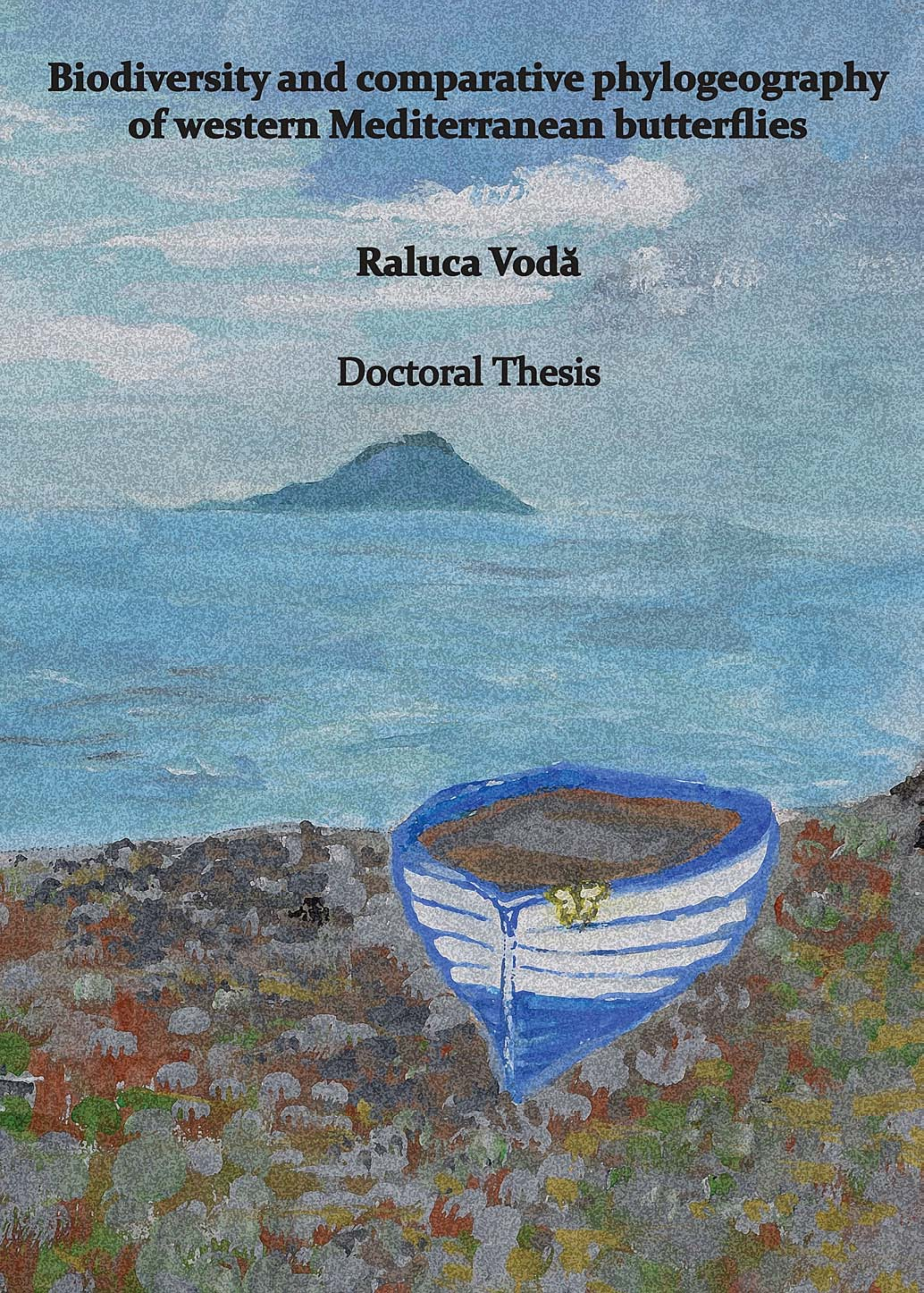
Biodiversity and comparative phylogeography of western Mediterranean butterflies

Raluca Vodă

Doctoral Thesis



Institut de Biologia Evolutiva
(CSIC-UPF)

Universitat Autònoma de
Barcelona

Biodiversity and comparative phylogeography of western Mediterranean butterflies

Biodiversitat i filogeografia comparada de les papallones de l'oest de la Mediterrània

Biodiversidad y filogeografía comparada de las mariposas del oeste del Mediterráneo

– DOCTORAL THESIS –
Raluca Vodă

Bellaterra, 2015
Facultat de Biociències - Departament de Genètica i Microbiologia

Memòria presentada per optar al grau de doctor per la Universitat Autònoma de Barcelona Programa de Doctorat en Genètica

Vist-i-plau dels directors/tutor de la tesi

Director
Dr. Roger Vila Ujaldón
Institut de Biologia Evolutiva
(CSIC-UPF)
Animal Biodiversity and Evolution
Program

Director
Dr. Leonardo Dapporto
Institut de Biologia Evolutiva
(CSIC-UPF)
Animal Biodiversity and Evolution
Program

Tutor
Prof. Alfredo Ruíz Panadero
Universitat Autònoma de Barcelona,
Departament de Genètica i
Microbiologia

Author
Raluca Vodă
Institut de Biologia Evolutiva
(CSIC-UPF)
Animal Biodiversity and Evolution
Program

BIODIVERSITY AND COMPARATIVE PHYLOGEOGRAPHY OF WESTERN MEDITERRANEAN BUTTERFLIES

~Doctoral Thesis~

Raluca Vodă

Table of contents

Acknowledgments

A PhD student needs a lot of help – or maybe that was just me. Anyway, there are so many people and so many favourable events that helped me have four amazing years that changed my life completely.

First of all I would like to thank my wonderful family: **my parents**, who made me discover and love nature during our many 'field trips' together and who taught me the true values in life: be honest, work hard and be kind to people, and my sister and best friend, **Ana**, who always stood by me when life was tough and encouraged me through it all. Without you I would not be where I am today. I love you all so much!

I also want to thank my extended family (you know who you are) for their unconditioned care and support.

This thesis would not have been possible without the help of three very knowledgeable, generous, and patient people who helped me in my 'quest'.
I would like to thank the two greatest supervisors I could have ever wanted:
Roger, you gave me freedom to work in my own pace and you trusted me that I could do anything I put my mind to. Thank you for being a friend and not a boss, and for teaching me so many things! I will never forget the amazing field trips together, some of them real tests of our physical limits.

Leonardo, thank you so much for your guidance, friendship and for gradually turning me into an 'islomaniac'. Our endless butterfly conversations were such an amazing source of knowledge and inspiration, even though we often got 'lost in translation'… but definitely not while singing Bohemian Rhapsody.

Vlad, there are so many things I want to thank you for … You always inspire me and not only you make me better but you make me want to be better. Your patience, kindness and love make everything that I do possible.

I would like to thank **Prof. László Rákosy** who is the reason I started to love the beautiful insects I study, who gave me my first butterfly net and who encouraged and helped me in difficult times.

I would also like to thank **Prof. Tim Shreeve** for giving me the opportunity to have a great experience and a rewarding stay at Oxford Brookes University.

I am grateful to **Prof. Paul Hebert** and to the **Biodiversity Institute of Ontario**, Canada for a very fruitful collaboration.

Sylvain, thank you for the field trips and the thousands kilometres we travelled to reach all the wonderful places you know or that we discovered together, for your friendship, for sharing your entomological experience but also a beer at the end of a collecting day.

I would also like to thank **Sandra**, **Sabina**, **Diana**, **Alícia**, **Santi**, **Carme** for welcoming me with open arms when I first moved here, and for always helping and caring for me.

A big thank you to the administration at IBE: **Anna**, **Rita**, **Blanca**, **Emiliano**, and **Vicente**, for their patience and for helping me not to get lost in the bureaucratic world.

A big thank you to all my friends, old and new ...

To my colleagues and friends in the Butterfly Diversity and Evolution Lab:
Gerard, **Marga**, **Juan**, **Sergio**, **Josep**, **Claudia**, **Martha**, **Lucas**, and the newest member, **Camille**. I want to thank you for all the help you gave me during this time, for your friendship, support and for all the good times we shared.
Thanks also to 'the other group' in the lab: **Lain**, **Lidia**, **Anna**, **Miriam**, for such a nice co-existence.

Life as a PhD student would have been so much boring without you: **Marina,** for being such a great friend, **Joan**, **Elena**, **DAVID**, **David**, **Marc**, **Cristina**, **Nuria**, **Claudia**, **Anna P**., **Ana R.**, **Javi**, **Santi**, **Carlos**, **Helena** - for all the home dinners, outings and for so many great times we had together.

To my dear friends, **Andrea** and **Clara** with whom I experienced 'life in Barcelona'. I care for you both so much!

To the Italian 'butterfly family': **Stefano, Luca**, **Ale**, **Mattia**, **Leonardo** and **Jacopo**. I cannot wait for the next butterfly week!

I also want to thank my home friends, and in particular **Andreea** (for endless memories), **Gabi**, **Silvia**, **Natalia** for their wonderful friendship and to my lovely weird cat **Ariel** whom I miss a lot ...

I would also like to acknowledge the financial support of a four-year predoctoral scholarship from the Spanish Ministerio de Economía y Competitividad (AP- 2010-5409).

Many thanks to all the colleagues and institutions that supported my research in numerous ways, including by providing precious samples.

And finally, I am grateful to the beautiful creatures I study for saving my life – and this is not an overstatement!

Abbreviations

bp – base pairs
COI - cytochrome c oxidase subunit I
mya – million years ago
ENM – Ecological Niche Modelling
ETIB – Equilibrium Theory of Island Biogeography
LGM – Last Glacial Maximum
mm – millimetres
mtDNA – mitochondrial DNA
MSC – Messinian Salinity Crisis
nuDNA – nuclear DNA

1. Introduction

1.1 The study system: butterflies of the western Mediterranean

The study region

“*What is the Mediterranean? One thousand things at a time. Not just one landscape but innumerable landscapes. Not just one sea, but a succession of seas. Not just one civilization but many civilizations packed on top of one another. The Mediterranean is a very old crossroads. Since millennia, everything converged on it*.” (Braudel 1985)

The Romans called the Mediterranean sea ‘mare nostrum’, meaning ‘our sea’, but also ‘mediterraneus’, Latin for ‘the sea in the middle of the Earth’, while at the other end of the sea the Turkish referred to it as ‘Akdeniz’, ‘the white sea’ (Öztürk 2011), as for them the colour white represents ‘the west’ (DeLong & Martinson 2012). More recently, Blondel et al. (2010) proposed that a suitable name would be the ‘sea-among-the-mountains’, because of the almost complete ring of mountains that encircles the Mediterranean basin.

