Accad. Entomol. 53, 89–111.

Biondi, M., 1990. Elenco commentato dei crisomelidi alticini della fauna italiana. Fragm. Entomol. Romana 22, 109–183.

Biondi, M., 1988. Contributo alla conoscenza dei Crisomelidi Alticini dei Monti Nebrodi e Peloritani (Sicilia Nord Orientale) e descrizione di una nuova specie di *Longitarsus* (Coleoptera, Chrysomelidae, Alticinae). Animalia 15, 129–149.

Biondi, M., Di Casoli, M.T., 1996. Il popolamento a Coleoptera Chrysomelidae del massiccio del Gran Sasso. Analisi.

Biondi, M., Urbani, F., D’alessandro, P., 2013. Endemism patterns in the Italian leaf beetle fauna (Coleoptera, Chrysomelidae). ZooKeys 177.

D’alessandro, P., Biondi, M., 2007. Analisi zoogeografica della crisomelidofauna appenninica legata ai piani altitudinali montano, subalpino e alpino (Coleoptera, Chrysomelidae). Biogeogr. J. Integr. Biogeogr. 28.

Farina, L., 2018. Note corologiche e tassonomiche su *Longitarsus springeri* Leonardi, 1975, con segnalazione di una forma cromatica inedita dei Monti Sibillini (Coleoptera, Chrysomelidae). Bull. Société Entomol. Fr. 123, 185–192.

Harold, A.S., Mooi, R.D., 1994. Areas of endemism: definition and recognition criteria. Syst. Biol. 43, 261–266.

Latella, L., Ruffo, S., Stoch, F., 2005. Il progetto Ckmap (Checklist e distribuzione della fauna italiana): metodologia di lavoro e tecniche informatiche. Checkl. E Distrib. Della Fauna Ital. Mem. Mus. Civ. Storia Nat. Verona 2, 15–20.

Leonardi, C., 2007. Dati inediti sul genere *Psylliodes* Latreille, con descrizione di quattro nuove specie mediterranee (Coleoptera Chrysomelidae). Atti Della Soc. Ital. Sci. Nat. E Mus. Civ. Storia Nat. Milano 148, 161–240.

Leonardi, C., 1975a. Le *Psylliodes* appenniniche del museo civico di storia naturale di verona (Coleoptera Chrysomelidae).

Leonardi, C., 1975b. Note su alcuni *Longitarsus* con descrizione di due nuove specifiee citazione di quattro entità nuove per la fauna europea (Coleoptera Chrysomelidae). Atti Della Soc. Ital. Sci. Nat. E Mus. Civ. Storia Nat. Milano 116, 199–217.

Nadein, K.S., 2008. Review of the *Psylliodes* gibbosus species group, with descriptions of two new species 34.

Warchałowski, A., 2013. The palaearctic Chrysomelidae identification keys, Volume 1 and 2. Natura optima dux Foundation.

CHAPTER 2

Integrative taxonomy and systematics of target species

Morphological techniques

As a first morphological assessment, the specimens collected in the sampling localities were identified and sorted at the genus level by Prof. Maurizio Biondi. A subsequent investigation was carried out on all the localities of known presence of the target species, where the collected specimens were identified at the species level. Morphological identification took place with the aid of a Leica M205C binocular microscope and using the diagnostic characters present mainly on the aedeagus and spermatheca. About 85% of the approximately 2,000 individuals collected were identified to species level on a morphological basis.

Molecular techniques and assessment

Molecular assessment was performed by analyzing DNA sequence data of selected gene fragments. To this end, three different methods of DNA extraction were applied on the specimens: (i) a destructive method, using the whole specimen for DNA extraction, (ii) a non- destructive method, separating the head-prothorax portion of the animal from the rest of the body using an entomological pin and dipping the two parts directly in lysis buffer and proteinase K, and (iii) using only the three left legs of each specimen. However, the non-invasive method, precisely because it allows to reference voucher of the specimen to be maintained, was used more often than the invasive method (about 95%). In all cases, total DNA extraction was performed with a standard high-salt protocol (Sambrook et al., 1989). Samples treated with the non-invasive method were recovered at the end of the lysis process, and the two parts of the animal were reassembled on an entomological card point. The importance of maintaining a reference voucher to which one or more DNA sequences can be associated is fundamental, especially when studying complex taxa with a still uncertain systematics.

For the molecular assessment, mitochondrial and nuclear markers were used. Selected mitochondrial markers were two fragments of the cytochrome c oxidase I (*cox1*) gene, namely the barcoding fragment and the 3' fragment. Among the nuclear markers, Wingless (*Wg*) was selected as it is widely used in the phylogeny and phylogeography of Chrysomelids, and three of the markers used by Gikonyo (2021) for the reconstruction of the phylogeny of *Psylliodes*, namely the Carbamoylphosphate synthase gene (*CAD*), the Crossveinless 2 gene (*Cv2*), the Methyl methanesulfonate sensitivity protein 22-like protein gene(*MMS22*), the Dorsal-ventral patterning protein Short gastrulation gene (*Sog*) and the Rad 50 protein gene (*Rad50*). The selection criteria were: (i) the highest evolutionary rates among Coleoptera according to OrthoDB (Kriventseva et al., 2019) and (ii) the successful amplification in the PCR reaction. Information on primers, fragment length, mixture and cycle are provided chapter by chapter.

Molecular and morphological mismatch in *Longitarsus* and *Psylliodes*: improving molecular and morphological tool for the systematics of target species

The morphological and molecular delimitation of the target species was a fundamental and preliminary step to the phylogeographic study. Thanks to this step, during the initial phase of the PhD project, three main issues related to the mismatch between molecular and morphological identification were pointed out and resolved, in particular:

- First, within the genus *Longitarsus*, we observed that the public database of DNA sequences (GenBank and BOLD) had several problems of species misidentification and taxonomic incongruences. This makes difficult to implement a molecular identification of species by comparing the DNA sequences of a target gene (*cox1*) of a sampled individual with deposited sequences in public database with a DNA barcoding approach. This molecular method of specimens identification uses a short segment of DNA from a specific standardized gene that is compared to a database of known sequences from morphologically identified specimens. To overcome species misidentification and taxonomic incongruences problem it was necessary to perform a critical taxonomic assessment of reference DNA sequence database. This first study allowed increasing the reliability of the DNA barcoding tools for the identification of taxonomically complex groups such as this genus of Alticini.
- Second, during the initial steps of molecular screening of the target species *P. ruffoi*, we observed that some samples, morphologically identified as *P. ruffoi*, had molecular characters (i.e. DNA sequences) typical of the related species *P. kiesenwetteri*. The two species live in close sympatry in different localities of their range. This finding raised doubts about the discriminatory power of morphological characters described for this species group and required the use of an integrative (morphological and molecular) approach to resolve the relationships between these two species and their identification.
- Third, during the molecular assessment of the Maiella population of *P. biondii*, we observed that this population show a large genetic distance with other populations of *P. biondii*, suggesting they are possibly two distinct species. A certain degree of morphological distinctiveness was already pointed out for Maiella population of *P. biondii* (Leonardi, 2007). However, these differences were no longer analyzed due to the lack of material. The phylogeographic approach implemented in this thesis allowed clarifying the degree of phylogenetic divergence between these two lineages of high altitude *Psylliodes* endemic to Central Apennines.

In the Part I of the thesis, I will investigate how the integrated use of molecular and morphological tools allowed solving the three reported cases of systematic uncertainty. This work was preparatory to the phylogeography part, which is explored in the Part II of the thesis.

References

Gikonyo, M., 2021. The evolution of host plant use and adaptation to Brassicaceae in the leaf beetle genus Psylliodes.

Kriventseva, E.V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F.A., Zdobnov, E.M., 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res. 47, D807–D811.

Leonardi, C., 2007. Dati inediti sul genere *Psylliodes* Latreille, con descrizione di quattro nuove specie mediterranee (Coleoptera Chrysomelidae). Atti Della Soc. Ital. Sci. Nat. E Mus. Civ. Storia Nat. Milano 148, 161–240.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.

CHAPTER 3

Case study I

Sharpening the DNA barcoding tool through a posteriori taxonomic validation: the case of *Longitarsus* flea beetles (Coleoptera: Chrysomelidae)

Introduction

DNA barcoding is a molecular method of specimen identification using a short segment of DNA from a specific standardized gene which is compared against a database of known sequences from morphologically identified specimens. Therefore, the intent of the DNA barcoding is to standardize the large-scale screening use of one or more reference genes in order to assign unknown individuals to species (Moritz and Cicero, 2004; DeSalle, 2006; Goldstein and DeSalle, 2011). The two key premises on which barcoding is based are: i) the nucleotide sequence used is characterized by a genetic divergence between close species that exceeds variation within the species; ii) the presence of a comprehensive sequences library obtained from individuals reliably determined at species level by expert taxonomists (Hebert et al., 2003; Floyd et al., 2010). Respecting these premises, DNA barcoding promises to be a fast and useful tool for taxonomists and a cost-effective system through which non-specialist can assign unidentified specimens to known species (Schindel and Miller, 2005).

In recent years, large-scale DNA barcoding studies has been performed on various groups of animals and have generated an enormous amount of cytochrome oxidase I (*cox1*) barcodes, which are usually stored in GenBank® and the official Barcode of Life database (BOLD) (Floyd et al., 2002; Barrett and Hebert, 2005; Hajibabaei et al., 2006). The association of such amount of sequences to taxa is a challenging step of these studies, especially for extraordinarily diverse group such as insects (Kvist, 2013). Indeed, for the most diverse orders (e.g. Diptera, Hymenoptera and Coleoptera), correct identification of species requires a high number of taxonomists, each one specialized on a single family or part thereof. Species-level identification represents a great challenge in some hyper-diverse and widespread genera, for which many taxonomists, each one with a long-standing taxonomic specialization on a regional fauna, might be required (Hamilton et al., 2010; Bacher, 2012). This implies that broad-based DNA barcoding studies should ideally recruit hundreds of specialized taxonomists, but this is not feasible. Thus, a certain degree of misidentification is inherent to these studies, and can be anticipated in species-rich taxa with difficult taxonomy (Ratnasingham and Hebert, 2013).

The inclusion of wrongly identified sequences into the reference databases undermines one of the key premises of the DNA barcoding tool and reduces its accuracy (Collins and Cruickshank, 2012; Packer et al., 2018). Indeed, these misidentified sequences generate taxonomic inconsistencies, either because they fix a wrong species tag, if they represent species new to BOLD (or any other reference database), or because they will cause incongruence with data that already exists in these databases. Taxonomic inconsistencies within reference databases can be considered as *extrinsic errors* of the DNA-barcoding tool and can be afterwards detected, revised and corrected. For this purpose, non-invasive methods of DNA extraction that allow to keep reference samples in collection for further morphological validation (Phillips and Simon, 1995; Huber, 1998; Rowley et al., 2007; Castalanelli et al., 2010; Porco et al., 2010; Vink et al., 2012; Schilthuizen et al., 2015) and high-quality metadata associated to submitted sequences (e.g. voucher type, date of collection, geographic coordinates, ecological information, images, etc.) are fundamental requirements for *a posteriori* revisions of the identification (Bergsten et al., 2012; Bortolus, 2008; Meyer and Paulay, 2005). Revisions can be directly targeted to the instances of taxonomic inconsistency that occur in large dataset, previously identified through bioinformatic analyses. A variety of tools is available for detecting taxonomic inconsistencies both before and after deposition in the global barcode library. A first approach is based on threshold clustering and assumes that the intra-specific nucleotide variability of sequences does not exceed a certain distance value, otherwise sequences are flagged as belonging to different species (Hebert et al., 2004; Meier et al., 2006; Ratnasingham and Hebert, 2007, 2013). This method is implemented in the BOLD platform with a standard threshold of 1% (Ratnasingham and Hebert, 2007). However, there is no a priori reason to assume a threshold with a prescribed limit (DeSalle et al., 2005; Rubinoff, 2006; Vogler and Monaghan, 2007). The recognition of a “boundary" among species will vary considerably due to differences in rate of nucleotide substitution and speciation time (Monaghan et al., 2009; Fujita et al., 2012). Establishing robust thresholds for species delimitation is a key component of the barcoding process. Therefore, use of software and protocols to generate an optimized threshold directly from the data is a more effective procedure (Meyer and Paulay, 2005; Brown et al., 2012; Puillandre et al., 2012; Virgilio et al., 2012). Other approaches aimed at detecting taxonomic inconsistencies in reference databases consider topological incongruence between the taxonomic and the phylogenetic tree as an indication that some of the sequences might be mislabeled (Kozlov et al., 2016; McDonald et al., 2012). Furthermore, the automated tool TAxCI, that combines multiple approaches for flagging and filtering inconsistent cases of specimen’s taxonomy has been recently developed (Rulik et al., 2017).

Accuracy of the DNA barcoding also depends on the extent of the so-called barcoding gap, i.e. the separation between intraspecific variation and interspecific divergence estimated on the basis of the selected DNA marker, e.g. the mitochondrial *cox1* gene fragment in the case of animals (Vrijenhoek, 1994). However, poorly established taxonomy (Ebach et al., 2006) as well as many biological processes including recent speciation (Maddison and Knowles, 2006; Weber et al., 2019), species-level polyphyly (Funk and Omland, 2003), interspecific hybridization (Arnold, 1993; Wilson and Bernatchez, 1998; Mastrantonio et al., 2016), horizontal gene transfer mediated by bacterial endosymbionts (Smith et al., 2012; Klopfstein et al., 2016) make mtDNA groups poorly predictive of species boundaries thus affecting the accuracy of the DNA barcoding tool. These circumstances in which the molecular identification tool loses sensitivity, even in the presence of an error-free dataset, can be considered as an *intrinsic error* of the barcoding method. Identifying those areas of the dataset in which this type of error is present allows us to know which are those species or species groups that need to be analyzed with more powerful integrative approaches to delimit, discover and identify species (Will et al., 2005).

Here we used the hyper-diverse and taxonomically complex genus *Longitarsus* Latreille (Coleoptera, Chrysomelidae, Galerucinae, Alticini) as a case study to assess the accuracy of the DNA barcoding tool following several optimization procedures. Alticini is a tribe of small to medium-sized Coleoptera named ‘flea beetles’ because of their ability to jump due to the presence of a metafemoral extensor tendon in the swollen hind femora (Nadein and Betz, 2016). *Longitarsus* is the most abundant genus among flea beetles, with over 700 species distributed in all zoogeographical regions. Larvae and adult feed respectively on roots and leaves of plants of different angiosperm families, with levels of trophic specialization ranging from strictly monophagous to widely polyphagous (Salvi et al., 2019a). Members of the genus are small-sized, with body length generally 2 to 4 mm. They can be recognized mainly by the co-occurrence of elongate first metatarsomere, exceeding half-length of hind tibia, confuse elytral punctuation, and absence of dorsal pubescence. Many species of this flea beetle genus are often part of morphologically homogenous species groups displaying striking similarities in external morphology, so a careful examination of the internal anatomic structures, mainly aedeagus and spermatheca, are also required to group specialists for reliable species identification (Biondi and D’Alessandro, 2008).

The main aims of this study are: (i) to identify taxonomic inconsistencies within the *cox1* reference database available for *Longitarsus* using barcoding gap analysis, inclusive threshold specimen identification analysis, and tree topology methods; (ii) to implement *a posteriori* taxonomic revision of ambiguous and incorrect sequences using newly generated sequences identified by *Longitarsus* specialized taxonomists and metadata obtained for sequences already in these databases; and (iii) to assess the effect of these bioinformatic and taxonomic procedures on the identification efficacy of the DNA barcoding tool. Furthermore, by resolving the *extrinsic errors* within the reference database of *Longitarsus* during steps (i) and (ii), we will identify the cases where *intrinsic errors* due to taxonomic uncertainty or specific biological processes are likely to occur, thus providing directions for integrative taxonomic and evolutionary studies on this group.

Materials and methods

Sample collection and morphological identification

Specimens analyzed in this chapter were collected and identified on a morphological basis as explained in Chapter 2. For each identified species we selected from 3 to 4 specimens, from the same locality, for DNA extraction. Among these specimens, before the DNA extraction, one specimen was mounted on an entomological card point with aedeagus or spermatheca after dissection and photomicrographs were taken using a Leica DFC500 camera and the Zerene Stacker software version 1.04. Scanning electron micrographs were taken using a Hitachi TM-1000camera (Figs 9,10).

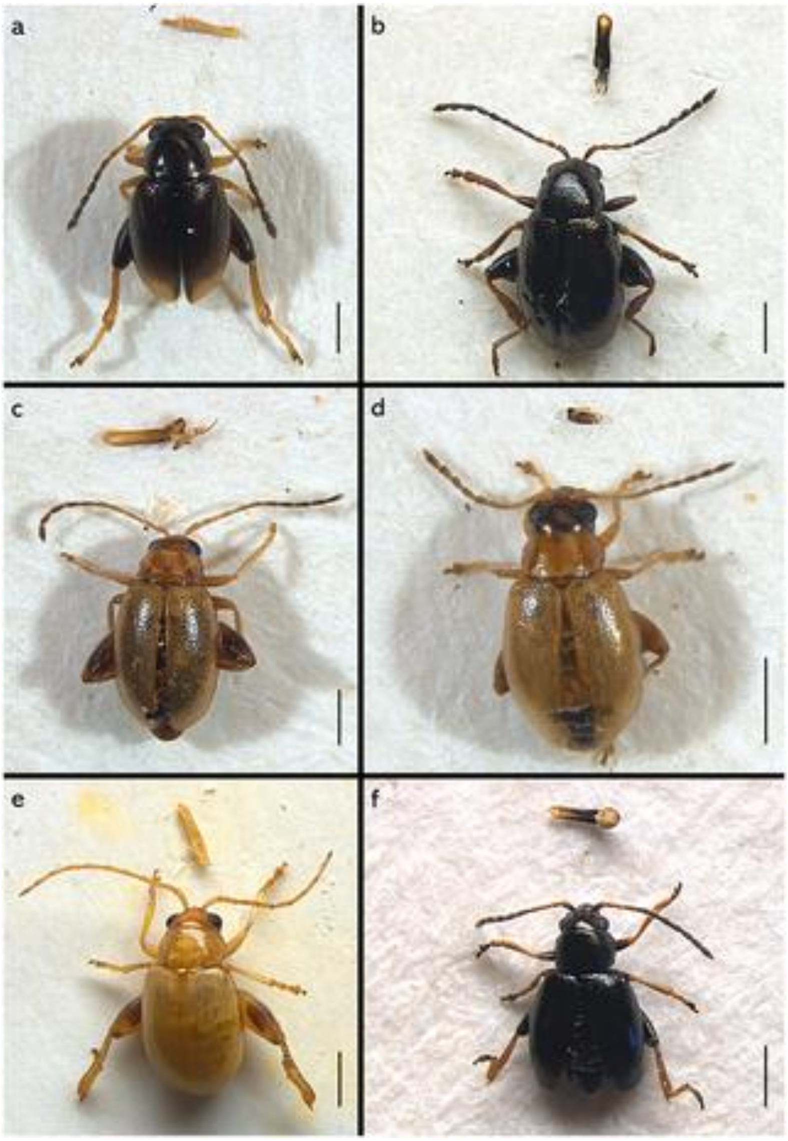


Fig.9 Photographs of voucher specimens of some of the species sequenced in this study (see S1–S3 Figs for photographs of the remaining species). Habitus and aedeagus or spermatheca of (a) *Longitarsus aeneicollis* ♂; (b) *L. corynthius metallescens* ♂; (c) *L. ballotae* ♂; (d) *L. pratensis* ♀; (e) *L. candidulus* ♂; (f) *L. anchusae* ♂. Scale bar 0.5 mm.

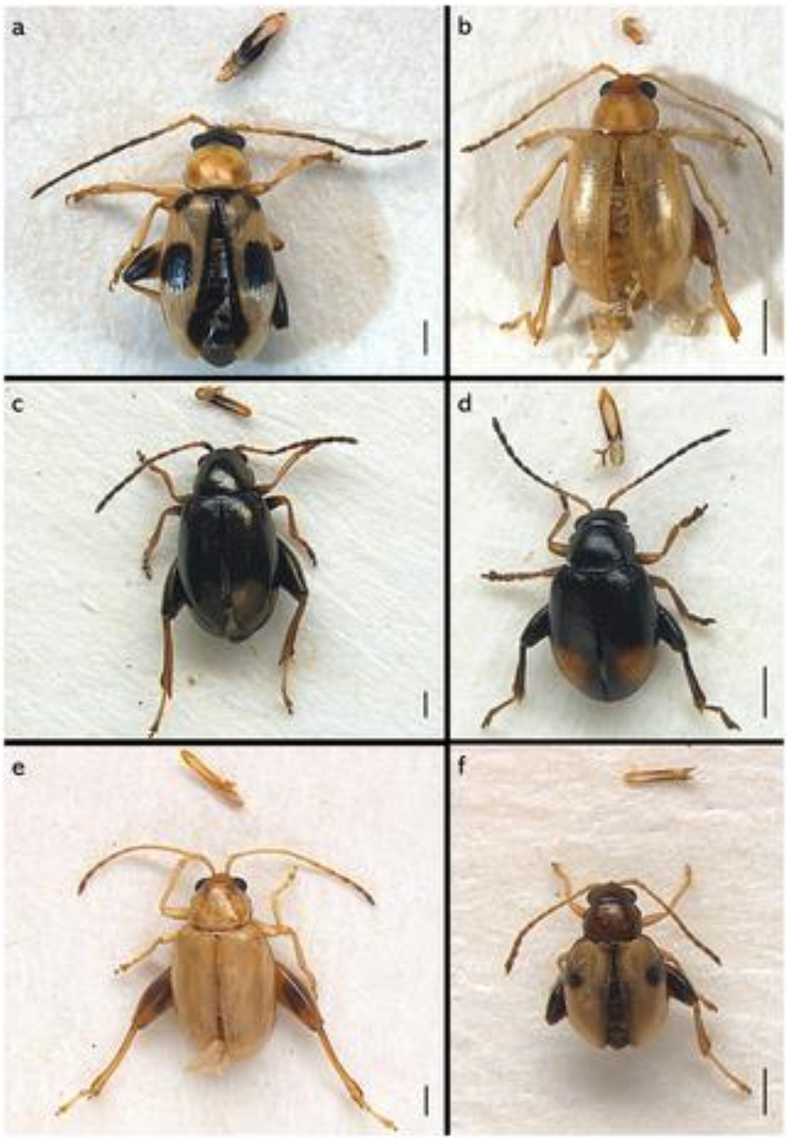


Fig.10 Photographs of voucher specimens of some of the species sequenced in this study (see S1–S3 Figs for photographs of the remaining species). Habitus and aedeagus or spermatheca of (a) *Longitarsus isoplexidis* ♂; (b) *L. pellucidus* ♀; (c) *L. echii* ♂; (d) *L. holsaticus* ♂; (e) *L. foudrasi* ♂; (f) *L. lateripunctatus* ♂. Scale bar 0.5 mm.

DNA extraction, amplification, and sequencing

DNA of specimens of two target species was extracted using non-invasive DNA extraction methods described in Chapter 1. The standard barcode region of the mitochondrial *cytochrome c oxidase I* (*cox1*) gene (658 bp) was amplified by PCR using the primers specifically designed for *Longitarsus* Lon-LCO-F (5’-CTC AGC CAT TTT ACC GAA TAA ATG-3’) and LonHCO-R (5’-GGA TTT GGI ATA ATT TCY CATA TTG-3’) (Salvi et al., 2019a). Amplification was carried out in a total volume of 25µl, with 12,5µl of BioMix™ 2x (Bioline Ltd, London, UK), 0.5 µl of each primer (10mM), 0.5 µl of BSA, and 1 µL (~40 ng) of DNA template. PCR cycling conditions for *cox1* followed (Salvi et al., 2019b). Successful amplification was determined by gel electrophoresis and PCR products were purified and sequenced by an external service (Genewitz, UK). The obtained chromatograms of each sequence were manually edited and assembled into a consensus sequence using Geneious R8 (Biomatters Ltd., Auckland, New Zealand).

Reference sequence dataset building

We built a non-redundant database including all sequences of cytochrome genes of *Longitarsus* available in the public repositories of GenBank and BOLD (data updated to 12/08/2019). We downloaded 1372 sequences from GenBank and 1433 sequences from BOLD. For sequence mining we use “Longitarsus cytochrome” as search query in the GenBank nucleotide database, and “Longitarsus” as search query in the Public Data Portal of BOLD. We eliminated all retrieved sequences that were not identified to species level (94 sequences from GenBank and 16 from BOLD). We used the *duplicated ()* function (Becker et al., 1998) of R studio to dereplicate the dataset by removing sequences having identical GenBank accession number. Before removing redundant sequences, we checked cases in which the same GenBank accession number was associated with two different specific names in BOLD and GenBank®. In these cases, we retained the more recently updated name. The non-redundant cytochrome sequences database built following this procedure includes 1429 sequences, to which we added 117 newly generated *cox1* sequences for 32 *Longitarsus* species, for a total of 1546 sequences. To select only those sequences corresponding to *cox1* barcoding fragment, we assembled the 1546 sequences using the *map to reference* option in Geneious R8, and we trimmed the assembled dataset, using as reference the *cox1* sequence that we generated for *Longitarsus pratensis* voucher ‘6c’ using standard *cox1* barcode primers (Vrijenhoek, 1994). Afterwards, a multiple sequence alignment was performed with MAFFT v.7 using the FFT-NS-1 progressive method algorithm (Katoh et al., 2002) and we eliminated two sequences that were shorter than 300 base pairs (bp). The final *cox1* dataset used for downstream analyses included 1502 sequences representing 78 *Longitarsus* species.

Sequences’ taxonomy assessment analyses

The R library *ape* v5.3 (Paradis and Schliep, 2019) was used to calculate a pairwise distance matrix of intraspecific and interspecific genetic distance using the Kimura-two parameters (K2P) substitution model (Kimura, 1980) with the pairwise deletion option. With the R package *spider* v1.5.0 (Brown et al., 2012) we performed the Barcoding Gap analyses (Meyer and Paulay, 2005) by estimating two statistics for each individual sequence in the dataset: (i) the *maximum intraspecific distance* (i.e., the maximum value of genetic distance between each sequence of the dataset with sequences of the same named species) and (ii) the *minimum interspecific distance* (i.e., the minimum value of genetic distance between each sequence of the dataset with sequences of different named species). When the difference between the maximum intraspecific distance and the minimum interspecific distance is equal to or less than zero, it means that there is the absence of a barcoding gap. For each species, we counted the instances of absence of the barcoding gap using a linear model and a kernel density estimate (KDE) developed in R library *ggplot2* (Wickham, 2016). We set the KDE method using a gaussian kernel function and the default smoothing bandwidth parameter of *ggplot2*. To assess the effect of limited sampling on the barcoding gap analyses we plotted the number of absences of barcoding gap against the total number of sequences available for each species (Meyer and Paulay, 2005; Wiemers and Fiedler, 2007).

The distance threshold is a key parameter for barcoding analyses. We performed a threshold optimisation analysis in *spider* in order to calculate the value of genetic distance which reduces the number of identifications error. The best threshold was identified with the *localMinima* function that is based on the concept of the barcoding gap and identifies a dip in the density of genetic distances as a transition between intra- and inter-specific distances. This function does not require prior knowledge of species identity to get an indication of potential threshold values. To reduce the negative effects of poor taxon coverage, we removed singletons (i.e., species represented by a single sequence) from the dataset (Meier et al., 2006; Virgilio et al., 2010).

The efficiency of molecular identification, before and after threshold optimization and singleton removal, was assessed using two methods: Best Close Match analyses (Meier et al., 2006) and the *TaxCI* pipeline developed by Rulik et al. (2017). The first method compares each sequence with the other sequences included in the dataset and checks if the smallest genetic distance (i.e. best match *sensu* Meier (2006)) are between sequences tagged with the same species name. The TaxCI method identifies taxonomic inconsistencies based on tree topology. For this analysis a Neighbour-Joining (NJ) tree was inferred using the K2P model in MEGA7 (Kimura, 1980).

Finally, we performed a taxonomic revision of all those sequences identified as wrong or ambiguous in the previous steps. The taxonomic revision was based on: (i) comparison with our newly generated reference sequence for 34 *Longitarsus* species; (ii) available metadata associated with sequences deposited in BOLD; (iii) newly generated metadata for voucher specimens kindly loaned to us by authors of sequences deposited in BOLD. Following this *a posteriori* taxonomic revision, incorrect identifications were corrected and the identification accuracy of the barcoding tool based on the resulting reference database was assessed using the same procedures described above (Barcoding Gap analyses, threshold optimization analyses, Best Close Match and *TaxCI* analyses). We used the ANOVA to test for differences between correct species identification ratio obtained with (i) the original reference dataset, (ii) the dataset with the optimised threshold, and (iii) the final reference dataset after the taxonomic revision. The ANOVA test was performed in the R library *clusterSim* after the data were normalized by quotient transformation (x/mean) (Walesiak and Dudek, 2020).

Result

Barcoding gap analysis and tree topology methods show that DNA barcoding accuracy and identification efficacy of the non-redundant database of *cox1* sequences (*Original dataset*) were improved after threshold optimization (*Optimal Threshold Dataset*) and *a posteriori* taxonomic revision (*Final dataset*).

Original dataset

The *cox1 Longitarsus* dataset includes 1502 sequences unevenly distributed among 78 species, which corresponds to ~ 11.1% of the described species diversity of the genus. The most represented species was *L. ordinatus* (Foudras, 1860) with 167 sequences, whereas 32 species were represented by less than 5 sequences. We found an overlap between the distribution of intra- and inter-specific pairwise K2P distances, resulting in the absence of an evident barcode gap in the *Longitarsus* datasets (Fig. 3). Intraspecific K2P distance values ranged from 0 to 17.7% (mean=1.4%). The maximum intraspecific value was observed among three sequences belonging to *L. erro* Horn, 1889, all collected in Canada. Interspecific K2P distances ranged from 0 to 27.6% (mean=15.4%). A value of interspecific distance equal to 0 was found in several comparisons between sequences tagged as different species. The difference between the *maximum intraspecific distance* and the *minimum interspecific distance* calculated for each sequence resulted in 573 cases in which the barcoding gap is not present (38.2% of sequences) (Fig. 11).

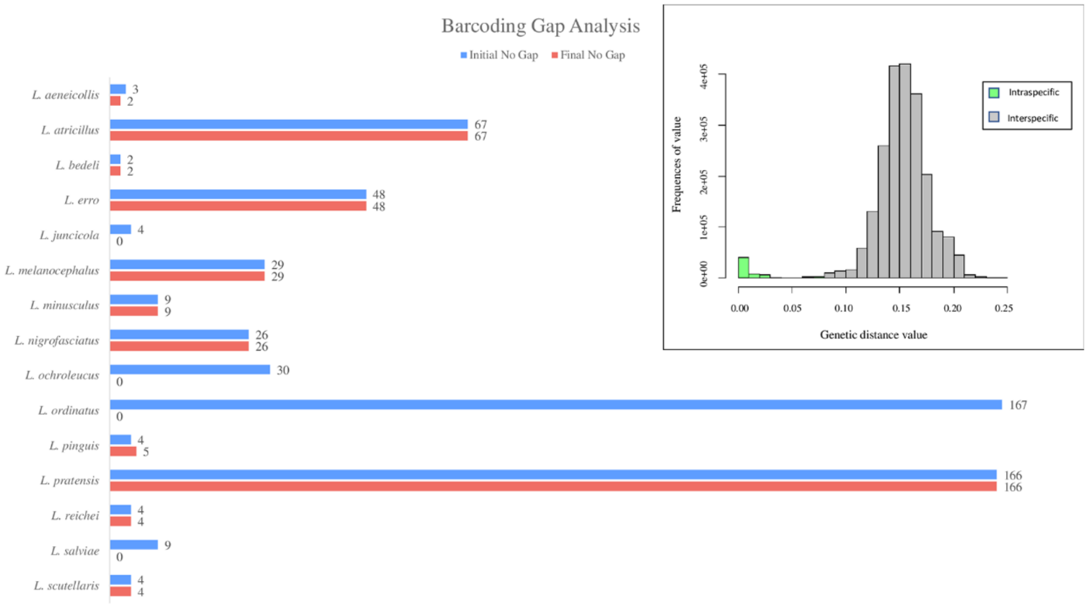


Fig.11 Results of the barcoding gap analyses for the original dataset and the final dataset. The number of absences of barcoding gap for each species is reported.

The identification of a barcoding gap can be biased by reduced sampling of nucleotide variability at an inter- and intra-specific level (Meyer and Paulay, 2005; Wiemers and Fiedler, 2007). In this study, the presence of the barcoding gap was not associated to the number of sequences per species (Fig. 12). The linear regression model shows an increase both in the absence (adjusted R^2=0.9989, *p*-value=2.2e-16) and in the presence (adjusted R^2=0.9828, *p*-value=2.2e-16) of the barcoding gap with the increase in the number of sequences per species (Fig. 12B). Both the number of absences and presences of the barcoding gap have a higher density estimate in species represented by ~ 10 sequences (Fig. 12A).

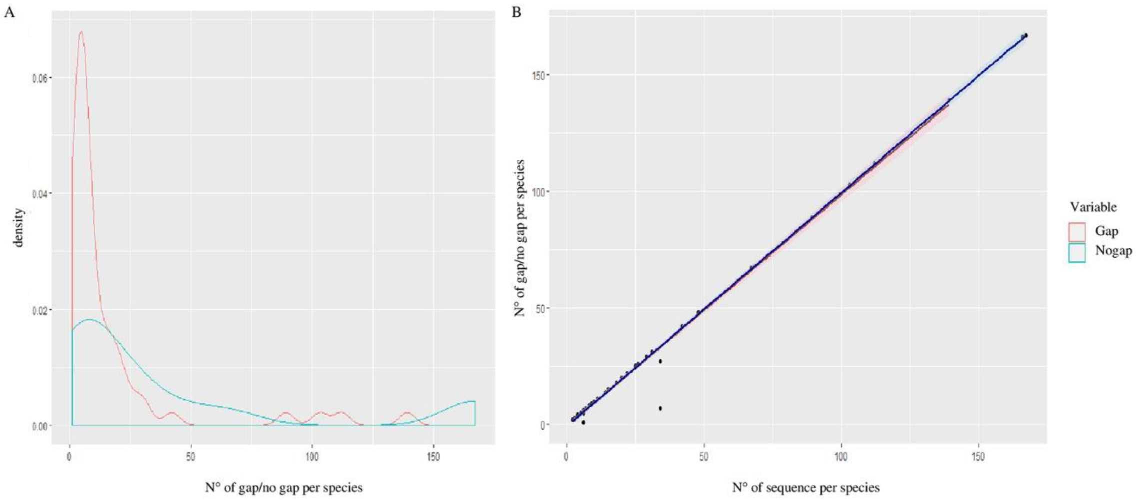


Fig.12 Effect of sequence sampling on the barcoding gap analyses. (a) Kernel Density Plots showing the distribution of instances of absence of the barcoding gap over the total number of sequences available for each species. (b) Linear regression models showing the association between the total number of sequences available for each species and the number instances of presence (red) and absence (blue) of barcoding gap.

The barcoding identification efficiency on the original *Longitarsus* dataset evaluated through the best close match analysis, with the default 1% distance threshold, resulted in 95.2% of correct identification (1431 out of 1502). Remaining sequences resulted in: 36 ambiguous sequences, presenting more than one species as the closest match or within the distance threshold; 10 incorrect sequences, which present a different species as their closest match and 25 no id sequences that do not have a close match within the given threshold. It should be noted that all the ambiguous and incorrect sequences represent cases in which the barcoding gap is not present (Tab. 3). Results of the TaxCI analysis are overall in line with the other analyses and identified 28 heterospecific cluster of which 13 have sequences found in more than one cluster. Furthermore, TaxCI identified some new cases of inconsistent identification as reported in Tab.4 (see also Supplementary S6).

			Original Dataset				**Optimal Threshold Dataset**				**Final Dataset**			
Species	**Min inter dist**	**Max intra dist**	**Correct**	**Ambiguous**	**Incorrect**	**No id**	**Correct**	**Ambiguous**	**Incorrect**	**No id**	**Correct**	**Ambiguous**	**Incorrect**	**No id**
Longitarsus atricillus	0.2	6.3	64	0	1	2	66	0	1	0	66	0	0	0
Longitarsus bedeli	0.2	0.9	0	0	2	0	0	0	2	0	0	0	2	0
Longitarsus brisouti	13.8	1.4	2	0	0	1	3	0	0	0	3	0	0	0
Longitarsus isoplexidis	12.0	1.7	4	0	0	1	5	0	0	0	5	0	0	0
Longitarsus juncicola	0	0.6	1	2	1	0	1	2	1	0	42	0	0	0
Longitarsus minusculus	6.8	9.0	8	0	0	1	9	0	0	0	9	0	0	0
Longitarsus nigrocillus	9.8	2.8	2	0	0	1	3	0	0	0	3	0	0	0
Longitarsus obliteratus	10.4	4.8	14	0	0	1	15	0	0	0	15	0	0	0
Longitarsus ochroleucus	2,8	15.2	29	0	0	1	29	0	1	0	29	0	0	0
Longitarsus ordinatus	0	8.8	152	15	0	0	152	15	0	0	129	0	0	0
Longitarsus parvulus	10.6	10.9	31	0	0	1	31	0	0	1	31	0	0	1
Longitarsus pinguis	11.8	13.7	5	0	0	1	5	0	0	1	5	0	0	1
Longitarsus pratensis	0	8.5	145	18	1	2	146	18	2	0	146	18	2	0
Longitarsus refugiensis	15.2	3.3	2	0	0	2	4	0	0	0	4	0	0	0
Longitarsus reichei	0	5.1	3	0	1	0	3	0	1	0	3	0	1	0
Longitarsus salviae	11.5	17.2	8	0	0	1	8	0	0	1	8	0	0	0
Longitarsus scutellaris	0	0.9	0	1	3	0	0	1	3	0	0	1	3	0
Longitarsus succineus	9.2	2.0	14	0	0	1	15	0	0	0	15	0	0	0

Tab.3 Results of the best close match analyses. Taxonomic inconsistency for each species are reported as minimum intraspecific (Min inter dist) and maximum interspecific (Max intra dist) genetic distance, and the number of correct, ambiguous, incorrect, and non-identify (No id) sequences, for the original dataset, the optimal thresholds dataset, and final dataset.

	Original Dataset				**Optimal Threshold Dataset**				**Final Dataset**			
Species	**tci**	**cl.het**	**sp.split**	**other.homog**	**tci**	**cl.het**	**sp.split**	**other.homog**	**tci**	**cl.het**	**sp.split**	**other.homog**
L. aeneicollis	0	8	8	0	0	8	0	0	0	8	0	0
L. apicalis	0	1	1	0	0	0	0	0	0	0	0	0
L. atricillus	67	67	67	0	67	67	67	0	67	67	0	0
L. bedeli	2	2	2	0	2	2	0	0	2	2	0	0
L. curtus	0	0	0	0	0	2	0	0	0	0	0	0
L. erro	48	48	0	48	48	45	45	0	48	45	45	0
L. exsoletus	0	89	0	89	0	0	0	0	0	0	0	0
L. juncicola	4	4	4	0	4	4	4	0	0	0	0	0
L. kutscherae	0	3	3	0	0	3	0	0	0	3	0	0
L. lateripunctatus	0	5	0	5	0	0	0	5	0	0	0	5
L. lycopi	0	15	0	15	0	0	0	15	0	0	0	15
L. melanocephalus	34	34	8	26	34	34	0	0	34	34	0	0
L. minusculus	9	9	0	9	9	0	0	9	9	0	0	9
L. nasturtii	0	0	0	0	0	3	0	0	0	3	0	0
L. nigrocillus	0	3	0	3	0	0	0	0	0	0	0	0
L. nigrofasciatus	26	26	0	26	26	0	0	26	26	0	0	26
L. obliteratus	0	15	0	15	0	0	0	0	0	0	0	0
L. ochroleucus s.str.	25	25	24	0	025	25	25	0	24	24	0	0
L. ochroleucus lindbergi	5	5	5	0	0	5	0	0	5	5	0	0
L. ordinatus	167	167	15	152	167	167	38	0	0	0	0	0
L. parvulus	0	32	0	32	0	0	0	32	0	0	0	32
L. pellucidus	0	22	0	22	0	0	0	0	0	0	0	0
L. pinguis	0	6	0	6	0	0	0	6	0	0	0	6
L. pratensis	166	166	151	15	166	166	166	0	166	166	0	0
L. refugiensis	0	4	0	4	0	0	0	0	0	0	0	0
L. reichei	4	4	1	0	4	4	4	0	4	4	0	0
L. salviae	9	9	0	9	9	9	0	9	0	0	0	0
L. scutellaris	4	4	4	0	4	4	4	0	4	4	0	0
L. suturellus	0	7	0	7	0	0	0	0	4	4	0	0
L. tabidus	0	25	0	25	0	0	0	0	0	0	0	0

Tab.4 Results of the TaxCI analyses. Taxonomic inconsistency for each species are reported as the number of: individuals of a given species not grouped as monophylum (tci), individuals of heterogeneous distance-based cluster (cl.het), individuals of a species found in more than one cluster (sp.split) and all individuals of a species in a homogeneous cluster with members in at least one other homogeneous cluster.

Optimal Threshold Dataset

The optimal distance threshold for *Longitarsus* was estimated at 5.4% (Fig. 12). Ten singleton sequences have been identified and removed [*L. apicalis* (Beck, 1817), *L. fallax* Weise, 1888, *L. fulgens* (Foudras, 1860), *L. linnaei* (Duftschmid, 1825), *L. nanus* (Foudras, 1860), *L. niger* (Koch, 1803), *L. nigripennis* Motschulsky, 1866, *L. rubellus* (Foudras, 1860), *L. saulicus* Gruev & Döberl, 2005 and *L. vilis* Wollaston, 1864]. Once the optimal threshold has been set and the singletons removed, the barcoding identification efficiency of *Longitarsus* evaluated through the best close match analysis, increased to 96% of correct identification (1442 out of 1492). Remaining sequences resulted in 36 ambiguous sequences, 11 incorrect sequences and a net decrease in the number of no id sequences (3 sequences) (Tab. 3). Consistent with the other analyses, following threshold optimization, TaxCI results show a decrease in heterospecific clusters, from 28 of the original dataset, to 16 cases (Tab. 4 and Supplementary S7).

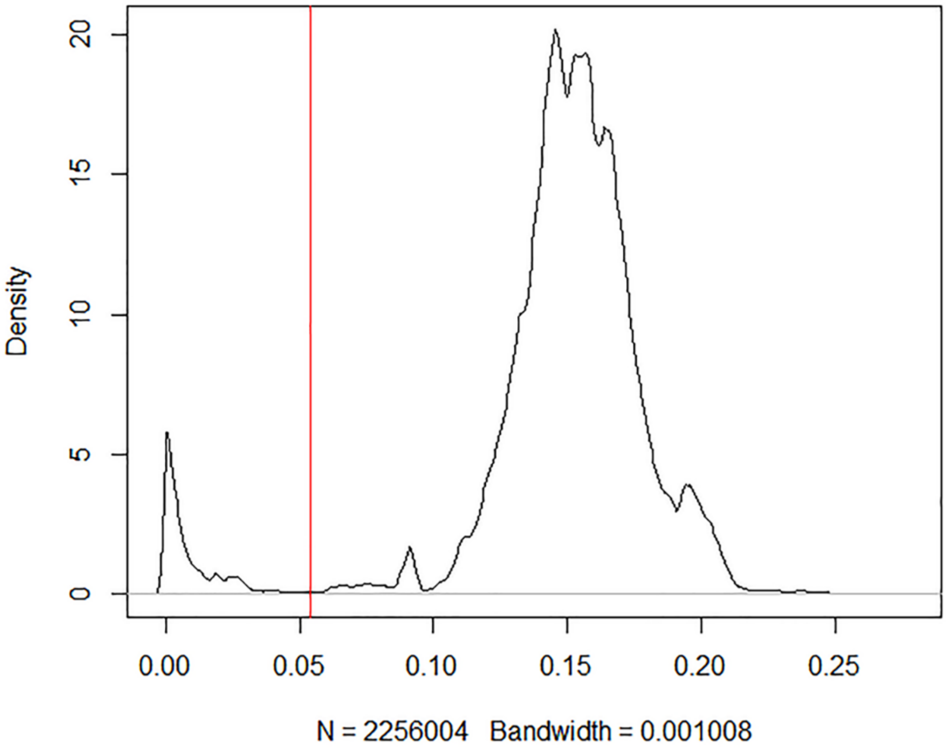


Fig.12 Results of the threshold optimisation analysis. The best threshold is identified as the dip in the density of genetic distances that indicates a transition between intra- and inter-specific distances. The optimised threshold is indicated by the red vertical line.

Final dataset

As a last step, available voucher material relative to ambiguous and incorrect identification sequences was assessed by Maurizio Biondi to confirm or not the identification error. Thanks to this procedure we were able to identify at least 69 incorrect specimen identifications. Among these, based on morphological assessment of voucher materials we identified several *L. juncicola* (Foudras, 1860) that had been wrongly identified as *L. ordinatus* (Foudras, 1860).Three outlier sequences belonging to *L. atricillus*, *L. salviae* Gruev, 1975 and *L. ochroleucus* (Marsham, 1802) were removed from the final dataset because they failed taxonomic validation. These three sequences have a large genetic divergence relative to conspecific sequences (*L. atricillus*: 6.3%; *L. ochroleucus*: 15,2%; *L. salviae*: 17,2%). The sequences of *L. ochroleucus* and *L. salviae* do not cluster with conspecific sequences, but rather they form singletons (a single branch with no affinity to other species) suggesting ambiguous identification (*sensu* Meier et al. (2006)); the sequence of *L. atricillus* clusters within an allospecific clade (within the *L. aeneicollis* clade), suggesting a misidentification (*sensu* Meier et al. (2006)). On the other hand, all remaining sequences of *L. atricillus*, *L. salviae*, *L. ochroleucus* and *L. aeneicollis* form well-defined and homogeneous clusters and their identification was validated by our sequenced vouchers. After the taxonomic revision step, barcoding gap analysis, best close match analysis and TaxCI analysis were repeated to verify if the elimination and correction of erroneous sequences improved the identification accuracy of the barcoding tool. The difference between the *maximum intraspecific distance* and the *minimum interspecific distance* calculated for the remaining 1489 sequences indicates 361 cases in which the barcoding gap is not present (24% of sequences) (Fig. 11). Overall, there was an average reduction in the absence of barcode gaps per species (24%). Furthermore, the barcoding identification efficiency, evaluated through the best close match analysis, resulted in 98.1% of correct identifications (1460 out of 1489). Most of the ambiguous and incorrect sequences (93%) belong to the *L. pratensis* group. The remaining cases regard two sequences of *L. bedeli* Uhagon, 1887; this species according to (Baselga et al. (2013), at the mitochondrial DNA level is not differentiated from *L. atricillus* (Linnaeus, 1761) (Tab. 3). Also TaxCI analysis found an improvement in the taxonomic consistency of the final dataset, with 13 heterospecific clusters. The *L. pratensis* group represented 48% of the sequences belonging to non-monophyletic species. Seven species are identified as heterogeneous distance-based clusters by TaxCI: *L. aeneicollis*, *L. atricillus*, *L. bedeli*, *L. erro*, *L. kutscherae* (Rye, 1872), *L. melanocephalus* (Geer, 1775) and *L. nasturtii* (Fabricius, 1793) (Tab. 4 and Supplementary S8). The significant increase of sequences per species identified as correct after taxonomic revision was confirmed by ANOVA results. While the increase of correct identifications from the original dataset to the dataset with the set threshold (*F*-value =2.138, *p*-value=0.159) was not significant, the increase of correct identifications from the original dataset to the final dataset was statistically significant (*F*-value =3.38, *p*-value<0.05) (Fig. 13).

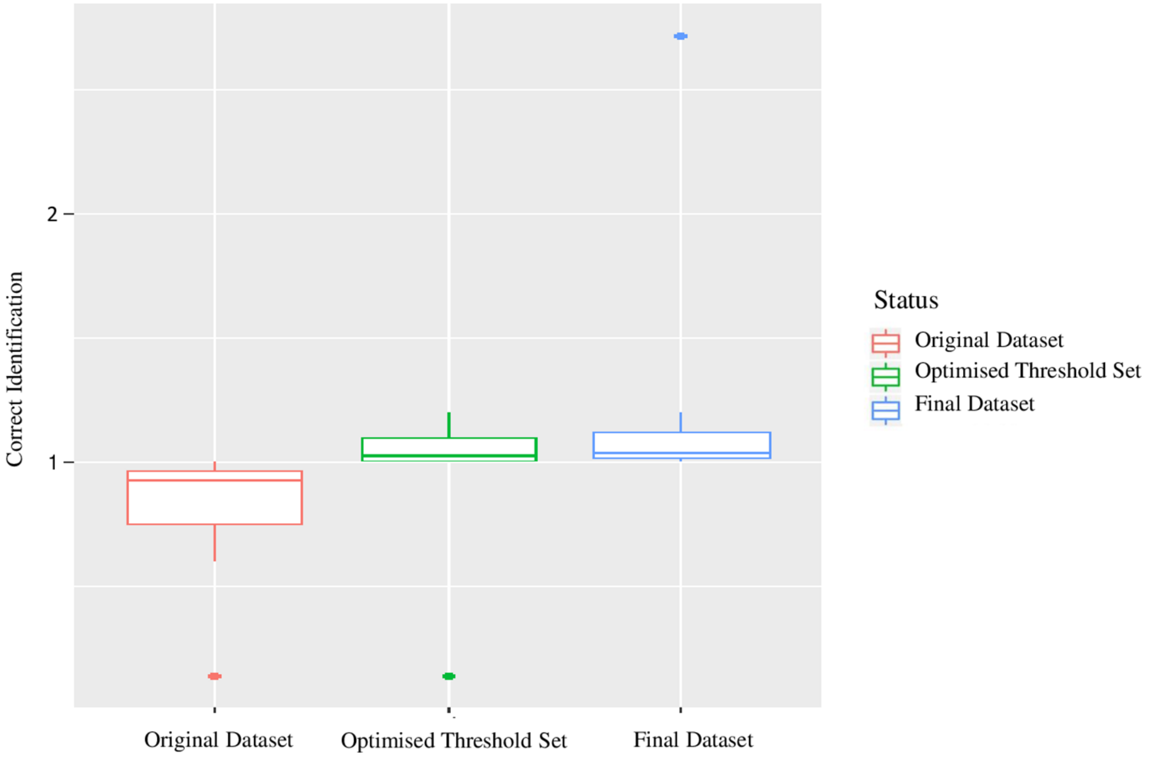


Fig.13 Results of the ANOVA analyses comparing the correct species identification ratio of the original dataset (red), the optimised threshold dataset (green), and the final dataset (blue).

Discussion

DNA barcoding is a molecular tool for species identification, and like for any tool, it is essential to know its potential as much as its limits. In this study we focused on identifying errors that affect the accuracy of DNA barcoding, distinguishing between tool *extrinsic errors,* i.e. those relative to the quality of the reference dataset, and *intrinsic errors,* i.e. those due to all those biological processes that generate a mismatch between mtDNA groups and species boundaries, thus making the barcoding tool unreliable in identifying specimens to the species level. While *intrinsic errors* to be solved require an integrative taxonomic and evolutionary study approach, which goes beyond the idea of barcoding as identification tool, *extrinsic errors* are due to human mistakes and can be corrected much more easily. Thus, in a reference dataset free of *extrinsic errors* it would be easy to spot species identification inconsistencies that require further taxonomic research.

In this study we identified the *extrinsic errors* occurring in the available *cox1* sequence dataset of the taxonomically complex genus *Longitarsus*. Barcoding gap analyses of this dataset showed several instances of overlap between intra- and interspecific genetic distance within this genus. The use of an *ad hoc* distance threshold, optimised for this dataset, resulted in an improvement of the quality of identification in agreement with previous studies (Lefébure et al., 2006; Puillandre et al., 2012; Sonet et al., 2013). However, the use of an optimal threshold did not significantly reduce the taxonomic uncertainty of the barcoding tool that was mostly associated to the *extrinsic errors* occurring in the reference datasets. These kinds of errors were readily identified using bioinformatic pipelines such as TaxCI, amended through a taxonomic revision carried out by the group specialist, and implemented in the reference database. The correct assignment of these misidentified sequences significantly increased the barcoding identification accuracy up to 98.1%. This identification rate is comparable to that found in other studies on Alticini (Coral Şahin et al., 2019) or Chrysomelidae (Magoga et al., 2018), showing the utility of DNA barcoding as molecular identification tool of taxonomically diverse groups.

Once the *extrinsic errors* were removed, we have been able to identify those areas of the dataset affected by *intrinsic errors*. Taxonomic uncertainty within the *L. pratensis* species group account for 93% of such *intrinsic errors*. This group is represented in the dataset by *L. pratensis* (Panzer, 1794), *L. scutellaris* (Rey, 1874) and *L. reichei* (Allard, 1860). All the analyses showed that specimens assigned to these species are genetically undifferentiated one each other. The high morphological similarity of these species and their sympatric distribution makes them extremely difficult to identify (Gruev and Merkl, 1992; Warchałowski, 2013). Species boundaries within this group are not well defined due to the lack of a comprehensive and integrative taxonomic assessment combining morphological and molecular approaches. The remaining *intrinsic errors* identified in this study regard the species pair *L. atricillus* and *L. bedeli* and has been already discussed in a previous work (Baselga et al., 2013). These two species show morphological differences on elytral coloration and female genitalia, with no morphological intermediates, but mtDNA does not detect any distinguishable phylogenetic structure that allows to separate these two species (Baselga et al., 2013). Also in this case an integrative approach will be required to reach a firm taxonomic conclusion; the use of multiple nuclear loci would allow disentangling lineage sorting or mitochondrial introgression as the processes responsible for the observed mitochondrial pattern.

On the other hand, we identified different species that, despite being monophyletic in the TaxCI analyses, are characterized by (i) a high intraspecific divergence such as *L. pinguis* Weise, 1888, *L. parvulus* (Paykull, 1799), *L. lateripunctatus* Rosenhauer, 1856 and *L. lycopi* (Foudras, 1860); or by (ii) a low interspecific divergence such as between, *L. nasturtii* and *L. erro*, *L. atricillus* and *L. aeneicollis*. (i) In *L. pinguis* high genetic distance was observed between specimens collected in northern Italy (Lombardia region) and specimens from central Italy. As for *L. parvulus*, high genetic distance is found between the unique Greek specimen and specimens from central-western Europe. In *L. lateripunctatus* a high genetic distance is observed between specimens from the opposite sides of the Apennine mountains in central Italy. For these three species the high genetic distance seems to be associated to a geographic structure. Instead, genetic variation within *L. lycopi*, does not seem to have geographical structure. (ii) The species within the two pairs *L. nasturtii*/*L. erro* and *L. atricillus*/*L. aeneicollis* form two reciprocally monophyletic sister clades with limited genetic distance, suggesting a recent divergence. This analysis also confirms the monophyly of *L. ochroleucus lindbergi* (Madar, 1963) within the *Longitarsus ochroleucus* clade, supporting the validity of this subspecies that is endemic to Madeira (Portugal).

Moreover, TaxCI results identified some non-monophyletic species that deserve taxonomic attention. Sequences belonging to *L. kutscherae* (Rye, 1872) are nested within the clade of *L. melanocephalus*. These two species are morphologically very similar and have a sympatric distribution (Warchałowski, 2013). However, due to the absence of metadata associated to these sequences, we were unable to verify whether this phylogenetic pattern is due to an incorrect specimen identification, to a poorly established taxonomy of these species, or because of any biological processes causing *intrinsic* type errors. *L. minusculus* (Foudras, 1860), *L. nigrofasciatus* (Goeze, 1777) and *L. erro* are polyphyletic species. All these species present a large distribution and the reasons for the absence of monophyly can be manifold and should be explored.

The importance of a dataset free of *extrinsic error* for the accuracy of the DNA barcoding tool cannot be overstated. Depositing only high-quality sequences correctly annotated with correct species names in public repositories would be the "golden standard" and is crucial for keeping the global barcode library functional and reliable (Floyd et al., 2002; Hajibabaei et al., 2006; Rulik et al., 2017). The high number of taxonomists required to avoid any error in morphological identification of species should not be an impediment of large-scale DNA barcoding campaigns (Adamowicz et al., 2019), especially in a period of risk for biodiversity that calls for a rapid assessment of species identification. On the other hand, this should not coincide with the risk of large but low-quality data production, thus it is fundamental maintaining a standard that allows a posteriori verification of identifications via morphological analysis (Ebach and Holdrege, 2005; Porco et al., 2010; Lehmann et al., 2017). In this regard we proved the efficacy of a non-invasive DNA extraction protocol that allows successful amplification of the barcoding gene fragment in flea beetle specimens as small as 1.5 mm. Using this non-invasive extraction methods has allowed us to maintain a reference voucher sample for future taxonomic assessments.

In conclusion, results of this study show that while taxonomic inconsistencies in reference sequence databases greatly affect the DNA barcoding accuracy, they can be readily identified using bioinformatic pipelines, and resolved through a posteriori re-assessment by an expert taxonomist based on available metadata, vouchers, or newly generated sequences (Lipscomb et al., 2003; Stoeckle, 2003; Ebach and Holdrege, 2005). Once again, this study underlines the key role of taxonomists in any step of the DNA barcoding pipeline, from the initial association of DNA sequences with morphologically identified species to the *a posteriori* revision of the inconsistencies identified in the reference database. Furthermore, such step of taxonomic revision on existing data allows identifying hot research areas for *Longitarsus* taxonomy, further corroborating the intimate link between the accuracy of the DNA barcoding tool and taxonomic knowledge.

Reference

Adamowicz, S.J., Boatwright, J.S., Chain, F., Fisher, B.L., Hogg, I.D., Leese, F., Lijtmaer, D.A., Mwale, M., Naaum, A.M., Pochon, X., 2019. Trends in DNA barcoding and metabarcoding. Genome 62, v–viii.

Arnold, J., 1993. Cytonuclear disequilibria in hybrid zones. Annu. Rev. Ecol. Syst. 24, 521–553.

Bacher, S., 2012. Still not enough taxonomists: reply to Joppa et al. Trends Ecol. Evol. 27, 65–66.

Barrett, R.D., Hebert, P.D., 2005. Identifying spiders through DNA barcodes. Can. J. Zool. 83, 481–491.

Baselga, A., Gómez-Rodríguez, C., Novoa, F., Vogler, A.P., 2013. Rare Failures of DNA Bar Codes to Separate Morphologically Distinct Species in a Biodiversity Survey of Iberian Leaf Beetles. PLoS ONE 8, e74854. https://doi.org/10.1371/journal.pone.0074854

Becker, R., Chambers, J., Wilks, A., 1998. The new S language: A Programing Environment for Data Analysis and Graphics (revised ed.).

Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M.T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G.N., Ribera, I., Nilsson, A.N., Barraclough, T.G., Vogler, A.P., 2012. The Effect of Geographical Scale of Sampling on DNA Barcoding. Syst. Biol. 61, 851–869. https://doi.org/10.1093/sysbio/sys037

Biondi, M., D'alessandro, P., 2008. Taxonomical revision of the *Longitarsus* capensis species-group: An example of Mediterranean-southern African disjunct distributions (Coleoptera: Chrysomelidae). Eur. J. Entomol. 105.

Bortolus, A., 2008. Error Cascades in the Biological Sciences: The Unwanted Consequences of Using Bad Taxonomy in Ecology. AMBIO J. Hum. Environ. 37, 114–118. https://doi.org/10.1579/0044-7447(2008)37[114:ECITBS]2.0.CO;2

Brown, S.D., Collins, R.A., Boyer, S., Lefort, M., Malumbres-Olarte, J., Vink, C.J., Cruickshank, R.H., 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol. Ecol. Resour. 12, 562–565.

Castalanelli, M.A., Severtson, D.L., Brumley, C.J., Szito, A., Foottit, R.G., Grimm, M., Munyard, K., Groth, D.M., 2010. A rapid non-destructive DNA extraction method for insects and other arthropods. J. Asia-Pac. Entomol. 13, 243–248.

Collins, R.A., Cruickshank, R.H., 2012. The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. n/a-n/a. https://doi.org/10.1111/1755-0998.12046

Coral Şahin, D., Magoga, G., Özdikmen, H., Montagna, M., 2019. DNA Barcoding as useful tool to identify crop pest flea beetles of Turkey. J. Appl. Entomol. 143, 105–117.

DeSalle, R., 2006. Species Discovery versus Species Identification in DNA Barcoding Efforts: Response to Rubinoff. Conserv. Biol. 20, 1545–1547. https://doi.org/10.1111/j.1523-1739.2006.00543.x

DeSalle, R., Egan, M.G., Siddall, M., 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos. Trans. R. Soc. B Biol. Sci. 360, 1905–1916. https://doi.org/10.1098/rstb.2005.1722

Ebach, M.C., Holdrege, C., 2005. More Taxonomy, Not DNA Barcoding. BioScience 55, 823. https://doi.org/10.1641/0006-3568(2005)055[0823:MTNDB]2.0.CO;2

Ebach, M.C., Williams, D.M., Morrone, J.J., 2006. Paraphyly is bad taxonomy. Taxon 55, 831–832.

Floyd, R., Abebe, E., Papert, A., Blaxter, M., 2002. Molecular barcodes for soil nematode identification. Mol. Ecol. 11, 839–850. https://doi.org/10.1046/j.1365-294X.2002.01485.x

Floyd, R., Lima, J., deWaard, J., Humble, L., Hanner, R., 2010. Common goals: policy implications of DNA barcoding as a protocol for identification of arthropod pests. Biol. Invasions 12, 2947–2954. https://doi.org/10.1007/s10530-010-9709-8

Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27, 480–488. https://doi.org/10.1016/j.tree.2012.04.012

Funk, D.J., Omland, K.E., 2003. Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421

Goldstein, P.Z., DeSalle, R., 2011. Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description. BioEssays 33, 135–147. https://doi.org/10.1002/bies.201000036

Gruev, B., Merkl, O., 1992. To the geographic distribution of the *Longitarsus* pratensis-group (Coleoptera, Chrysomelidae: Alticinae). Folia Ent Hung 52, 15–20.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W., Hebert, P.D., 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proc. Natl. Acad. Sci. 103, 968–971.

Hamilton, A.J., Basset, Y., Benke, K.K., Grimbacher, P.S., Miller, S.E., Novotný, V., Samuelson, G.A., Stork, N.E., Weiblen, G.D., Yen, J.D., 2010. Quantifying uncertainty in estimation of tropical arthropod species richness. Am. Nat. 176, 90–95.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M., 2004. Identification of Birds through DNA Barcodes. PLoS Biol. 2, e312. https://doi.org/10.1371/journal.pbio.0020312

Huber, J.T., 1998. The importance of voucher specimens, with practical guidelines for preserving specimens of the major invertebrate phyla for identification. J. Nat. Hist. 32, 367–385. https://doi.org/10.1080/00222939800770191

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Klopfstein, S., Kropf, C., Baur, H., 2016. Wolbachia endosymbionts distort DNA barcoding in the parasitoid wasp genus *Diplazon* (Hymenoptera: Ichneumonidae). Zool. J. Linn. Soc. 177, 541–557.

Kozlov, A.M., Zhang, J., Yilmaz, P., Glöckner, F.O., Stamatakis, A., 2016. Phylogeny-aware identification and correction of taxonomically mislabeled sequences. Nucleic Acids Res. 44, 5022–5033.

Kvist, S., 2013. Barcoding in the dark?: A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. Mol. Phylogenet. Evol. 69, 39–45. https://doi.org/10.1016/j.ympev.2013.05.012

Lefébure, T., Douady, C.J., Gouy, M., Gibert, J., 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. Mol. Phylogenet. Evol. 40, 435–447. https://doi.org/10.1016/j.ympev.2006.03.014

Lehmann, A., Devriese, H., Tumbrinck, J., Skejo, J., Lehmann, G.U.C., Hochkirch, A., 2017. The importance of validated alpha taxonomy for phylogenetic and DNA barcoding studies: a comment on species identification of pygmy grasshoppers (Orthoptera, Tetrigidae). ZooKeys 679, 139–144. https://doi.org/10.3897/zookeys.679.12507

Lipscomb, D., Platnick, N., Wheeler, Q., 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends Ecol. Evol. 18, 65–66. https://doi.org/10.1016/S0169-5347(02)00060-5

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55, 21–30.

Magoga, G., Sahin, D.C., Fontaneto, D., Montagna, M., 2018. Barcoding of Chrysomelidae of Euro-Mediterranean area: efficiency and problematic species. Sci. Rep. 8, 1–9.

Mastrantonio, V., Porretta, D., Urbanelli, S., Crasta, G., Nascetti, G., 2016. Dynamics of mtDNA introgression during species range expansion: insights from an experimental longitudinal study. Sci. Rep. 6, 30355. https://doi.org/10.1038/srep30355

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., Hugenholtz, P., 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 6, 610–618.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L., 2006. DNA Barcoding and Taxonomy in Diptera: A Tale of High Intraspecific Variability and Low Identification Success. Syst. Biol. 55, 715–728. https://doi.org/10.1080/10635150600969864

Meyer, C.P., Paulay, G., 2005. DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLoS Biol. 3, e422. https://doi.org/10.1371/journal.pbio.0030422

Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J.G., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated Species Inventory on Madagascar Using Coalescent-Based Models of Species Delineation. Syst. Biol. 58, 298–311. https://doi.org/10.1093/sysbio/syp027

Moritz, C., Cicero, C., 2004. DNA Barcoding: Promise and Pitfalls. PLoS Biol. 2, e354. https://doi.org/10.1371/journal.pbio.0020354

Nadein, K., Betz, O., 2016. Jumping mechanisms and performance in beetles. I. Flea beetles (Coleoptera: Chrysomelidae: Alticini). J. Exp. Biol. 219, 2015–2027.

Packer, L., Monckton, S.K., Onuferko, T.M., Ferrari, R.R., 2018. Validating taxonomic identifications in entomological research. Insect Conserv. Divers. 11, 1–12. https://doi.org/10.1111/icad.12284

Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528.

Phillips, A.J., Simon, C., 1995. Simple, Efficient, and Nondestructive DNA Extraction Protocol for Arthropods. Ann. Entomol. Soc. Am. 88, 281–283. https://doi.org/10.1093/aesa/88.3.281

Porco, D., Rougerie, R., Deharveng, L., Hebert, P., 2010. Coupling non-destructive DNA extraction and voucher retrieval for small soft-bodied Arthropods in a high-throughput context: the example of Collembola. Mol. Ecol. Resour. 10, 942–945. https://doi.org/10.1111/j.1755-0998.2010.2839.x

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation: ABGD, AUTOMATIC BARCODE GAP DISCOVERY. Mol. Ecol. 21, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Ratnasingham, S., Hebert, P.D., 2013. A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. PloS One 8, e66213.

Ratnasingham, S., Hebert, P.D., 2007. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). Mol. Ecol. Notes 7, 355–364.

Rowley, D.L., Coddington, J.A., Gates, M.W., Norrbom, A.L., Ochoa, R.A., Vandenberg, N.J., Greenstone, M.H., 2007. Vouchering DNA-barcoded specimens: Test of a nondestructive extraction protocol for terrestrial arthropods. Mol. Ecol. Notes 7, 915–924.

Rubinoff, D., 2006. Essays: Utility of Mitochondrial DNA Barcodes in Species Conservation: DNA Barcodes and Conservation. Conserv. Biol. 20, 1026–1033. https://doi.org/10.1111/j.1523-1739.2006.00372.x

Rulik, B., Eberle, J., Mark, L., Thormann, J., Jung, M., Köhler, F., Apfel, W., Weigel, A., Kopetz, A., Köhler, J., Fritzlar, F., Hartmann, M., Hadulla, K., Schmidt, J., Hörren, T., Krebs, D., Theves, F., Eulitz, U., Skale, A., Rohwedder, D., Kleeberg, A., Astrin, J.J., Geiger, M.F., Wägele, J.W., Grobe, P., Ahrens, D., 2017. Using taxonomic consistency with semi-automated data pre-processing for high quality DNA barcodes. Methods Ecol. Evol. 8, 1878–1887. https://doi.org/10.1111/2041-210X.12824

Salvi, D., D'Alessandro, P., Biondi, M., 2019a. Host plant associations in Western Palaearctic *Longitarsus* flea beetles (Chrysomelidae, Galerucinae, Alticini): a preliminary phylogenetic assessment. ZooKeys 856, 101.

Salvi, D., Maura, M., Pan, Z., Bologna, M.A., 2019b. Phylogenetic systematics of *Mylabris* blister beetles (Coleoptera, Meloidae): a molecular assessment using species trees and total evidence. Cladistics 35, 243–268.

Schilthuizen, M., Vairappan, C.S., Slade, E.M., Mann, D.J., Miller, J.A., 2015. Specimens as primary data: museums and 'open science.' Trends Ecol. Evol. 30, 237–238. https://doi.org/10.1016/j.tree.2015.03.002

Schindel, D.E., Miller, S.E., 2005. DNA barcoding a useful tool for taxonomists. Nature 435, 17–17. https://doi.org/10.1038/435017b

Smith, M.A., Bertrand, C., Crosby, K., Eveleigh, E.S., Fernandez-Triana, J., Fisher, B.L., Gibbs, J., Hajibabaei, M., Hallwachs, W., Hind, K., Hrcek, J., Huang, D.-W., Janda, M., Janzen, D.H., Li, Y., Miller, S.E., Packer, L., Quicke, D., Ratnasingham, S., Rodriguez, J., Rougerie, R., Shaw, M.R., Sheffield, C., Stahlhut, J.K., Steinke, D., Whitfield, J., Wood, M., Zhou, X., 2012. Wolbachia and DNA Barcoding Insects: Patterns, Potential, and Problems. PLoS ONE 7, e36514. https://doi.org/10.1371/journal.pone.0036514

Sonet, G., Jordaens, K., Nagy, Z.T., Breman, F.C., De Meyer, M., Backeljau, T., Virgilio, M., 2013. Adhoc: an R package to calculate ad hoc distance thresholds for DNA barcoding identification. ZooKeys 329.

Stoeckle, M., 2003. Taxonomy, DNA, and the Bar Code of Life. BioScience 53, 796. https://doi.org/10.1641/0006-3568(2003)053[0796:TDATBC]2.0.CO;2

Vink, C.J., Paquin, P., Cruickshank, R.H., 2012. Taxonomy and irreproducible biological science. BioScience 62, 451–452.

Virgilio, M., Backeljau, T., Nevado, B., Meyer, M.D., 2010. RCesoeamrchpartaicrleative performances of DNA barcoding across insect orders 10.

Virgilio, M., Jordaens, K., Breman, F.C., Backeljau, T., De Meyer, M., 2012. Identifying Insects with Incomplete DNA Barcode Libraries, African Fruit Flies (Diptera: Tephritidae) as a Test Case. PLoS ONE 7, e31581. https://doi.org/10.1371/journal.pone.0031581

Vogler, A.P., Monaghan, M.T., 2007. Recent advances in DNA taxonomy. J. Zool. Syst. Evol. Res. 45, 1–10. https://doi.org/10.1111/j.1439-0469.2006.00384.x

Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3, 294–9.

Walesiak, M., Dudek, A., 2020. The choice of variable normalization method in cluster analysis. Educ. Excell. Innov. Manag. A 2025, 325–340.

Warchałowski, A., 2013. The palaearctic Chrysomelidae identification keys, Volume 1 and 2. Natura optima dux Foundation.