But the Mediterranean means much more than a sea – it is a landscape, a climate, a biome, an ancient culture, a lifestyle (King et al. 1997); it is “*one of the richest and most complex regions on Earth*” (Blondel et al. 2010).

Geology

The formation of the Mediterranean basin was a long process that started around 250 mya during the Mesozoic era, and it is the result of a series of powerful and complex movements of tectonic plates (Rosenbaum et al. 2002a, Blondel et al. 2010). At the beginning of the Jurassic period the ancient supercontinent Pangaea started to split into two smaller continents, Gondwana to the south and Laurasia to the north. The two landmasses continued to fragment throughout the Mesozoic, giving rise to the continents and oceans we know today. North America and Europe started to move apart from Africa around 165 mya as a consequence of the seafloor spreading that created the Atlantic Ocean (Blondel 2010). The opening of the Atlantic and the northeastern movement of the Afro-Arabian plate towards a fixed Eurasia led to the formation of a large ocean basin in between the two plates, extending from the western Mediterranean to the eastern Himalayas (Coward & Dietrich 1989). This ancient ocean called Tethys was the predecessor of the present Mediterranean, Caspian, Aral and Black seas (Suess 1893, Faccenna et al. 2001, Rosenbaum et al. 2002a).

Following its northeastern movement, Africa first collided with Eurasia during the Oligocene, generating the uplifting of several mountain chains, including the Alps and the Pyrenees, as well as the consumption of the former Tethys Ocean (Hsü 1971, Coward & Dietrich 1989, Jolivet & Faccenna 2000, Faccenna et al. 2001, Blondel 2010). The Alpine orogenesis occurred in two distinct episodes, for the Eastern and Western Alps throughout the Cretaceous and for the Central Alps throughout the Tertiary (Pfiffner 1992). During the orogenesis of the Alps several other tectonic processes played a major role in the formation of the western Mediterranean basin. Large-scale horizontal extensions associated with subduction rollback (the backward migration of the subducting plate), back-arc extension (the extension of an arc crust as a result of the sinking of the subducting slab) and accretion events were the driving mechanisms in the tectonic evolution of this region (Rehault et al. 1985, Malinverno & Ryan 1986, Royden 1993, Lonergan & White 1997, Jolivet & Faccenna 2000, Rosenbaum et al. 2002a, Rosenbaum et al. 2002b, van Hinsbergen et al. 2014).

Although reconstructions of the western Mediterranean region are widely similar (van Hinsbergen et al. 2014), there are still debates regarding the initiation age of the subduction between Africa and Iberia (Faccenna et al. 2001, Rosenbaum et al. 2002a, Handy et al. 2010) and about the lateral extent of the western Mediterranean subduction zone (Lonergan and White 1997, Rosenbaum et al. 2002a, Rosenbaum & Lister 2004, Schettino & Turco, 2006).

The convergent movement of the African-Arabian plate towards Eurasia started to slow down around 30 mya (Jolivet & Faccenna 2000, McQuarrie & Hinsbergen 2013). In the absence of a sufficient convergence to support subduction rollback, the overriding plate underwent extension from west to east, leading to the opening of several basins in the Mediterranean (Rosenbaum et al. 2002b). Moreover, these large-scale extensions were responsible for the breakup, drift and rotation of different continental blocks that were formerly connected to mainland (Rosenbaum et al. 2002a). This is the case of Corsica, Sardinia, the Balearic Islands, the Kabylies, Calabria and the internal part of the Rif-Betic Cordillera. During the Oligocene these fragments were attached to Europe, between northern Spain and southern France, where today the Gulf of Lion is located (Ricou et al. 1986, Lonergan & White 1997, Rosenbaum et al. 2002a) (fig. 1). Corsica, Sardinia and Calabria started to rotate counterclockwise during the opening of the Ligurian and Tyrrhenian seas (fig. 1 b, c) and collided with the Apennines around 18 mya (fig. 1 d, e, f). The Apennines incorporated in the subduction area and stopped further eastward subduction rollback (Rosenbaum et al. 2002a). The Balearic Islands started to rotate clockwise when the Valencia Trough opened (Pares et al. 1992) (fig. 1 b). Around 21 mya the Kabylies separated from the Balearics and due to southward rollback of the subduction zone they started to move towards the south (fig. X 1 c, d), until they collided and accreted to the territory of today's Algeria (Tricart et al. 1994) (fig. 1 e, f, g). The reconstruction of the Rif-Betic position is still debated but according to the model proposed by Lonergan & White (1997) and Rosenbaum et al. (2002a), the origin of this block is the Gulf of Lion (fig. 1 a). After separation from mainland the Rif-Betic fragments started to drift southwestwards until they reached the passive margin of north Africa (fig. 1 b, c, d) and continued to move towards the Gibraltar strait (Rosenbaum et al. 2002a). Around 10 mya they accreted to Iberia and Africa, respectively, followed by the cessation of back-arc extension and rollback subduction in this area (Lonergan & White 1997, Rosenbaum et al. 2002a) (fig. 1 e, f, g, h).

All these tectonic processes and others, like earthquakes and volcanic activity, had an immense influence not only on the geology of the Mediterranean basin, but also on the distribution of animals and plants and in particular they formed the basis for the observed patterns of endemicity in this region (Blondel et al. 2010).

The Messinian Salinity Crisis – a trilogy of the Mediterranean

An event that is often advocated as one of the main drivers for the emergence of a distinctive fauna and flora in the Mediterranean is the Messinian Salinity Crisis (MSC). This short but extreme episode occurred in the late Miocene, following the progressive closure of the Gibraltar strait, the only marine gateway that connected the Mediterranean sea with the Atlantic ocean (Hsü et al. 1973,

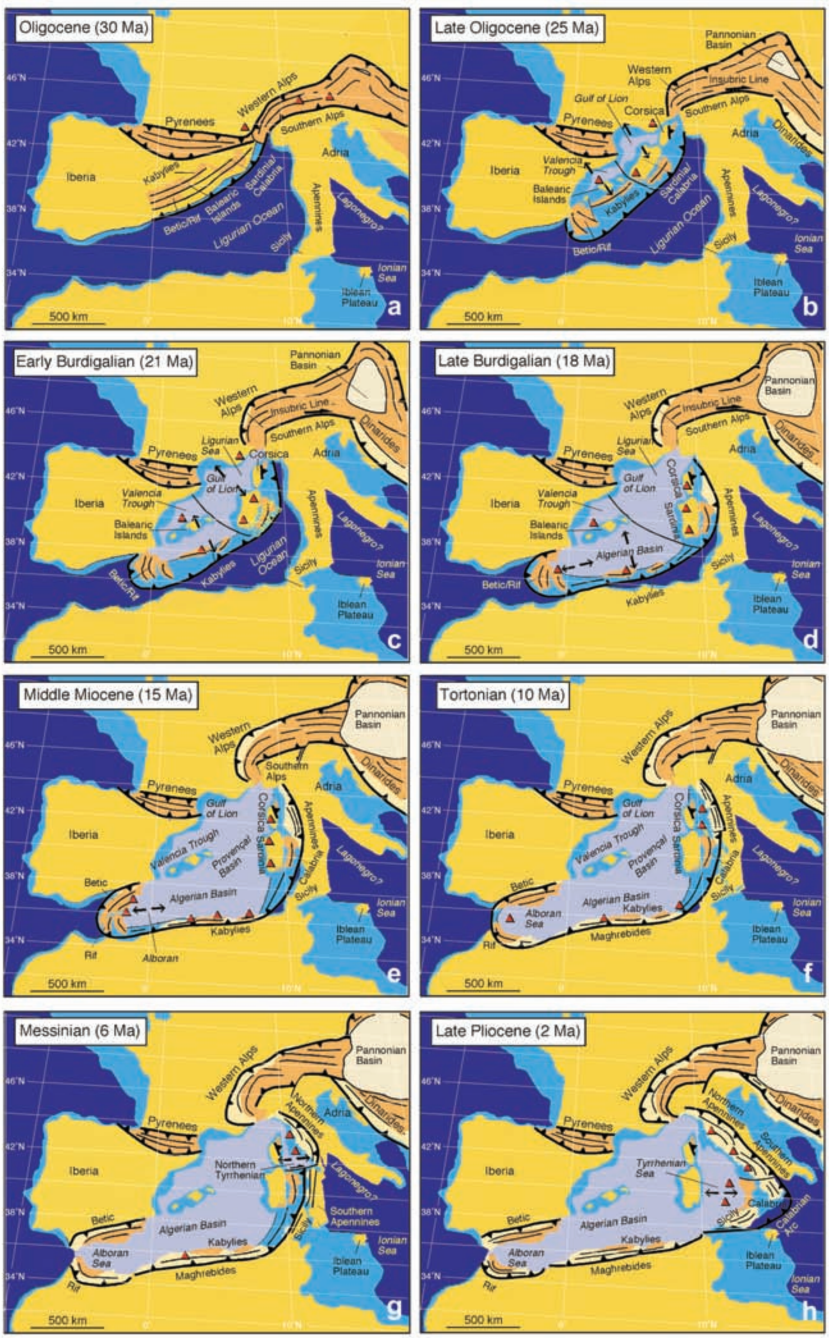


Figure 1 Illustration of the different key events in the formation and evolution of the western Mediterranean basin, between 30 and 2 mya. [Extracted from Rosenbaum et al. 2002a]

Duggen et al. 2003, Ryan 2011). Although the causes of the closure are still not fully understood, the isolation was enough for the Mediterranean sea to dry up almost completely (Hsü et al. 1973, Krijgsman et al. 1999, Duggen et al. 2003, Blondel et al. 2010). This event started about 5.9 mya and ended some 600 000 years later, around 5.3 mya, although the chronology is still debated (Krijgsman et al. 1999, Roveri & Manzi 2006, Rouchy & Caruso 2006, Blondel et al. 2010, Cosentino 2013, Manzi et al. 2013).