Weber, A.A.-T., Stöhr, S., Chenuil, A., 2019. Species delimitation in the presence of strong incomplete lineage sorting and hybridization: Lessons from *Ophioderma* (Ophiuroidea: Echinodermata). Mol. Phylogenet. Evol. 131, 138–148. https://doi.org/10.1016/j.ympev.2018.11.014

Wickham, H., 2016. ggplot2: elegant graphics for data analysis. springer.

Wiemers, M., Fiedler, K., 2007. Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). Front. Zool. 4, 8. https://doi.org/10.1186/1742-9994-4-8

Will, K.W., Mishler, B.D., Wheeler, Q.D., 2005. The Perils of DNA Barcoding and the Need for Integrative Taxonomy. Syst. Biol. 54, 844–851. https://doi.org/10.1080/10635150500354878

Wilson, C.C., Bernatchez, L., 1998. The ghost of hybrids past: fixation of arctic charr (Salvelinus alpinus) mitochondrial DNA in an introgressed population of lake trout (S. namaycush). Mol. Ecol. 7, 127–132.

Supporting information

This chapter of the thesis was publish on PlosOne: https://doi.org/10.1371/journal.pone.0233573

All supporting information are accessible on https://doi.org/10.1371/journal.pone.0233573

List of cox1 *Longitarsus* sequences retrieved from GenBank and BOLD and used in this study.

Showing 1/11: pone.0233573.s001.pdf

S1 Table. List of cox1 *Longitarsus* sequences retrieved from GenBank and BOLD and used in this study.

10.1371/journal.pone.0233573.s001

S2 Table. List of specimens sequenced in this study.

For each specimen are reported voucher info, place and date of collection, coordinates, collectors, BOLD and GenBank accession numbers. All specimens have been identified by Maurizio Biondi (University of L'Aquila).

10.1371/journal.pone.0233573.s002

S1 Fig. Photographs of habitus and aedeagus or spermatheca of (a) *Longitarsus albineus* ♂; (b) *L. exsoletus* ♂; (c) *L. juncicola* ♂; (d) *L. ochroleucus lindbergi* ♂; (e) *L. laureolae* ♂; (f) *L. luridus* ♂. Scale bar 0.5 mm.

10.1371/journal.pone.0233573.s003

S2 Fig. Photographs of habitus and aedeagus or spermatheca of (a) *Longitarsus ordinatus* ♂; (b) *L. nigrofasciatus* ♂; (c) *L. melanocephalus* ♂; (d) *L. pinguis* ♂; (e) *L. parvulus* ♂; (f) *L. rectilineatus* ♂. Scale bar 0.5 mm.

10.1371/journal.pone.0233573.s004

S3 Fig. Photographs of habitus and aedeagus or spermatheca of (a) *Longitarsus salviae* ♂; (b) *L. strigicollis* ♂; (c) *L. springeri* ♂; (d) *L. succineus* ♂; (e) *L. tabidus* ♂; (f) *L. zangherii* ♂. Scale bar 0.5 mm.

10.1371/journal.pone.0233573.s005

S4 Fig. Linear regression models showing (a) the relationships between yield of DNA extraction and time of permanence in alcohol of the specimens, tp; and (b) the body size of the specimens, bs. For each species two specimens were selected and DNA extracted either with the invasive method, IM, or with the non-invasive method, NIM (bs-IM: R2 = 0.1053, p-value = 0.1407; bs-NIM: R2 = 0.0155, p-value = 0.581; tp-IM: R2 = 0.07261, p-value = 0.2252; tp-NIM: R2 = 0.1191, p-value = 0.1156).

10.1371/journal.pone.0233573.s006

S5 Fig. Comparison of DNA extraction yield, amplification and sequencing success between the invasive DNA extraction method, IM, and the non-invasive DNA extraction method, NIM. (a) ANOVA results comparing the final amount of the DNA extracted with the two methods. (b) Amplification success and (c) Sanger sequencing success of the cox1 gene fragment using DNA templates obtained with the two DNA extraction methods. Bands obtained with DNA templates extracted with the NIM are marked on the electrophoretic gel. Sequencing success rate for PCR products obtained using DNA templates extracted with the two methods: IM, 90% good, 10% poor and 0% failed; NIM, 91% good, 6% poor and 2% failed.

10.1371/journal.pone.0233573.s007

S6 Fig. Tree output from TaxCI analysis on the original dataset.

10.1371/journal.pone.0233573.s008

S7 Fig. Tree output from TaxCI analysis on the optimised threshold dataset.

10.1371/journal.pone.0233573.s009

S8 Fig. Tree output from TaxCI analysis on the final dataset.

10.1371/journal.pone.0233573.s010

S1 Raw images.

10.1371/journal.pone.0233573.s011

CHAPTER 4

Case study II

Cryptic, sibling, or neither of the two? Systematics and taxonomic identification of *Psylliodes ruffoi and P. kiesenwetteri* (Chrysomelidae, Galerucinae, Alticini) based on a combined morphological and phylogeographic assessment

Introduction

Species discovery and description is crucial for documenting biodiversity patterns and understanding processes underlying the diversification of organisms on Earth (Agapow et al., 2004; Schlick-Steiner et al., 2010). For centuries comparative morphology has played a fundamental role in the taxonomy of plants and animals (Wheeler, 2007; Padial et al., 2010). However, relying only on morphological characters presents shortcomings that, in many cases, prevent the delimitation and, consequently, the identification of species (Xiao et al., 2010; Montagna et al., 2017). Indeed, various biological processes such as recent origin of species (Egea et al., 2016), extensive intraspecific variability (Memon et al., 2006), hybridization (Mallet, 2007) or convergent evolution (Nevo, 2001), make morphological characters not predictive of species boundaries (Bickford et al., 2007). In the last twenty years, the increasing use of integrative taxonomic approaches, which combine different lines of evidence (e.g. genetic, morphological, ecological) and methodologies (e.g. phylogenetic inference, ecological models), has led to significant advances in species discovery and delimitation (Sites Jr and Marshall, 2003; Dayrat, 2005; Schlick-Steiner et al., 2010).

Irrespective of the approach used for species delimitation, species description on the basis of intrinsic characters, typically morphological, is a well-established taxonomic practice (ICZN, 1999; Bauer et al., 2011). This procedure provides a comparative framework for the identification of specimens and is particularly effective when the characters used have a discrete, not-overlapping, variation across closely related species and when the examined material is representative of the intraspecific variation at these characters. However, given the complex biological reality of species these expectations are not always met. For example, complexes of species which are morphologically very similar, challenge this approach due to the limited availability (or the lack in the case of cryptic species) of diagnostic variation to serve for species discrimination (Bickford et al., 2007). In other cases, characters used for the species diagnosis, and assumed to have a discrete variation, when further assessed across a wider geographic and population scope they revealed a pattern of continuous variation that is overlapping with closely related species. This is especially likely for wide-ranging species whose descriptions are based on a few individuals from a low number of localities covering a limited portion of the species range (Tessens et al., 2021). This is quite a common case for species described in the old times or from poorly sampled areas (McBride et al., 2009; Darwell and Cook, 2017; Deng et al., 2019; Morek et al., 2019). Therefore, to overcome these potential difficulties and uncertainties in species discrimination, not only an integrative taxonomic approach combining independent characters is required but also a comprehensive assessment of the geographic variation of these characters.

While combining molecular and morphological data is now a well-established taxonomic practice, it is far less common for these data to be assessed across comprehensive geographical scales. In this respect, a phylogeographic approach combined with morphological assessment offers several advantages as it allows delimiting independent evolutionary lineages and their ecogeographic distributions and delimit the spatial boundaries of intraspecific morphological variability (Raxworthy et al., 2007), offering further support for the delimitation of species (Espíndola et al., 2016; Sites Jr and Marshall, 2003; Struck et al., 2018; Templeton, 2001; Yeates et al., 2011). Moreover, this approach allows for a backward morphological taxonomy process to re-evaluate the morphology in light of phylogenetic lineages and to identify, even in complexes of species considered cryptic, morphological differences that have not previously emerged (Blanquer and Uriz, 2008; Ramasindrazana et al., 2011; Martinsson et al., 2015; Leavitt et al., 2016; Grebennikov, 2019; Jossart et al., 2021).

The case of *P. kiesenwetteri* and *P. ruffoi* (Coleoptera, Chrysomelidae) offers an ideal experimental setting for testing the utility of a combined integrative and phylogeographic approach for taxonomic delimitation and identification of morphologically similar and geographically overlapping species. These two species belongs to the *Psylliodes gibbosa* complex, a group of species distributed in the Mediterranean region, which has a long and eventful taxonomic history (Leonardi, 1975; Nadein, 2008). The member species of this complex are generally distinguishable only based on subtle morphological differences, mainly in the shape of the spermatheca and aedeagus and, in general, no single character taken individually can be considered sufficient for species identification (Leonardi, 1975). For this reason, the discrimination of species using a morphological approach is challenging, and misidentifications are frequent by inexperienced specialists. Whitin this species complex, populations belonging to *P. ruffoi* have been confused for a long time with *P. gibbosa* or *P. kiesenwetteri* (Leonardi, 1975). Morphological characters of *P. ruffoi* are somewhat intermediate between *P. kiesenwetteri* and *P. gibbosa*, with several traits being shared among all three species. Not only morphological variation seems overlapping among these species but also their geographic range show a wide area of overlap in south Italy. While recent molecular data support the phylogenetic distinction between *P. gibbosa* and *P. kiesenwetteri* (Gikonyo, 2021), the phylogenetic relationship between *P. kiesenwetteri* and *P. ruffoi* is still unclear. *Psylliodes ruffoi* and *P. kiesenwetteri* can be found in close sympatry along the mountain range of the south Apennines in association with semi-arid meadows in clearings of xerophilous woods. In locations where they are syntopic, it is very common to find individuals of both species on the same plant. Overlapping morphological variation and geographic range with strict syntopy on several mountains, cast the question on whether these taxa represent two species or a single polymorphic species. The clarification of their systematics and distinctiveness also has conservation implications since *Psylliodes ruffoi* represent an endemic species to southern Italy with a narrower range than *P. kiesenwetteri*.

The main aims of this study are to assess genetic and morphological differentiation between the two species in sympatric and allopatric populations, test for the hypothesis of a cryptic or sibling species complex, and assess the diagnostic value of morphological characters for their discrimination. By extending the taxon set to closely related species of the *Psylliodes gibbosa* complex we also want to provide a multi-locus phylogenetic framework for their systematics. Finally, the utility of a combined phylogeographic and morphological approach for the systematic assessment of species complex is discussed.

Materials and methods

Study System

The flea beetle genus *Psylliodes* Latreille comprises over 200 species worldwide (Gikonyo et al. 2019). Adults are easily distinguished from other flea beetle genera based on their 10-segmented antennae and tarsi inserted pre-apically on the metatibia of the hind legs. Most *Psylliodes* species have a restricted host plant range (35% are monophagous and 51% are oligophagous), and only 14% are polyphagous. Of all *Psylliodes* species with known host plants, 50% are specialised on Brassicaceae, followed by 13% feeding on Poaceae, 10% on Solanaceae and 10% on Fagaceae (Gikonyo et al. 2019).

The species-group of *Psylliodes gibbosa* was proposed by Leonardi (1970) based on the shape of the spermatheca, frontal tubercles and frontal lateral sulci, including the species *P. gibbosa* Allard, *P. kiesenwetteri* Kutschera (reported as *P. latifrons* Weise) and *P. inflata* Reiche. Later, Leonardi (1975) attributed to this species-group also the new species *P. ruffoi*, from southern Italy and Sicily, and *P. gougeleti* Allard, from the Iberian peninsula. Based on the morphological analysis of Nadein (2008), the *Psylliodes gibbosa* group includes eight species with a West Palaearctic distribution, divided into two subgroups: the *P. gibbosa* subgroup, composed of *P. gibbosa*, *P. kiesenwetteri*, *P. gougeleti*, and *P. ruffoi* Leonardi, occupies the northern part of the range and is present in France and in the Iberian, Italian, Balkan and Anatolian peninsulas; the *P. inflata* subgroup, composed of *P. inflata, P. fageli* Bechyné, *P. tenuidentatus* Nadein and *P. ridendus* Nadein, with a more southern distribution, is extended in North Africa, southern Mediterranean islands and in Middle East, from Morocco to Iran. Moreover, the Palaeartic species *Psylliodes cucullata* (Illiger), which was already considered similar to the *P. gibbosa* group based on morphological character (Nadein, 2008), has been confirmed to be a species belonging to the group on the basis of a recent molecular study (Gikonyo, 2021).
Four species belonging to the *Psylliodes gibbosa* group are present in Italy, namely: *P. gibbosa, P. kiesenwetteri, P. ruffoi* and *P. inflata*. While the latter is present only in Sicily and Sardinia, *Psylliodes gibbosa*, *P. kiesenwetteri*, and *P. ruffoi*, as reported above, share part of their range of distribution, with a close sympatry in the southern part of the peninsula (Fig. 14).

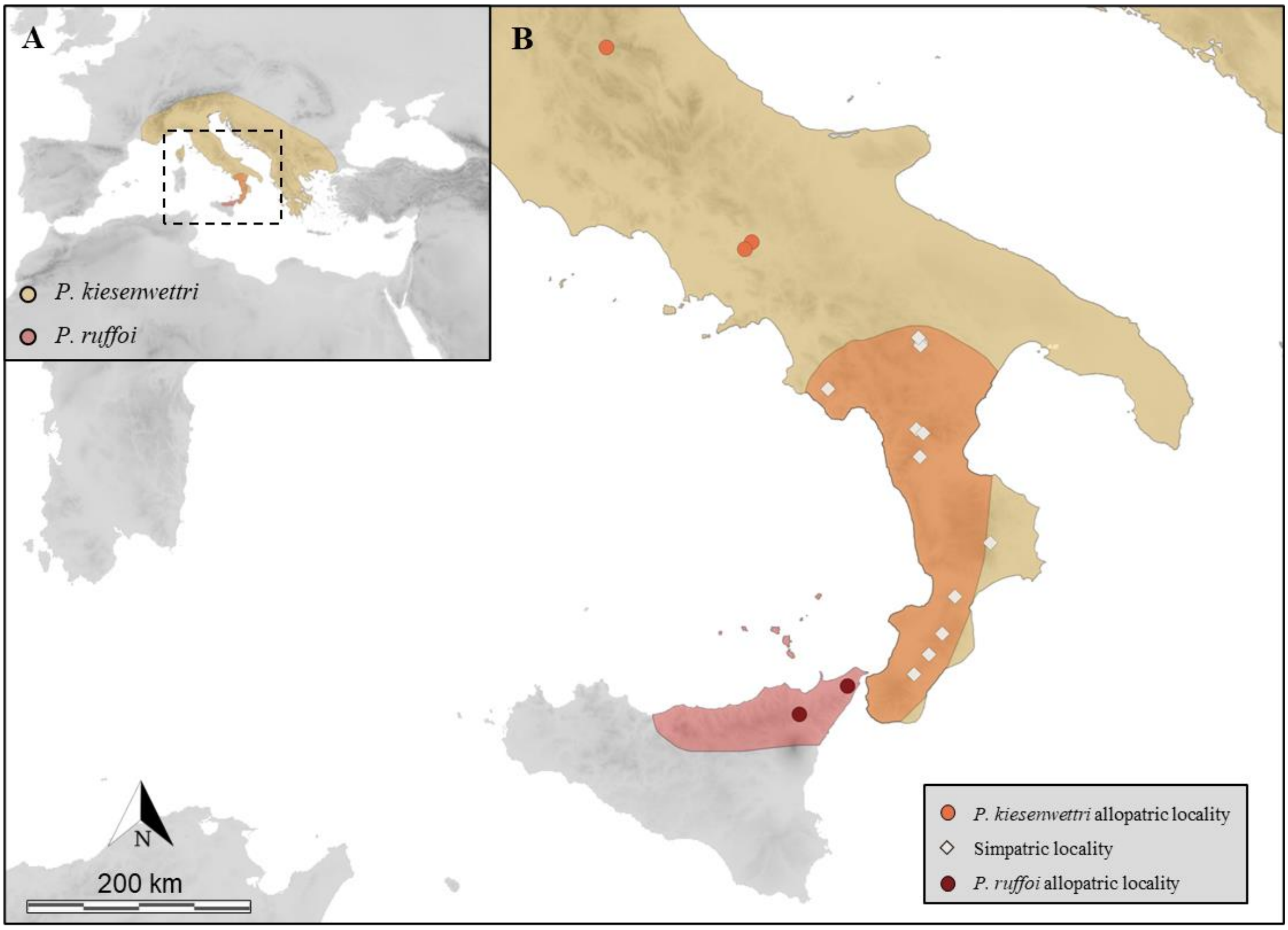


Fig.14 The geographical distribution of the two species in Europe (A), in yellow the range of *P. kiesenwetteri* and in red that of *P. ruffoi*. The dashed box is the focus on the study area (B), which shows the sampling locations in sympatry and allopatry of the two species.

Sampling

Specimens of *P. ruffoi* and *P. kiesenwetteri* were collected from 15 different locations in central and southern Italy along a transect of about 720 km, in order to cover the entire range of *P ruffoi* including sites of sympatry with *P kiesenwetteri* as well as in areas of allopatry between the two species. Sampling design was based on historic distribution data (Latella et al., 2005); material from entomological collections and bibliographic records (Baviera and Biondi, 2015; Biondi, 1988; Leonardi, 1975, 1970; Urbani et al., 2015). Specimens were collected from their host plant by sweep net and the aid of aspirator and then stored in 95% ethanol.

DNA Extraction, Amplification, and Sequencing

Total genomic DNA of 130 specimens was extracted using the protocol presented in Chapter 2. For these specimens we amplified the standard barcode region of the mitochondrial *cytochrome c oxidase I* (*cox1*) gene. For selected individuals of *P. ruffoi* and *P. kiesenwetteri*, one from sympatric and one from allopatric locality, and for one representant of *P. gibbosa* from Morocco, one additional mitochondrial gene fragment, the 3’ end fragment of *cox1*, and five protein-coding single copy nuclear genes from Gikonyo et al. (2021) were amplified: Carbamoylphosphate synthase (*CAD*), Crossveinless 2 (*Cv2*), Methyl methanesulfonate-sensitivity protein 22-like (*MMS22*), Rad 50 protein (*Rad50*) and Short gastrulation (*Sog*). These nuclear markers were selected based on their high evolutionary rate (Gikonyo et al. 2021). Two fragments of the mitochondrial gene cytochrome oxidase subunit I (*cox1*) were amplified using the conditions described in Salvi et al. (2019): the standard barcode region, using the universal primers LCO1490 and HC02198 (Folmer et al., n.d.), and the 3' fragment, using primers C1-J-2183 and TL2-N-3014 (Simon et al., 1994). The five protein-coding single copy nuclear genes were amplified using primers and amplification conditions described in Gikonyo (2021). Successful amplification was determined by gel electrophoresis and PCR products were purified and sequenced by an external service (Genewitz, UK). The obtained chromatograms of each sequence were manually edited and assembled into a consensus sequence using Geneious Prime 2021 (Biomatters Ltd., Auckland, New Zealand).

Molecular Species Delimitation

Newly generated sequences of the standard barcode region of *cox1* were aligned with MAFFT v7.450 using the G-INS-I progressive method algorithm (Katoh et al., 2002) together with one sequence of *P. kiesenwetteri* from Croatia with GenBank accession number MW254865 (Gikonyo et al., 2021).

Sequence divergences were calculated using the Kimura two parameter (K2P) distance model to explore the extent of overlap between intraspecific variation and interspecific divergence (Kimura, 1980; Nei and Kumar, 2000). A neighbor-joining (NJ) tree of K2P distances was created to provide a graphic representation of the patterning of divergence between species and a reference framework for species delimitation. In this tree building methods *Psylliodes vehemens* Wollaston, 1854 was used as outgroup based on phylogenetic affinity (Gikonio et al., 2021). The NJ tree was constructed using MEGAX (Kumar et al., 2018) with Kimura-two parameters (K2P) (Kimura, 1980). The bootstrap support for a NJ tree for the branch nodes was estimated through bootstrapping with 1000 replicates.

We performed two species delimitation analyses to infer the number of species clusters and their correspondence to morphospecies: (i) a nucleotide distance with thresholds estimated *ad hoc* on the data set and (ii) Assemble Species by Automatic Partitioning (ASAP) (Puillandre et al., 2021). Nucleotide distance-based methods resulting in high efficiency for species delimitation on Chrysomelidae (Magoga et al., 2021). The pairwise nucleotide distance matrices required for *ad hoc* nucleotide distance threshold methods were estimated using the R library *ape* v5.3 (Paradis and Schliep, 2019). A pairwise distance matrix of intraspecific and interspecific genetic distance was calculated using the K2P substitution model with the pairwise deletion option. With the R package *spider* v1.5.0 (Brown et al., 2012) we performed a threshold optimization analysis using the *localMinima* function. Finally, we cluster nucleotide sequences with the *ad hoc* threshold using the function *tclust* . ASAP analyses were run using the program web-interface (https://bioin fo.mnhn.fr/abi/publi c/asap); K2P was selected as nucleotide substitution model and other parameters were left as default.

Dataset and Sample preparation for morphological analysis

Morphological characters were assessed on a comprehensive set of specimens collected in localities from both areas of sympatry between *P. kiesenwetteri* and *P. ruffoi* and allopatric areas. Examined material consists of 20 ♂♂ and 20 ♀♀ for each species, previously extracted for DNA with non-invasive method. These samples were dissected and subsequently dried and mounted on an entomological card with visible median lobe of the aedeagus or spermatheca, this last one included in Euparal. The specimens were examined, measured, and dissected using a Leica M205C stereomicroscope. Photographs were taken using a Leica DMC5400 camera and composed using Zerene Stacker software version 1.04. Scanning electron micrographs were taken using a Hitachi TM- 1000.

Morphological assessment

First, we used the diagnostic characters proposed by Leonardi (1975) and Nadein (2008) for a preliminary species identification of the collected specimens. Specifically, the characters used were shape of the median lobe of the aedeagus and spermatheca, as they are more reliable for the determination of species. Second, we performed a comprehensive morphometric analysis based on a wide set of characters in order to assess (i) the morphological variation of the target species, and (ii) the diagnostic power of character-sets to correctly assign specimens to molecularly delimited species. Fifteen morphometric variables were selected as predictors: length of elytrae (LE), width of elytrae (WE), length of pronotum (LP), width of pronotum (WP), length of antennae (LAN), length of median lobe of aedeagus (LAED); length of spermatheca (LSP); LE/LP; WE/WP; WP/LP; WE/LE; LE+LP; LAN/(LE+LP); LE/LAED; LE/LSP. No data standardization or normalization were performed for these measures. Terminology follows D'Alessandro et al. (2016) for the median lobe of aedeagus, and (Furth and Suzuki, 1992) for the spermatheca. A forward stepwise discriminant function analysis was performed separately on males and females (Tabachnick and Fidell, 1989), with a P-level to enter = 0.05, was performed. To assess the discrimination power of the 15 morphometric variables considered for the four groups analyzed (*kiesenwetteri* ♂♂, *kiesenwetteri* ♀♀, *ruffoi* ♂♂, *ruffoi* ♀♀) and to compute the relative discriminant functions, a Canonical Analysis was performed. Analysis was performed using the package NCSS version 11 for Windows.

Phylogenetic analyses

To investigate the phylogenetic relationship between the Western Paleartic species *P. kiesenwetteri* and the Italian endemic *P. ruffoi* and with the other representatives of the *Psylliodes gibbosa* group, we used a multi-locus approach based on DNA sequences of the 7 gene fragments amplified on representative individuals of each species or retrieved from GenBank from Gikonyo et al., 2021 (Tab. 5).

	Cox1	**CAD**	**Cv2**	**MMS22**	**Rad50**	**Sog**
Psylliodes cucullata	MW254859	MW254447	MW254623	MW254682	-	MW254741
Psylliodes gibbosa	MW254858	MW254446	MW254622	MW254681	MW254799	MW254740
Psylliodes inflata	MW254862	MW254450	MW254626	MW254685	MW254802	MW254744
Psylliodes kiesenwetteri	MW254865	MW254453	MW254629	MW254688	MW254805	MW254747
Psylliodes vehemens	MW254889	MW254477	MW254653	MW254712	MW254829	MW254771

Tab.5 GenBank accession numbers of the sequences used in this study.

Sequences of each gene were aligned separately with MAFFT v7.450 using the G-INS-I progressive method algorithm. Then, we build a concatenated sequence alignment and inferred phylogenetic relationships using both Maximum likelihood (ML) and Bayesian Inference (BI) methods using *Psylliodes vehemens* Wollaston, 1854 as an outgroup (Gikonyo et al., 2021). ML trees were inferred in IQ-TREE 1.6.12 (Nguyen et al., 2015) using the W-IQ-TREE webserver (Trifinopoulos et al., 2016). The best substitution models of each partition in our concatenated matrix of all genetic markers were determined by the ModelFinder module, including flexible rate heterogeneity across sites models (Kalyaanamoorthy et al., 2017), based on the Bayesian Information Criterion. We used the Edge Linked partition model to allow each partition to have its own evolutionary rate. Branch support was assessed by 1000 replicates of ultrafast bootstrapping (UFboot) (Hoang et al., 2018; Minh et al., 2013) and SH-like approximate likelihood ratio test (SH- aLRT) (Guindon et al., 2010). BI analyses were performed with two independent MCMC runs, with four chains each, on Mr Bayes v3.2.6 (Ronquist et al., 2012) for 10 million generations with a burn-in of 1 million generations (10%). Trees were sampled every 1000 generations and Tracer v1.7 (Rambaut et al., 2018) was used to assess convergence. FigTree v1.3.1 (Rambaut and Drummond, 2009) was used to depict the trees.

Results

Molecular species delimitation and preliminary morphological identification

We obtain 126 *cox1* sequences: 86 sequences for *P. kiesenwetteri* and 40 sequences for *P. ruffoi*. The NJ tree inferences produced a topology with two distinct clades (both with 100% bootstrap support) coherent with the two species delimitation analysis (Fig. 15). The frequency distributions of pairwise K2P distances estimated using the R library *ape* v5.3 and ASAP highlighted the existence of a clear barcoding gap in the *cox1* dataset. Intraspecific K2P distance values ranged from 0 to 4.1% (mean= 0.5%) and interspecific K2P distances ranged from 14.3% to 16.4 % (mean= 15.5%). The optimal distance threshold identify by *localMinima* function was estimated at 9%. With these threshold value set as distance cut off for clustering, *tclust* function divided the sequences into two clusters, coherently with the two clusters delimited by the best ASAP-score partitions (asap-score: 1.00; P-val (rank): 1.00e-05 (1); W (rank): 1.19e-03 (1)).

The preliminary assignment of collected individuals to *P. kiesenwetteri* and *P. ruffoi* achieved through the assessment of shape of the median lobe of the aedeagus and spermatheca as described by Leonardi (1975) and Nadein (2008) revealed multiple instances of mismatch, especially in immature specimens, between morphological identification and molecular species delimitation.

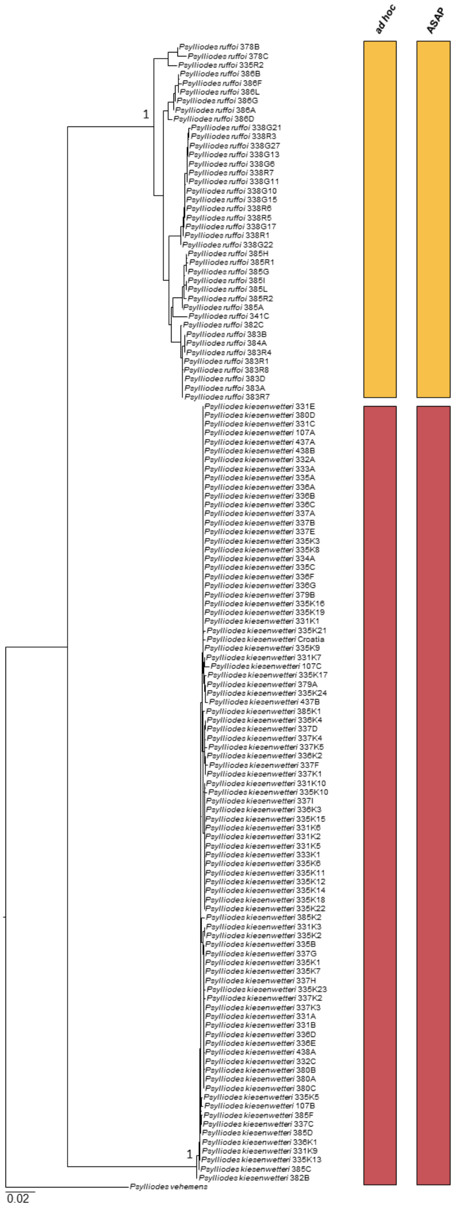


Fig.15 The NJ tree with schematic representation of results of species delimitation methods compared (ASAP, *ad hoc*). In yellow the *P. ruffoi* group and in red the *P. kiesenwetteri* group identify by the two species delimitation methods.

Morphometric results and new morphological diagnostic characters

Morphometric analyses based on discriminant function shows that the variables LAED (for ♂♂), LSP (for ♀♀), LE/LAED (for ♂♂), LE/LSP (for ♀♀), LAN, and LAN/(LE+LP) discriminate between the two species with highly significant support (Tab. 6).

	Lambda	F	*p*-value
LSP	91.03	385.75	<0.01
LE/LSP	81.91	172.10	<0.01
LAN/(LE+LP)	81.36	110.62	<0.01
LAN	79.11	95.96	<0.01
LAED	33.14	18.34	<0.01
LE/LAED	9.27	10.24	<0.05

Tab.6 Discriminant Stepwise Analysis for males and females: variables in the model, Wilk's Lambda, F to enter and *p*-level.

The classification matrix returns 100% of corrected attributions for every species analysed (Tab. 7). In addition, squared Mahalanobis distances matrix (SMD) suggests that the following pairs are well discriminated: ♂♂ *ruffoi* - ♂♂ *kiesenwetteri* (SMD = 52.46), and ♀♀ *ruffoi* - ♀♀ *kiesenwetteri* (SMD = 115.66).

	%	*kiesenwetteri* ♂	*Kiesenwetteri* ♀	*ruffoi* ♂	*ruffoi* ♀
kiesenwetteri ♂	100	20	0	0	0
kiesenwetteri ♀	100	0	20	0	0
ruffoi ♂	100	0	0	20	0
ruffoi ♀	100	0	0	0	20

Tab.7 Discriminant Stepwise Analysis: classification matrix for males and females. Rows: observed classifications; columns: predicted classifications.

The first two functions identify by the Canonical Analysis (CV1 and CV2), representing 98.6% of total explained variance, were considered, and the group centroids are reported in (Fig. 16). The first discriminant function accounts for 94.3% of EV and allows to easily discriminate the males of *kiesenwetteri* from those of *ruffoi*. The second function (4.3% of EV) is mainly useful to discriminate the females of the two species.

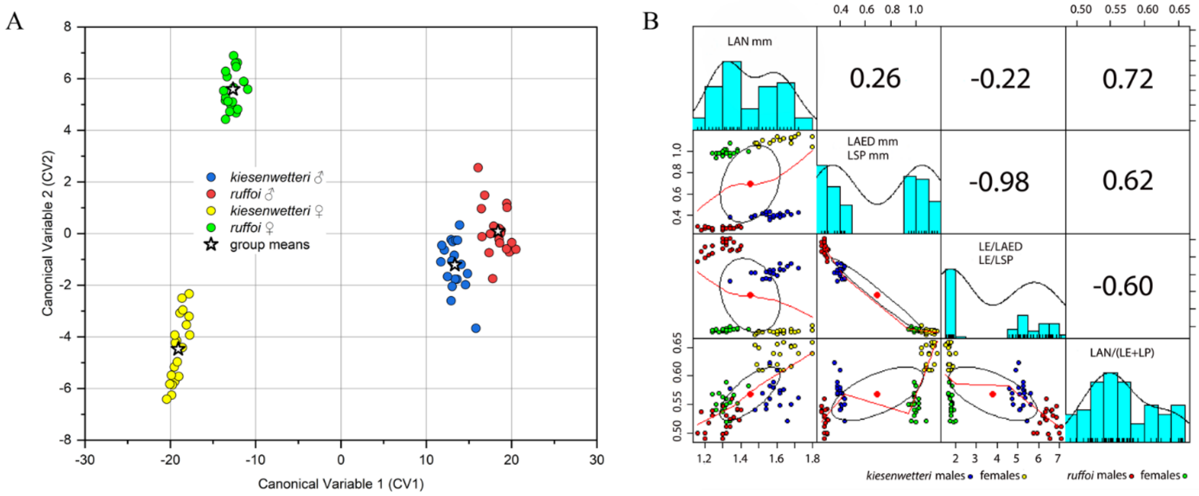


Fig.16 (A) Discriminant Stepwise Analysis: scatterplots (CV1 by CV2) of the Canonical Variates Analysis for males and females of *Psylliodes kiesenwetteri* and *P. ruffoi*. (B) Discriminant Stepwise Analysis: scatterplot matrix of the variables entered into the model and relative Pearson correlation coefficients.

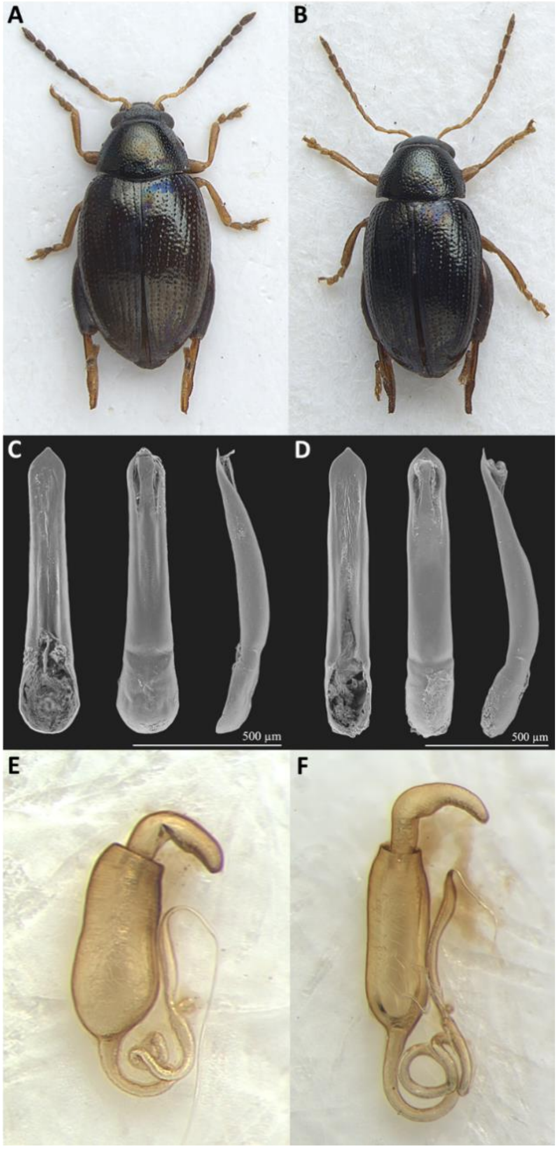


Fig.17 (A) *Psylliodes kiesenwetteri*, habitus; (B) *Psylliodes ruffoi*, habitus; (C) *Psylliodes kiesenwetteri*, median lobe of the aedeagus; (D) *Psylliodes ruffoi*, median lobe of the aedeagus; (E) *Psylliodes kiesenwetteri*, spermatheca; (F) *Psylliodes ruffoi*, spermatheca.

Phylogenetic relationship

Phylogenetic analyses based on ML and BI methods gave consistent results and identified the same two highly supported clades (BS = 100, PP = 1) within the *Psylliodes gibbosa* complex: clade A including *P. ruffoi*, *P. gibbosa* from Morocco and *P. infalta* from Turkey, and clade B composed by *P. cucullata* from Spain and *P. kiesenwetteri* (Fig. XX). Within clade A, *P. ruffoi* is sister of the subclade A1 of *P. gibbosa* and *P. infalta* (BS = 100, PP = 1). Two specimens of *Psylliodes ruffoi*, one from Sicily and one from Calabria, for a supported clade (subclade A2; BS = 100, PP = 1). Within Clade B, the three specimens of *P. kiesenwetteri* from Calabria, Abruzzo and Croatia formed a well-supported clade (subclade B1; BS = 100, PP = 1) (Fig. 18).