Because the Mediterranean is almost 90% fed with water from the Atlantic Ocean, an obstruction of the narrow strait of Gibraltar would determine a decrease of the sea level with more than one meter each year (Blondel et al. 2010). It probably took less than 1000 years for the Mediterranean to have a major sea-level drop of more than 1500 m (Ryan & Cita 1978, Suc 1984). This dry-up was followed by an accumulation of very large and thick seabed salt deposits (Hsü et al. 1973, Lofi et al. 2005, Gaullier et al. 2010) and by massive erosions of the basin's margins and the formation of deep subaerial canyons. Several submerged gorges still exist in places where rivers such as the Nile or the Rhône were feeding the Mediterranean (Blondel et al. 2010, Lofi et al. 2011). For instance, a gorge 900 m below sea level was found near Marseille, where the Rhône flowed into the Mediterranean (Blondel et al. 2010).

The MSC had three main stages, characterized by different paleoenvironmental conditions (Roveri et al. 2014). First, evaporites precipitated in shallow basins then passed to the deepest depocentres during the second stage, and finally the Mediterranean transformed into a brackish water lake in the last "Lago di Mare" phase, which was characterized by large-scale environmental fluctuations (Roveri et al. 2014).

At the beginning of the Pliocene a series of tectonic events broke the land bridge between Africa and Europe and the Mediterranean basin started to refill rapidly (Blanc 2002, Loget & Van Den Driessche 2006). It is believed that when the waters of the Atlantic Ocean started to submerge in the Mediterranean, huge cascades of up to 4000 m high were formed, although evidence for these waterfalls was not yet found (Blondel et al. 2010).

The MSC was defined as an ecological crisis driven by a combination of glacio-eustatic and tectonic events that had a tremendous impact on the marine and terrestrial faunas and dramatic consequences for the subsequent geological history of the Mediterranean region and for the salinity of all the oceans (Roveri et al. 2006, Manzi et al. 2013, Roveri et al. 2014).

Most probably, many marine organisms could not survive in such extreme conditions of dryness and salinity, while others could have persisted in deep areas or near large rivers (Por 2009, Calvo 2015). During the MSC many regions that were separated by the sea for million years got into contact, like Corsica with Sardinia, Morocco with Spain, Tunisia with Sicily, or the Balearic Islands with Spain (Rögl & Steininger 1983, De Jong 1998), but probably also the reverse occurred and several areas that acquired favourable conditions for terrestrial life remained isolated between the salty beds. Several studies suggested that this short period of connection between lands could be an explanation for the vicariance patterns found in many species that are now separated by sea, including butterflies. There is vicariance between lineages occurring in Morocco and Iberia, like in the case of *Zerynthia rumina* (Nazari & Sperling 2007), several members of the genus *Elphinstonia* (Back et al. 2005), or different lineages and species from the genus *Melanargia* (Nazari et al. 2010). Another case that could be explained by the MSC is the one of *Coenonympha corinna*, an endemic butterfly that occurs on several Mediterranean islands. It seems that the ancestor of this butterfly colonized Sardinia, Corsica and the surrounding islands from the Maghreb during the MSC and remained isolated when the waters returned (Kodandaramaiah & Wahlberg 2009). MSC was probably the most spectacular event during the Cenozoic era (Blondel et al. 2010) and is a logical hypothesis for the origin of the several endemic butterflies in the Mediterranean islands.

Glaciations

Climate has a major impact on driving global biogeographic patterns and fast changes in species distributions after recent climate change have been documented for various taxa (Parmesan 1999, Devictor et al. 2012 for individual butterflies species and communities). Fossil evidence and current phylogeographic patterns show that the Pleistocene ice ages, in particular, had a tremendous impact on the past distribution of species and their effects are still visible in the current distribution of species and genetic lineages (Hewitt 1993, 1996, 1999, 2000, Taberlet et al. 1998). It has been postulated that changes in the Earth's orbit, coupled with changes in atmosphere composition, have led to Quaternary climatic oscillations (Hays et al. 1976) in cycles of 100 000, 41 000 and 19 - 23 000 years (Taberlet & Cheddadi 2002, Hewitt 2000). Thus, over the past 2 my glacial and interglacial periods alternated, with cold intervals being much longer than warm ones (Blondel et al. 2010) (fig. 2). There were about ten major ice ages with warm intervals in the last one my (Hewitt 2011). In fact Davis (1976) asserted that the Quaternary was "*a cold epoch interrupted periodically by catastrophic warmer events – the brief interglacials with climate similar to that of today*".

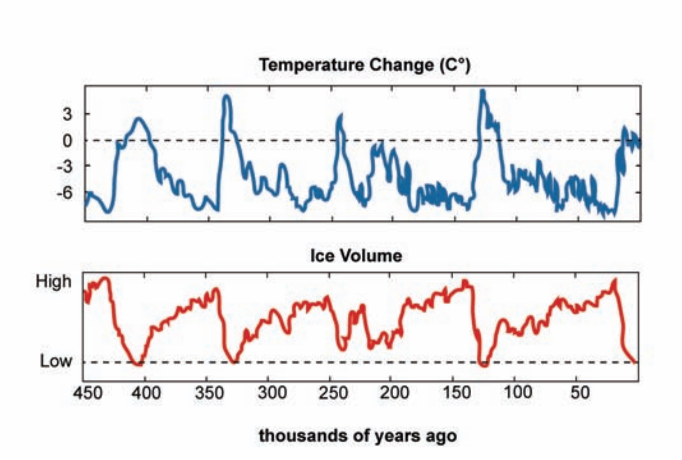


Figure 2 Climatic oscillations (temperature change and ice volume) in the past 450 000 years. [Modified from the graph produced by Robert A. Rohde, University of California, Berkeley]

In the Mediterranean, although the modern climate was established around 2.8 mya, the Pleistocene climatic fluctuations shifted repeatedly and severely affected species, which had to move back and forth in search of suitable habitats (Blondel et al. 2010).

The toughest Pleistocene conditions for species survival in northern and central Europe were during the Last Glacial Maximum (LGM), when ice covered the European continent as low as Warsaw (Poland) and Norwich (UK) (Hewitt 2011). This 'compressed' plant and animal biota to the southern regions, in particular in the Iberian, Italian and Balkan Peninsulas, which played a fundamental role for the survival, diversification and subsequent recolonization of the north (Hewitt 2000, Schmitt 2007). For a long time the Maghreb was not recognized among the main refugia, but recent studies revised its high level of endemicity (Husemann et al. 2014) and presented evidence for refugial-expansion phenomena between north Africa and Europe, at least for butterflies (Habel et al. 2005, 2008, 2009).

Less drastic climatic changes occurred in the dozen Mediterranean islands, which also represented important places for species survival in the Mediterranean (Masini et al. 2008, Médail & Diadema 2009, Dapporto 2010).

During interglacial periods, the populations that were isolated (and presumably diverged in allopatry) in different refugia, expanded their distribution towards the north (Taberlet et al. 1998, Hewitt, 1999). Advances of molecular techniques and good fossil records helped researchers establish the limits of refugial ranges and also to infer possible colonisation routes. Although each species has its own colonisation particularities, Hewitt (2000) described three main paradigms of glacial contraction and interglacial expansion of European biotas (fig. 3).

Each of the main refugial areas had multiple minor sub-refugia and contributed differently to the recolonization by various species (Provan & Bennett 2008, Hewitt 2011). Lineages from the Balkans are the most frequent post-glacial colonizers of central Europe, followed by Iberian and with much fewer numbers the Italian ones (Hewitt 2000, 2004). The later two were probably less frequent because the Alps and Pyrenees mountains impeded their dispersal. However, recent evidence indicates that Hewitt's contraction/expansion models might be too simplistic to depict the more complex biogeographic histories. For instance, genetic data show that fossil individuals of *Ursus arctos* (the model species for the ‘Bear Paradigm’, fig. 3) from Iberia belong to the eastern lineage and specimens collected in the Carpathian mountains belong to the western one (Valdiosera et al. 2007). Thus, the phylogeography of this species, and several others, has been shaped by complex and multiple invasions and retractions, sometimes impeding allopatric speciation. Direct assessments of past movements of invertebrates are even more complicated, largely because for most taxa there are no fossil data and past movements are usually reconstructed indirectly, by modelling present distributions and genetic variability. For butterflies, several studies that combined these approaches reconstructed the evolutionary, immigration and extinction events of certain taxa and a fourth possible paradigm was proposed – the ‘Buttefly Paradigm’ (Habel et al. 2005). This new colonization route was based on the pattern found in *Melanargia galathea/lachesis/lucasi* complex and it postulated that Iberian lineages were not able to expand more than southern France, therefore central and northern Europe was colonized by

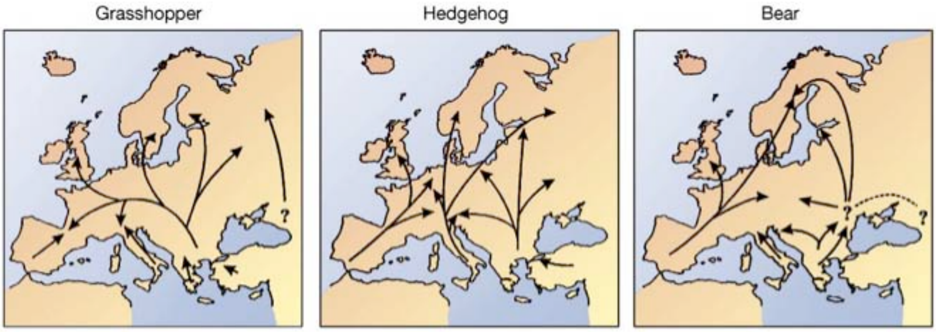


Figure 3 The three main postglacial colonization routes, based on DNA differences for the grasshopper (*Chorthippus parallelus*), the hedgehog (*Erinaceus europeus/concolor*) and bear (*Ursus arctos*). [Extracted from Hewitt 2000]

Italian and Balkan lineages. This pattern was also retrieved in other two European butterfly species, *Polyommatus bellargus* and *P. coridon/hispana* (Schmitt & Seitz 2001, Schmitt et al. 2005a). However, the inferred colonization events of several other species did not match the 'Butterfly Paradigm'. In the case of *Polyommatus icarus* and several species in the Satyrinae subfamily it has been shown that different genetic lineages repeatedly invaded and presumably replaced the original ones throughout their distribution range (Dincă et al. 2011a, Dapporto & Bruschini 2011, Dapporto et al. 2012). These studies also highlighted that many times the secondary invasions were stopped at sea straits, explaining a larger concentration of relict species on islands (Masini et al. 2008 for fossil evidence, Dapporto et al. 2011, 2012 for butterflies).