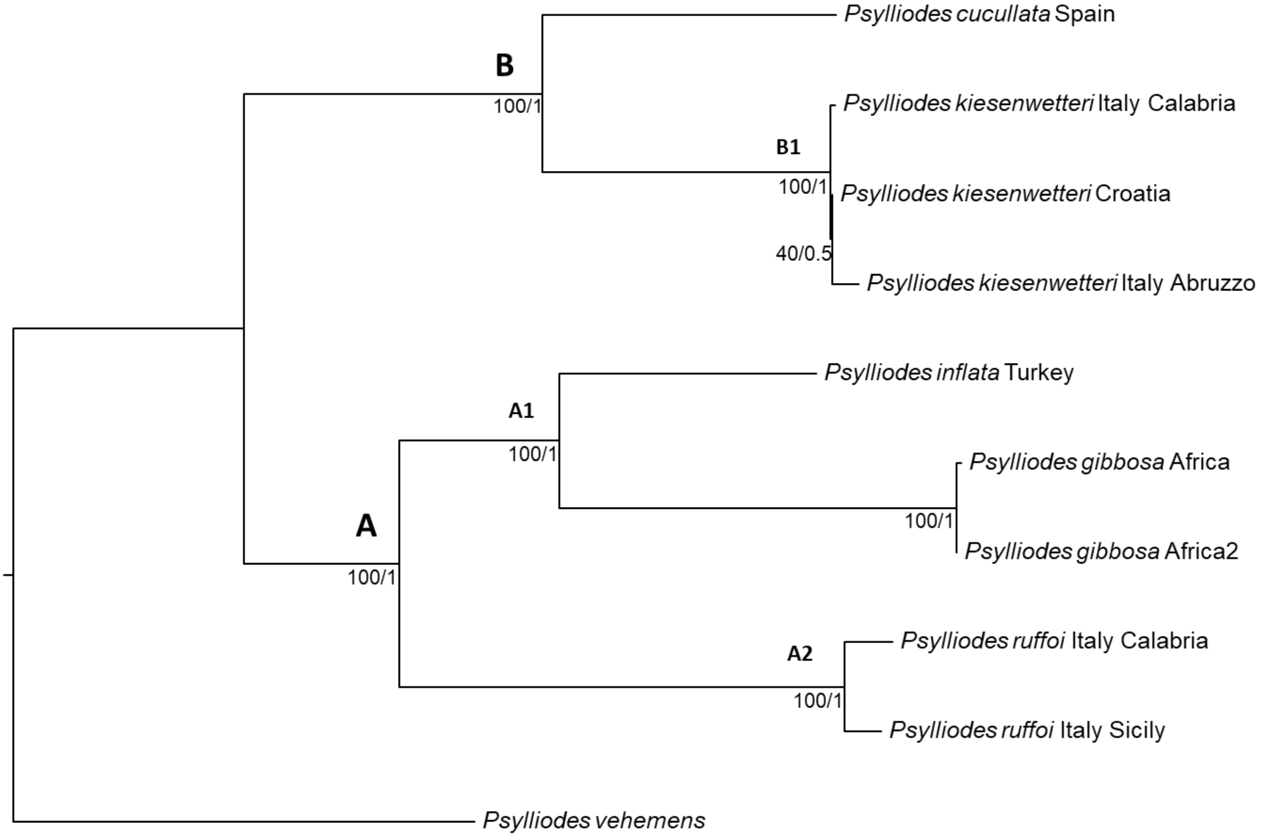


Fig.18 Maximum Likelihood and Bayesian phylogenetic tree of *Psylliodes gibbosus* complex based on concatenated DNA sequences of *cox1* and 5 nuclear genes. In correspondence of nodes are reported the bootstrap support (BS, left) from Maximum Likelihood and posterior probability (PP, right) from Bayesian analysis. Locality of the specimens are reported on the right of the name of species.

Discussion

In many animal groups the identification and description of new species is still based on morphological characters. The exclusive use of these characters can, however, lead to the wrong delimitation of the species when the interspecific morphological variability is absent, as in cryptic species, or the intraspecific morphological variability is too large. Precisely for this reason, the use of a comprehensive phylogeographic survey is the only way to evaluate the congruence between morphological and genetic variation of species, allowing to have a complete view of the interspecific limits of such variability. The case of *P. kiesenwetteri* and *P. ruffoi* demonstrated how the use of an integrative taxonomy approach within a phylogeographic framework allowed to evaluate the intra and interspecific morphological and genetic variability, leading to the delimitation of the lineages corresponding with the two species and the assessment of the discrimination capacity of morphological characters. In particular, this approach pointed out how traditional morphological characters might lead to ambiguous species identification, unless combined with additional characters in a solid statistical framework. This combined with the phylogenetic pattern allows the delimitation of two well distinct evolutionary and taxonomic units.

Through a classic morphological assessment, some male of *P. kiesenwetteri* have been identified as *P. ruffoi*. From a morphological point of view, *P. ruffoi* is considered similar to *P. kiesenwetteri* by Leonardi (1975) in its original description, while according to Nadein (2008) *P. ruffoi* has an intermediate morphology between *P. gibbosus* and *P. kiesenwetteri*. The lack of discrete characters and the high morphological variability makes it very complex to morphologically identify these species. In particular, the results obtained highlighting how the variability of the aedeagus makes the two species analyzed cryptic relative to this character alone. This result, however, could also be influenced by the presence of individuals of the two species at different stages of adult development, which could lead, in particular in the aedeagus, to greater variability of this character. Indeed, the phylogenetic, species delimitation and morphometric analyses based on multiple characters, coherently identify two evolutionary and taxonomic units. The analysis included specimens of *P. ruffoi* and *P. kiesenwetteri* from 15 different locations in central-southern Italy including the sympathetic and allopatric sites of the two species. This allowed us to have a more precise idea of the genetic and morphological variability of the two species. Analyses of neighbor-joining trees have recovered two well-supported clades. The two identified clades are also supported by the two species delimitation analyses (ASAP and *tclust*). The level of divergence at the *cox1* gene fragment shows a clear gap between the intra and the interspecific levels, which allows for a straightforward molecular identification of these species. Furthermore, the interspecific *cox1* divergence between *P. kiesenwetteri* and *P. ruffoi* is higher than that found in other studies on Alticini (Coral Şahin et al., 2019) or Chrysomelidae (Magoga et al., 2018). Morphometric analysis also identified two groups clearly distinguished by different variables, specifically LAED, LE / LAED (for ♂♂) , LSP (for ♀♀), LE / LSP (for ♀♀), LAN and LAN / (LE + LP) (for ♂♂ and ♀♀). This result, obtained from the analysis of 40 mature adults of both species from different locality in sympatry and parapatry, demonstrates how genital characters can discriminate between the two species. This result strengthens the hypothesis that the incorrect assignment of species of some specimens was due to the presence of immature adults, but also that a *quantitative* morphometric analysis, which further investigates the possible morphological differences between different species, is a more reliable method than a classic *qualitative* morphological assessment (Mutanen and Pretorius, 2007).

The Bayesian and Maximum Likelihood phylogenetic analysis of the *Psylliodes gibbosa* group demonstrates that *P. kiesenwetteri* and *P. ruffoi* are two distinct, and highly divergent, species, and these two species are not sisters. In fact, *P. ruffoi* appears to belong to a clade (clade A) in which it is sister of *P. gibbosa* from Morocco and *P. inflata*; whereas *Psylliodes kiesenwetteri* belong to a distinct clade (clade B) in which it is sister of *P. cucullata*.

While phylogenetic results demonstrate that *P. ruffoi* and *P. kiesenwetteri* are two distinct species, and are not sister, the morphometric approach demonstrate that these two species are not cryptic either. These results confirm that the use of an integrated taxonomy approach is useful for untangling complex cases such as that of the two species examined here. What initially appeared to be a pair of cryptic, possibly sibling, species, turned out to be two distinct species neither sibling, nor cryptic. These differences were present both for the samples present in sympatric and in allopatric locality, showing how even by sampling a wider morphological and genetic variability, the distinction of the two taxa is maintained. Finally, it is confirmed that a morphometric approach, supported by a robust molecular analysis, is more reliable in discriminating the species than through a nonmetric, subjective, visual comparisons of morphological traits.

This study allowed better defining the systematic and the variability of *P. ruffoi* and *P. kiesenwetteri* and provide an integrative and geographically explicit framework that can be applied to other putative cryptic species complex with overlapping distribution.

References

Agapow, P., Bininda-Emonds, O.R.P., Crandall, K.A., Gittleman, J.L., Mace, G.M., Marshall, J.C., Purvis, A., 2004. The Impact of Species Concept on Biodiversity Studies. Q. Rev. Biol. 79, 161–179. https://doi.org/10.1086/383542

Amano, T., Lamming, J.D.L., Sutherland, W.J., 2016. Spatial Gaps in Global Biodiversity Information and the Role of Citizen Science. BioScience 66, 393–400. https://doi.org/10.1093/biosci/biw022

Bauer, A.M., Parham, J.F., Brown, R.M., Stuart, B.L., Grismer, L., Papenfuss, T.J., Böhme, W., Savage, J.M., Carranza, S., Grismer, J.L., 2011. Availability of new Bayesian-delimited gecko names and the importance of character-based species descriptions. Proc. R. Soc. B Biol. Sci. 278, 490–492.

Baviera, C., Biondi, M., 2015. The Alticini (Coleoptera: Chrysomelidae, Galerucinae) of Sicily: recent records and updated checklist. Atti Della Accad. Peloritana Dei Pericolanti-Cl. Sci. Fis. Mat. E Nat. 93, 2.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22, 148–155.

Biondi, M., 1988. Contributo alla conoscenza dei Crisomelidi Alticini dei Monti Nebrodi e Peloritani (Sicilia Nord Orientale) e descrizione di una nuova specie di *Longitarsus* (Coleoptera, Chrysomelidae, Alticinae). Animalia 15, 129–149.

Blanquer, A., Uriz, M.-J., 2008. 'A posteriori'searching for phenotypic characters to describe new cryptic species of sponges revealed by molecular markers (Dictyonellidae: Scopalina). Invertebr. Syst. 22, 489–502.

Brown, S.D., Collins, R.A., Boyer, S., Lefort, M., Malumbres-Olarte, J., Vink, C.J., Cruickshank, R.H., 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol. Ecol. Resour. 12, 562–565.

Coral Şahin, D., Magoga, G., Özdikmen, H., Montagna, M., 2019. DNA Barcoding as useful tool to identify crop pest flea beetles of Turkey. J. Appl. Entomol. 143, 105–117.

D'Alessandro, P., Samuelson, A., Biondi, M., 2016. Taxonomic revision of the genus *Arsipoda* Erichson, 1842 (Coleoptera, Chrysomelidae) in New Caledonia. Eur. J. Taxon. 1–61.

Darwell, C.T., Cook, J.M., 2017. Cryptic diversity in a fig wasp community—morphologically differentiated species are sympatric but cryptic species are parapatric. Mol. Ecol. 26, 937–950.

Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85, 407–417.

Deng, J., Guo, Y., Cheng, Z., Lu, C., Huang, X., 2019. The prevalence of single-specimen/locality species in insect taxonomy: an empirical analysis. Diversity 11, 106.

Egea, E., David, B., Choné, T., Laurin, B., Féral, J.-P., Chenuil, A., 2016. Morphological and genetic analyses reveal a cryptic species complex in the echinoid *Echinocardium cordatum* and rule out a stabilizing selection explanation. Mol. Phylogenet. Evol. 94, 207–220. https://doi.org/10.1016/j.ympev.2015.07.023

Espíndola, A., Ruffley, M., Smith, M.L., Carstens, B.C., Tank, D.C., Sullivan, J., 2016. Identifying cryptic diversity with predictive phylogeography. Proc. R. Soc. B Biol. Sci. 283, 20161529. https://doi.org/10.1098/rspb.2016.1529

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., n.d. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates Mol Mar Biol Biotechnol. 1994; 3: 294–9.

Furth, D.G., Suzuki, K., 1992. Character correlation studies of problematic genera of Alticinae in relation to Galerucinae (Coleoptera: Chrysomelidae). Presented at the Proceedings of the third international symposium on the Chrysomelidae, Beijing, Citeseer, pp. 116–135.

Gikonyo, M., 2021. The evolution of host plant use and adaptation to Brassicaceae in the leaf beetle genus *Psylliodes*.

Grebennikov, V.V., 2019. First DNA analysis of pill scarabs (Coleoptera: Hybosoridae: Ceratocanthinae) reveals multiple paraphyly of Afrotropical Philharmostes. Eur. J. Entomol. 116.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

International Commission on Zoological Nomenclature, International Union of Biological Sciences, 1999. International code of zoological nomenclature: International Commission on Zoological Nomenclature. International Trust for Zoological Nomenclature.

Jossart, Q., Kochzius, M., Danis, B., Saucède, T., Moreau, C.V., 2021. Diversity of the Pterasteridae (Asteroidea) in the Southern Ocean: a molecular and morphological approach. Zool. J. Linn. Soc. 192, 105–116.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547.

Latella, L., Ruffo, S., Stoch, F., 2005. Il progetto Ckmap (Checklist e distribuzione della fauna italiana): metodologia di lavoro e tecniche informatiche. Checkl. E Distrib. Della Fauna Ital. Mem. Mus. Civ. Storia Nat. Verona 2, 15–20.

Leavitt, S.D., Esslinger, T.L., Divakar, P.K., Crespo, A., Lumbsch, H.T., 2016. Hidden diversity before our eyes: delimiting and describing cryptic lichen-forming fungal species in camouflage lichens (Parmeliaceae, Ascomycota). Fungal Biol. 120, 1374–1391.

Leonardi, C., 1975. Le *Psylliodes* appenniniche del museo civico di storia naturale di verona (Coleoptera Chrysomelidae).

Leonardi, C., 1970. Materiali per uno studio filogenetico del genere *Psylliodes* (Coleoptera Chrysomelidae). Atti Della Soc. Ital. Sci. Nat. E Mus. Civ. Storia Nat. Milano 110, 201–223.

Magoga, G., Fontaneto, D., Montagna, M., 2021. Factors affecting the efficiency of molecular species delimitation in a species-rich insect family.

Magoga, G., Sahin, D.C., Fontaneto, D., Montagna, M., 2018. Barcoding of Chrysomelidae of Euro-Mediterranean area: efficiency and problematic species. Sci. Rep. 8, 1–9.

Mallet, J., 2007. Hybrid speciation. Nature 446, 279–283.

Martinsson, S., Rota, E., Erséus, C., 2015. Revision of Cognettia (Clitellata, Enchytraeidae): re-establishment of Chamaedrilus and description of cryptic species in the sphagnetorum complex. Syst. Biodivers. 13, 257–277.

McBride, C.S., Van Velzen, R., Larsen, T.B., 2009. Allopatric origin of cryptic butterfly species that were discovered feeding on distinct host plants in sympatry. Mol. Ecol. 18, 3639–3651. https://doi.org/10.1111/j.1365-294X.2009.04309.x

Memon, N., Meier, R., Manan, A., Su, K.F.-Y., 2006. On the use of DNA sequences for determining the species limits of a polymorphic new species in the stink bug genus Halys (Heteroptera: Pentatomidae) from Pakistan. Syst. Entomol. 31, 703–710. https://doi.org/10.1111/j.1365-3113.2006.00350.x

Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.

Montagna, M., Kubisz, D., Mazur, M.A., Magoga, G., Ścibior, R., Kajtoch, Ł., 2017. Exploring species-level taxonomy in the *Cryptocephalus flavipes* species complex (Coleoptera: Chrysomelidae). Zool. J. Linn. Soc. 179, 92–109.

Morek, W., Stec, D., Gąsiorek, P., Surmacz, B., Michalczyk, Ł., 2019. Milnesium tardigradum Doyère, 1840: The first integrative study of interpopulation variability in a tardigrade species. J. Zool. Syst. Evol. Res. 57, 1–23.

Mutanen, M., Pretorius, E., 2007. Subjective visual evaluation vs. traditional and geometric morphometrics in species delimitation: a comparison of moth genitalia. Syst. Entomol. 32, 371–386.

Nadein, K., 2008. Review of the Psylliodes gibbosus species group, with descriptions of two new species. Koleopterol. Rundsch. 78, 333–366.

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford University Press, USA.

Nevo, E., 2001. Evolution of genome-phenome diversity under environmental stress. Proc. Natl. Acad. Sci. 98, 6233–6240. https://doi.org/10.1073/pnas.101109298

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. Front. Zool. 7, 1–14.

Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528.

Puillandre, N., Brouillet, S., Achaz, G., 2021. ASAP: assemble species by automatic partitioning. Mol. Ecol. Resour. 21, 609–620. https://doi.org/10.1111/1755-0998.13281

Ramasindrazana, B., Goodman, S.M., Schoeman, M.C., Appleton, B., 2011. Identification of cryptic species of *Miniopterus* bats (Chiroptera: Miniopteridae) from Madagascar and the Comoros using bioacoustics overlaid on molecular genetic and morphological characters. Biol. J. Linn. Soc. 104, 284–302.

Rambaut, A., Drummond, A., 2009. FigTree version 1.3. 1 [computer program]. Website Httptree Bio Ed Ac Uksoftwarefigtree Last Accessed Oct. 21, 2017.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901.

Raxworthy, C.J., Ingram, C.M., Rabibisoa, N., Pearson, R.G., 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. Syst. Biol. 56, 907–923.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Salvi, D., Maura, M., Pan, Z., Bologna, M.A., 2019. Phylogenetic systematics of *Mylabris* blister beetles (Coleoptera, Meloidae): a molecular assessment using species trees and total evidence. Cladistics 35, 243–268.

Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E., Crozier, R.H., 2010. Integrative Taxonomy: A Multisource Approach to Exploring Biodiversity. Annu. Rev. Entomol. 55, 421–438. https://doi.org/10.1146/annurev-ento-112408-085432

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Sites Jr, J.W., Marshall, J.C., 2003. Delimiting species: a Renaissance issue in systematic biology. Trends Ecol. Evol. 18, 462–470.

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K.-H., Liow, L.H., Nowak, M.D., 2018. Finding evolutionary processes hidden in cryptic species. Trends Ecol. Evol. 33, 153–163.

Tabachnick, B., Fidell, L., 1989. Using Multivariate Statistics.(Second Edn. ed.).

Templeton, A.R., 2001. Using phylogeographic analyses of gene trees to test species status and processes. Mol. Ecol. 10, 779–791.

Tessens, B., Monnens, M., Backeljau, T., Jordaens, K., Van Steenkiste, N., Breman, F.C., Smeets, K., Artois, T., 2021. Is ‘everything everywhere’? Unprecedented cryptic diversity in the cosmopolitan flatworm *Gyratrix hermaphroditus*. Zool. Scr. 50, 837–851. https://doi.org/10.1111/zsc.12507

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235.

Urbani, F., D’Alessandro, P., Frasca, R., Biondi, M., 2015. Maximum entropy modeling of geographic distributions of the flea beetle species endemic in Italy (Coleoptera: Chrysomelidae: Galerucinae: Alticini). Zool. Anz.- J. Comp. Zool. 258, 99–109.

Wheeler, Q.D., 2007. Invertebrate systematics or spineless taxonomy. Zootaxa 1668, 11–18.

Xiao, J.-H., Wang, N.-X., Li, Y.-W., Murphy, R.W., Wan, D.-G., Niu, L.-M., Hu, H.-Y., Fu, Y.-G., Huang, D.-W., 2010. Molecular approaches to identify cryptic species and polymorphic species within a complex community of fig wasps. PLoS One 5, e15067.

Yeates, D.K., Seago, A., Nelson, L., Cameron, S.L., Joseph, L., Trueman, J.W., 2011. Integrative taxonomy, or iterative taxonomy? Syst. Entomol. 36, 209–217.

CHAPTER 5

Case study III

A phylogeographic and morphological assessment of the *Psylliodes springeri* species complex (Chrysomelidae, Alticini) uncovers new endemic diversity from the Majella mountains (southern Apennine)

Introduction

Mountain are among the areas with the greatest environmental heterogeneity (Rahbek et al., 2019). The complexity of these environments has produced a marked effect on the evolutionary processes, making mountains reservoirs of biodiversity and one of the richest environments in endemic species (Hewitt, 1996; Schmitt, 2007; Kenyeres et al., 2009; Svenning et al., 2015; Smyčka et al., 2017; Amori et al., 2019).

In Italy, the Apennine range remains relatively under-studied as regards the biodiversity and phylogeographic patterns, and this is especially true for invertebrates. The Apennines is the southernmost mountain system in Europe, crossing the Italian peninsula for over thousand kilometers from north to south. The geological conformation of this mountain range and its position made it the scene for an eventful history of evolutionary and biogeographic processes. As a consequence, the Apennine, despite a relatively small size, represent a hot spot of diversity on a par with more extensive mountain ranges such as Alps and Pyrenees (Canestrelli et al., 2006, 2008; Bogdanowicz et al., 2015). In particular, the central sector of the Apennines consists of highest peaks, characterized by a geological history largely modeled by the Quaternary glaciations (Giraudi and Frezzotti, 1997; Giraudi et al., 2004). These characteristics have made this area of the Apennines a refugia for various cold adapted species and it is among the mountain areas that host the highest number of species in the entire Italian peninsula area (Maiorano et al., 2006; Biondi et al., 2013; Urbani et al., 2015; Menchetti et al., 2021).

Since the second half of the twentieth century, several new species of the genus *Psylliodes* Latreille have been described from the Apennine (Leonardi, 1975; Leonardi, 2007; Biondi and D'Alessando, 2017). Among these, two new species are associated to high-altitude environments of the central Apennines: *Psylliodes springeri*, micro endemic of the valley of the Pilato Lake, in the Sibillini mountains, and *P. biondii*, with a wider distribution that comprehends almost all the high peaks of the central Apennine mountain ranges, from the Gran Sasso massif to the Matese mountains. Both these species have a very similar ecology, live in the same high mountain environments (above 1500 m of altitude) and share the same host plants, such as Brassicaceae of the genus *Isatis* and *Erysimum*. These species are also morphologically similar one each other, and with the alpine species *P. picipes* (Leonardi, 2007). However, despite the morphological affinity between the species within this complex, a high morphological variability was observed within *P. biondii*. The populations of the northern part of the range of this species are distinguishable from those of the southern part.

Indeed, the individuals of the Maiella massif, even if considered as *P. biondii* by Leonardi (2007) they have been considered by this same author as requiring further investigations given their morphological peculiarities. However, the evolutionary (and taxonomic) significance of these subtle morphological differences is not clear, and it remains to be investigated whether the peculiarities of the Majella population represent local phenotypes or instead an independent evolutionary lineage relative to the widespread *P. biondii*.

To tackle these questions, we implemented an integrative taxonomic approach based on a comprehensive assessment of the morphological variation and the phylogenetic structure of the P. *springeri*/*P. biondii* complex across the Apennines. The main aims of this study are to: (i) established a systematic framework for the *P. springeri* complex and identify evolutionary and taxonomic units and (ii) among them, identify those that are suitable for the phylogeographic investigation of this thesis project.

Materials and Methods

Sample collection and morphological identification

Specimens analyzed in this chapter were collected and identified on a morphological basis as explained in Chapter 2.

DNA extraction

Total genomic DNA was extracted using non-invasive DNA extraction method as described in Chapter 1. Six different gene fragments were amplified through polymerase chain reaction (PCR). Two fragments of the mitochondrial gene cytochrome oxidase subunit I (*cox1*) were amplified from 104 individuals using the conditions described in Salvi et al. (2019): the standard barcode region, using the universal primers LCO1490 and HC02198 (Folmer et al., n.d.), and the 3' fragment, using primers C1-J-2183 and TL2-N-3014 (Simon et al., 1994). Three protein-coding single copy nuclear genes were amplified for a subsample of individuals representing the main mitochondrial clades: Carbamoylphosphate synthase (*CAD*), Crossveinless 2 (*Cv2*), and Rad 50 protein (*Rad50*), using primers and amplification conditions described in Gikonyo (2021). Successful amplification was determined by gel electrophoresis and PCR products were purified and sequenced by an external service (Genewitz, UK).

The obtained chromatograms of each sequence were edited and assembled into a consensus sequence using Geneious Prime 2021 (Biomatters Ltd., Auckland, New Zealand). Sequences of each gene were aligned separately using MAFFT v7.450 with the G-INS-I progressive method algorithm (Katoh et al., 2002) and then used for the downstream phylogenetic and specie delimitation analysis.

Species Delimitation

Species delimitation analyses were performed using the standard barcode fragment of *cox1*. First, we estimated the pairwise nucleotide distance matrices, required for *ad hoc* nucleotide distance threshold methods, using the R library *ape* v5.3 (Paradis and Schliep, 2019), with the K2P substitution model and the pairwise deletion option. With the R package *spider* v1.5.0 (Brown et al., 2012) we performed a threshold optimization analysis. The best threshold was identified with the *localMinima* function and used to cluster sequences with the function *tclust* of package *spider*. Second, we applied the Assemble Species by Automatic Partitioning (ASAP; Puillandre et al., 2021) approach, using the program web-interface (https://bioin fo.mnhn.fr/abi/publi c/asap); with K2P as nucleotide substitution model and other parameters left as default (Puillandre et al., 2021).

Phylogenetic analyses

Phylogenetic tree analyses were performed on the concatenated sequence alignment of the two fragment of *cox1* using the Maximum likelihood (ML) method. ML trees were inferred in IQ-TREE 1.6.12 (Nguyen et al., 2015) using the W-IQ-TREE webserver (Trifinopoulos et al., 2016). The best substitution models of each partition in our concatenated matrix of the two gene fragments were determined by the ModelFinder module, including flexible rate heterogeneity across sites models (Kalyaanamoorthy et al., 2017), under the Bayesian Information Criterion. We used the Edge Linked partition model to allow each partition to have its own evolutionary rate. Branch support was assessed by 1000 replicates of ultrafast bootstrapping (UFboot; Hoang et al., 2018; Minh et al., 2013) and SH-like approximate likelihood ratio test (SH- aLRT; Guindon et al., 2010). FigTree v1.3.1 (Rambaut and Drummond, 2009) was used to depict the trees using a mid-point root.

Phylogenetic networks were used to infer genealogies of each nuclear marker and to assess the occurrence of nuclear haplotype sharing between mitochondrial lineages. Haplotype phase of nuclear sequences was determined using the PHASE algorithm (Stephens et al., 2001; Stephens and Donnelly, 2003) as implemented in DnaSP v5 (Librado and Rozas, 2009), with 1000 initial iterations discarded as burn-in, 1 as thinning interval, and 1000 post-burnin iterations. The phylogenetic relationships among phased haplotypes were inferred through the median-joining distance method (Bandelt et al., 1999) using the PopArt 1.7 software (Leigh and Bryant, 2015).

Results

Phylogenetic and species delimitation analysis

We obtained 104 sequences for both the two fragment of *cox1* (the standard barcode region of 608 bp and the 3’ fragment of 767 bp), 45 sequences for *Psylliodes springeri*, 39 sequences for *P. biondii* and 20 for the population from Maiella. Results of the species delimitation analyses based on ASAP and *tcclust* discriminate

P. springeri and *P. biondii* and the Maiella specimens as three separate clusters. In particular, both the first (1.50) best ASAP score and *tcclust* with the optimal distance threshold (3%) identified three congruent clusters (Fig. 19).

Phylogenetic analyses based on ML recovered the three species cluster inferred by ASAP as three supported clades: the *P. springeri* clade (SH- aLRT = 100; UFboot = 100), the *P. biondii* clade (SH- aLRT = 99.4; UFboot = 100), and the Maiella clade (SH- aLRT = 99.5; UFboot = 100). The two latter clades have a sister relationship (Fig. 19).

For the three nuclear markers we obtained 204 phased sequences, 74 sequences of *CAD*, 74 of *Cv2* and 56 of *Rad50*. All three mitochondrial clades identified by the ML tree were represented for the three nuclear markers with at least 14 sequences (and up to 46 sequences). The phylogenetic relationships between haplotypes of each nuclear marker, as inferred by the median-joining approach, show a clear segregation of the haplotypes belonging to the three taxa, with a complete lack of haplotype sharing (Fig. 20). Haplogroups representing distinct taxa are separated by at least 7 mutational steps in any nuclear marker.

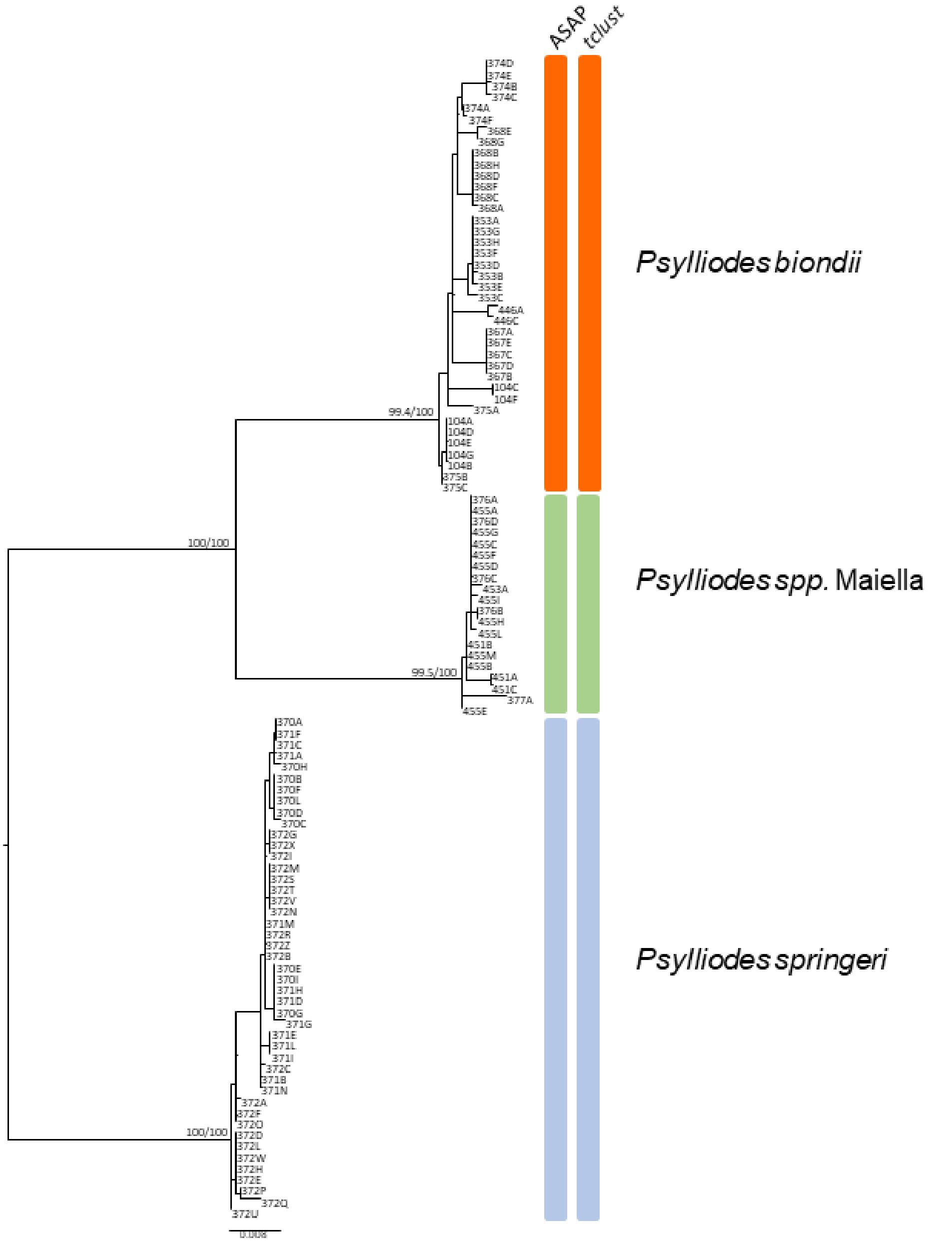


Fig. 19 Maximum Likelihood tree of *Psylliodes springeri* complex based on two concatenated fragments of c*ox1*. In correspondence of nodes are reported the bootstrap support (SH- aLRT on the left; UFboot on the right) from Maximum Likelihood. The bar plot represents the result of species delimitation analyses based on the c*ox1* barcode fragment. Results of species delimitation analyses using ASAP and *tclust*.

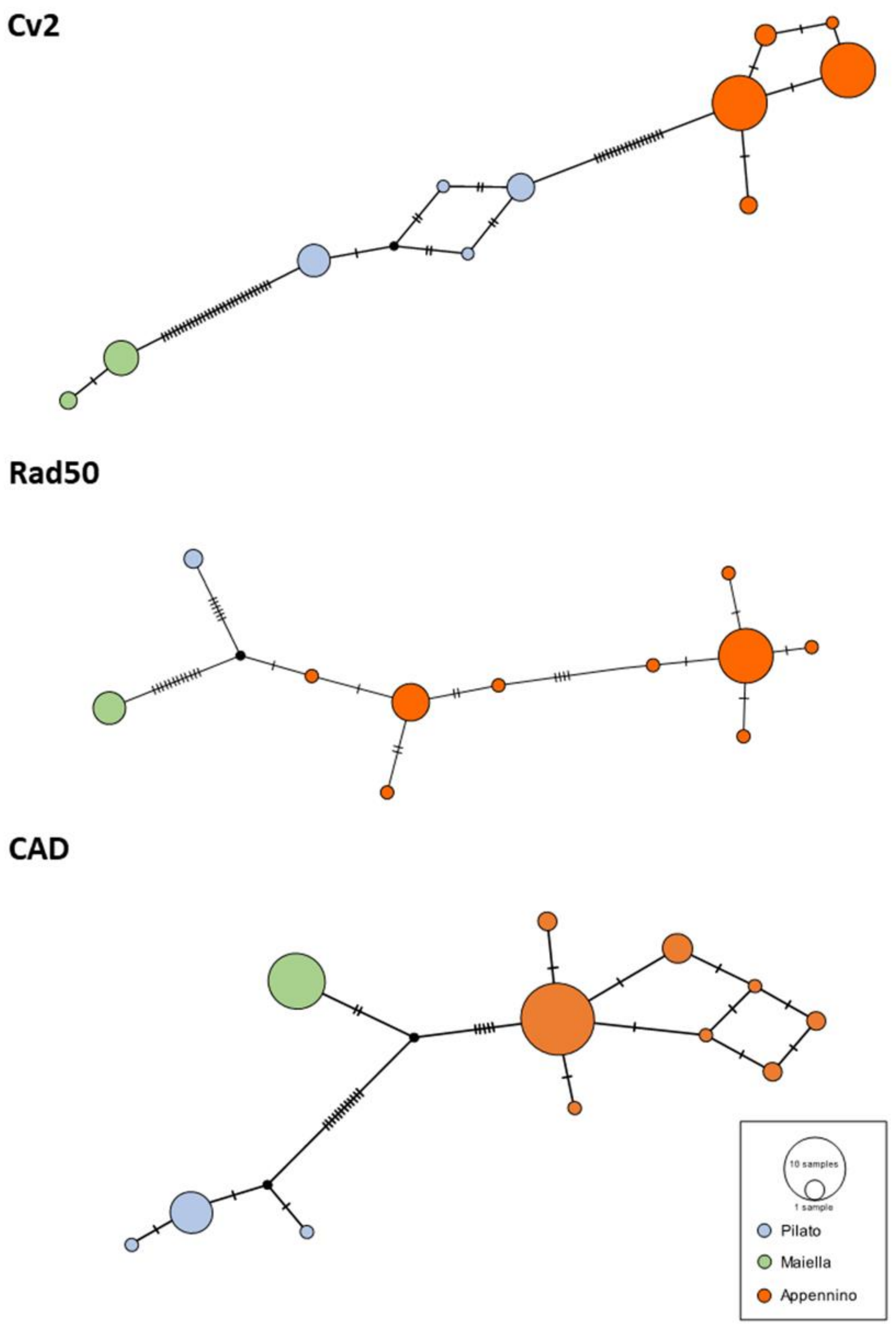


Fig.20 Haplotype networks of *Psylliodes springeri* complex based on three nuclear markers. The network of each nuclear marker reveals a clear lack of haplotype sharing among the three mitochondrial lineages. In blue the haplotypes of *P. springeri* from Pilato Lake, in green the haplotype of the Maiella lineages and in orange the haplotypes of *P. biondii* from different localities of Central Apennine.