Recent phylogeographic studies are adding new evidence to the complex recolonization history and show that unknown glacial refugia, called 'cryptic refugia', existed outside the Mediterranean basin (Provan & Bennett 2008). The great amount of lineages that expanded towards the north encountered along the way other different lineages that survived in these smaller refugia and produced mixed biotas, often with the formation of hybrid zones (Hewitt 1998, 2004). A recurrent finding is that these contact zones are in most cases very narrow, suggesting that dispersal of individuals belonging to one lineage in the area occupied by the other lineage is hampered by several possible factors, such as density dependent phenomena, competition, reproductive interference or adaptation to local environmental settings (Waters et al. 2011).

Besides climatic oscillations, there was another phenomenon associated with Quaternary glaciations. During glacial periods the sea water level lowered with a maximum of about 100 m. In these conditions many present-day islands were connected to continents, others considerably extended their surface and some islands that today are below sea level emerged from the waters (Whittaker & Fernández-Palacios 2007). The land connections were a perfect route for biotic exchange between areas that are currently no longer in contact. In the Mediterranean, the circum-Sicilian islands and most of the islands in the Tuscan Archipelago were connected with the Italian Peninsula during the Pleistocene sea regression (Dapporto & Cini 2007, Fattorini 2009, 2011). Research done on various groups of organisms occurring on these islands (butterflies, chrysids, tenebrionids, reptiles and molluscs) provided evidence that some species colonized the islands from the Italian mainland, during the temporary land connection. Some of them probably underwent subsequent extinctions after the LGM, as a result of the decrease in area and habitat loss, but others persisted and evolved into endemic species (Fattorini 2009, 2011).

The western Mediterranean today

Geologic and climatic events concurred into shaping the western Mediterranean basin (fig. 4). Today this region is formed by a mosaic of landscapes, from mountains as high as 4000 m, fertile hilly areas, semi-arid steppes, coastal lands, dozens of superb islands of different origin, to extended deserts that are bordering the south (Box 1). All these environments are tremendously heterogenous and have marked local variations in climate, topography or soil types. The western Mediterranean is also considered as a 'tension zone' (Raven 1964), situated in-between two deserts: a hot one to the south (Sahara) and a cold one to the north (northern Scandinavia). All these, combined with its geographical position at the interface between two continents – Africa and Europe – produced a very distinctive flora and fauna (Blondel et al. 2010).

Before the Quaternary the Mediterranean was largely covered by deciduous and partly evergreen trees, some which persisted until the present, accompanied by several tropical and subtropical arboreal taxa that went extinct probably due to glaciations and the MSC (Sadori et al. 2013). During glacial periods the most widespread vegetation types were the steppe and grasslands, while the interglacials were characterised by deciduous or evergreen forests (Sadori et al.

Figure 4 An approximate delimitation (pink shade) of the western Mediterranean basin, based on the delineation made by Blondel et al. (2010). The main mountain chains are highlighted.

2013). The present day Mediterranean vegetation has been shaped by climatic oscillations since the second half of the present interglacial and by human impact. Today, the most characteristic and widespread vegetation type is the *maquis* or *matorral*, with species from the genera *Quercus, Erica, Juniperus, Myrtus, Olea, Phillyrea,* and *Pistacia* (Tomaselli 1977). Other characteristic Mediterranean species belonging to the genera *Arbutus, Calluna, Ceratonia, Chamaerops*, and *Larus* are tropical relicts from forests that were dominant two million years ago (Moreira et al. 2012). Shrubland with *Artemisia, Lavandula, Rosmarinus, Salvia* and *Thymus* are also a common occurrence in the semi-arid, lowland, and coastal regions, forming the typical *garrigue* or *phrygana* (Moreira et al. 2012).

The Mediterranean has been a meeting ground where taxa of various origins interacted, hybridized, or speciated (Blondel et al. 2010). The diversity of plants and animals is therefore composed by a mixture of species that either evolved *in situ* or dispersed from other regions recently or in the distant past (Blondel et al. 2010). The Mediterranean has a particularly high diversity of flowering plants – more than 25 000 are native to this region (Vogiatzakis et al. 2006, Médail 2008) and more than half of them are endemic (Cuttelod et al. 2008). This, as well as the rich diversity and endemism of other groups of organisms, explain that the Mediterranean is considered one of the 35 biodiversity hotspots in the world (Fig. 5) (Myers et al. 1999, Mittlemeier et al. 2004, http://www.conservation.org/How/Pages/Hotspots.aspx).

The climate of the Mediterranean is dominated by cool and wet winters and hot, dry summers, while rainfall ranges from as little as 100 mm to as much as 4000 mm in some areas (Blondel et al. 2010). Snowfalls are rare in the Mediterranean, except on the mountains, but periods of frost can be common and cause high mortality in plants and animals, especially if they occur after a warmer period.

The Mediterranean is governed by a wide range of winds that have major impacts on living organisms. In winter, the most preponderant winds are the *mistral*, blowing from the northwest, the *tramontane*, blowing from the north, *bora* from north-north-east and the *gregale*, a north-eastern wind that is especially dreaded in Malta island for its powerful force. During warmer climate the *sirocco* takes over from Africa towards the Mediterranean sea, although in north

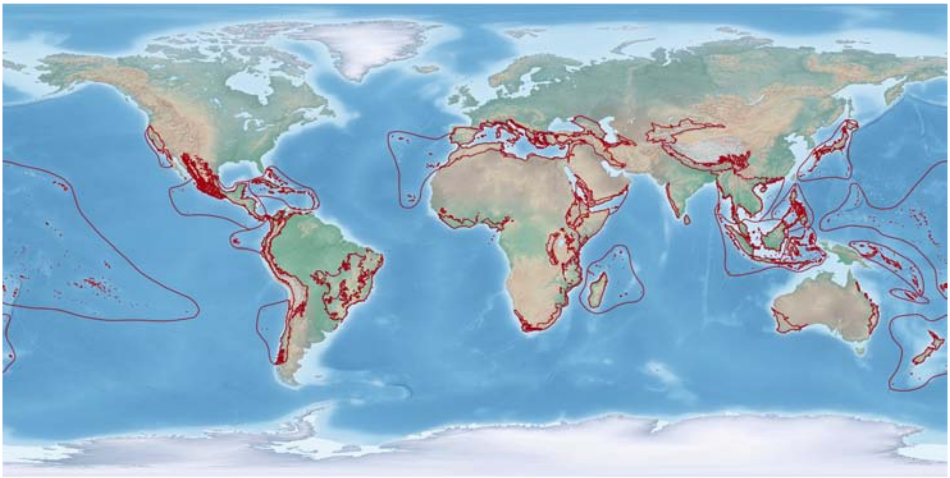


Figure 5 The 35 biodiversity hotspots around the world defined by Conservation International. An area is considered as hotspot if it has at least 1500 vascular plants (> 0.5 percent of the world's total) as endemics and 30% or less of its original natural vegetation. [Based on the data provided by the Critical Ecosystem Partnership Fund at http://www.cepf.net/]

Africa it is known under different names (e.g. *chilli* in Tunisia and south Algeria). The northwestern part of the Mediterranean basin is also hit by humid winds such as the *levant*, from the east and the *ponant* from the west, which can rise in winter and become very violent (Blondel et al. 2010).

Not only natural events are continuously transforming the Mediterranean but also humans (Box 1). Various civilizations have greatly impacted the ecosystems of the Mediterranean Basin in the last 8000 years by deforestation, intensive livestock grazing, fires and infrastructure development, especially on the coastal areas (Sundseth 2009). The agricultural lands, evergreen woodlands and *maquis* habitats that dominate today's landscapes are in good part the result of these anthropogenic disturbances. Although human impact was mostly negative, several activities such as grazing and fires, maintained species richness since in their absence forests would have taken over many open habitats, thus reducing local biodiversity. This is particularly the case for butterflies, since most species fly in more or less open areas and only a minority are confined to dense forests.

There is no doubt that the rich diversity in the Mediterranean has been generated after a long history of tectonic events, by a mosaic landscape and by the interaction with human history, but perhaps the factor that had the highest impact on the Mediterranean, as we know it today, was the climate of the western Palearctic during the Pliocene and Pleistocene.

Box 1 Interesting facts about the Mediterranean

- The Mediterranean Sea is the warmest deep sea on the planet;
- The Mediterranean is also the largest inland sea in the world and the only large sea surrounded by three continents: Europe, Africa and Asia;
- The Mediterranean Basin has one of the largest archipelagos in the world, with almost 12 000 islands and islets;
- The Mediterranean is one of the most studied regions of the world for fauna and flora since the earliest emergence of philosophy;
- The Mediterranean region is the number one touristic destination in the world.

Butterflies as a model system

Butterflies have been used for centuries to investigate many biological queries and have arguably become one of the most used model systems in biology (Watt & Boggs 2003, Dennis 2009, 2012). There are many attributes that make them particularly suitable for research in diverse fields such as genetics, ecology, developmental biology, physiology, evolution, population dynamics and biological conservation (Dennis 2009, Roe et al. 2010). Butterflies have been studied for more than 250 years and thus they provide a unique data resource that is unrivalled in the insect realm both in terms of geographic resolution and time scales. As a consequence they have a very well known taxonomy that provides the necessary framework for various types of research (Settele et al. 2009). Butterflies are also one of the groups that are currently leading country-based monitoring schemes (e.g. Pollard & Yates 1994, Stefanescu 2000, van Swaay et al. 2008,) as well as the evaluation and implementation of the DNA barcoding initiative (Hebert et al. 2003a).