Discussion

The high peaks of the central Apennines served as a refuge for various cold adapted species and are among the areas with the highest number of endemic species within the Italian peninsula. There are many groups of insects that present endemic species or relict lines in this area. For example, for the Lepidoptera, Apennines functioned as an ex situ refugium, as indicated by the presence of several cold-adapted species and because mountain areas host the richest butterfly communities of the Peninsula-Sicily center (Menchetti et al., 2021). The same also applies to the Orthoptera (Kenyeres et al., 2009) and several families of beetles such as Tenebrionidae (Fattorini, 2010), Carabidae (Brandmayr et al., 2003) and Chrysomelidae (Biondi et al., 2013). Specifically, the latter family has several cryophilic species in the central Apennines beyond those studied in this work, such as *Longitarsus springeri*, *Luperus fiorii* and *Oreina sibylla*. However, the systematics and genetic diversity of many of these species has not yet been investigated, leaving many questions about the evolutionary processes that took place in the central Apennines unanswered.

This study documented cryptic diversity within the *Psylliodes springeri* complex and identified the populations from the Maiella mountains as a separate evolutionary unit. The molecular species delimitation analysis based on mitochondrial data clearly indicate that this population is not conspecific of *P. biondii* and that genetic differentiation among these two lineages is similar to the interspecific distance observed between *P. biondii* and *P. springeri* and consistent with values reported for other species of Alticini (Magoga et al., 2018) or Chrysomelidae (Magoga et al., 2021). This pattern is strongly supported also by nuclear genealogies that showed a complete lineage sorting among *P. springeri, P. biondii* and the Maiella lineage (Fig. 20). These three lineages do not share any haplotype both in fast evolving nuclear markers such as *Cv2* as well as in markers with lower substitution rates such as *Rad50* and *CAD*. From a morphological point of view the Majella lineage is more similar to *P. biondii*. This is in agreement with molecular data as these two species are siter in the phylogenetic tree (Fig. 2).

The three lineages show an allopatric distribution and a high level of divergence, which suggest a diversification process mediated by geographic isolation in distinct mountains during a prolonged time. Such a scenario is consistent with the ecology of these beetles that are strongly associated to high altitude environments, have a strong association with host plant and low dispersal ability. Similar patterns can be found in other insects living in high altitude environments of the Italian peninsula, such as in the case of leaf beetles of the genus *Oreina* (Borer et al., 2010), in several Alpine butterflies (Haubrich and Schmitt, 2007; Huemer and Hebert, 2011) and in the caddisfly, *Drusus discolor* (Pauls et al., 2006). In these cases the divergence between endemic strains of distinct mountain ranges seems to be due to survival in refugial habitats during the Pliocene period (Schmitt, 2009). In the case of the Apennine *Psylliodes* analyzed in this study, comparisons with the closely related Alpine species *Psylliodes picipes* will allow a deeper understanding of the evolutionary and biogeographic history of this group of altitude flea beetles as well as to clarifying the taxonomic status of the Maiella population.

This study demonstrates that our knowledge on the Apennine biotas is still far from being exhaustive and that the high mountain environments within this ranges are key target for biodiversity assessments. Our comprehension of the evolutionary and biogeographic process underlying the assembly of the Appenine biota is therefore highly limited and will require further research. In this respect, the integrative morphological and phylogenetic approach used in this study (and in the previous chapter) proved to be effective even for the study of species complex with extensive, and shared, morphological variation. In the case of the *Psylliodes biondii* complex, this study allowed disentangling the three evolutionary units that are part of this complex. Among them the *P. biondii* populations from Central Apennines (excluding the Maiella population) represent a cohesive phylogenetic unit that will be selected for the phylogeographic investigations of the Part II of this thesis project.

References

Amori, G., Boitani, L., Milana, G., Maiorano, L., Luiselli, L., 2019. Endemism and diversity in European montane mammals: macro-ecological patterns. Biol. J. Linn. Soc. 128, 225–237.

Bandelt, H.-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Biondi, M., D’alessandro, P., 2017. Psylliodes urbaniae: a new species from Central Apennines (Coleoptera: Chrysomelidae: Galerucinae: Alticini). Onychium 121–130.

Biondi, M., Urbani, F., D’alessandro, P., 2013. Endemism patterns in the Italian leaf beetle fauna (Coleoptera, Chrysomelidae). ZooKeys 177.

Bogdanowicz, W., Hulva, P., Černá Bolfíková, B., Buś, M.M., Rychlicka, E., Sztencel-Jabłonka, A., Cistrone, L., Russo, D., 2015. Cryptic diversity of Italian bats and the role of the Apennine refugium in the phylogeography of the western Palaearctic. Zool. J. Linn. Soc. 174, 635–648.

Borer, M., Alvarez, N., Buerki, S., Margraf, N., Rahier, M., Naisbit, R.E., 2010. The phylogeography of an alpine leaf beetle: Divergence within Oreina elongata spans several ice ages. Mol. Phylogenet. Evol. 57, 703–709.

Brandmayr, P., Pizzolotto, R., Scalercio, S., Algieri, M., Zetto, T., 2003. Diversity patterns of carabids in the Alps and the Apennines, in: Alpine Biodiversity in Europe. Springer, pp. 307–317.

Brown, S.D., Collins, R.A., Boyer, S., Lefort, M., Malumbres-Olarte, J., Vink, C.J., Cruickshank, R.H., 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol. Ecol. Resour. 12, 562–565.

Canestrelli, D., Cimmaruta, R., Costantini, V., Nascetti, G., 2006. Genetic diversity and phylogeography of the Apennine yellow-bellied toad Bombina pachypus, with implications for conservation. Mol. Ecol. 15, 3741–3754.

Canestrelli, D., Cimmaruta, R., Nascetti, G., 2008. Population genetic structure and diversity of the Apennine endemic stream frog, *Rana italica*–insights on the Pleistocene evolutionary history of the Italian peninsular biota. Mol. Ecol. 17, 3856–3872.

Crandall, M.C.D.P.K., Clement, M., Posada, D., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1660.

Fattorini, S., 2010. Use of insect rarity for biotope prioritisation: the tenebrionid beetles of the Central Apennines (Italy). J. Insect Conserv. 14, 367–378.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., n.d. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates Mol Mar Biol Biotechnol. 1994; 3: 294–9.

Gikonyo, M., 2021. The evolution of host plant use and adaptation to Brassicaceae in the leaf beetle genus Psylliodes.

Giraudi, C., Ehlers, J., Gibbard, P., 2004. The Apennine glaciations in Italy. Quarternary Glaciat. Chronol. Part Eur. Dev. Quat. Sci. 2, 215–224.

Giraudi, C., Frezzotti, M., 1997. Late Pleistocene glacial events in the central Apennines, Italy. Quat. Res. 48, 280–290.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Haubrich, K., Schmitt, T., 2007. Cryptic differentiation in alpine-endemic, high-altitude butterflies reveals down-slope glacial refugia. Mol. Ecol. 16, 3643–3658.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247–276.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Huemer, P., Hebert, P.D., 2011. Cryptic diversity and phylogeography of high alpine Sattleria—a case study combining DNA barcodes and morphology (Lepidoptera: Gelechiidae). Zootaxa 2981, 1–22.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kenyeres, Z., Rácz, I.A., Varga, Z., 2009. Endemism hot spots, core areas and disjunctions in European Orthoptera. Acta Zool. Cracoviensia-Ser. B Invertebr. 52, 189–211.

Leigh, J.W., Bryant, D., 2015. popart: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116.

Leonardi, C., 2007. Dati inediti sul genere *Psylliodes* Latreille, con descrizione di quattro nuove specie mediterranee (Coleoptera Chrysomelidae). Atti Della Soc. Ital. Sci. Nat. E Mus. Civ. Storia Nat. Milano 148, 161–240.

Leonardi, C., 1975. Le P*sylliodes* appenniniche del museo civico di storia naturale di Verona (Coleoptera Chrysomelidae).

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Magoga, G., Fontaneto, D., Montagna, M., 2021. Factors affecting the efficiency of molecular species delimitation in a species-rich insect family.

Magoga, G., Sahin, D.C., Fontaneto, D., Montagna, M., 2018. Barcoding of Chrysomelidae of Euro-Mediterranean area: efficiency and problematic species. Sci. Rep. 8, 1–9.

Maiorano, L., Falcucci, A., Boitani, L., 2006. Gap analysis of terrestrial vertebrates in Italy: priorities for conservation planning in a human dominated landscape. Biol. Conserv. 133, 455–473.

Menchetti, M., Talavera, G., Cini, A., Salvati, V., Dincă, V., Platania, L., Bonelli, S., Balletto, E., Vila, R., Dapporto, L., 2021. Two ways to be endemic. Alps and Apennines are different functional refugia during climatic cycles. Mol. Ecol. 30, 1297–1310.

Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528.

Pauls, S.U., Lumbsch, H.T., Haase, P., 2006. Phylogeography of the montane caddisfly *Drusus discolor*: evidence for multiple refugia and periglacial survival. Mol. Ecol. 15, 2153–2169.

Puillandre, N., Brouillet, S., Achaz, G., 2021. ASAP: assemble species by automatic partitioning. Mol. Ecol. Resour. 21, 609–620. https://doi.org/10.1111/1755-0998.13281

Rahbek, C., Borregaard, M.K., Colwell, R.K., Dalsgaard, B., Holt, B.G., Morueta-Holme, N., Nogues-Bravo, D., Whittaker, R.J., Fjeldså, J., 2019. Humboldt's enigma: What causes global patterns of mountain biodiversity? Science 365, 1108–1113.

Rambaut, A., Drummond, A., 2009. FigTree version 1.3. 1 [computer program]. Website Httptree Bio Ed Ac Uksoftwarefigtree Last Accessed Oct. 21, 2017.

Salvi, D., Maura, M., Pan, Z., Bologna, M.A., 2019. Phylogenetic systematics of *Mylabris* blister beetles (Coleoptera, Meloidae): a molecular assessment using species trees and total evidence. Cladistics 35, 243–268.

Schmitt, T., 2009. Biogeographical and evolutionary importance of the European high mountain systems. Front. Zool. 6, 1–10.

Schmitt, T., 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. Front. Zool. 4, 1–13.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Smyčka, J., Roquet, C., Renaud, J., Thuiller, W., Zimmermann, N.E., Lavergne, S., 2017. Disentangling drivers of plant endemism and diversification in the European Alps–A phylogenetic and spatially explicit approach. Perspect. Plant Ecol. Evol. Syst. 28, 19–27.

Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162–1169.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989.

Svenning, J.-C., Eiserhardt, W.L., Normand, S., Ordonez, A., Sandel, B., 2015. The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. Annu. Rev. Ecol. Evol. Syst. 46, 551–572.

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235.

Urbani, F., D'Alessandro, P., Frasca, R., Biondi, M., 2015. Maximum entropy modeling of geographic distributions of the flea beetle species endemic in Italy (Coleoptera: Chrysomelidae: Galerucinae: Alticini). Zool. Anz. - J. Comp. Zool. 258, 99–109. https://doi.org/10.1016/j.jcz.2015.08.002

Part II

PHYLOGEOGRAPHY OF ITALIAN ENDEMIC ALTICINI

CHAPTER 6

Refugia within refugia – assessing the link between Late Quaternary climate stability and spatial patterns in endemism and genetic divergence in the Italian peninsula

Introduction

Quaternary climatic oscillations have played an important role in shaping the current geographic distribution of species and their genetic diversity (Hewitt, 2004a, 2004b). During these period of strong climatic and environmental changes, species have undergone contractions and expansions of ranges following different routes and patterns based on their ecology, environmental tolerances, and adaptations (Stewart et al., 2010). Therefore, multiple and diverse species response to quaternary climatic changes have been documented, and different refugial areas have been identified for different species in which they have survived during unfavorable climatic periods. In Europe, temperate taxa were forced to repeated cycles of retreat within refugia during the glaciations and underwent range expansion during the subsequent interglacial phases, whereas alpine and cryophiles species experienced expansion during glacial periods and contractions during post-glacial phases (Hewitt, 1996, 2000). The three European peninsulas as Iberia, Italy and the Balkans, served as important southern Quaternary glacial refugia, due to their geographical position and topographical characteristics. In addition to the greater diversity of species, these peninsulas also retain greater intraspecific genetic diversity (Hewitt, 1996; Taberlet et al., 1998; Hewitt, 1999, 2000). This diversity, however, cannot always be explained by the classic scenario with three southern Mediterranean refuges. Multiple studies revealed more complex scenarios called "refugia-within-refugia" (Gómez and Lunt, 2007) according to which, within each southern peninsular refugium there were more distinct refugia areas. This implies that the main factors that shaped the genetic diversity of current species would be the allopatric differentiation that took place in these "refugia-within-refugia" settings, followed by demographic expansions of populations with secondary contacts and consequent mixing of distinct lineages (Bella et al., 2007; Canestrelli et al., 2007; Grill et al., 2009; Abellán and Svenning, 2014).

As for the Italian peninsula, the complex scenario of the "refugia-within-refugia" has been documented in the genetic patterns of various species of the Italian Apennines (Canestrelli et al., 2007; Vega et al., 2010; Mezzasalma et al., 2018). This mountain range, as already mentioned in Chapter 4, crosses the entire Italian peninsula. The Apennine landscapes is heterogeneous and extremely complex as it alternate mountains with peaks over 2000 m to valleys, with areas of abrupt transition from high peaks to sea level. Furthermore, despite the Apennines being one of the southernmost ranges in Europe, during the glacial periods, various mountain ranges were covered by glaciers (Giraudi and Frezzotti, 1997; Giraudi et al., 2004). All these characteristics have made these mountains a reservoir of specific diversity and intraspecific genetic diversity, allowing the coexistence of different ecological groups, from species adapted to high mountain environments, to more thermophilic species linked to the temperate belt. Precisely for this reason, the Apennines are a perfect laboratory to study how past climate changes have influenced and structured the current genetic diversity in the different ecosystems present on this mountain range.

Despite the diversity of environments and species, however, most of the phylogeographic studies carried out in the Apennines so far include a few vertebrate species linked to temperate environments. Little is known on patterns and processes relative to non-temperate species and on invertebrate groups. Alticini have a high number of endemic species in the central-southern Apennines, with a fairly homogeneous distribution between the different environments present in this region (D'Alessandro and Biondi, 2007; Biondi et al., 2013; Urbani et al., 2015). In fact, as reported in Chapter 1, there are endemic species linked to high mountain environments, such as *Longitarsus springeri* and *Psylliodes biondii*, species of temperate environments, such as *Longitarsus laureolae* and *Longitarsus zangherii*, and species of xerophilous environments, such as *Psylliodes ruffoi*. In this chapter, we investigate the genetic structure of the populations of different target species selected for the PhD project, through a phylogeographic and phylogenetic analysis. The main aim is to study the historical processes that have shaped current patterns of genetic diversity within different species and to compare how species with different ecological requirements have responded to Quaternary climate change. We will also try to verify whether the current intraspecific diversity of the target species is comparable with a scenario of multiple refugia and whether the observed patterns are similar between ecologically similar species.

Materials and Methods

Sample collection and morphological identification

Specimens analyzed in this chapter were collected and identified on a morphological basis as explained in Chapter 2. Geographic information on samples used in this study are reported in table Tab. 8.

Genus	Specie	ID Locality	Region	Locality	Coordinates	Altitude	COI-5'	COI-3'
Longitarsus	*springeri*	393	Marche (AP)	Lago di Pilato	42°49'24.7"N13°16'00.0"E	2054 m	4	4
Longitarsus	*springeri*	395	Marche (AP)	Lago di Pilato	42°49'36.6"N13°15'56.2"E	1958 m	4	4
Longitarsus	*springeri*	24	Abruzzo (AQ)	Gran Sasso	42°26'55"N13°32'24"E	2140 m	4	4
Longitarsus	*springeri*	318	Abruzzo (AQ)	Gran Sasso	42°27'12.29"N13°33'32.11"E	2293 m	4	4
Longitarsus	*springeri*	324	Abruzzo (AQ)	Monte Camicia	42°26'22.24"N13°43'4.49"E	2560 m	5	5
Longitarsus	*springeri*	325	Abruzzo (AQ)	Monte Camicia	42°26'26.46"N13°43'29.03"E	2345 m	2	2
Longitarsus	*springeri*	317	Abruzzo (AQ)	Monte Velino	42°9'31.53"N13°21'45.37"E	2153 m	4	4
Longitarsus	*springeri*	403	Abruzzo (AQ)	Punta Trento	42°10'01.5"N 13°24'07.1"E	2198 m	3	3
Longitarsus	*springeri*	323	Abruzzo (AQ)	Monte Sirente	42°9'0.05"N13°36'6.94"E	2232 m	5	5
Longitarsus	*springeri*	322	Abruzzo (AQ)	Monte Sirente	42°8'49.90"N13°36'32.46"E	2276 m	6	6
Longitarsus	*springeri*	391	Lazio (FR)	Monte Viglio	41°53'16.88"N13°22'29.06"E	2065 m	6	6
Longitarsus	*springeri*	398	Abruzzo (CH)	Maiella, Monte Focalone	42°06'47.9"N14°06'57.9"E	2537 m	2	2
Longitarsus	*springeri*	399	Abruzzo (CH)	Maiella, Monte Focalone	42°06'01.5"N14°06'32.9"E	2602 m	2	2
Longitarsus	*springeri*	401	Abruzzo (CH)	Maiella, Secondo Portone	42°05'55.9"N14°05'39.6"E	2537 m	2	2
Longitarsus	*springeri*	400	Abruzzo (PE)	Maiella, Monte Amaro	42°05'33.4"N14°04'58.4"E	2639 m	1	1
Longitarsus	*springeri*	321	Abruzzo (AQ)	Monte Marsicano	41°48'7.22"N13°52'3.10"E	2232 m	5	5
Longitarsus	*springeri*	405	Abruzzo (AQ)	Monte Meta	41°41'30.9"N 13°56'44.3"E	1925 m	1	1
Longitarsus	*springeri*	407	Lazio (FR)	Monte Meta	41°41'20.0"N 13°56'15.5"E	2217 m	4	4
Longitarsus	*springeri*	319	Molise (CB)	Matese, La Gallinola	41°26'3.10"N14°25'17.80"E	1923 m	9	9
Longitarsus	*springeri*	320	Molise (CB)	Matese, La Gallinola	41°26'4.60"N14°25'32.60"E	1721 m	1	1
Longitarsus	*zangherii*	366	Emilia-Romagna (FC)	Foreste Casentinesi, Campigna	43°52'14.7"N11°44'35.8"E	1087 m	10	10
Longitarsus	*zangherii*	369	Marche (AP)	Sentiero per lago di Pilato da Foce	42°51'03.8"N13°15'45.6"E	1376 m	6	6
Longitarsus	*zangherii*	373	Marche (AP)	Montemoncao	42°52'10.9"N13°20'10.6"E	941 m	2	2
Longitarsus	*zangherii*	365	Marche (AP)	Faete	42°46'05.6"N13°18'03.0"E	767 m	4	4
Longitarsus	*zangherii*	364	Lazio (RI)	Poggio D'api	42°43'31.4"N13°18'25.7"E	1032 m	3	3
Longitarsus	*zangherii*	357	Abruzzo (TE)	Vallinquina	42°43'54.1"N13°32'08.2"E	856 m	2	2

Longitarsus	*zangherii*	355	Abruzzo (TE)	Lago di Sbraccia_1	42°43'29.3"N13°32'44.6"E	1148 m	4	4
Longitarsus	*zangherii*	356	Abruzzo (TE)	Lago di Sbraccia_2	42°43'13.5"N13°32'44.3"E	1131 m	2	2
Longitarsus	*zangherii*	276	Abruzzo (TE)	Ceppo_1	42°40'17.4"N13°27'55.9"E	1341 m	2	2
Longitarsus	*zangherii*	285	Abruzzo (TE)	Frazione Alvelli	42°39'59.17"N13°29'28.56"E	889 m	4	4
Longitarsus	*zangherii*	281	Abruzzo (TE)	Ceppo_2	42°39'45.6"N13°27'46.8"E	1412 m	5	5
Longitarsus	*zangherii*	345	Abruzzo (TE)	Pietracamela	42°31'17.2"N13°33'51.9"E	1201 m	1	1
Longitarsus	*zangherii*	47	Abruzzo (TE)	Prati di Tivo_1	42°30'26.40"N13°33'14.46"E	1344 m	1	1
Longitarsus	*zangherii*	348	Abruzzo (TE)	Prati di Tivo_2	42°30'09.7"N13°33'41.7"E	1450 m	6	6
Longitarsus	*zangherii*	349	Abruzzo (TE)	Presso Fonte del Peschio	42°27'31.7"N13°40'11.6"E	1387 m	6	6
Longitarsus	*zangherii*	351	Abruzzo (PE)	Colle Mesole	42°26'49.7"N13°45'11.3"E	1422 m	6	6
Longitarsus	*zangherii*	352	Abruzzo (PE)	Vado di Sole	42°24'07.5"N13°47'33.0"E	1591 m	6	6
Longitarsus	*zangherii*	360	Abruzzo (PE)	Bosco di Lama Bianca	42°05'39.5"N14°02'30.6"E	1323 m	5	5
Longitarsus	*zangherii*	359	Abruzzo (AQ)	Valle di Fonte Romana	42°02'52.6"N14°02'52.7"E	1244 m	6	6
Longitarsus	*zangherii*	361	Abruzzo (AQ)	Bosco di Sant'Antonio	41°58'10.7"N 14°01'05.4"E	1377 m	4	4
Longitarsus	*zangherii*	358	Abruzzo (CH)	Quarto di Santa Chiara	41°56'54.8"N14°05'03.4"E	1377 m	6	6
Longitarsus	*zangherii*	363	Abruzzo (AQ)	Coppo dell'Orso	41°50'15.5"N14°02'46.8"E	1480 m	6	6
Longitarsus	*laureolae*	362	Abruzzo (AQ)	Bosco di Sant'Antonio	41°56'54.8"N14°05'03.4"E	1377 m	10	10
Longitarsus	*laureolae*	387	Calabria (CS)	Lungro, Santuario S. Maria del Monte	39°45'04.6"N16°05'32.7"E	1241 m	10	10
Longitarsus	*laureolae*	390	Calabria (CZ)	Monte Mancuso	39°00'38.3"N16°12'34.8"E	1066 m	8	8
Longitarsus	*laureolae*	327	Calabria (RC)	Aspromonte, Piano della Limina	38°25'10.3"N16°10'36.0"E	955 m	8	8
Longitarsus	*laureolae*	388	Calabria (RC)	Aspromonte, Zervò	38°13'19.4"N15°59'49.4"E	1187 m	7	7
Longitarsus	*laureolae*	330	Sicilia (ME)	Bosco di Malabotta	37°58'19.3"N15°03'14.4"E	1248 m	1	1
Longitarsus	*laureolae*	310	Sicilia (ME)	Bosco di Malabotta	37°58'22.8"N15°03'04.5"E	1238 m	8	8
Longitarsus	*laureolae*	328	Sicilia (ME)	Rocche dell'Argimusco	37°59'35.0"N15°02'10.9"E	1151 m	1	1
Longitarsus	*laureolae*	329	Sicilia (ME)	Nebrodi, Monte Soro	37°55'31.8"N14°40'25.4"E	1632 m	2	2
Longitarsus	*laureolae*	389	Sicilia (ME)	Nebrodi, Villa Miraglia	37°54'51.8"N14°39'18.3"E	1520 m	8	8
Longitarsus	*laureolae*	447	Sicilia (AG)	Nebrodi, Portella dell'Obolo	37°53'57.8"N 14°30'05.6"E	1506 m	8	8
Psylliodes	*biondii*	367	Abruzzo (AQ)	Monte Corvo	42°28'29.7"N13°29'51.5"E	2387 m	5	5
Psylliodes	*biondii*	375	Abruzzo (TE)	Sella dei due Corni	42°28'29.6"N13°33'43.8"E	2517 m	3	3

Psylliodes	*biondii*	104	Abruzzo (AQ)	Gran Sasso, Passo della Portella	42°26'56.1"N13°32'24.4"E	2246 m	7	7
Psylliodes	*biondii*	368	Abruzzo (AQ)	Pizzo Trento	42°09'48.1"N13°25'45.3"E	2054 m	8	8
Psylliodes	*biondii*	353	Abruzzo (AQ)	Neviera del Sirente	42°08'13.4"N13°37'51.6"E	1800-2000 m	8	8
Psylliodes	*biondii*	374	Abruzzo (AQ)	Sopra Aremogna	41°49'02.6"N14°00'50.5"E	2111 m	6	6
Psylliodes	*biondii*	446	Molise (IS)	Monte Miletto	41°26'59.8"N 14°22'19.9"E	2047 m	2	2
Psylliodes	*ruffoi*	335	Basilicata (MT)	Accettura, Piccole Dolomiti Lucane	40°32'46.5"N 16°05'45.2"E	484 m	3	3
Psylliodes	*ruffoi*	378	Campania (SA)	Parco Nazionale del Cilento, San Biase	40°12'04.8"N 15°17'59.8"E	723 m	3	3
Psylliodes	*ruffoi*	381	Calabria (CZ)	Girifalco	38°47'27.6"N 16°25'01.7"E	777 m	1	1
Psylliodes	*ruffoi*	338	Calabria (VV)	Parco Regionale delle Serre, Ninfo	38°32'07.7"N 16°18'19.2"E	1035 m	11	11
Psylliodes	*ruffoi*	382	Calabria (RC)	Piano della Limina	38°23'33.8"N 16°11'17.4"E	820 m	2	2
Psylliodes	*ruffoi*	383	Calabria (RC)	Piano della Limina	38°25'25.5"N 16°09'41.2"E	940 m	7	7
Psylliodes	*ruffoi*	339	Calabria (RC)	Aspromonte	38°22'55.2"N 16°09'39.0"E	899 m	1	1
Psylliodes	*ruffoi*	384	Calabria (RC)	Piano della Limina	38°22'55.2"N 16°09'39.0"E	899 m	5	5
Psylliodes	*ruffoi*	385	Calabria (RC)	Aspromonte, vista sui due mari	38°15'20.7"N 16°03'27.0"E	1011 m	9	9
Psylliodes	*ruffoi*	386	Sicilia (ME)	Pizzo Impegna	38°10'27.2"N 15°28'22.3"E	874 m	11	11
Psylliodes	*ruffoi*	340	Sicilia (ME)	Fiumedinisi e Monte Scuderi	38°02'32.0"N 15°21'31.7"E	495 m	1	1
Psylliodes	*ruffoi*	341	Sicilia (ME)	Bosco di Malabotta	37°58'22.8"N 15°03'04.5"E	1238 m	3	3
Psylliodes	*ruffoi*	342	Sicilia (ME)	Bosco di Malabotta	37°58'41.8"N 15°02'55.1"E	1238 m	1	1
Psylliodes	*ruffoi*	343	Sicilia (ME)	Rocche dell'Argimusco	37°59'35.0"N 15°02'10.9"E	1151 m	2	2

Tab.8 Information on collection sites and sequence data for the 5 target species used in this study.

DNA extraction and amplification

DNA of specimens of target species was extracted using non-invasive DNA extraction methods describe in Chapter 1. Two fragments of the cytochrome oxidase subunit I (*cox1*) gene were amplified through polymerase chain reaction (PCR): the standard barcode region was amplified using universal primers LCO1490 and HC02198 (Folmer et al., n.d.) and the 3' fragment using primers C1-J-2183 and TL2-N-3014 (Simon et al., 1994). Amplification conditions were those described in Salvi et al. (2019). Successful amplification was determined by gel electrophoresis and PCR products were purified and sequenced by an external service (Genewitz, UK). Sequence data information on samples used in this study are reported in table Tab. 8.

DNA sequence data analysis

The obtained chromatograms of each sequence were manually edited and assembled into a consensus sequence using Geneious Prime 2021 (Biomatters Ltd., Auckland, New Zealand). Sequences of two gene fragments were aligned separately using MAFFT v7.450 with the G-INS-I progressive method algorithm (Katoh et al., 2002), concatenated in a single matrix and then used for the following analysis. Phylogenetic relationships were inferred using the Maximum Likelihood (ML) method using IQ-TREE 1.6.12 (Nguyen et al., 2015) as implemented in the W-IQ-TREE webserver (Trifinopoulos et al., 2016). The best substitution models of each partition in our concatenated matrix of all genetic markers were determined by the ModelFinder module, including flexible rate heterogeneity across sites models (Kalyaanamoorthy et al., 2017), based on the Akaike information criterion. We used the Edge Linked partition model to allow each partition to have its own evolutionary rate. Branch support was assessed by 1000 replicates of ultrafast bootstrapping (UFboot) (Hoang et al., 2018; Minh et al., 2013) and SH-like approximate likelihood ratio test (SH- aLRT) (Guindon et al., 2010). FigTree v1.3.1 (Rambaut and Drummond, 2009) was used to depict the trees.

The genealogical relationships between haplotypes were also estimated by a phylogenetic network approach, using both the statistical parsimony algorithm described by Templeton et al. (1992), and implemented in the software tcs 1.13 (Crandall et al., 2000) and the median-joining distance method (Bandelt et al., 1999) using the PopArt 1.7 software (Leigh and Bryant, 2015).

To estimate the divergence time of major mitochondrial lineages, we performed a Bayesian MCMC analysis in Beast 1.8.2 (Drummond and Rambaut, 2007). BEAUTI 1.8.2 (part of the Beast 1.82 package) was used to create the XML format input files for BEAST. The HKY + I + G model of nucleotide substitution was selected for the two *cox1* gene fragments. Calibration was based on a strict molecular clock model and a lognormal prior (mean in real space = 0.0168, stdev = 0.06) on the 3' fragment of the *cox1* clock rate based on the available substitution rate estimated by Papadopoulou et al. (2010) in Coleoptera, already used for Chrysomelidae (Brunetti et al., 2019; Kubisz et al., 2020; Gómez-Zurita and Cardoso, 2021). Clock models were unlinked between the two mitochondrial gene fragments whereas the tree topology was linked. A coalescent model was used, in particular an Extended Bayesian Skyline Plot (Heled and Drummond, 2008), with a linear model. Default values were used for the other priors. MCMC chains were of 300 million generations, sampling every 30,000 generations. TRACER 1.7 was used to verify the posterior distribution and the effective sample sizes (ESSs) from the MCMC output (Rambaut et al., 2018). TreeAnnotator v1.8.4 was used to summarize tree data according to the “mean height”, and the first 25% of trees were discarded to represent the “burn-in” period, which ended well after the stationarity of the chain likelihood values had been established. The tree and divergence time are displayed in FigTree 1.4.2 (Rambaut and Drummond, 2009).

Results

Longitarsus laureolae

The ML tree and haplotype network reconstructions (TCS and MJ) were congruent and revealed seven reciprocally monophyletic groups (haplogroups) within *L. laureolae* whose geographical distribution is shown in Fig. 21b (see also Fig. S9). The Sicilian haplogroup has a high diversity with 14 haplotypes (about half of the 27 identified haplotypes in the species) with the populations from Villa Miraglia and Portella dell’Olobo having the highest number of haplotypes (four haplotypes in Villa Miraglia and five haplotypes in Portella dell’Olobo). In Calabria, there are four haplogroups: two in the Aspromonte massif (one in the southern region and another one in the northern region); one haplogroup in the southern Sila (Mount Mancuso), and one in the southern Pollino (Lungro). In Abruzzo we find two haplogroups and a high diversity, with six haplotypes present in Maiella. In general, the divergence between haplogroups is low, with one or at most two substitutions. The haplogroups of the Limina plane are an exception, as they have six substitution steps from the closest haplotypes. The parsimony analysis done with TCS did not split the haplotypes into subnetworks (Fig. 21a)

The divergence time between major mitochondrial lineages is shown as a chronogram in Fig. 21. The oldest split within the entire group was estimated to have occurred in the early Middle Pleistocene (170 thousand years ago, Kya) between the northern clade (the most recent common ancestor (TMRCA) 85 Kya), including haplogroups from the Limina plain to Abruzzo, and the southern clade (TMRCA 105 Kya), with the Sicilian and southern Calabria haplogroups. The subsequent splits among the haplogroups fell within the period of the last glaciation (i.e from 80Kya to 20Kya; Fig. 21b).

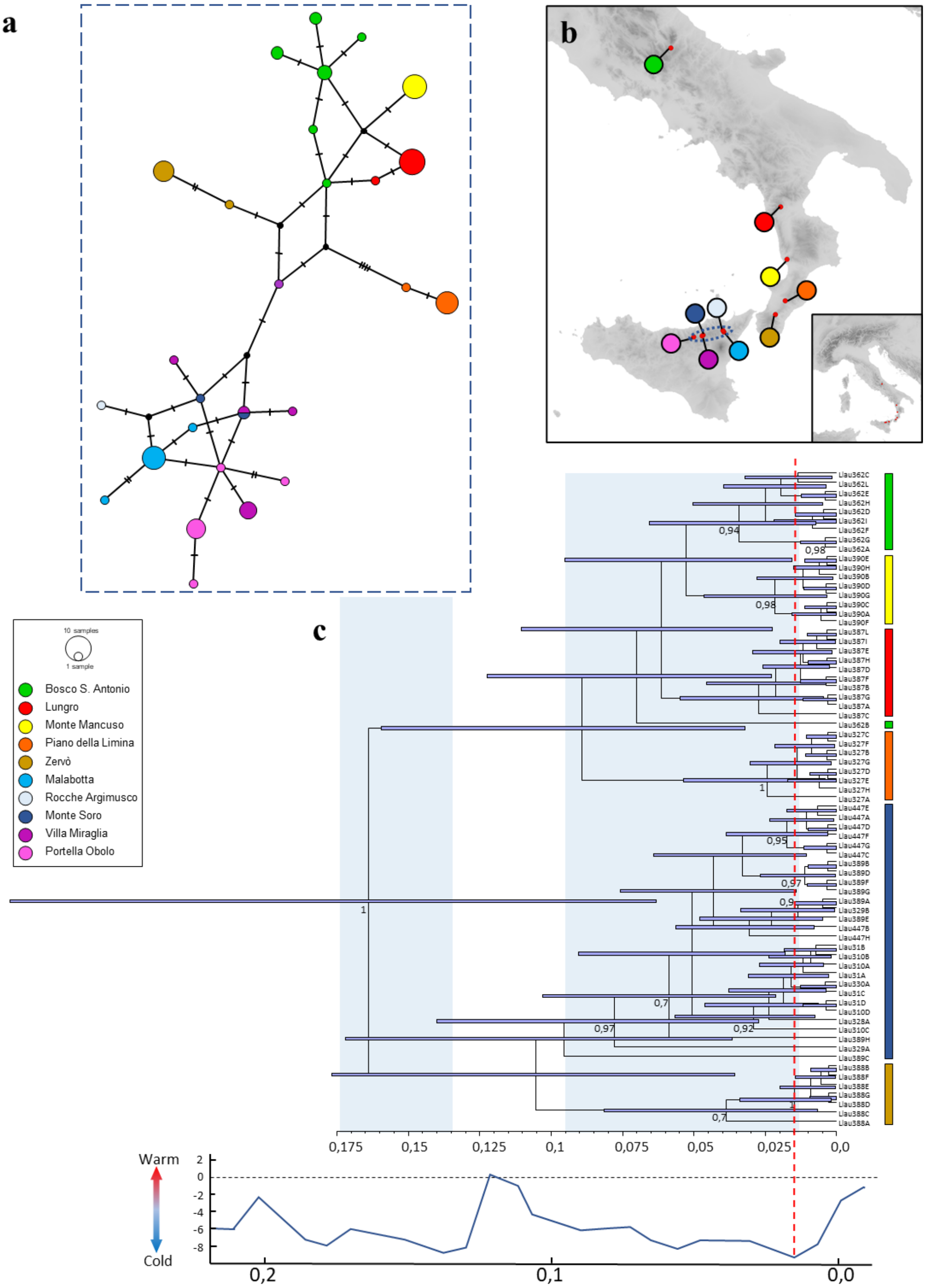


Fig. 21 (a) *Longitarsus laureolae* MJ haplotype network based on the two *cox1* gene fragments and with TCS sub-networks marked with a dashed line (b) Geographic distribution of haplotypes. (c) Time-calibrated BEAST phylogenetic tree of the major mitochondrial lineages of *Longitarsus laureolae*. Nodes with a posterior probability > 0.7 are reported. On each nodes the 95% High Posterior Density interval (HPD95) of node age is reported as a blue bar. Time scale in millions of years shown. The red dashed line highlights the Last Glacial Maximum (c. 18,000 years ago).