They have a very high ecological importance, not only because of the high species number and percentage of biomass in the ecosystems, but also because they act as herbivores, pollinators and food for insectivores. Their diurnal activity and conspicuous appearance also make them relatively easy to observe in the field and collect data. It is because of this that butterflies have been extensively surveyed since the XVIIIth century and at present numerous specimens are preserved in vast museum collections around the world (Settele et al. 2009). This amount of data facilitates studies focused on tracking changes in community distribution and abundance throughout long periods of time (Parmesan et al. 1999, Devictor et al. 2012). Due to their phytophagous life-style, with the larvae often feeding on a single or few genera of plants, and to their high dispersal capabilities, butterflies respond very fast to perturbations in their environment (e.g. fluctuations in temperature, microclimate, humidity) (Kremen et al. 1993). Recently, it has been demonstrated that their climatic debt is among the smallest known since, due to their winged adults, they can easily track suitable habitats

(Devictor et al. 2012). For these reasons they represent reliable bioindicators to study the impact of habitat loss and fragmentation, as well as the effects of climate change (Brown 1991, Kremen et al. 1993, Fagua 1999, Dennis 2012).

Besides systematics, the analysis of the distribution of butterflies is relevant for eco-evolutionary research. For instance, the Mediterranean, with its well-defined glacial refugia, allows for clear working hypotheses regarding the formation of taxa and their climatic preferences (Hewitt 2000, Schmitt, 2007).

Given that they are probably the most popular group among insects, butterflies have become a flagship group for the conservation of invertebrates and also umbrella species whose targeted conservation is beneficial to other threatened organisms (Settele et al. 2005, 2009).

Butterflies are not only important for science, they are popular for the general public as well, promoting a more positive perception on invertebrates in general (Kellert 1993). People like to see them in nature, to photograph or rear them, some even travel to other countries for 'butterfly watching', a form of ecotourism that is beneficial for many developing countries.

Below, examples of several species that can be regarded as models for various lines of research are briefly presented.

Probably the best-known model species is *Danaus plexippus* and its fantastic yearly migration between North America and Mexico, where an impressive number of adults overwinter. This made the species a notorious model to study invertebrate migration, reproductive physiology, community management, overwinter biology, ecotourism, etc. (CEC 2008). Another famous example of insect migration is that of the cosmopolitan species *Vanessa cardui* (Petrakis & Legakis 2005, Stefanescu et al. 2007).

Bicyclus anynana is a tropical species that displays phenotypic traits associated with seasonal changes in the environments of eastern Africa. Such plasticity, linked to the fact that the species is easily reared in the laboratory, represents fertile ground for the fields of evolutionary genetics and developmental biology studies (Brakefield 1998, Brakefield & French 1999, Brakefield et al. 2009, Liénard 2014, Wasik et al. 2014).

Butterflies in the genus *Heliconius* display both Müllerian and Batesian mimicry and they are mostly studied for speciation, genetics of wing patterns and hybridization. Recently, they are also becoming primary models for genomic studies (Benson 1972, Mavárez 2006, *Heliconius* Genome Consortium 2012, Martin et al. 2013).

In Europe, several Nymphalidae species (*Melitaea cinxia, Euphydryas aurinia, Boloria aquilonaris* and *B. eunomia*) occur in metapopulations and they are intensively studied for the ecology and dynamics of natural populations (Hanski et al. 1994, Warren 1994, Hanski 1999, Mouson et al. 1999, Goffart et al. 2001, Brunzel 2002, Wahlberg et al. 2002, Anthes et al. 2003, Hula et al. 2004, Baguette et al. 2011). Among them, *M. cinxia* is an internationally recognized model system for metapopulation biology (see Hanski 1999 for a review).

But probably some of the most intensively studied butterflies in Europe are the species in the genus *Maculinea* (placed within *Phengaris* by Fric et al. 2007). Their highly specialized life cycle, with larvae that are parasites of ant nests, converted them into a model for studying the evolution of parasitism (Als et al. 2004). All the European taxa in the *Maculinea* genus are highly endangered as a result of habitat loss and they are flagship species for conservation (Elmes & Thomas 1992, Thomas 1994, Settele & Kühn 2009). *Maculinea* butterflies have also been selected as one of the three priorities in butterfly conservation by the World Conservation Union, together with *Ornithoptera alexandrae* (the largest butterfly in the world) and the Mexican roosts of *D. plexippus* (Thomas & Settele 2004).

Another group that is increasingly attracting attention is the genus *Leptidea*, in particular *Leptidea sinapis*. For a long time this species was considered as common and widespread, until recent morphologic, genetic, karyological, and behavioural research revealed that it actually represents a triplet of cryptic species: *L. sinapis, L. reali* and *L. juvernica* (Lorković 1993, Martin et al. 2003, Dincă et al. 2011b, 2013). In addition to cryptic diversity, the *Leptidea* genus is a striking example of chromosomal instability in organisms with holocentric chromosomes. The available data indicate that the within-genus chromosome number variability is very pronounced and ranges from 2n=51 (*L. reali*) to 2n=208 (*L. duponcheli*). Furthermore, several species within the genus display intraspecific chromosome number variability, with the most notable case being that of *L. sinapis*. Chromosome numbers within this species, ranging from 2n=56 to 2n=106, represent the widest known intraspecific chromosome number variability excluding cases of polyploidy (Lukhtanov et al. 2011). Recent karyological studies also revealed a curious sex determination system with 3-4 W and 3-4 Z chromosomes in three species of *Leptidea*, which is unique in the Lepidoptera and which could also have played a role in the diversification of the genus (Šíchová et al. 2015). All these recent discoveries render *Leptidea* as an emerging system for the study of speciation and chromosomal evolution.

The phylogeography of *Maniola jurtina* has been probably the first and the best studied among European species. Due to a striking diversification of the genitalia structures and to a corresponding variation in allozyme patterns (Thomson 1987) *M. jurtina* has been used as a model to better understand how species shifted their distribution as a consequence of the glaciations and what were the effects on their genetic structure (e.g. Schmitt et al. 2005b, Dapporto et al. 2009, 2011, Habel et al. 2009, Dapporto & Bruschini 2011, Kreuzinger et al. 2015). However, due to contrasting results obtained among different studies, the phylogeographic history of this species is still not fully elucidated and requires further research.

The butterflies of the western Mediterranean

Although the taxonomy of European butterflies is very well studied, still there is a lack of a single widely accepted species composition for the butterflies of this region. The total number of European species varies between 440 and 489 (Karsholt & Razowski 1996, Kudrna 2002, van Swaay et al. 2010, Karsholt & Nieukerken 2013) with most species, including endemics, cumulated in the mountainous areas (fig. 6).

Currently in the western Mediterranean there are about **335 species** of butterflies (superfamily Papilionoidea) (based on: Karsholt & Nieukerken 2013, Tennent 1996, Tarrier & Delacre 2008, Dapporto 2009, Tshikolovets 2011, Dincă et al. 2011a, 2011b).

The majority of invertebrates occurring in the Mediterranean, including butterflies, are of Palearctic or Holarctic origin, some are North African elements and very few are of tropical descent (e.g. *Charaxes jasius*, *Lampides boeticus*, *Leptotes pirithous*, *Danaus chrysippus*, *Danaus plexippus, Borbo borbonica*, for butterflies) (Larsen 1986, Casevitz-Weulersee 1992, Blondel et al. 2010). Mediterranean butterflies belong to six families: Papilionidae, Pieridae, Lycaenidae, Nymphalidae, Riodinidae and Hesperiidae. The family Riodinidae is mainly Neotropical, with only one species in Europe (Espeland et al. 2015).

Some families are more abundant in particular areas than others. If compared with the mean proportions for Europe plus north Africa, there is a surplus of Lycaenidae in Iberia, of Pieridae in north Africa and the Mediterranean islands, of Hesperiidae in Iberia, particularly in the Algarve region (Portugal), and a deficit of Satyrinae outside the mountain areas (Dennis et al. 1995).

The number of species declines with latitude, whereas the relative abundance of Nymphalidae species becomes relatively higher as latitude increases (Dennis et al. 1995). Moreover, butterfly species richness displays a bell-shaped relationship with altitude: maximum diversity is found at mid mountain, as is the case with plants and other organisms (Gutiérrez et al., in prep.).

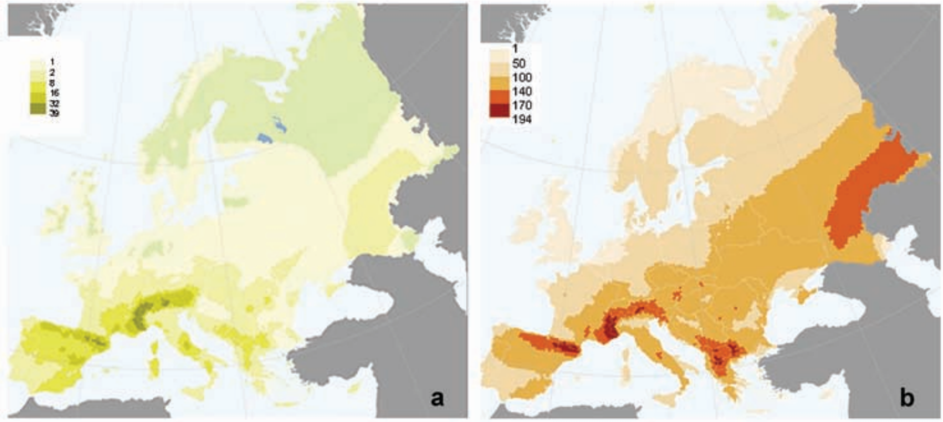


Figure 6 Species richness (a) and the distribution of endemic species (b) in European butterflies, according to the IUCN red list of butterflies (Van Swaay et al. 2010). The figures in the legends represent the number of species and endemic species, respectively, per 864 square km hexagon [Maps extracted and modified from Van Swaay et al. (2010)]

Currently, 89 species of endemic butterflies have been identified for this region (fig. 7). The European and the African sides of the western Mediterranean are equally contributing to the number of endemics (around 33% each), while island endemics represent a lower but still important fraction (around 13%). Interestingly, the endemic species that are shared by two or more of these three main areas (Europe, north Africa and islands), turned out to be much fewer (no more than 20% of the total). This suggests that the speciation processes operating in allopatry among these three main blocks are the main engine for endemicity in the western Mediterranean.

As stated before, butterflies are one of the best-studied groups of invertebrates and the comprehensive knowledge we have about them today reflects more than two centuries of research, done by both professional and amateur lepidopterists. For much of this period the research was mostly focused on the classification of taxa based on external (e.g. de Jong 1972) and internal (genitalia) morphology (e.g. Warren 1936, Higgins 1975) and on documenting their distribution with increasing precision (Kudrna 2002). Since the mid-twentieth century, karyological data started to be used as a tool for butterfly systematics (Lorković 1941, de Lesse 1960, Lorković 1990, Lukhtanov et al. 2005, Kandul et al. 2007), and in the last three decades allozyme variation and DNA sequencing became commonly used in taxonomy and systematics as well (Wahlberg & Saccheri 2007,

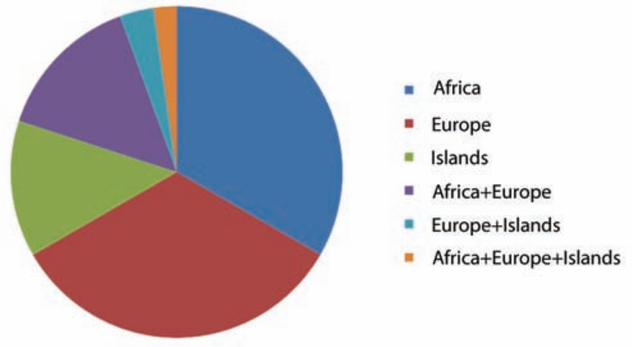


Figure 7 The number of endemic species occurring in each of the main areas in the western Mediterranean.

Habel et al. 2010, Dincă et al. 2013, Talavera et al. 2013, Tóth et al. 2014). Currently there are many books documenting various aspects for butterflies, such as their biology (Vane-Wright & Ackery 1989), ecology (Settele et al. 2009), conservation (Settele et al. 2008, van Swaay et al. 2010) or distribution at a large scale (Kudrna 2015). There is also a growing number of books dealing with the butterfly fauna from specific countries or regions (e.g. Henriksen & Kreutzer 1982, Geiger 1987, Stoltze 1996, Lafranchis 2000, Settele et al. 2000, Asher et al. 2001, Beneš et al. 2002, Maravalhas 2003, Eliasson et al. 2005, Tarrier & Delacre 2008, Pamperis 2009, García-Barros et al. 2013).

The western Mediterranean butterflies are usually studied in a broader European context, but most studies focus only on certain taxa (Weingartner et al. 2006, Kodandaramaiah & Wahlberg 2009, Dincă et al. 2011a, Dincă et al. 2013) and the markers used show many times discrepant results (Schmitt et al. 2003, Habel et al. 2005, Dapporto 2010, Nazari et al. 2010, Dincă et al. 2011a). Several studies examine the island butterfly communities, but they are based only on presence data (Dennis et al. 2008, Dapporto & Dennis 2009, Fattorini 2009) or focus on insular endemics (Cianchi et al. 2003, Cesaroni et al. 1994, Grill et al. 2007). The biggest flaw of the studies targeting the taxonomy, distribution and the ecology of the butterflies from the western Mediterranean region is the general lack of data and analyses of the Maghreb region (e.g. Kudrna et al. 2011, 2015, Devictor et al. 2012, but see Dennis et al. 1991, 1995, 1998, Husemann et al. 2014). Nevertheless, the high fraction of endemics (fig. 7, Husemann et al. 2014) and the presence of the southern limit of distribution for many species (a fundamental proxy for thermal tolerance) in the Maghreb are key elements for understanding the Mediterranean butterfly fauna and for assessing conservation strategies under the current global warming threat.

Although the European butterfly fauna is one of the best studied among invertebrates, thanks to the development of molecular techniques we are still discovering new cryptic species (Dincă et al. 2011a, 2011b, 2015, Nazari et al. 2010, Zinetti et al. 2012, Tóth et al. 2014,).

Cryptic species can represent a very important fraction of biodiversity because they may provide an unbalanced contribution to community diversification. In recent years several studies showed that cryptic species are most often closely related taxa, and they tend not to occur in sympatry until their diversification reached a certain level (Pigot & Etienne 2015). For the Mediterranean fauna there are studies showing that particular cryptic species display a chequered distribution, meaning that they are widely distributed, but never or rarely co-exist on islands. The mechanism maintaining mutual exclusive patterns of distribution are still largely unknown and many phenomena have been advocated to explain them (Waters 2011), but testing their relative importance remains a difficult task. Understanding the patterns and the forces maintaining chequered distribution patterns could help us understand the evolution of species and the formation of diversity, not only in the Mediterranean but at the global scale.

Conservation in the ‘Anthropocene epoch’

Contrary to the current perception, it is suggested that Earth’s biodiversity has never been as rich and varied in the last 500 million years as it is now (Valentine 1970, Signor 1990, Rozensweig 1995, Dirzo & Raven 2003). But biodiversity is now also facing rates of extinction as never before, with levels between 100 to 1000 times higher than before humans conquered the Planet (Pimm et al. 1995, 2014, Carlton et al. 1999, Barnosky et al. 2011). Even though the predictions of current rates of extinction vary widely (May 2002, Stork 2010, Costello et al. 2013, Mora et al. 2013, Costello 2015), we are undoubtedly losing biodiversity at an alarming pace. A recent study by Urban (2015) suggests that under business-as-usual policies climate change will accelerate extinctions and one in six species could become extinct by the end of the century. That means loosing more than 20 species every day. Scientists are widely accepting that we are heading into a sixth mass extinction event, one that is comparable to the “Big Five” extinctions of the geological past, when up to 96% of species disappeared (Pimm & Brooks 1997, May 2002, Rosenzweig 2005, Wake & Vredenburg 2008, Stork 2010, Barnosky et al. 2011, Kolbert 2014). The main difference is that this extinction would be mostly the result of human activities (May 2002, Kolbert 2014, Pimm et al. 2014).

A critical part in understanding the magnitude and the impact of the ‘extinction crisis’ is to have precise estimates about how many species live on Earth and how they are distributed. But we still do not know how many species there are and the global estimates vary greatly (see Table 1). According to these figures there could be between 1.84 and 100 million species, although some scientists consider that it is impossible to determine a real range based on the data we have (Scheffers et al. 2012). Recent studies are claiming that the lack of convergence for the estimates of global species richness would be decreased by using estimates of uncertainty (Caley et al. 2014, Fisher et al. 2015) and currently, more and more studies are beginning to take them into account (Mora et al. 2011, Costello et al. 2013,

Table 1 The evolution of the estimated numbers of described non-microbial diversity (species) and of the global number of species from 1982 until present, and the percentage of undescribed species. The superscripts represent the sources of information, which are listed below; the numbers in the square brackets represent the confidence interval and the ones in round brackets represent the range; the * symbol stands for estimates obtained only for arthropod species.

[1]Erwin 1982, [2]Raven 1983, [3]May 1988, [4] Stork & Gaston 1990, [5]Hodkinson and Casson 1991, [6]Hammond 1992, [7]Stork 1993, [8]Hawksworth & Kalin-Arroyo 1995, [9]Tangley 1997, [10]Groombridge & Jenkins 2002, [11]Dirzo & Raven 2003, [12]Chapman 2006, [13]Champman 2009, [14]Hamilton et al. 2010, [15]Carbayo & Marques 2011, [16]IUCN 2011, [17]Hamilton et al. 2013, [18]Costello 2015, [19]Stork et al. 2015.

Year	No. of described species (million)	Total no. of species (million)	Undescribed biodiversity (%)
1982[1]		30	
1983[2]	1.5	3-5	62,5
1988[3]	< 2	10-50	80-96
1990[4]		4.9-6.6	
1991[5]		1.84–2.57	
1992[6]	1.7	12.5	86,4
1993[7]	1.4-1.8	< 10	82-86
1995[8]	1.75	13.6	87
1997[9]	1.5-1.8	100	98,2-98,5
2002[10]	1.75	14	87.5
2003[11]		~ 7	
2006[12]	1.786	8-9	77,6-80,1
2009[13]	1,9	11	> 82
2010[14]	1.6-1.7	2.5-3.7	34-55,4
2011[15]	1.4	6.8	> 79
2011[16]	> 1.2	~ 8.7, ±1.3	86
2013[17]		6.1 [3.6-11.4]*	
2015[18]	1.5	5	70
2015[19]	1.5	6.8 (5.9-7.8)*	> 77

Hamilton et al. 2013, Fisher et al. 2015). But even in the absence of relatively certain estimates, most analyses suggest that approximately 80% of the global biodiversity is still undescribed (Table 1).

Other major challenges that hinder us from obtaining good estimates are the frequent synonymies (the description of one species more than once) (Scheffers et al. 2012, Costello et al. 2013), the difficulty of discriminating certain species based on morphology alone (Hebert et al. 2004, Bickford et al. 2006, Smith et al.

2008, Dincă et al. 2015) or even differences of opinion if a species is "good" or "bad" (Mallet & Descimon 2009), and the fact that many undiscovered species are small, difficult to find and have a restricted geographic range (Pimm et al. 2014). Moreover, the number of taxonomists who could identify and describe new species is decreasing (Godfray 2002, although see Costello et al. 2013) and their labour force is inefficiently distributed, with only a third working on invertebrates (May 2002), which are at least 90% of the total number of species (Wilson 1992).

And how many of the known species are threatened? According to the IUCN Red Lists of Threatened Species there are almost 22,800 species that are threatened in the present, more than double the number in 1996. These calculations were made for only 77,340 species (about 5% of the total described species) of the major groups of organisms.

Currently there are no specific assessments for the degree of endangerment of the butterflies in the Mediterranean region, but there are relatively recent evaluations of IUCN at European level, which are based on national assessments made by butterfly experts. According to these evaluations most of the threatened species are located in central and eastern Europe, although only 21% of these evaluations are reported to be good (fig. 8). Consequently, these potentially subjective evaluations made by researchers should be complemented by other less biased assessments.