Longitarsus zangherii

The ML tree and haplotype network reconstructions (TCS and MJ) were congruent and revealed six reciprocally monophyletic haplogroups within *L. zangherii* whose geographical distribution is shown in Fig. 22b (see also Fig. S10). The five haplogroups are divided into two main clades. Clade A has one haplogroup in the area of the Sibillini mountains and in the northern area of Laga mountains and one in the southern part of the Gran Sasso (Vado di Sole). Clade B has four haplogroups: the northernmost haplogroup of the Casentinesi forest, one in south area of Laga and Gran Sasso with the highest number of haplotypes (15 different haplotypes) and, separated by the Tirino and Pescara valleys, the two southern haplogroups of the Maiella massif and Aremogna plain. The divergence between haplogroups ranges from a minimum of three substitutional steps between the two haplogroups of the Maiella, to a maximum of 50 substitutional steps between the haplogroups of Gran Sasso and the haplogroups of the Sibillini mountains and the northern area of Laga mountains. The parsimony analysis done with TCS identified three subgroups (Fig. 22a).

The divergence time between major mitochondrial lineages is shown as a chronogram in Fig. 22c. The oldest split within the entire group was estimated to have occurred in the late Early Pleistocene (1.65 Mya) between the northern clade A and clade B. In clade A, the subsequent split among the Vado di Sole haplogroup and the Sibillini and Northern Laga haplogroups (TMRCA 110 Kya) occurred in the penultimate glacial period (250 Kya). The subsequent splits among the haplogroups of clade A fell within the period of the last glaciation (between 80 Kya and 40 Kya). In clade B, the split event among northern haplogroup of the Casentinesi forests and the remaining central-south haplogroups occurred in the late Middle Pleistocene (630 Kya). The subsequent split event among the haplogroup of Gran Sasso and Monti della Laga and Maiella-Aremogna haplogroups occurred in the Middle Pleistocene (450 Kya). Finally, even in clade B, the subsequent splits among the haplogroups were estimate to the period of the last glaciation (between 80 Kya and 40 Kya) (Fig. 22c).

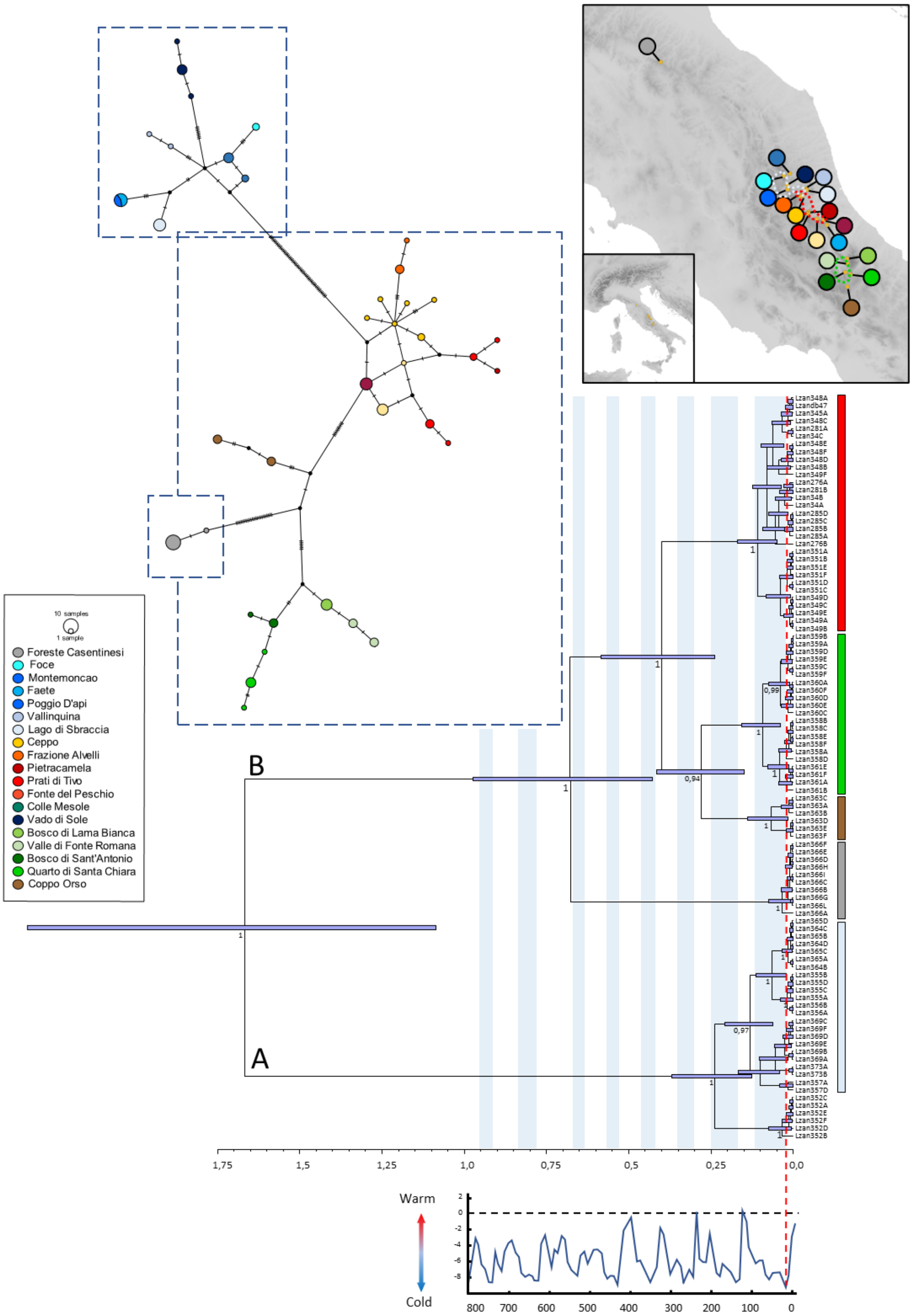


Fig.22 (a) *Longitarsus zangherii* MJ haplotype network based on the two *cox1* gene fragments and with TCS sub-networks marked with a dashed line (b) Geographic distribution of haplotypes. (c) Time-calibrated BEAST phylogenetic tree of the major mitochondrial lineages of *Longitarsus zangherii*. Nodes with a posterior probability > 0.7 are reported. On each nodes the 95% High Posterior Density interval (HPD95) of node age is reported as a blue bar. Time scale in millions of years shown. The red dashed line highlights the Last Glacial Maximum (c. 18,000 years ago).

Psylliodes ruffoi

The ML tree and haplotype network reconstructions (TCS and MJ) were congruent and revealed eight reciprocally monophyletic groups (haplogroups) within *P. ruffoi* whose geographical distribution is shown in Fig.23b (see also Fig.S11). In Sicily there are four haplogroups and three in Calabria, in the Aspromonte massif. In the northernmost part of the range there is a haplogroup in the Piccole Dolomiti Lucane and an haplogroup in Cilento massif. In general, each haplogroup presents a high level of diversity, with a mean of five haplotypes each. The divergence between the various haplogroups is marked, with a minimum of ten substitutional. The parsimony analysis done with TCS split the haplotypes into six subnetworks (Fig. 23a).

The divergence time between major mitochondrial lineages is shown as a chronogram in Figure (Fig. 23c). The oldest split within the entire group was estimated to have occurred in the late Early Pleistocene (1,15 Mya) between northern haplogroups and Calabrian/Sicilian haplogroups. The subsequent split among the Sicilian and Calabrian haplogroups took place in the early Middle Pleistocene (750 Kya). In Sicilian clade (TMRCA 630 Kya) subsequent splits occurred in the late Middle Pleistocene (from 370 Kya to 160 Kya). In Calabrian clade (TMRCA 610 Kya) subsequent split among Mount Scudieri clade and southern Aspromonte took place in Middle Pleistocene (550 Kya) (Fig. 23b).

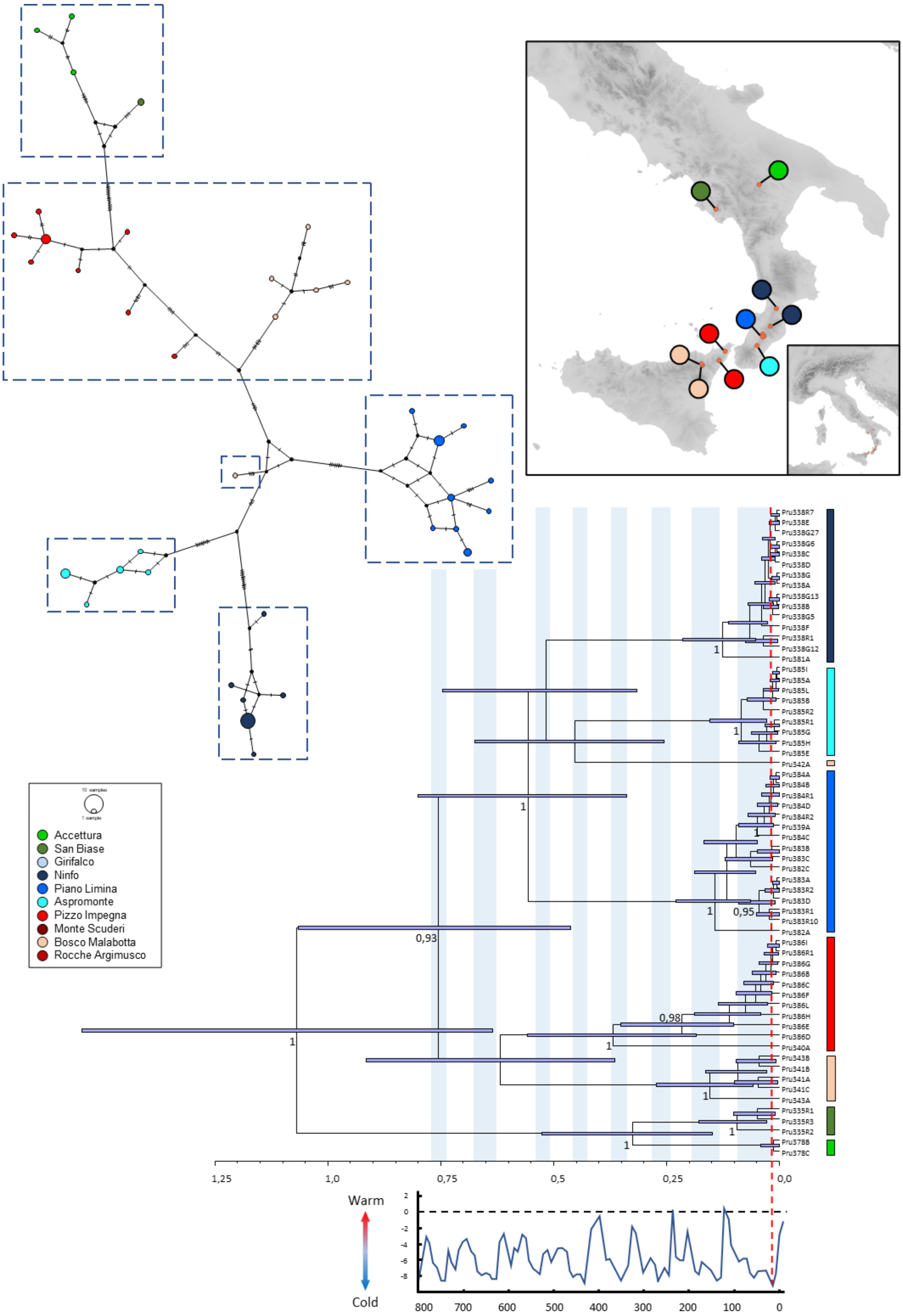


Fig.23 (a) *Psylliodes ruffoi* MJ haplotype network based on the two *cox1* gene fragments and with TCS sub-networks marked with a dashed line (b) Geographic distribution of haplotypes. (c) Time-calibrated BEAST phylogenetic tree of the major mitochondrial lineages of *Psylliodes ruffoi*. Nodes with a posterior probability > 0.7 are reported. On each nodes the 95% High Posterior Density interval (HPD95) of node age is reported as a blue bar. Time scale in millions of years shown. The red dashed line highlights the Last Glacial Maximum (c. 18,000 years ago).

Longitarsus springeri

The ML tree and haplotype network reconstructions (TCS and MJ) were congruent and revealed ten reciprocally monophyletic haplogroups within *L. laureolae* whose geographical distribution is shown in Fig. 24b (see also Fig. S12). In this species, each haplogroup is associated with a single mountain group. The Matese and Sirente haplogroup has the highest diversity with six and five haplotypes respectively (representing about half of the of the 28 identified haplotypes in the species). The divergence between haplogroups ranges from a minimum of three substitutions and a maximum of 35. The TCS parsimony analysis split haplotypes into four subnetworks (Fig. 24a).

The divergence time between major mitochondrial lineages is shown as a chronogram in Figure (Fig. 24c). The oldest split within the entire group was estimated to have occurred in the Early Pleistocene (700 Kya) between the haplogroup of Monte Viglio and all the other haplogroups. The subsequent splits among the Matese haplogroup and the Lake Pilate haplogroup occurred in the Middle Pleistocene (from 430 Kya to 380 Kya). Finally, the subsequent splits of the haplogroups of Abruzzo took place in the Middle Pleistocene (250 Kya) (Fig. 24c).

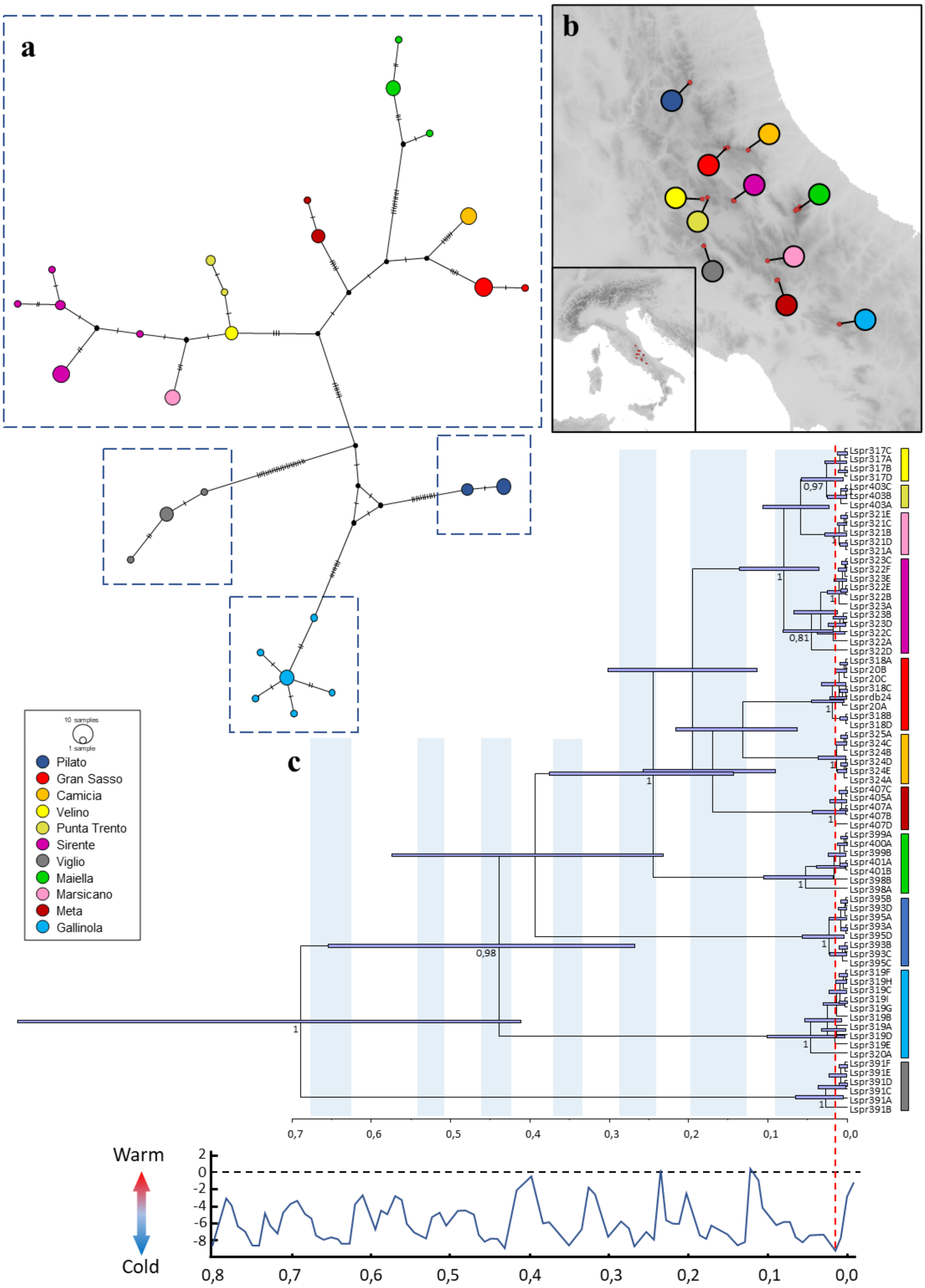


Fig.24 (a) *Longitarsus springeri* MJ haplotype network based on the two *cox1* gene fragments and with TCS sub-networks marked with a dashed line (b) Geographic distribution of haplotypes. (c) Time-calibrated BEAST phylogenetic tree of the major mitochondrial lineages of *Longitarsus springeri*. Nodes with a posterior probability > 0.7 are reported. On each nodes the 95% High Posterior Density interval (HPD95) of node age is reported as a blue bar. Time scale in millions of years shown. The red dashed line highlights the Last Glacial Maximum (c. 18,000 years ago).

Psylliodes biondii

The ML tree and haplotype network reconstructions (TCS and MJ) were congruent and revealed ten reciprocally monophyletic haplogroups within *P. biondii* whose geographical distribution is shown in Fig. 25b (see also Fig. S13). In this species each mountain system has its own unique haplogroups. The Pizzo Trento and Aremogna have two haplogroups each and a total of nine different haplotypes, reflecting a high diversity. In the Gran Sasso (Passo della Portella and Sella dei Due Corni) there are three different haplogroups. The Sirente, Monte Corvo and Monte Miletto have one haplogroups each. The divergence between haplogroups is of at least five substitutions. The TCS parsimony analysis did not split the haplotypes into subnetworks (Fig. 25a).

The divergence time between major mitochondrial lineages is shown as a chronogram in Fig. 25c. The oldest split within the entire group was estimated to have occurred in the late Middle Pleistocene (320 Kya). The subsequent splits leding to the separation of the current haplogroups occurred during the Late Pleisotcene ice ages (from 260 Kya to 160 Kya) (Fig. 25c).

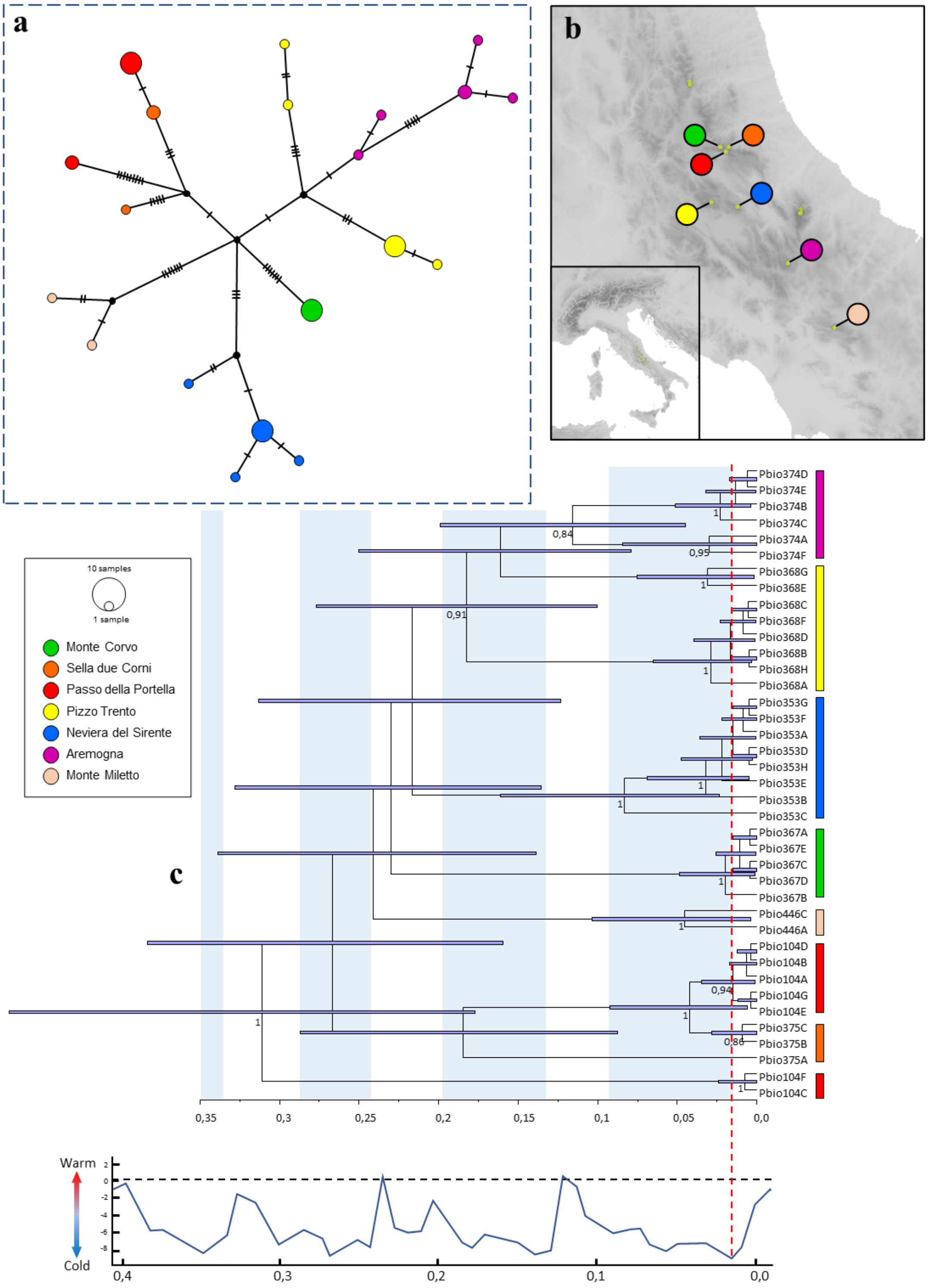


Fig.25 (a) *Psylliodes biondii* MJ haplotype network based on the two *cox1* gene fragments and with TCS sub-networks marked with a dashed line (b) Geographic distribution of haplotypes. (c) Time-calibrated BEAST phylogenetic tree of the major mitochondrial lineages of *Psylliodes biondii*. Nodes with a posterior probability > 0.7 are reported. On each nodes the 95% High Posterior Density interval (HPD95) of node age is reported as a blue bar. Time scale in millions of years shown. The red dashed line highlights the Last Glacial Maximum (c. 18,000 years ago).

Discussion

The five studied species are associated to three different environments of the central Apennines: (i) temperate forests (beech forests), (ii) mesophilic environments (mixed woods or oak) and (iii) high altitude environments. As such one might expect shared response to past climatic changes between species with similar ecological requirements. In contrast we observed mainly idiosyncratic patterns underlying individualistic species response, with marked differences in terms of geographic distribution and of time of divergence of intra-specific lineages.

L. laureolae and *L. zangherii* are the two species related to beech forest environments and present a portion of their ranges in syntopy. Despite this, the phylogeographic patterns of the two species are very different. In fact, *L. laureolae* has a low divergence between the various haplogroups suggesting a recent diversification between the main lineages, with the most ancestral split dating back less than 200 Kya. This splitting led to the formation of two main lineages, a southern one with the Sicilian haplogroups and the Zervò haplogroup (southern part of Calabria), and a northern one in which the remaining haplogroups of Italian peninsula are present. All the subsequent cladogenetic events, that led to the formation of remaining lineages of *L. laureolae*, occurred during the last glaciation, suggesting that the last glacial phase had an important impact on the genetic diversity of this species. The geographic distribution of lineages indicates an important role of the Strait of Messina and the Catanzaro plain as prominent biogeographical barriers that promoted the allopatric diversification of the main lineages of this species. This pattern corroborate previous finding of studies on southern temperate fauna (Canestrelli and Nascetti, 2008; Stöck et al., 2008; Mezzasalma et al., 2018). On the contrary, *L. zangherii* presents much more ancient splitting events, with the oldest event dated to about 1.60 Mya. Regarding differences in time of divergence, we emphasize that due to the intrinsic inaccuracies of the molecular clock (Ayala, 1997; Gibbons, 1998; Ayala, 1999; Welch and Bromham, 2005; Ho and Larson, 2006), caution is needed when using molecular data to date historical processes. The estimates reported are therefore to be considered indicative of approximate time windows of the inferred events. However, given the shared calibration approach significant differences between species in terms of time of intraspecific cladogenetic events can be considered as indicative of processes occurring during distinct climatic periods of the Plio-Pleistocene. The oldest split within *L. zangherii* led to the separation of the two main clades (clades A and B). The same clades then had a different evolutionary story. The clade A has undergone the main events of diversification in the last c 250 Kya, whereas within the clade B the separation of lineages leading to the main haplogroups occurred between 600 and 700 Kya. However, also for this species, a high number of split events clusters during the period of the last glaciation, giving rise to the current genetic diversity. It should be emphasized that both the clades have a lineage that is isolated with respect to the others and that is interposed within the range of the other clade. Given the origin of the two clades in the Early Pleistocene, this peculiar phylogeographic pattern is probably the result of multiple glacial cycles that occurred in the last 2 Mya. Following multiple glacial-interglacial cycles repeated events of contraction and expansion took place, resulting in extinctions and colonization processes that generated the particular pattern we see today. Furthermore, the disjoint range observed in both *L. laureolae* and *L. zangherii*, with the isolated populations of *L. laureolae* in Abruzzo and of *L. zangherii* in the Casentinesi region, is probably due to the absence of environmentally suitable areas along the Apennines. The ecological barrier represented by the lower altitude of the Umbrian-Marche (Bologna, 1994) and Campania sectors (D'Alessandro and Biondi, 2007) may have played an important role in determining the isolation, and the consequent divergence, between the populations of the northern, central and southern sectors.

P. ruffoi is a more mesophilic species, often found in clearing environments of mixed woods or oak. While *P. ruffoi* has a distribution pattern very similar to that of *L. laureolae*, it shares the phylogeographic pattern with *L. zangherii*. In fact, also *P. ruffoi* shows ancient splits, occurring since 1 Mya separating the northernmost lineages from the Sicilian/Calabrian ones. The subsequent splitting events leading to current haplogroups date back between 750 Kya and 500 Kya. The Sicilian and Calabrian haplogroups represent two separate lineages originated about 750 Kya, likely because of the geographical barrier of the Strait of Messina. However, unlike *L. laureolae* in which the Sicilian and Calabrian haplogroups present relatively low divergences, in this species the divergence that exists between the different lineages is more ancient. In Calabria, for example, within the region of the Serre Calabresi and Aspromonte it is possible to identify three lineages with an origin around 500 Kya. Paleogeographic reconstructions for the Calabria region support a scenario of allopatric fragmentation at different time scales during the Plio – Pleistocene (Ghisetti, 1979; Ghisetti and Vezzani, 1981; Tortorici, 1981; Caloi et al., 1989; Bonfiglio et al., 2002). In this period, in fact, the main mountain massifs of Calabria suffered repeated and prolonged isolation due to marine introgression. The sea basins located between these massifs ultimately emerged asynchronically, due to uplifting tendency of the area, until the Late Pleistocene. These events have left detectable footprints in in the distribution of species, as documented by the paleontological record, and in the pattern of intraspecific genetic diversity (Caloi et al., 1989; Bonfiglio et al., 2002; Podnar et al., 2005; Canestrelli et al., 2006, 2007; Vega et al., 2010). Also in Sicily the two haplogroups identified within *P. ruffoi* show an ancient divergence dating back to about 600 Kya, thus suggesting that the Peloritani and Nebrodi mountains played a similar role of biogeographic barrier like the Calabrian massifs (Bonfiglio et al., 2002). Unlike many temperate species, *P. ruffoi* show an ancient diversification that largely predates the last glacial period.

The two cryophilic species analyses in this study, *L. springeri* and *P. biondii*, have a sympatric distribution across the same mountain systems and often in the same localities. However, they have distinct and non-overlapping periods of activity and are almost never found together. Despite that, the phylogeographic patterns observed in these two species are quite similar to each other, with a recent intraspecific diversification that took place during the last glaciation. In *L. springeri* the splits that led to the formation of the main clades almost all occurred before the last interglacial (before 130 Kya) with the only exception of the lineages present in the Sirente-Velino and Monte Marsicano group, which appear to have a more recent separation linked to the last glacial period. In general, the lineages of the Abruzzo mountains appear to have a greater genetic affinity, while the lineages with greater divergence are those of the haplogroups of Monte Viglio, Matese and Sibillini. Populations on each mountain relief have its own haplotype, with a low haplotypic diversity (on average one, maximum two haplotypes). Also in *P. biondii* the formation of the main clades took place before the last interglacial and each mountain hosts an endemic haplogroup. In this species, the haplotypic diversity of each haplogroup is greater than in *L. springeri*. Both these two cryophilic species show a strong phylogeographic structure into distinct mountain peaks, suggesting a scenario of isolation due to the fragmentation of high altitude habitats during Pleistocene climatic oscillations. Indeed, the areas of the central Apennines have undergone strong environmental changes between the interglacial and glacial periods, with the presence of glaciers in many of the mountain systems present (Giraudi and Frezzotti, 1997; Giraudi et al., 2004; Hughes et al., 2006). This could explain the present patterns, in which ancient haplotypes are still present on each mountain thanks to in situ survival on small ice-free islands of habitat surrounded by the ice-sheet, so-called nunataks or sky islands (Janetschel, 1956; Stehlik, 2000; Knowles, 2001; Tribsch and Schönswetter, 2003; McCormack et al., 2009; Lohse et al., 2011).

This study allowed to shed light on the evolutionary processes that have affected the Apennine biota, with a focus on understudied fauna such as insects. We found that species associated to temperate and mesophilic environments show different evolutionary histories. This finding corroborates that species often have individualistic response to environmental changes (Stewart et al., 2010; Hornsby and Matocq, 2012). As regards the two cryophilic species, a similar pattern was observed for the two species, which seems to confirm a process of isolation in suitable mountain areas typical of the so-called Sky islands. In general, for all species it can be observed that, regardless of when the split events of the main lineages occurred, each mountain complex has its own unique haplogroups, confirming the importance of the Apennines and in general of mountain systems as catalysts of evolutionary processes and unique reservoirs of biodiversity.

References

Abellán, P., Svenning, J.-C., 2014. Refugia within refugia–patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. Biol. J. Linn. Soc. 113, 13–28.

Ayala, F.J., 1999. Molecular clock mirages. Bioessays 21, 71–75.

Ayala, F.J., 1997. Vagaries of the molecular clock. Proc. Natl. Acad. Sci. 94, 7776–7783.

Bandelt, H.-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bella, J., Serrano, L., Orellana, J., Mason, P., 2007. The origin of the *Chorthippus parallelus* hybrid zone: chromosomal evidence of multiple refugia for Iberian populations. J. Evol. Biol. 20, 568–576.

Biondi, M., Urbani, F., D'Alessandro, P., 2013. Endemism patterns in the Italian leaf beetle fauna (Coleoptera, Chrysomelidae). ZooKeys 177.

Bologna, M.A., 1994. Considerazioni sul popolamento animale e vegetale dell'Appennino umbro-marchigiano in conclusione del XXIX Congresso della Società Italiana di Biogeographia. Biogeographia, Lav. Soc. Ital. Biogeog. (n.s.), 17 (1993): 321-365.

Bonfiglio, L., Mangano, G., Marra, A.C., Masini, F., Pavia, M., Petruso, D., 2002. Pleistocene calabrian and sicilian bioprovinces. Geobios 35, 29–39.

Brunetti, M., Magoga, G., Iannella, M., Biondi, M., Montagna, M., 2019. Phylogeography and species distribution modelling of *Cryptocephalus barii* (Coleoptera: Chrysomelidae): is this alpine endemic species close to extinction? ZooKeys 856, 3.

Caloi, L., Malatesta, A., Palombo, M., 1989. Biogeografia della Calabria meridionale durante il Quaternario. Accad. Peloritana Dei Pericolanti Cl. FF MM NN 67, 261–278.

Canestrelli, D., Cimmaruta, R., Costantini, V., Nascetti, G., 2006. Genetic diversity and phylogeography of the Apennine yellow-bellied toad *Bombina pachypus*, with implications for conservation. Mol. Ecol. 15, 3741–3754.

Canestrelli, D., Cimmaruta, R., Nascetti, G., 2007. Phylogeography and historical demography of the Italian treefrog, *Hyla intermedia*, reveals multiple refugia, population expansions and secondary contacts within peninsular Italy. Mol. Ecol. 16, 4808–4821.

Canestrelli, D., Nascetti, G., 2008. Phylogeography of the pool frog *Rana* (*Pelophylax*) *lessonae* in the Italian peninsula and Sicily: multiple refugia, glacial expansions and nuclear–mitochondrial discordance. J. Biogeogr. 35, 1923–1936.

Crandall, M.C.D.P.K., Clement, M., Posada, D., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1660.

D'Alessandro, P., Biondi, M., 2007. Analisi zoogeografica della crisomelidofauna appenninica legata ai piani altitudinali montano, subalpino e alpino (Coleoptera, Chrysomelidae). Biogeogr. J. Integr. Biogeogr. 28.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., n.d. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates Mol Mar Biol Biotechnol. 1994; 3: 294–9.

Ghisetti, F., 1979. Evoluzione neotettonica dei principali sistemi di faglie della Calabria centrale. Boll. Della Soc. Geol. Ital. 98, 387–430.

Ghisetti, F., Vezzani, L., 1981. Contribution of structural analysis to understanding the geodynamic evolution of the Calabrian arc (Southern Italy). J. Struct. Geol. 3, 371–381.

Gibbons, A., 1998. Calibrating the mitochondrial clock. Science 279, 28–29.

Giraudi, C., Ehlers, J., Gibbard, P., 2004. The Apennine glaciations in Italy. Quarternary Glaciat. Chronol. Part Eur. Dev. Quat. Sci. 2, 215–224.

Giraudi, C., Frezzotti, M., 1997. Late Pleistocene glacial events in the central Apennines, Italy. Quat. Res. 48, 280–290.

Gómez, A., Lunt, D.H., 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula, in: Phylogeography of Southern European Refugia. Springer, pp. 155–188.