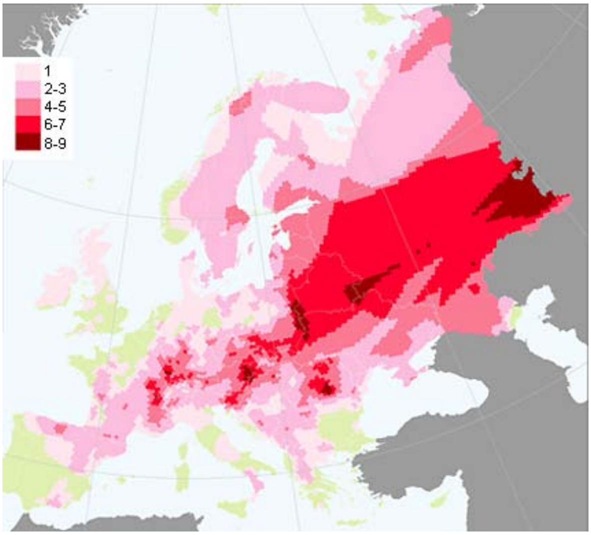


Figure 8 Distribution of threatened butterflies in Europe according to the IUCN red list of butterflies (Van Swaay et al. 2010). The figures in the legend represent the number of threatened species per 864 square km hexagon [Maps extracted and modified from Van Swaay et al. (2010)]

Conservation biogeography defined as "*the application of biogeographical principles, theories, and analyses, being those concerned with the distributional dynamics of taxa individually and collectively, to problems concerning the conservation of biodiversity*", promises to fill these drawbacks (Wittaker et al. 2005). The approach promoted by the field of conservation biogeography can provide highly detailed information for conservation strategies in order to preserve biodiversity from individual populations to entire biotas. However it usually requires large amount of data such as occurrence records, species ecological traits and/or genetic information. Nevertheless, due to the ever-increasing amount of datasets available for European and Mediterranean butterflies, the conservation biogeography approach starts to be a feasible alternative to assess evidence-based conservation strategies.

1.2 Biodiversity and Biogeography

Species assembly and community structure

A community is an assemblage of species that co-occur in space and time. Understanding how communities form and why they have different degrees of similarities/dissimilarities among them are essential goals for ecology (Begon et al. 2006). The classic concept of the assembly rules argued that the composition of species in a local community is mainly filtered from a regional species pool by dispersal capabilities, environmental constraints and species interactions (Diamond 1975, Begon et al. 2006, Belyea & Lancaster 1999) (fig. 9). This view, retrieved in the Equilibrium Theory of Island Biogeography (ETIB) (MacArthur & Wilson 1967) or in the neutral theory of biodiversity (Hubbell 2001), considers the regional species pool of the community as an independent entity that is unchangeable during the period of community assembly and that over short temporal scale, communities are mostly determined by metacommunity dynamics (Mittelbach & Schemske 2015). However, this approach does not consider the fact that species pools are dynamic and that communities are assembled over large spatial scales and/or across evolutionary time. Therefore, a more realistic model should take into account other processes such as speciation, local adaptation, dispersal and extinction (Ricklefs 2008, Pigot & Etienne 2014, Mittelbach & Schemske 2015).

Phenomena operating at the geological time scale had a fundamental importance in shaping the butterfly communities in the Mediterranean (Dennis et al. 1991, Cesaroni 1994, Schmitt 2007, Dapporto 2010, Dincă et al. 2011a). This region experienced a complex paleogeographic and paleoclimatic history and represented a generally suitable environment throughout the Quaternary climatic oscillations, when most living organisms were wiped-out from large parts of Europe (Taberlet et al. 1998, Hewitt 1999, Schmitt 2007). The long-term persistence of species over the Mediterranean islands and mainland permitted the occurrence of many processes operating at evolutionary scale that left an evident fingerprint in current butterfly communities (Dennis et al. 1991, Weingartner et al. 2006, Schmitt 2007, Kodandaramaiah et al. 2009, Leneveu et al. 2009, Fattorini 2009, Habel et al. 2010, Dincă et al. 2011a).

Islands represent a discrete, manageable and replicable system, which makes them an excellent setting to study and understand the community assembly processes (Warren et al. 2015). It is no wonder that most of the studies that reconstruct the processes determining species richness and composition in communities are carried out on island systems since they represent a more simple case that is subsequently used to understand the more complex mainland communities (MacArthur & Wilson 1967, Levins 1969, Fernández-Palacios & Whittaker 2010, Hanski 2010). In Alfred Russel Wallace's words: "*... it is not too much to say that when we have mastered the difficulties presented by the peculiarities of island life we shall find it comparatively easy to deal with the more complex and less clearly defined problems of continental distribution...* "

Islands, being discrete portions of land, allow precise measurements of their area, isolation and of the number of species occurring there. The number of species living on an island depends on a multitude of factors. In particular, three main island characteristics determine species richness: area, isolation and the species occurring at the nearest faunistic sources (MacArthur & Wilson 1967, Kalmar & Currie 2006, Whittaker & Fernández-Palacios 2007, Dennis et al. 2008, Kreft et al. 2008).

The correlation between the number of species and the size of the area they inhabit is one of the best-assessed phenomena in ecology. The first description of the species-area relationship (SAR) belongs to H.G. Watson (1835), but it was formulated mathematically by Arrhenius more than eight decades later (1921). Several subsequent studies demonstrated that the number of species increases with area, according to a curvilinear power function:

eq. 1 $S = CA^z$

where S is the number of species, A is the the size of the area and C and z are two parameters describing the shape of the curve and varying among organisms and island systems (fig. 9 a) (MacArthur & Wilson 1967, Lomolino 2000, Whittaker &

Fernández-Palacios 2007). After a logarithmic transformation the SAR becomes linear (fig. 9 b):

eq. 2 $$logS = zlogA+logC$$

A linear relationship allows for a simpler computation of the parameters.

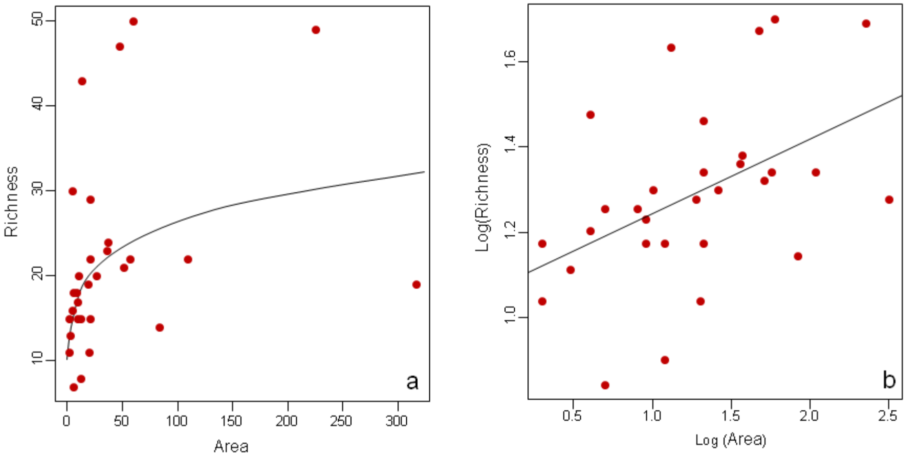


Figure 9 The relationship between island area and species richness for the butterflies occurring on the western Mediterranean islands smaller than 400 km^2, using untransformed values (a) and log-transformed values (b). The untransformed values produce a non-linear fit, as expected for a power function, becoming linear in the transformed relationship. The fit is far from being perfect and island area only explains about 20% of the variance. This is due to the fact that several other factors (mostly isolation and source richness) are at play in determining species richness.

Two main aspects explain why area influences species richness. First, larger islands usually encompass a higher number of different resources and environments, thus providing suitable habitats for more species (Whittaker & Fernández-Palacios 2007, Hortal et al. 2009). This is reinforced by the fact that larger islands often display a more pronounced altitudinal gradient, providing additional potential for the existence of different environments. Second, species richness is determined by the area *per se*, a hypothesis that is strongly rooted in community ecology and in the Equilibrium Theory of Island Biogeography (ETIB) (MacArthur & Wilson 1967, Ricklerfs & Lovette 1999). The ETIB hypothesizes that island communities contain a dynamic set of species, and their composition changes due to continuous phenomena of extinction and immigration. When there are few species on an island, immigration is more likely to occur than extinction, but extinction probabilities rapidly increase with increasing richness. As a result, richness attains an equilibrium point, different for different islands, and each island will harbour an almost constant number of species, even though the community composition can change over time. Larger areas have higher carrying capacities, thus hosting larger populations. Consequently, the extinction probability of the species is reduced, shifting the equilibrium point towards a higher richness (MacArthur & Wilson 1967).

For similar reasons, isolation is another well-known determinant for island species richness, and the more isolated an island, the lower the possibility for species to colonize it and to provide new individuals to the existing populations (the so called rescue-effect, Brown & Kodric-Brown 1977) thus increasing their extinction probability. Consequently, according to the ETIB, a reduced immigration rate results in reduced richness (MacArthur & Wilson 1967). Isolation also lowers or decreases to zero the gene flow between island populations and others in the vicinity, producing two opposite effects. First, a reduction in gene flow will increase the inbreeding rate and, mostly in small areas, the population will become less variable and more prone to extinction (Templeton 1990, Keller & Waller 2002). On the other hand, the founder effect, a higher genetic drift and adaptation to the unique environment of an island may result in the formation of endemic species (Emerson 2002, Coyne & Orr 2004, Whittaker & Fernández-Palacios 2007).

Areas richer in species provide more immigrants to nearby islands. As island richness is a balance between extinction and colonization, richer faunistic sources will determine higher richness on islands (Dennis & Shreeve 1997, Dapporto & Dennis 2008). Also climate has an important influence in determining richness at sources and on islands, and it is well known that different taxonomic groups have higher richness in particular climates, and butterflies make no exception (Zapata et al. 2003, Turner & Hawkins 2004, Hawkins et al. 2004).

All these factors, island area, altitude, isolation, source richness and climate, are variables operating in ecological (contemporary) time to determine the species that live and survive on each island. However, populations can persist on islands for very long periods and are subjected to processes that occur over evolutionary time (Whittaker & Fernández-Palacios 2007). In some cases, species colonized islands in ancient times, when island characteristics were different from those in the present, and survived (and possibly evolved) there until the present (Cronk 1992, Emerson 2002, Whittaker & Fernández-Palacios 2007). The most important changes of island conditions during geological time are due to: i) tectonic movements resulting in the separation of islands from continents (Briggs 1987, 2003), ii) the lowering of the sea level during glacial periods, reducing islands isolation or producing a temporal connection of islands with mainland or other islands (Fernández-Palacios et al. 2015b), iii) volcanic activity, which accrues oceanic islands but that may also erase their entire biodiversity (Thornton 1997, 2001, Whittaker et al. 2008) and iv) the erosion of islands resulting in a reduction of their area and altitude, and consequently in their species richness (Whittaker et al. 2008). Differences in climate also occurred in the past but they simultaneously affected both islands and mainland populations, thus producing less contrasting results. Several species, mostly those characterized by low dispersal capability (e.g. snails, non-flying insects, mammals, reptiles, amphibians), that colonized islands during glaciations, often persisted until recent times (Fattorini 2002, 2009, Hausdorf & Hennig 2005, Papadopoulou & Knowles 2015). In some extreme cases, small populations occurring on islands were even less prone to extinction than mainland ones, since isolation allowed them to avoid competitors, predators or diseases carried by parasitoids and microorganisms (Masini et al. 2008, Dapporto et al. 2012). A classical example is provided by the separation of Australia from Antarctica, which favoured the survival of marsupials by blocking the invasion of the placental mammals. Islands that were in contact with mainland acquired many species in the past and some of them survived until today as relicts, enriching the observed communities (Whittaker & Fernández-Palacios 2007). Besides the impact that contemporary and historical factors have on island characteristics, the effect of human activities largely varies among islands producing multiple alterations to the original environment (e.g. deforestation, fragmentation, overgrazing, different types of agriculture, urbanization, introduction of invasive alien species, etc.). The human impact has a strong influence in determining island richness, and is usually directed towards reducing the number of species (Whittaker & Fernández-Palacios 2007).