Gómez-Zurita, J., Cardoso, A., 2021. Molecular systematics, higher-rank classification and Gondwanan origins of Cryptocephalinae leaf beetles. Zool. Scr. 50, 592–615.

Grill, A., Amori, G., Aloise, G., Lisi, I., Tosi, G., Wauters, L.A., Randi, E., 2009. Molecular phylogeography of European *Sciurus vulgaris*: refuge within refugia? Mol. Ecol. 18, 2687–2699.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 1–15.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.

Hewitt, G.M., 2004a. Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 359, 183–195.

Hewitt, G.M., 2004b. The structure of biodiversity–insights from molecular phylogeography. Front. Zool. 1, 1–16.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. 68, 87–112.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247–276.

Ho, S.Y., Larson, G., 2006. Molecular clocks: when timesare a-changin'. TRENDS Genet. 22, 79–83.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Hornsby, A.D., Matocq, M.D., 2012. Differential regional response of the bushy-tailed woodrat (*Neotoma cinerea*) to late Quaternary climate change. J. Biogeogr. 39, 289–305.

Hughes, P., Woodward, J., Gibbard, P., 2006. Quaternary glacial history of the Mediterranean mountains. Prog. Phys. Geogr. 30, 334–364.

Janetschel, H., 1956. Das Problem der inneralpinen Eiszeitüberdauerung durch Tiere: ein Beitrag zur Geschichte der Nivalfauna. Springer.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Knowles, L.L., 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. Mol. Ecol. 10, 691–701.

Kubisz, D., Magoga, G., Mazur, M., Montagna, M., Ścibior, R., Tykarski, P., Kajtoch, Ł., 2020. Biogeography and ecology of geographically distant populations of sibling *Cryptocephalus* leaf beetles. Eur. Zool. J. 87, 223–234.

Leigh, J.W., Bryant, D., 2015. popart: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116.

Lohse, K., Nicholls, J.A., Stone, G.N., 2011. Inferring the colonization of a mountain range—refugia vs. nunatak survival in high alpine ground beetles. Mol. Ecol. 20, 394–408.

McCormack, J.E., Huang, H., Knowles, L.L., Gillespie, R., Clague, D., 2009. Sky islands. Encycl. Isl. 4, 841–843.

Mezzasalma, M., Di Febbraro, M., Guarino, F.M., Odierna, G., Russo, D., 2018. Cold-blooded in the Ice Age:“refugia within refugia”, inter-and intraspecific biogeographic diversification of European whipsnakes (Squamata, Colubridae, Hierophis). Zoology 127, 84–94.

Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol. Biol. Evol. 27, 1659–1672.

Podnar, M., Mayer, W., Tvrtković, N., 2005. Phylogeography of the Italian wall lizard, *Podarcis sicula,* as revealed by mitochondrial DNA sequences. Mol. Ecol. 14, 575–588.

Rambaut, A., Drummond, A., 2009. FigTree version 1.3. 1 [computer program]. Website Httptree Bio Ed Ac Uksoftwarefigtree Last Accessed Oct. 21, 2017.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901.

Salvi, D., Maura, M., Pan, Z., Bologna, M.A., 2019. Phylogenetic systematics of *Mylabris* blister beetles (Coleoptera, Meloidae): a molecular assessment using species trees and total evidence. Cladistics 35, 243–268.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Stehlik, I., 2000. Nunataks and peripheral refugia for alpine plants during quaternary glaciation in the middle part of the Alps. Bot. Helvetica 110, 25–30.

Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: individualistic responses of species in space and time. Proc. R. Soc. B Biol. Sci. 277, 661–671.

Stöck, M., Sicilia, A., Belfiore, N.M., Buckley, D., Lo Brutto, S., Lo Valvo, M., Arculeo, M., 2008. Post-Messinian evolutionary relationships across the Sicilian channel: mitochondrial and nuclear markers link a new green toad from Sicily to African relatives. BMC Evol. Biol. 8, 1–19.

Taberlet, P., Fumagalli, L., Wust-Saucy, A., Cosson, J., 1998. Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. 7, 453–464.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132, 619–633.

Tortorici, L., 1981. Analisi delle deformazioni fragili dei sedimenti postorogeni della Calabria settentrionale. Boll. Della Soc. Geol. Ital. 100, 291–308.

Tribsch, A., Schönswetter, P., 2003. Patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. Taxon 52, 477–497.

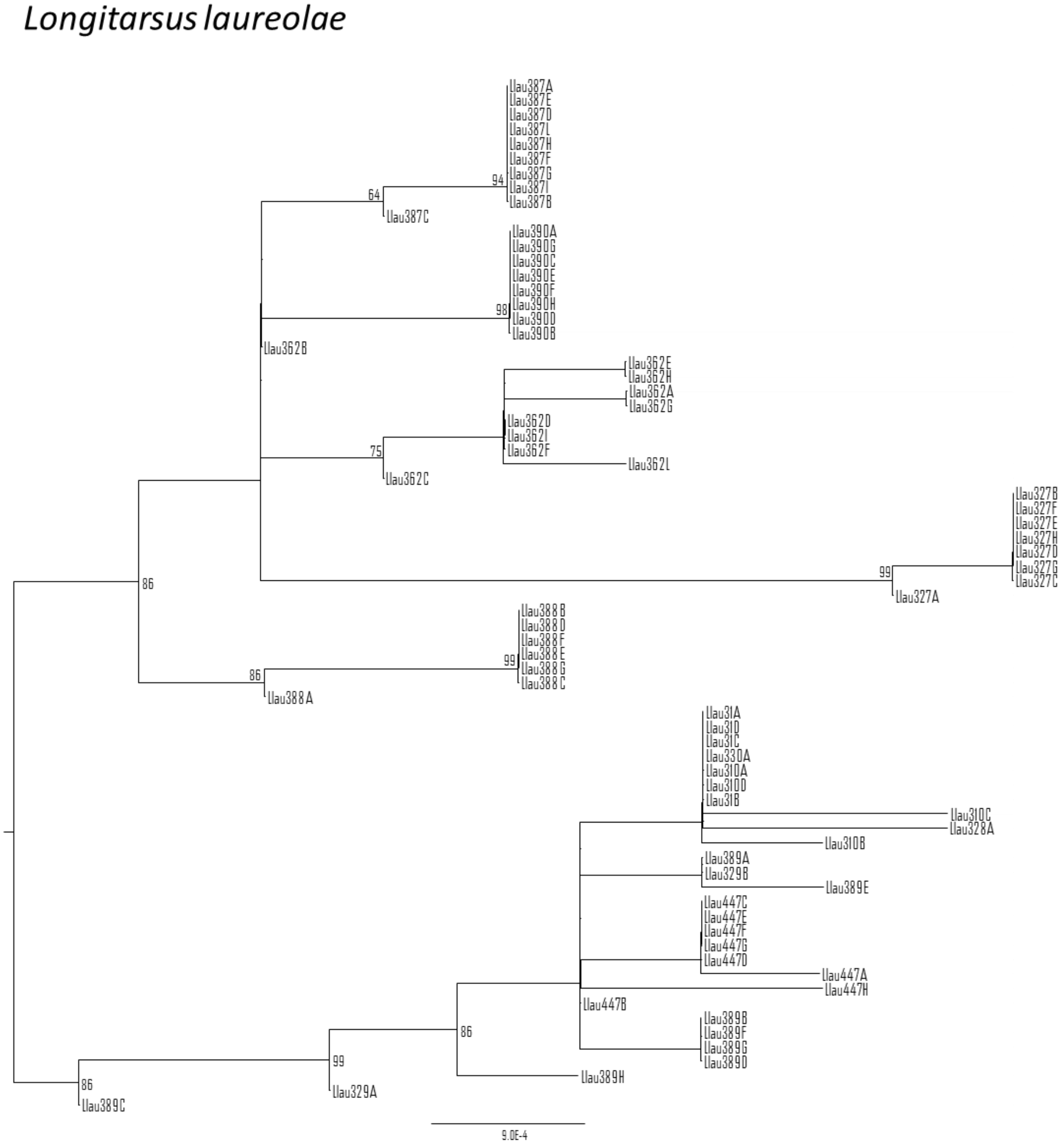
Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235.

Urbani, F., D’Alessandro, P., Frasca, R., Biondi, M., 2015. Maximum entropy modeling of geographic distributions of the flea beetle species endemic in Italy (Coleoptera: Chrysomelidae: Galerucinae: Alticini). Zool. Anz. - J. Comp. Zool. 258, 99–109. https://doi.org/10.1016/j.jcz.2015.08.002

Vega, R., Amori, G., Aloise, G., Cellini, S., Loy, A., Searle, J.B., 2010. Genetic and morphological variation in a Mediterranean glacial refugium: evidence from Italian pygmy shrews, *Sorex minutus* (Mammalia: Soricomorpha). Biol. J. Linn. Soc. 100, 774–787.

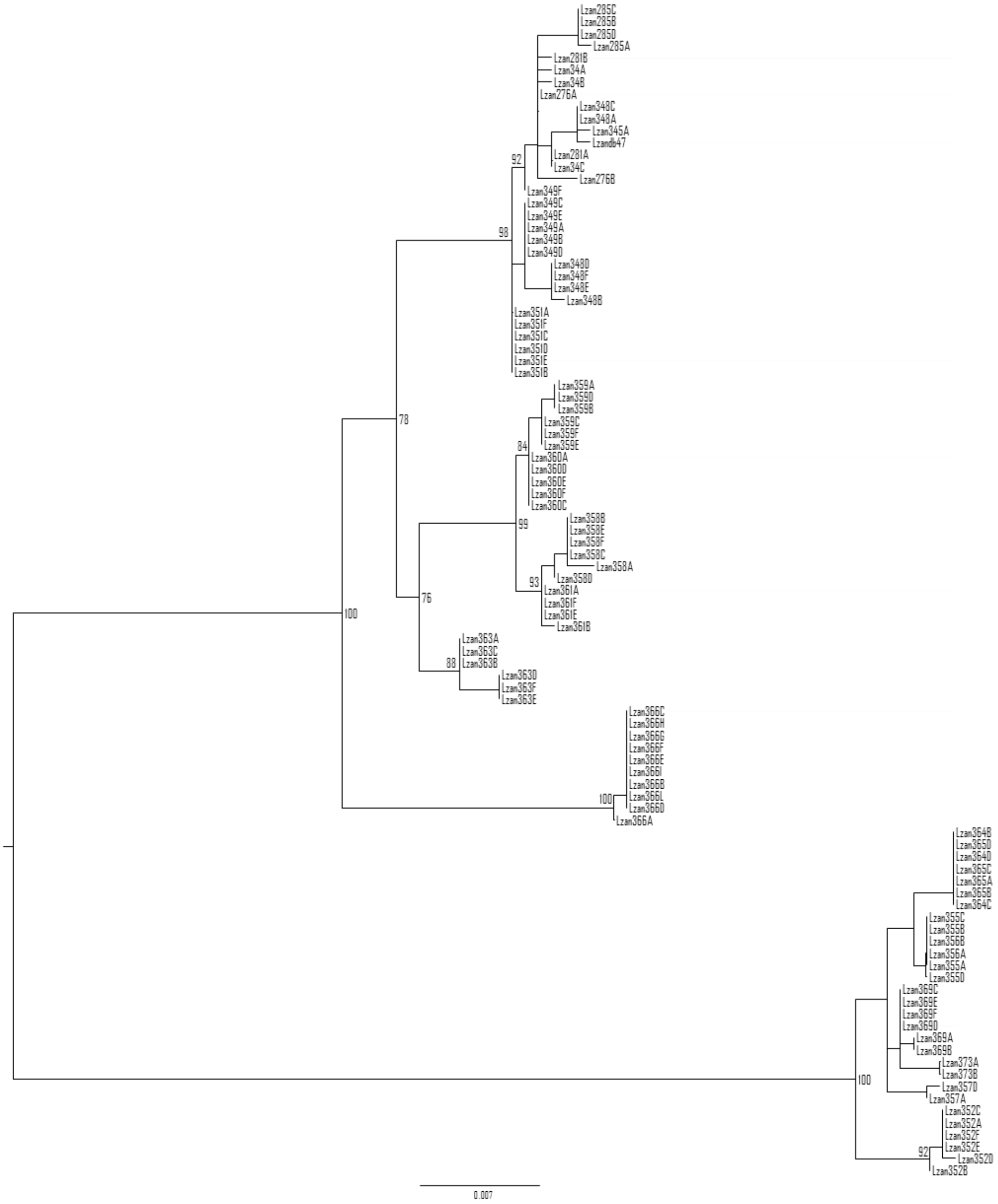
Welch, J.J., Bromham, L., 2005. Molecular dating when rates vary. Trends Ecol. Evol. 20, 320–327.

Supplementary



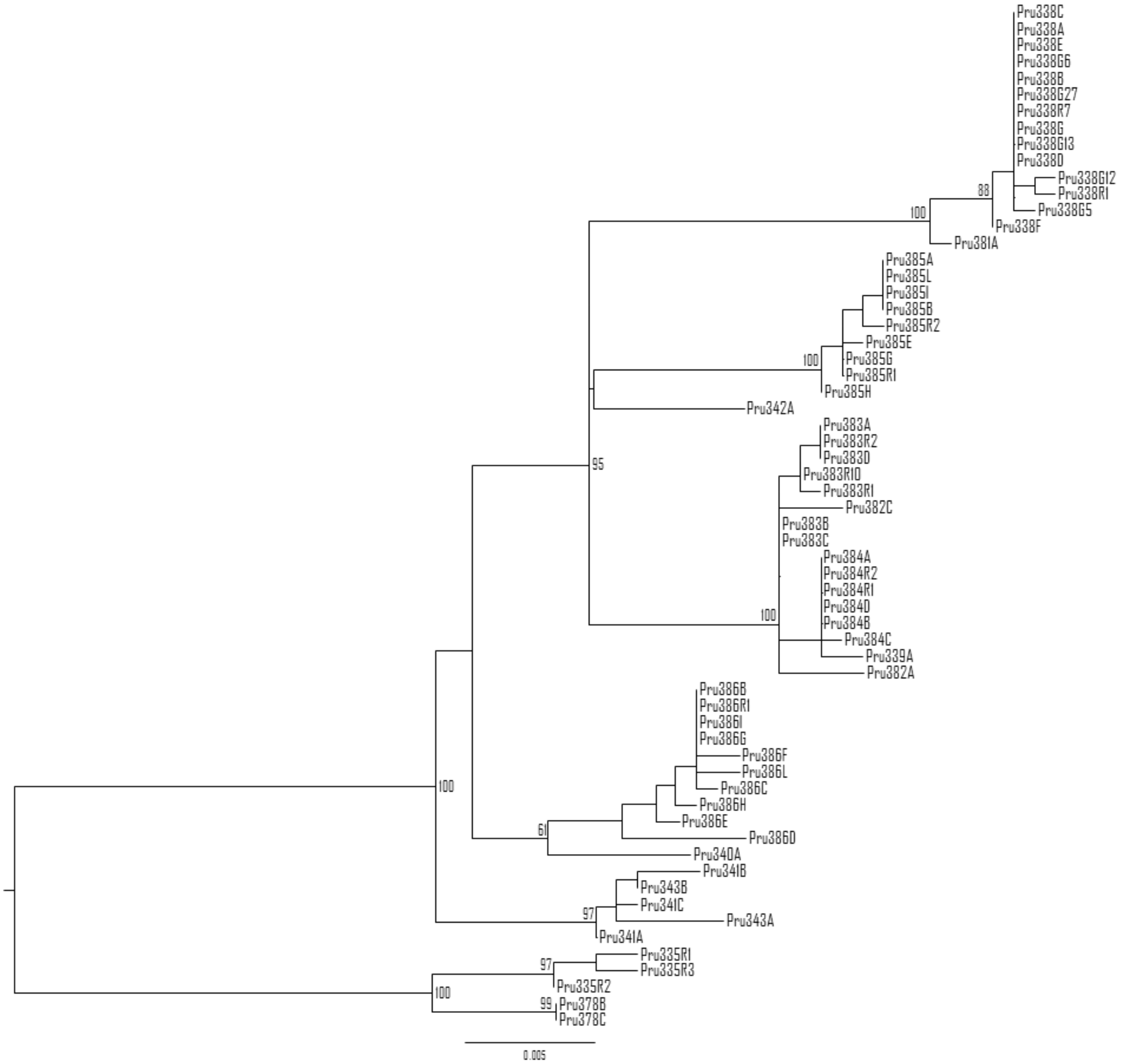
S9 Fig. Maximum Likelihood tree of *Longitarsus laureolae* based on two concatenated fragments of *cox1*. In correspondence of nodes are reported the bootstrap support (BS, left) from Maximum Likelihood. The nodes supported is reported only for BS>60.

Longitarsus zangherii

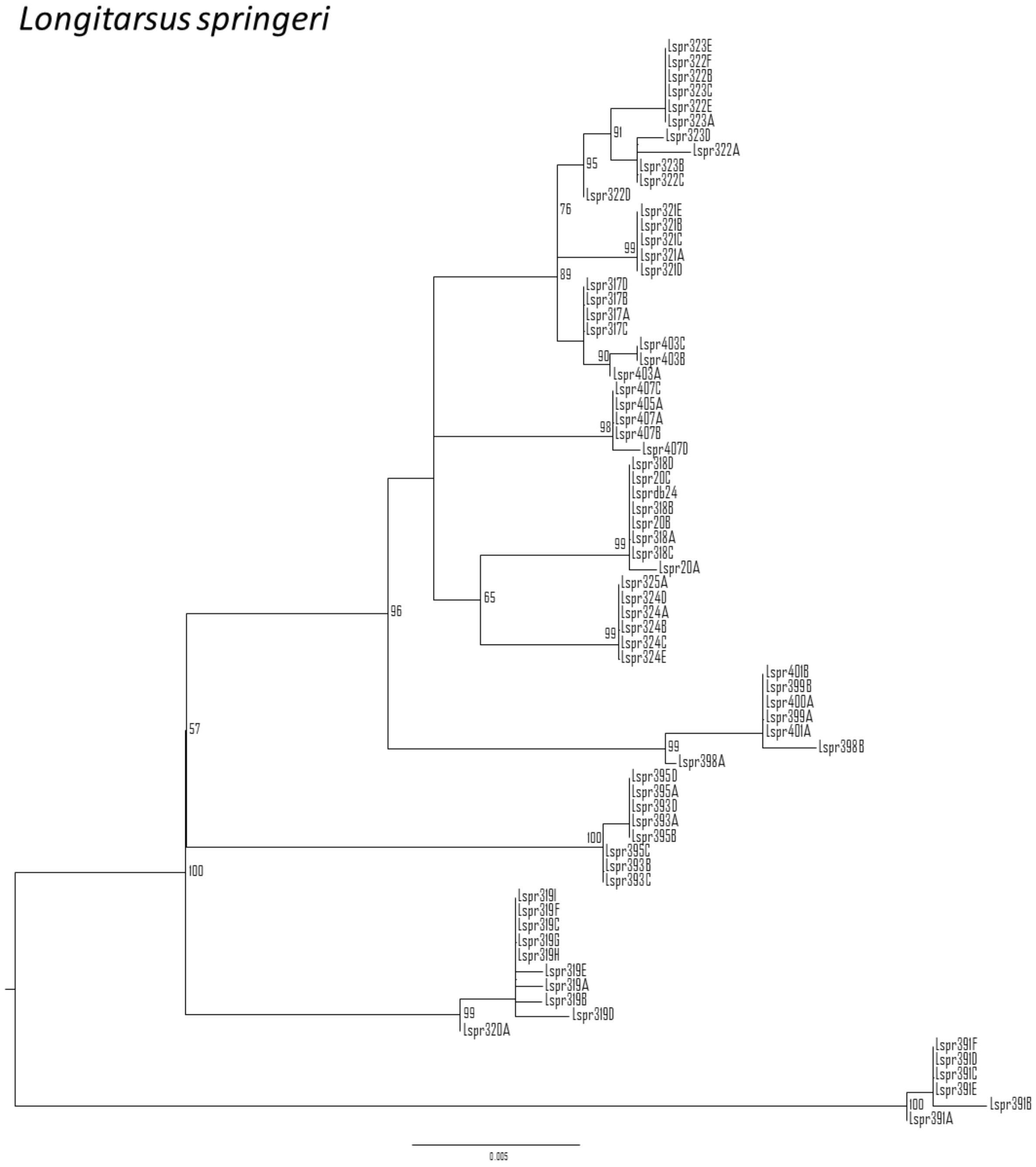


S10 Fig. Maximum Likelihood tree of *Longitarsus zangherii* based on two concatenated fragments of *cox1*. In correspondence of nodes are reported the bootstrap support (BS, left) from Maximum Likelihood. The nodes supported is reported only for BS>60.

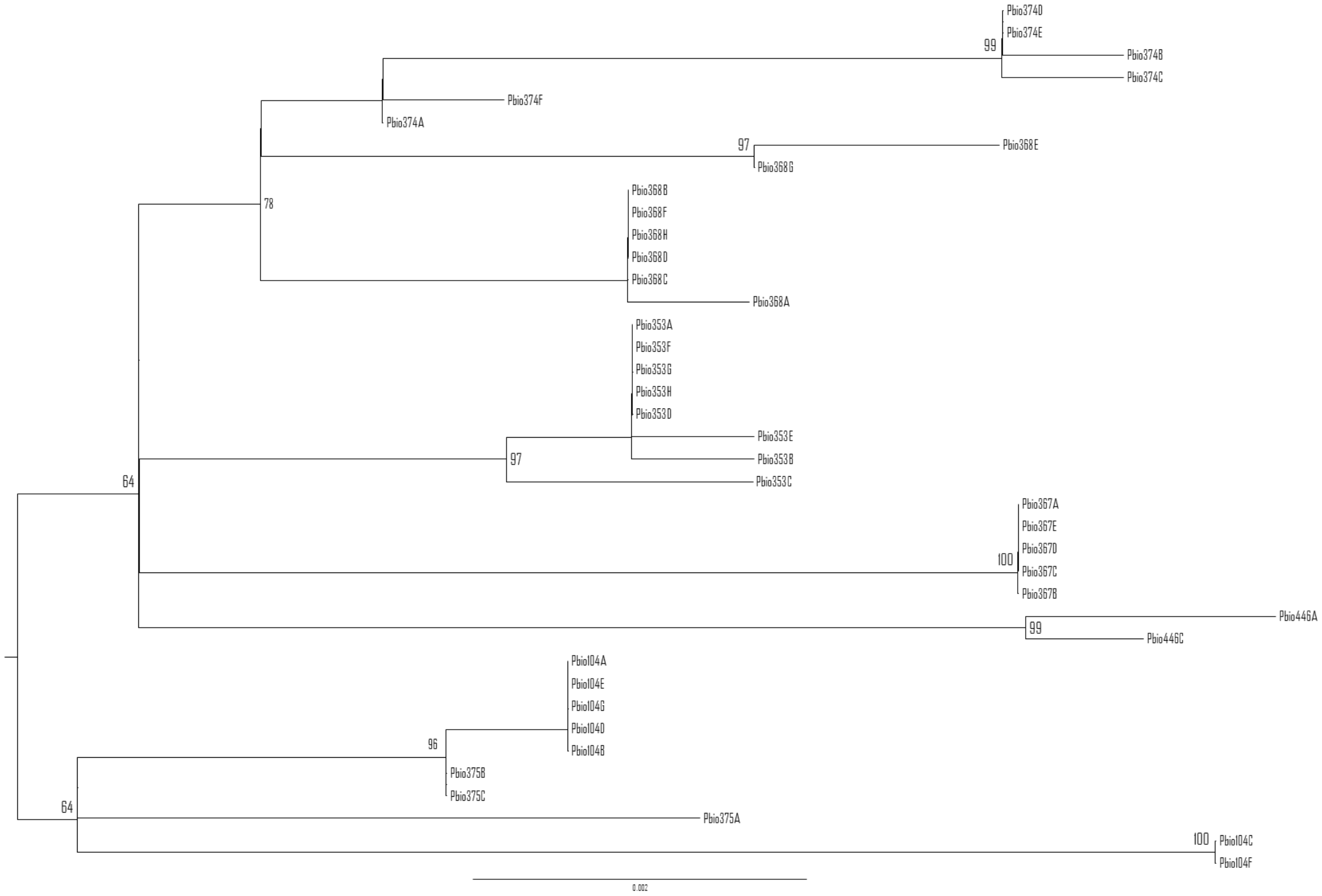
Psylliodes ruffoi



S11 Fig. Maximum Likelihood tree of *Psylliodes ruffoi* based on two concatenated fragments of *cox1*. In correspondence of nodes are reported the bootstrap support (BS, left) from Maximum Likelihood. The nodes supported is reported only for BS>60.



S12 Fig. Maximum Likelihood tree of *Longitarsus springeri* based on two concatenated fragments of *cox1*. In correspondence of nodes are reported the bootstrap support (BS, left) from Maximum Likelihood. The nodes supported is reported only for BS>60.



S13 Fig. Maximum Likelihood tree of *Psylliodes biondii* based on two concatenated fragments of *cox1*. In correspondence of nodes are reported the bootstrap support (BS, left) from Maximum Likelihood. The nodes supported is reported only for BS>60.

CHAPTER 7

Effect of Quaternary climatic oscillations on temperate and alpine biotas: insights from comparative phylogeography and historical demography of Italian endemic Alticini along the Apennine mountains

Introduction

In the previous Chapter we used a phylogeographic approach to analysed the relationships between intraspecific mitochondrial lineages. Results indicated that the genetic structure of the studied species was strongly influenced by the environmental changes induced by the glacial and interglacial cycles of the Quaternary and their interplay with Apennine topography. The different species, even those sharing a similar ecology and distribution, responded differently to these changes, each presenting its own pattern. The observed genetic structure is suggestive of cycles of range contraction towards refugia. Such spatial contraction processes have been accompanied by demographic contractions and bottlenecks. Some taxa underwent to population size decrease (Mapelli et al., 2012) even close to extinction (Nogués-Bravo et al., 2010), whereas others have remained relatively more stable by persisting in widespread suitable environments (Davis and Shaw, 2001) or migrating to more suitable areas (Hewitt, 2000; Davis and Shaw, 2001). Following climatic amelioration most of these organisms experience spatial and demographic expansion (Hewitt, 1996, 2000, 2004a, 2004b) .

The signature of these demographic changes can be detected from genetic data by applying coalescent models to the inference of intraspecific gene genealogies (Ho and Shapiro, 2011). The coalescent framework has contributed important information about the demographic history of species (Shapiro et al., 2004; Atkinson et al., 2008; Campos et al., 2010; Lorenzen et al., 2011). The reconstruction of historical trend of population size has improved our understanding of the factors that have affected past ecosystems, either climatic or anthropogenic, recent or ancient. Indeed, trends in effective population size (Ne) over time can be associated with climatic or geological phenomena that have impacted the size and structure of populations (Kingman, 1982; Avise et al., 1987, 1998; Edwards and Beerli, 2000; Wang, 2005). Demographic inference methods based on the coalescent usually assumes panmixia, i.e. the absence of population structure, although this is not a realistic assumption in many biological situations. A number of recent studies have investigated the effect of violating the panmixia assumption for inferring population size changes (Städler et al., 2009; Chikhi et al., 2010; Peter et al., 2010). These studies suggest that population structure can lead to erroneous conclusions about demographic changes in a population that in fact has remained stationary through time (Heller et al., 2013).

Furthermore, the use of a single locus, for example only mitochondrial markers, for demographic analysis could generate unreliable results. The non-recombinant nature of this genome and maternal inheritance means that it tracks only one genealogical realization within a population's history (Ho and Shapiro, 2011). Furthermore, the use of a single marker, if there has been a selection on that particular gene region, risks not being representative of the demographic history of the population (Brito and Edwards, 2009; Ruane et al., 2015). Since the last decade, it has been demonstrated how the use of multiple unrelated markers increases the reliability of the demographic reconstruction (Galtier et al., 2000; Heled and Drummond, 2008).

In this chapter we built on results of the previous chapter on the phylogeographic structure of intraspecific lineages of the studied species to reconstruct the historical demography in order to investigate the effect of past climate changes on the populations of the target species. To avoid problems due to population structure, demographic reconstructions were inferred separately for each lineage identified in the previous chapter. Finally, the results of the analyses will be seen in a comparative perspective between species and across environments.

Materials and Methods

Sample collection and morphological identification

Specimens analysed in this chapter were collected and identified on a morphological basis as explained in Chapter 1 and 2.

DNA extraction

DNA of specimens of target species was extracted using non-invasive DNA extraction methods described in Chapter 2. Six different gene fragments, both mitochondrial and nuclear, were amplified through polymerase chain reaction (PCR): two *cox1* fragments (sequences analysed in chapter 7); and four protein-coding nuclear genes were amplified for a subsample of individuals, Carbamoylphosphate synthase (*CAD*), Crossveinless 2 (*Cv2*), Rad 50 protein (*Rad50*) and Wingless (*Wg*). The first three nuclear markers were amplified using primers and amplification conditions described in Gikonyo (2021). The *Wg* was amplified using primers Wg550F and WgAbRZ-R (Wild and Maddison, 2008) and condition described in Riccieri et al. (2020). Successful amplification was determined by gel electrophoresis and PCR products were purified and sequenced by an external service (Genewitz, UK). The obtained chromatograms of each sequence were manually edited and assembled into a consensus sequence using Geneious Prime 2021 (Biomatters Ltd., Auckland, New Zealand). Sequences of each gene were aligned separately using MAFFT v7.450 with the G-INS-I progressive method algorithm (Katoh et al., 2002) and then used for the following analysis.

The gametic phase of nuclear sequences was determined using DnaSP v5 (Librado and Rozas, 2009). using the PHASE algorithm (Stephens et al., 2001; Stephens and Donnelly, 2003), with the initial 1000 iterations discarded as burn-in, 1 as thinning interval and 1000 post-burn-in iterations.

To investigate the demographic variations through time it was performed an Extended Bayesian Skyline Plot analysis (EBSP) implemented in BEAST 1.8.2 (Drummond and Rambaut, 2007). BEAUTI 1.8.2 (in the BEAST package) was used to create the XML format input files for BEAST. The EBSP model is based on the genetic information from multiple loci and it incorporates stochastic differences between gene genealogies in the estimation of the population parameter. As a further advantage over previous skyline methods, the EBPS approach is not constrained *a priori* to a specific model of historical population size change (Heled and Drummond, 2008; Ho and Shapiro, 2011; Drummond et al., 2012). This was ensured using a Poisson prior with mean $\lambda = 0.693$ for the number of population changes (ψ), which put an equal probability (50% prior weight) to a constant and to a nonconstant population size (Drummond et al., 2012). To allow for continuous population size dynamics, a piecewise linear model was used. Analyses were based on mtDNA and nuDNA and, separately for each intraspecific phylogenetic group identified in chapter 7, in order to avoid the potential biases generated by population structure (Heller et al., 2013). Tree models of the mtDNA fragments were linked, whereas substitution and clock models were unlinked across all genes. The HKY + I + G model of nucleotide substitution was selected. The EBSP analysis was run using strict clock models and the substitution rate was set on the rate estimated for the 3’ fragment of the *cox1* gene by Papadopoulou et al. (2010) and used as a reference to guide rate estimates for the other gene fragments. The calibration was implemented with a lognormal prior (mean in real space = 0.0168, stdev = 0.06). Markov chain length were set from 3×10^8 to 5×10^8 generations depending on the runs convergence assessed examining the effective sample size of parameters (ESSs) and a visual inspection of the likelihood with TRACER (Rambaut et al., 2018). Trees and parameters were sampled every 30,000 to 50,000 generations depending on the total number of generations. The EBS plots were generated using the Python scripts provided in Heled (2010). To assess the time frame at which the EBSP reconstructions have high reliability, we generated a detailed demographic function description following Heled (2010) and we inspected the number of demographic functions which had an X point at each interval of time.

Results

Longitarsus laureolae

For the lineage of *L. laureolae* from the southern Italian Peninsula we obtained 43 sequences for both the *cox1* fragments and 178 phased nuclear sequences (52 sequences of *CAD*, 46 sequences of *Cv2*, 42 sequences of *Rad50* and 38 sequences of *Wg*); for the Sicilian lineage: 23 sequences for both the two fragment of *cox1* and for the four nuclear markers we obtained 96 phased sequences (24 sequences of *CAD*, 32 sequences of *Cv2*, 32 sequences of *Rad50* and 8 sequences of *Wg*). The EBSP showed historical demographic trends with a marked population expansion for both the Sicilian and the Peninsular lineages, starting right after the LGM (15-20 Kya). Inspection of the number of demographic functions for each time interval of the EBSPs indicates that the demographic inferences can be deemed as reliable backward in time until c. 100 Kya the peninsular lineage and c. 80 Kya for Sicilian (Fig. 26).

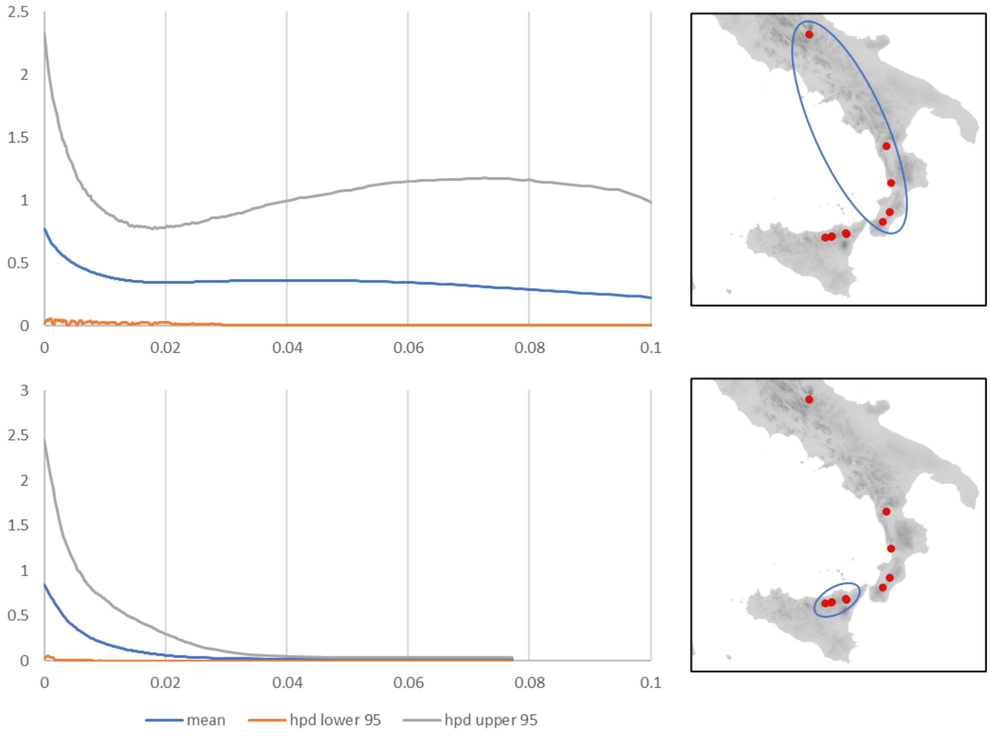


Fig. 26 Extended Bayesian skyline plots based on combined c*ox1*, *CAD*, *Cv2*, *Rad50* and *Wg* sequences for the main lineages identified within *L. laureolae*. Population assignment to lineages is based on results obtained in the previous chapter and is shown in the maps by the blue circles.

Longitarsus zangherii

For the lineage of *L. zangherii* we obtain for the lineages A 23 sequences for both the *cox1* fragments and 60 phased nuclear sequences (10 sequences of *CAD*, 28 sequences of *Cv2*, 6 sequences of *Rad50* and 16 sequences of *Wg*); (ii) for lineage B 58 sequences for both *cox1* fragment and 198 phased nuclear sequences (64 sequences of *CAD*, 60 sequences of *Cv2*, 22 sequences of *Rad50* and 52 sequences of *Wg*). The EBSP showed historical demographic trends with a constant population size for both the Northern and Central-South lineages. Inspection of the number of demographic functions for each time interval of the EBSPs indicates that the demographic inferences can be deemed as reliable backward in time until c. 60 Kya for both lineages (Fig. 27).

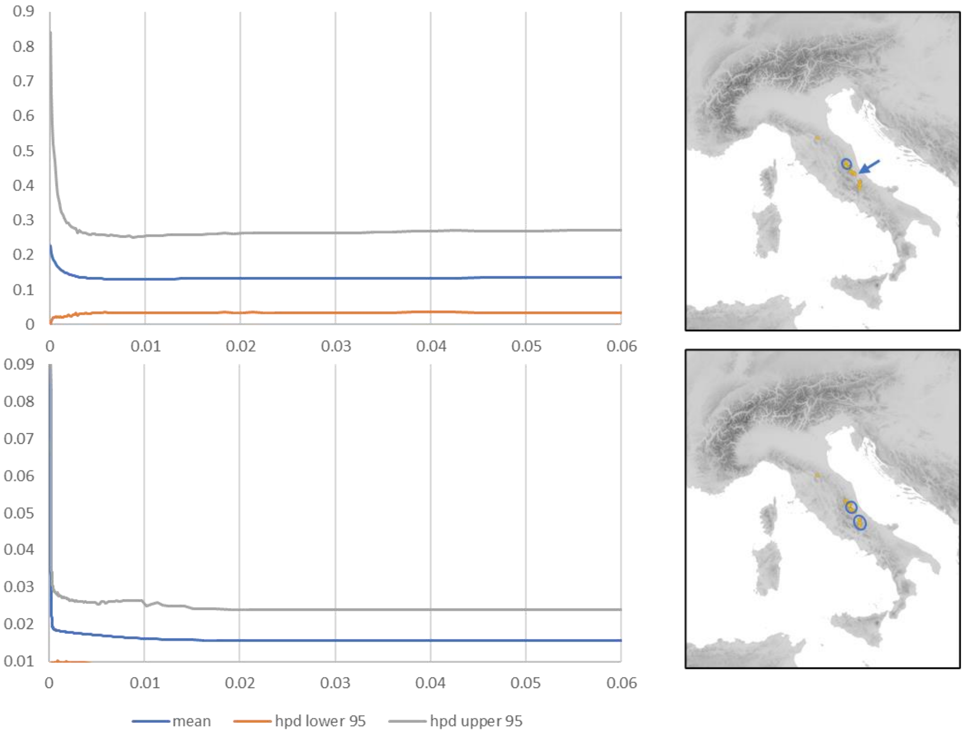


Fig.27 Extended Bayesian skyline plots based on combined *cox1*, *CAD*, *Cv2*, *Rad50* and *Wg* sequences for the main lineages identified within *L. zangherii*. Population assignment to lineages is based on results obtained in the previous chapter and is shown in the maps by the blue circles.

Psylliodes ruffoi

For the Calabrian lineage of *P. ruffoi* we obtained 40 sequences for both *cox1* fragment and 176 phased nuclear sequences (52 sequences of *CAD*, 48 sequences of *Cv2*, 32 sequences of *Rad50* and 44 sequences of *Wg*); for the Sicilian lineage, 17 sequences for both *cox1* fragment and 86 phased nuclear sequences (20 sequences of *CAD*, 26 sequences of *Cv2*, 12 sequences of *Rad50* and 28 sequences of *Wg*). The EBSP showed historical demographic trends with a weak population expansion for both the Sicilian and the Calabrian lineages, starting at the end of the last glacial period (45-20 Kya). Inspection of the number of demographic functions for each time interval of the EBSPs indicates that the demographic inferences can be deemed as reliable backward in time until c. 100 Kya for both lineages (Fig. 28).

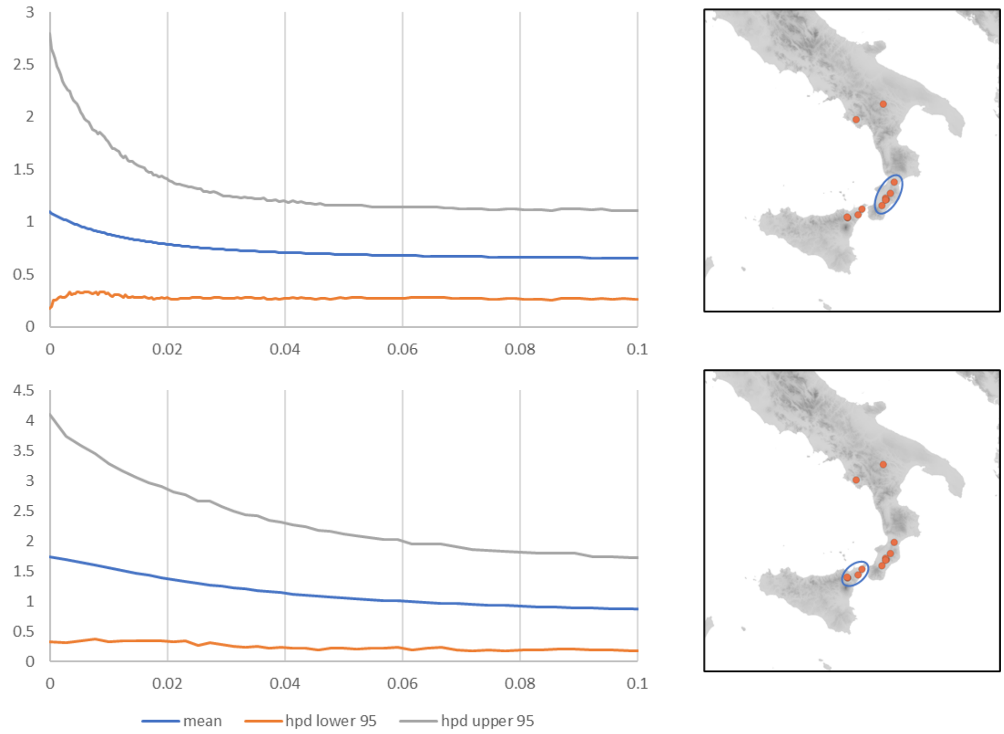


Fig.28 Extended Bayesian skyline plots based on combined *cox1*, *CAD*, *Cv2*, *Rad50* and *Wg* sequences for the main lineages identified within *P. ruffoi*. Population assignment to lineages is based on results obtained in the previous chapter and is shown in the maps by the blue circles.

Longitarsus springeri

For the lineage of *L. springeri* we obtain for the whole species 73 sequences for both *cox1* fragment and 282 phased nuclear sequences (70 sequences of *CAD*, 94 sequences of *Cv2*, 22 sequences of *Rad50* and 96 sequences of *Wg*); for Abruzzo lineages 49 sequences for both *cox1* fragment and 184 phased nuclear sequences (46 sequences of *CAD*, 56 sequences of *Cv2*, 14 sequences of *Rad50* and 68 sequences of *Wg*). The EBSPs showed relatively stationary historical demographic trends, with an imperceptible demographic decrease both for the whole species and for the Abruzzo subclade, starting from c. 50 and 25 Kya respectively. Inspection of the number of demographic functions for each time interval of the EBSPs indicates that the demographic inferences can be deemed as reliable backward in time until c. 400 Kya for both lineages (Fig. 29).

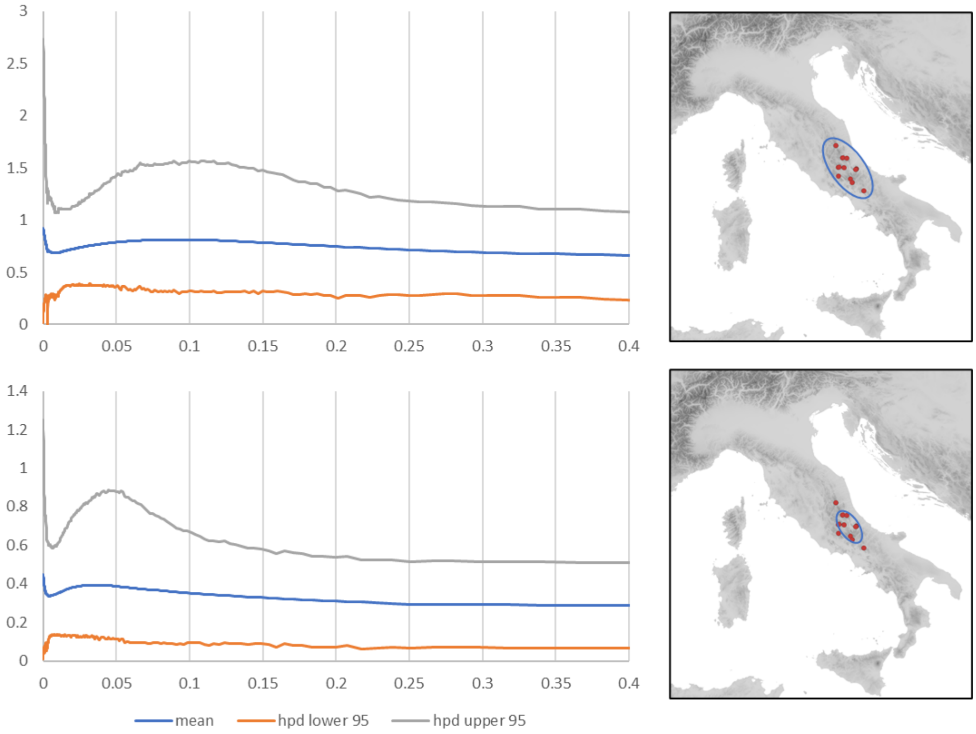


Fig.29 Extended Bayesian skyline plots based on combined c*ox1*, *CAD*, *Cv2*, *Rad50* and *Wg* sequences for the main lineages identified within *L. springeri*. Population assignment to lineages is based on results obtained in the previous chapter and is shown in the maps by the blue circles.

Psylliodes biondii

For the lineage of *P. biondii* we obtain 39 sequences for both *cox1* fragment and 154 phased nuclear sequences (44 sequences of *CAD*, 46 sequences of *Cv2*, 32 sequences of *Rad50* and 32 sequences of *Wg*). The EBSP showed historical demographic trends with a marked population expansion starting at c. 300 Kya. After this phase of expansion, the demographic trend changes, with a population decrease starting at c. 75 Kya. Inspection of the number of demographic functions for each time interval of the EBSPs indicates that the demographic inferences can be deemed as reliable backward in time until c. 500 Kya (Fig. 30).

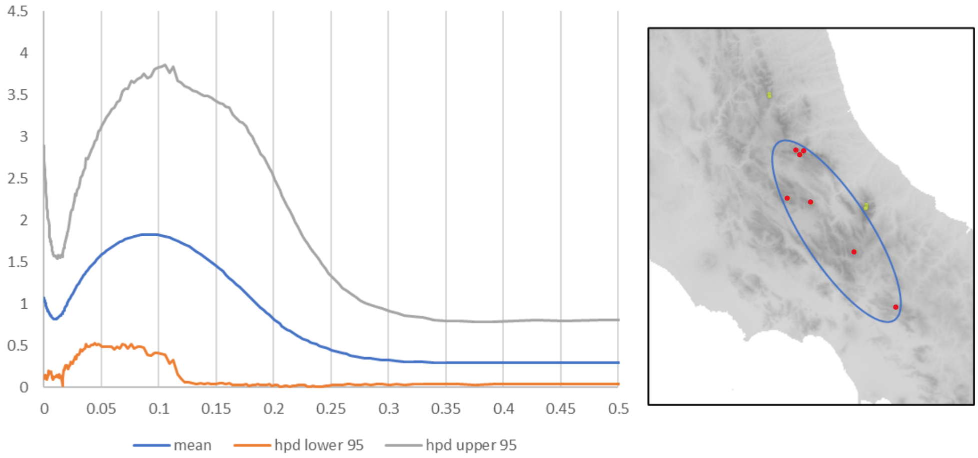


Fig.30 Extended Bayesian skyline plots based on combined *cox1*, *CAD*, *Cv2*, *Rad50* and *Wg* sequences for the main lineages identified within *P. biondii*. Population assignment to lineages is based on results obtained in the previous chapter and is shown in the maps by the blue circles.

Discussion

The climatic oscillations of the Pleistocene have strongly affected the range dynamics of temperate species, both by shifting their optimal bioclimatic conditions, and by influencing the geography of the physical availability of the habitat (Hewitt, 1996; Taberlet et al., 1998; Hewitt, 2000, 2004a; Schmitt, 2007). These processes had a strong impact on the genetic variation of the target species analyzed, as already seen above. In this study, we showed how these processes have also influenced the population demographic patterns. Results corroborate that also historical demographic trends indicate an individualistic response of species to past environmental changes.

As for the two temperate species, *L. laureolae* and *L. zangherii*, historical demographic reconstructions showed that changes in population size followed distinctive trends in the two species. Both the Sicilian and the Peninsular lineages of *L. laureolae* underwent a sudden postglacial expansion in population size, likely due to the post-glacial increase of habitat availability. In this period, the climatic amelioration led to an expansion of the temperate forest belts, which were fragmented into small isolated populations during the last glaciation (Blondel et al., 2010; Thompson, 2020). This increase in environmental availability is probably the reason for the demographic increase of this species, as already observed in other Mediterranean endemic temperate species (Salvi et al., 2016). In contrast, the two lineages of *L. zangherii* showed an historical demographic trend with a relatively stable population size. This species lives in the same environments as *L. laureolae* but has a more northern distribution. The observed historical demographic stability suggests that late Quaternary climatic oscillation had no discernible effect on population size in these lineages. Or, perhaps, the structure of the mountain system of the central part of the Apennines is such that it has allowed a large space of suitable environments even during the last glacial period. These hypotheses will deserve future studies at a smaller geographical scale.

Also for the mesophilic species *P. ruffoi*, the historical demographic reconstruction shows a relatively stable trend trough time, with a weak population expansion for both the Sicilian and the Calabrian lineages prior to the last interglacial. This is a pattern already seen in other Mediterranean species although it is usually observed in coastal or island species (Salvi et al., 2014; Senczuk et al., 2019). Another possibility is that this effect is given by the presence of a population structure internal to the two lineages since, as seen in the previous chapter, the Sicilian and Calabrian haplogroups of this species present very deep divergences. However, the analysis of individual haplogroups faces the limitation of small sample size and are thus not suitable to disentangling the effect of population structure on demographic reconstructions.

Finally, the two Alpine species, showed diverging historical demographic trends. For *L. springeri* the analyses considering the whole species, or the Abruzzo lineages gave similar results, with a not significant demographic decrease in both lineages since the beginning of the last interglacial. These results suggest that the temperature rise due to the interglacial had the effect of decreasing the population size of these lineages due to the reduction of suitable environment availability. This is in line with the sky island isolation pattern observed for this species in the previous chapter. A different case is that of *P. biondii* endemic to the central Apennine, which presents a marked population expansion starting at c. 300 ka followed by a population decrease starting at c. 75 ka. This pattern may indicate that the central Apennine region acted as a refugium for this species (see also Loughlin et al., 2008).

Historical demographic reconstructions fully corroborate results of the previous chapter and underline how the studied species had an individualistic response to environmental changes associated with Pleistocene climatic cycles. Some species, such as *L. laureolae* and *L. springeri*, seem to show demographic histories that can be explained by the habitat expansion and contraction driven by the climatic oscillations of the Pleistocene (Hewitt, 1996; Taberlet et al., 1998; Hewitt, 2000, 2004a; Provan and Bennett, 2008). On the other hand, the other species analysed show patterns that cannot be explained by the classic processes of expansion and contraction, and which probably derive from species-specific response that took place at a regional scale. Further studies will certainly be necessary to investigate these patterns. Our results allowed understanding how the Apennines have been the scene of multiple processes that have led species to follow distinctive evolutionary, demographic and biogeographic trajectories that gave rise to the high diversity of this biota.

References

Atkinson, Q.D., Gray, R.D., Drummond, A.J., 2008. mtDNA Variation Predicts Population Size in Humans and Reveals a Major Southern Asian Chapter in Human Prehistory. Mol. Biol. Evol. 25, 468–474. https://doi.org/10.1093/molbev/msm277

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu. Rev. Ecol. Syst. 18, 489–522.

Avise, J.C., Walker, D., Johns, G.C., 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proc. Biol. Sci. 265, 1707–1712. https://doi.org/10.1098/rspb.1998.0492

Blondel, J., Aronson, J., Bodiou, J.-Y., Boeuf, G., 2010. The Mediterranean Region: Biological Diversity in Space and Time. OUP Oxford.

Brito, P.H., Edwards, S.V., 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135, 439–455. https://doi.org/10.1007/s10709-008-9293-3

Campos, P.F., Willerslev, E., Sher, A., Orlando, L., Axelsson, E., Tikhonov, A., Aaris-Sørensen, K., Greenwood, A.D., Kahlke, R.-D., Kosintsev, P., Krakhmalnaya, T., Kuznetsova, T., Lemey, P., MacPhee, R., Norris, C.A., Shepherd, K., Suchard, M.A., Zazula, G.D., Shapiro, B., Gilbert, M.T.P., 2010. Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (Ovibos moschatus) population dynamics. Proc. Natl. Acad. Sci. 107, 5675–5680. https://doi.org/10.1073/pnas.0907189107

Chikhi, L., Sousa, V.C., Luisi, P., Goossens, B., Beaumont, M.A., 2010. The Confounding Effects of Population Structure, Genetic Diversity and the Sampling Scheme on the Detection and Quantification of Population Size Changes. Genetics 186, 983–995. https://doi.org/10.1534/genetics.110.118661

Davis, M.B., Shaw, R.G., 2001. Range shifts and adaptive responses to Quaternary climate change. Science 292, 673–679.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Edwards, ScottV., Beerli, P., 2000. Perspective: Gene Divergence, Population Divergence, and the Variance in Coalescence Time in Phylogeographic Studies. Evolution 54, 1839–1854. https://doi.org/10.1111/j.0014-3820.2000.tb01231.x

Galtier, N., Depaulis, F., Barton, N.H., 2000. Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. Genetics 155, 981–987.

Gikonyo, M., 2021. The evolution of host plant use and adaptation to Brassicaceae in the leaf beetle genus Psylliodes.

Heled, J., 2010. Extended Bayesian skyline plot tutorial. Available Tutor. East Bio Ed Ac UkTutorials.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 1–15.

Heller, R., Chikhi, L., Siegismund, H.R., 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. PloS One 8, e62992.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.

Hewitt, G.M., 2004a. Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 359, 183–195.

Hewitt, G.M., 2004b. The structure of biodiversity–insights from molecular phylogeography. Front. Zool. 1, 1–16.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247–276.

Ho, S.Y., Shapiro, B., 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Mol. Ecol. Resour. 11, 423–434.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kingman, J.F.C., 1982. The coalescent. Stoch. Process. Their Appl. 13, 235–248. https://doi.org/10.1016/0304-4149(82)90011-4

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Lorenzen, E.D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K.A., Ugan, A., Borregaard, M.K., Gilbert, M.T.P., Nielsen, R., Ho, S.Y.W., Goebel, T., Graf, K.E., Byers, D., Stenderup, J.T., Rasmussen, M., Campos, P.F., Leonard, J.A., Koepfli, K.-P., Froese, D., Zazula, G., Stafford, T.W., Aaris-Sørensen, K., Batra, P., Haywood, A.M., Singarayer, J.S., Valdes, P.J., Boeskorov, G., Burns, J.A., Davydov, S.P., Haile, J., Jenkins, D.L., Kosintsev, P., Kuznetsova, T., Lai, X., Martin, L.D., McDonald, H.G., Mol, D., Meldgaard, M., Munch, K., Stephan, E., Sablin, M., Sommer, R.S., Sipko, T., Scott, E., Suchard, M.A., Tikhonov, A., Willerslev, R., Wayne, R.K., Cooper, A., Hofreiter, M., Sher, A., Shapiro, B., Rahbek, C., Willerslev, E., 2011. Species-specific responses of Late Quaternary megafauna to climate and humans. Nature 479, 359–364. https://doi.org/10.1038/nature10574

Loughlin, S., Okabayashi, T., Honda, M., Kitazoe, Y., Kishino, H., Somboon, P., Sochantha, T., Nambanya, S., Saikia, P., Dev, V., 2008. Complex population history of two *Anopheles dirus* mosquito species in Southeast Asia suggest the influence of Pleistocene climate change rather that humanmediated effects. J Evol Biol 21, 1555–1569.

Mapelli, F., Mora, M., Mirol, P., Kittlein, M., 2012. Effects of Quaternary climatic changes on the phylogeography and historical demography of the subterranean rodent Ctenomys porteousi. J. Zool. 48–57. https://doi.org/10.1111/j.1469-7998.2011.00849.x

Nogués-Bravo, D., Ohlemüller, R., Batra, P., Araújo, M.B., 2010. Climate predictors of late quaternary extinctions. Evol. Int. J. Org. Evol. 64, 2442–2449. https://doi.org/10.1111/j.1558-5646.2010.01009.x

Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol. Biol. Evol. 27, 1659–1672.

Peter, B.M., Wegmann, D., Excoffier, L., 2010. Distinguishing between population bottleneck and population subdivision by a Bayesian model choice procedure. Mol. Ecol. 19, 4648–4660. https://doi.org/10.1111/j.1365-294X.2010.04783.x

Provan, J., Bennett, K.D., 2008. Phylogeographic insights into cryptic glacial refugia. Trends Ecol. Evol. 23, 564–571.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901.

Riccieri, A., Mancini, E., Salvi, D., Bologna, M.A., 2020. Phylogeny, biogeography and systematics of the hyper-diverse blister beetle genus *Hycleus* (Coleoptera: Meloidae). Mol. Phylogenet. Evol. 144, 106706.

Ruane, S., Torres-Carvajal, O., Burbrink, F.T., 2015. Independent Demographic Responses to Climate Change among Temperate and Tropical Milksnakes (Colubridae: Genus *Lampropeltis*). PLOS ONE 10, e0128543. https://doi.org/10.1371/journal.pone.0128543

Salvi, D., Bisconti, R., Canestrelli, D., 2016. High phylogeographical complexity within Mediterranean islands: insights from the Corsican fire salamander. J. Biogeogr. 43, 192–203. https://doi.org/10.1111/jbi.12624

Salvi, D., Schembri, P.J., Sciberras, A., Harris, D.J., 2014. Evolutionary history of the Maltese wall lizard *Podarcis filfolensis*: insights on the ‘Expansion–Contraction’model of Pleistocene biogeography. Mol. Ecol. 23, 1167–1187.

Schmitt, T., 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. Front. Zool. 4, 1–13.

Senczuk, G., Harris, D.J., Castiglia, R., Litsi Mizan, V., Colangelo, P., Canestrelli, D., Salvi, D., 2019. Evolutionary and demographic correlates of Pleistocene coastline changes in the Sicilian wall lizard *Podarcis wagleriana*. J. Biogeogr. 46, 224–237.

Shapiro, B., Drummond, A., Rambaut, A., Wilson, M., Matheus, P., Sher, A., Pybus, O., Gilbert, M., Barnes, I., Binladen, J., Willerslev, E., Hansen, A., Baryshnikov, G., Burns, J., Davydov, S., Driver, J., Froese, D., Harington, C., Keddie, G., Cooper, A., 2004. Rise and Fall of the Beringian Steppe Bison. Science 306, 1561–5. https://doi.org/10.1126/science.1101074

Städler, T., Haubold, B., Merino, C., Stephan, W., Pfaffelhuber, P., 2009. The Impact of Sampling Schemes on the Site Frequency Spectrum in Nonequilibrium Subdivided Populations. Genetics 182, 205–216. https://doi.org/10.1534/genetics.108.094904

Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162–1169.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989.

Taberlet, P., Fumagalli, L., Wust-Saucy, A., Cosson, J., 1998. Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. 7, 453–464.

Thompson, J.D., 2020. Plant Evolution in the Mediterranean: Insights for Conservation. Oxford University Press.

Wang, J., 2005. Estimation of effective population sizes from data on genetic markers. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 1395–1409. https://doi.org/10.1098/rstb.2005.1682

Wild, A.L., Maddison, D.R., 2008. Evaluating nuclear protein-coding genes for phylogenetic utility in beetles. Mol. Phylogenet. Evol. 48, 877–891.

CONCLUSION

Results of this thesis allowed documenting phylogeographic patterns and underlying evolutionary processes that have determined the high genetic structure and diversity of Apennine fauna. This mountain system that crosses from north to south the Italian peninsula for over a thousand kilometres is a well-known diversity hotspot on a par with more extensive mountain ranges such as the Alps and the Pyrenees (Canestrelli et al., 2006, 2008; Bogdanowicz et al., 2015). In particular, the central sector of the Apennines, characterized by a geological and environmental history shaped by the Quaternary glaciations (Giraudi and Frezzotti, 1997; Giraudi et al., 2004) has been a refuge for various cold-adapted species. As a consequence this region is among the mountain areas that host the largest number of species in the entire area of the Italian peninsula (Maiorano et al., 2006; Biondi et al., 2013; Urbani et al., 2015; Menchetti et al., 2021). However, the phylogeographic studies focused on the evolutionary processes that took place on these mountains have been mainly focused on vertebrate taxa. In this thesis, a fine scale phylogeographic study of the Apennine mountain systems was carried out for the first time using a comparative multi-taxa approach on five endemic species of Alticini beetles as a reference system (Coleoptera, Chrysomelidae, Galerucinae). This tribe, in terms of number of species and ecological diversity, proved to be an excellent study model to unveil the complexity of species response to past climatic and environmental changes, which was the main aim of this thesis.

Thanks to a carefully thought sampling design, it was possible, over the three years of the PhD, to cover the whole range of selected target species by sampling a large number of populations and, in some cases, also expanding the known range of species. Thanks to the joint study of molecular and taxonomic analyses, it was possible to solve some methodological and taxonomic problems related to the systematic complexity of the examined group. In particular, for the genus *Longitarsus*, the revision of the sequences from the public database of GenBank and BOLD, through *a posteriori* taxonomic revision of those erroneous or ambiguous sequences, it was possible to improve the effectiveness of the DNA barcoding tool and reach a reliable species identification based on molecular data. The use of an integrative taxonomy approach within a phylogeographic framework allowed to delimit the two species *P. ruffoi* and *P. kiesenwetteri* from a molecular and morphological point of view, clarifying their phylogenetic and taxonomic relationship. Finally, molecular analyses also allowed to identify a new lineage (possibly a new species) endemic to the Maiella mountains within the high-altitude species complex of *Psylliodes springeri*. These results not only demonstrate the importance of an integrative approach when addressing the study of complex groups such as that of the Alticini, but also allowed us to identify the evolutionary units on which to apply phylogeographic and demographic analyses to answer the main questions of this thesis project.

Phylogeographic analysis have shown that species have an individualistic response to environmental changes, showing different evolutionary histories and individualistic adaptations to local environmental factors (Hornsby and Matocq, 2012; Stewart et al., 2010). The geographic distribution and time of origin of mitochondrial lineages indicate a greater effect of the last glacial period on temperate species compared to species linked to mesophilic environments. In contrast, cryophilic species have undergone a process of isolation in suitable mountain areas that closely match the so-called Sky islands model. For each of these species, the reconstruction of the historical demography through the EBSP corroborates how these species had an individualistic response to the past glaciations. Some species, such as *L. laureolae* and *L. springeri*, appear to exhibit demographic histories that can be explained by the glacial contraction and postglacial expansion typical of temperate species (Hewitt, 1996; Taberlet et al., 1998; Hewitt, 2000, 2004; Provan and Bennett, 2008). The other species, on the other hand, showed different trends, which cannot be explained by the classic processes of expansion and contraction, and which have probably derived from species-specific processes occurring on a regional scale.

These results confirmed that the Apennines are an important reservoir of genetic and endemic diversity. It has been shown how the geographical and ecological heterogeneity of this region, in synergy with the climatic changes of the past, has led ecological similar species, currently living in close sintopy, to undergo distinct evolutionary and demographic processes under the same past environmental changes. This allowed us to add one more piece of knowledge on the evolutionary and demographic dynamics of understudied fauna of this region. This also allowed us to improve our understanding of the processes that generated a biodiversity hotspot within this region and that determined the assembly of its biota. These results are crucial for planning informed strategies for the long-term conservation and management of the Apennine biodiversity.

References

Biondi, M., Urbani, F., D’Alessandro, P., 2013. Endemism patterns in the Italian leaf beetle fauna (Coleoptera, Chrysomelidae). ZooKeys 177.

Bogdanowicz, W., Hulva, P., Černá Bolfíková, B., Buś, M.M., Rychlicka, E., Sztencel-Jabłonka, A., Cistrone, L., Russo, D., 2015. Cryptic diversity of Italian bats and the role of the Apennine refugium in the phylogeography of the western Palaearctic. Zoological Journal of the Linnean Society 174, 635–648.

Canestrelli, D., Cimmaruta, R., Costantini, V., Nascetti, G., 2006. Genetic diversity and phylogeography of the Apennine yellow-bellied toad *Bombina pachypus,* with implications for conservation. Molecular Ecology 15, 3741–3754.

Canestrelli, D., Cimmaruta, R., Nascetti, G., 2008. Population genetic structure and diversity of the Apennine endemic stream frog, *Rana italica*–insights on the Pleistocene evolutionary history of the Italian peninsular biota. Molecular Ecology 17, 3856–3872.

Giraudi, C., Ehlers, J., Gibbard, P., 2004. The Apennine glaciations in Italy. Quarternary glaciations–extent and chronology, part I: Europe. Developments in quaternary science 2, 215–224.

Giraudi, C., Frezzotti, M., 1997. Late Pleistocene glacial events in the central Apennines, Italy. Quaternary research 48, 280–290.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.

Hewitt, G.M., 2004. The structure of biodiversity–insights from molecular phylogeography. Frontiers in zoology 1, 1–16.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological journal of the Linnean Society 58, 247–276.

Hornsby, A.D., Matocq, M.D., 2012. Differential regional response of the bushy-tailed woodrat (*Neotoma cinerea*) to late Quaternary climate change. Journal of Biogeography 39, 289–305.

Maiorano, L., Falcucci, A., Boitani, L., 2006. Gap analysis of terrestrial vertebrates in Italy: priorities for conservation planning in a human dominated landscape. Biological conservation 133, 455–473.

Menchetti, M., Talavera, G., Cini, A., Salvati, V., Dincă, V., Platania, L., Bonelli, S., Balletto, E., Vila, R., Dapporto, L., 2021. Two ways to be endemic. Alps and Apennines are different functional refugia during climatic cycles. Molecular Ecology 30, 1297–1310.

Provan, J., Bennett, K.D., 2008. Phylogeographic insights into cryptic glacial refugia. Trends in ecology & evolution 23, 564–571.

Stewart, J.R., Lister, A.M., Barnes, I., Dalén, 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, A., Cosson, J., 1998. Comparative phylogeography and postglacial colonization routes in Europe. Molecular ecology 7, 453–464.

Urbani, F., D’Alessandro, P., Frasca, R., Biondi, M., 2015. Maximum entropy modeling of geographic distributions of the flea beetle species endemic in Italy (Coleoptera: Chrysomelidae: Galerucinae: Alticini). Zoologischer Anzeiger - A Journal of Comparative Zoology 258, 99–109. https://doi.org/10.1016/j.jcz.2015.08.002

ACKNOWLEDGEMENTS

Sono molte le persone che devo ringraziare per essermi state vicino e ad avermi supportato durante questa tesi e, senza le quali, non avrei potuto concludere questo percorso. Vorrei quindi ringraziare per primi i miei genitori, senza i quali non sarei qui, per ovvi motivi, ma anche perché mi hanno permesso di intraprendere il percorso universitario supportandomi in ogni momento, aiutandomi nelle scelte e lasciandomi sempre la tranquillità di poterle fare. Voglio poi ringraziare Giulia, che mi è stata affianco in ogni momento di questi anni, sia in quelli facili che in quelli difficili. Oltre al fatto che è stata di enorme supporto per i campionamenti fatti sulle alpi, dove tra l'altro ha raccolto il Sacro Alticino, e quindi la ringrazio per l'aiuto e l'enorme pazienza che ha avuto. Ringrazio il mio tutor Daniele Salvi, che mi propose all'inizio questo fantastico progetto, che mi ha aiutato a crescere professionalmente, in questi anni, come ricercatore e scienziato, e che mi ha supportato e guidato in tutte le fasi del progetto. Inoltre lo ringrazio perché, al di fuori dell'università, si è rivelato essere un amico oltre che un docente. Quindi grazie anche per le risate, le giornate di campo e i racconti davanti un bel bicchiere di vino. Un ringraziamento enorme va ovviamente al professor Maurizio Biondi, senza il quale questa tesi non sarebbe stata possibile. Lo ringrazio per il suo supporto, scientifico ed umano datomi in questi anni, che non è mai mancato anche nei momenti di difficoltà che ci sono stati, e lo ringrazio per avermi permesso di scoprire e conoscere la meravigliosa ed affascinante tribù degli Alticini, che indubbiamente porterò nel cuore per tutta la vita, a prescindere da che strada prenderò poi nel futuro. Ringrazio anche la Professoressa Paola D'alessandro, per il suo supporto sul quale ho potuto contare in ogni momento. Un grazie anche a Marco, alla sua infinita pazienza ed alla sua onnipresenza per risolvere qualsiasi difficoltà ed un grazie anche a Mirco per aver condiviso tanti momenti di lab e tante cene. Un grazie unico va poi a Lacemod. Arrivare in un nuovo posto di lavoro, una nuova città è sempre difficile, ma grazie a voi mi sono sentito a casa da subito. Ringrazio quindi tutti voi ragazzi e ragazze per questi magnifici anni passati assieme, tra gioie, risate, incazzature e difficoltà! Un grazie quindi a Walter, con il quale ho condiviso oltre che l'esame d'ammissione anche i tre anni di dottorato, un grazie a Cristina, la conterranea che non mi ha fatto sentire lo stacco con Roma Tre, un grazie a Mattia, fedele vicino di banco per tre lunghi anni oltre che ogni presente aiuto per ogni difficoltà, un grazie a Carmelo con il quale ho condiviso laboratorio e case, un grazie a Francesco, alla sua imprevedibilità e al suo impegno politico e sociale, un grazie a Giorgia, che nonostante non abbia seguito il nostro percorso è sempre stata presente, un grazie a Giulia per la sua simpatia e gentilezza, un grazie a Valterone, per le risate e le splendide giornate di montagna passate assieme sulla sua Maiella e, infine, ringrazio Michele, per l'aiuto datomi sul campo e al computer, per le mille avventure in montagna e tra i peggiori bar del teramano, per gli azzardi a Zervò e per avermi avvicinato alla meravigliosa arte dell'arrampicata, anche se io continuo a fuggirle. Un grazie va anche a Rossella, per le volte che ci ha accompagnato sul campo e per il buon cibo. Ringrazio anche Lorenzo, Davide ed Alessia con i quali, nonostante abbia condiviso pochi mesi di dottorato, hanno comunque fatto parte di questo mio percorso. Per chiudere con i ragazzi del Lacemod mi sono tenuto i ringraziamenti per Matteo, ultimo ma non meno importante! Un grazie particolare va a te, per l'aiuto sul campo, le notti passate in laboratorio e i ritorni a casa sotto coprifuoco, le pizze mangiate in ufficio e le

PCR lasciate overnight, radio due e i momenti di follia. Grazie di tutto questo e del bel gruppo di ricerca che abbiamo creato nei mesi che abbiamo lavorato assieme. Voglio anche ringraziare Lorenzo De Vitis, per il suo fondamentale supporto durante il mio primo anno di dottorato, per la profonda amicizia che si è creata tra un campionamento e l'altro, in particolare quella meravigliosa prima campagna svolta nel sud Italia. Per chiudere vorrei ringraziare il professor Marco Bologna, perché probabilmente senza le sue lezioni di zoologia non starei qui, vorrei ringraziare anche Sandro e Lucia per la loro vicinanza in questi anni ed infine vorrei ringraziare i miei famigliari, gli zii, i cugini ed i nonni, che mi sono sempre vicini e che hanno seguito questo mio percorso come gli altri della mia vita.