In the presence of so many factors influencing species richness it becomes evident that each island community provides particular contemporary, historical and anthropic settings, and that species richness is one of the features that confer individuality to island communities (Cronk 1999).

Nestedness and species replacement (turnover)

Besides the quantitative information provided by the number of species occurring on an island, their qualitative identification and comparisons among communities can provide fundamental information about the processes that concur in assembling island faunas (Begon et al. 2006, Baselga 2010, Dobrovolski et al. 2012).

A characteristic that is universally recognized is that when examined not only for richness, but also for their species composition, smaller island communities tend to represent an almost perfect subset of richer communities from nearby areas (Patterson & Atmar 1986, Atmar & Patterson 1993). This phenomenon is called nestedness and it is one of the best-known and widespread characteristics of island communities (Atmar & Patterson, 1993, Ulrich et al. 2009). To better understand this phenomenon, the communities can be compared with a series of Matryoshka dolls, in which smaller dolls perfectly fit into the larger ones.

The nestedness phenomenon implies that some species have higher probabilities than others to establish themselves in a community. This was not considered in the ETIB, which is a 'neutral theory', and considers that all the species existing in a region have the same probability to colonize an island or to become extinct there (Losos & Ricklefs 2010). This assumption can be useful to describe how island richness is generated, but observations of different species frequencies represent a strong proof that species are far from having the same probabilities of occurrence on islands. In fact, in almost all the insular systems, there are species occurring over all islands (or almost all) and species that tend to occupy a much smaller number of islands, usually the largest and less isolated ones (Ulrich et al. 2009, Dennis et al. 2012). To colonize an island a species must cross a sea barrier and establish a population by exploiting its resources, usually limited in both quality and quantity. Indeed, species do not have the same capability to fly over long distances and they also differ in their habitat requirements (Lomolino 1996, 2000).

Most recent neutral theories argue that more than species traits, the different relative frequency of species at source determines a different probability for them to occur on islands (Hubbel 2001, 2009). The more individuals occur at source the more likely they will disperse and establish island populations (Hubbel 2001, 2009).

A simple system based only on different dispersal capability of a series of butterfly species is represented in fig. 10. Four islands (I1-I4) are located at different distances from the mainland (M), which hosts five species. We can assume a simplified but realistic scenario, in which each species has a different dispersal capability (indicated by the coloured curved lines in fig. 10), being able to reach a maximum distance from mainland. We also assume that island populations are much smaller than mainland populations and thus cannot provide enough individuals to allow island-to-island colonization. It is clear that the purple species can reach all the islands, the green species all islands except I4 and so on. This model shows a perfectly nested structure of communities (fig 10, right side) where all the species occurring in a smaller community are also present in the larger ones.

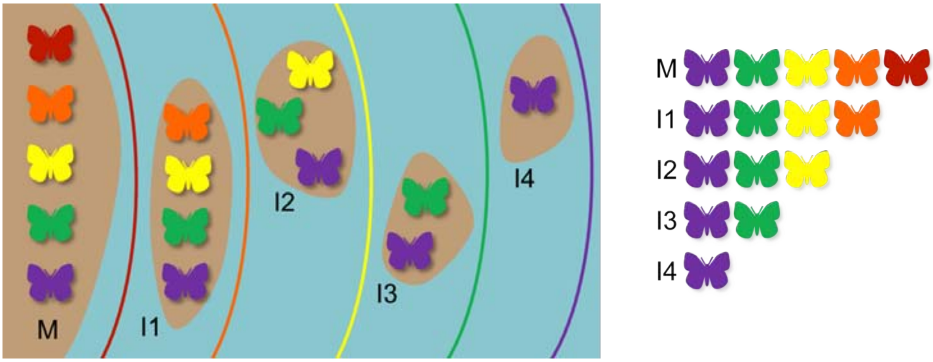


Figure 10 An example of butterfly communities based on different dispersal capabilities of five species, resulting in a perfectly nested pattern of island faunas. Coloured curved lines represent the dispersal limits for species with corresponding colours.

Similarly to a series of Matryoshka dolls, the individual traits of each community are limited to differences in size (represented here by species richness), and in a perfectly nested series of communities their individuality is linked to richness only. Therefore, simply knowing the relative frequency of the species at source and island richness it is possible to exactly predict the composition of each community. The occurrence of a highly nested pattern also provides essential information for biodiversity conservation (Whittaker & Fernández-Palacios 2007). In a situation like the one depicted in fig. 10, where dispersal predicts the occurrence, individuals of each species can often reach an island situated within their dispersal range and any population undergoing extinction would be in most cases readily re-established after a new colonization event. A key consequence is that, in a perfectly nested pattern, by preserving the richest area, all the sets of species and the biogeographical pattern of each island would be protected, since they would be easily re-established (Whittaker & Fernández-Palacios 2007). Paradoxically, in our example, only by protecting the mainland and ignoring the islands would seem enough to preserve the entire biodiversity contained in the system.

However, albeit communities often show a nested pattern, they are usually far from being perfectly nested (Matthews et al. 2015) because many factors operating in both ecological and historical time can disrupt the occurrence of such a pattern. To exemplify this phenomenon, another scenario is illustrated in fig. 11, in which the species maintained their differential dispersal capacity but two mainland areas and several other processes have been added. The existence of two faunistic sources would be the first factor disrupting the nestedness pattern since in most cases different areas host different species. Under this scenario only one species from M1 (the yellow species) is replaced in M2 by another species (the brown one) – a phenomenon described as species replacement or turnover. Presumably, the species responsible for turnover on mainland will occur on the islands closest to a particular mainland, depending on their distance from the two source areas. This in turn will generate turnover on islands. Notably, in a perfectly nested pattern as the one represented in fig. 10, there is no evidence of turnover. Species replacement is also generated by the evolution of new taxa and by relict populations. A speciation event has been hypothesized to occur on the smallest and most isolated island (I3) and the pale green species evolved into the dark green species. We also introduced the existence of a relict species (black), persisting on some islands after its extinction on mainland. Finally, the purple species lacks a key resource in I1 and the pale green became extinct on I5 because of a stochastic event. A more complex combination of realistic processes can transform the perfectly nested pattern represented in fig. 10, into a highly chequered distribution (fig. 11).

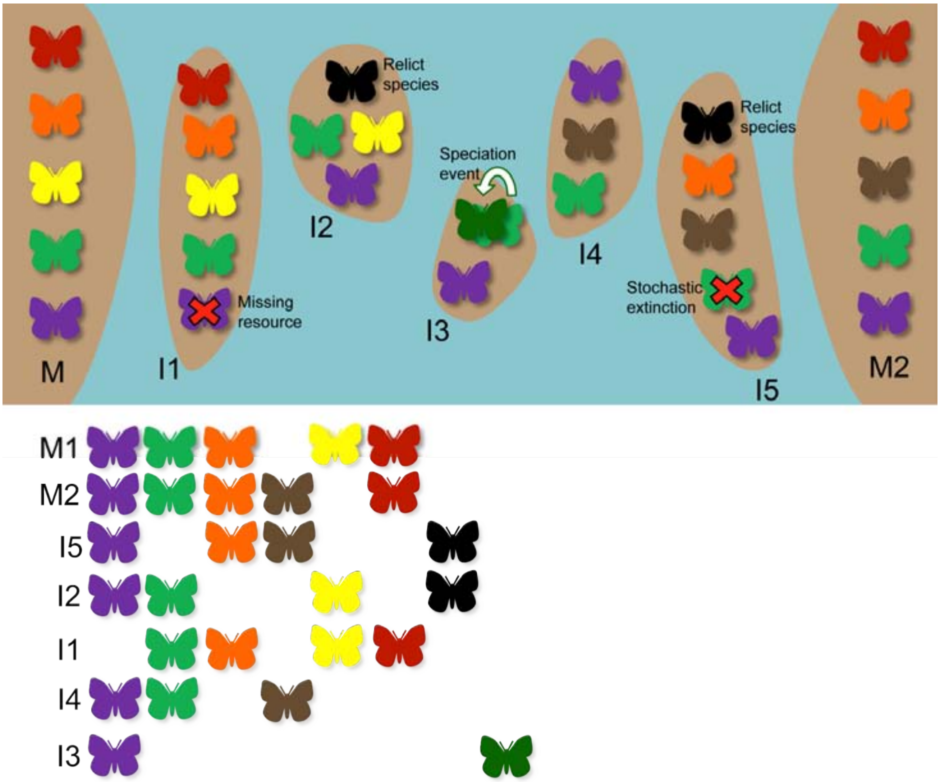


Figure 11 A scenario showing how factors such as different source faunas, relictuality, speciation and extinction events can disrupt the nested structure and create turnover and unpredictable and unique communities on islands.

A nestedness matrix for the butterflies of the western Mediterranean shows that the communities are significantly nested but with notable differences from a perfect pattern (fig. 12). The large occurrence of turnover suggests that complex eco-evolutionary processes operate in assembling the current island communities in this region.

These complex patterns of community structures on islands have profound implications for conservation biogeography. In fact, the particular composition of species for each island cannot be predicted solely based on the number of species occurring there. Under this scenario, each species on each island is the outcome of a unique story due to a combination of various influences, such as different sources, speciation events, presence of relicts, particular habitat structure, etc. All these factors are determining the presence or absence of a species, which

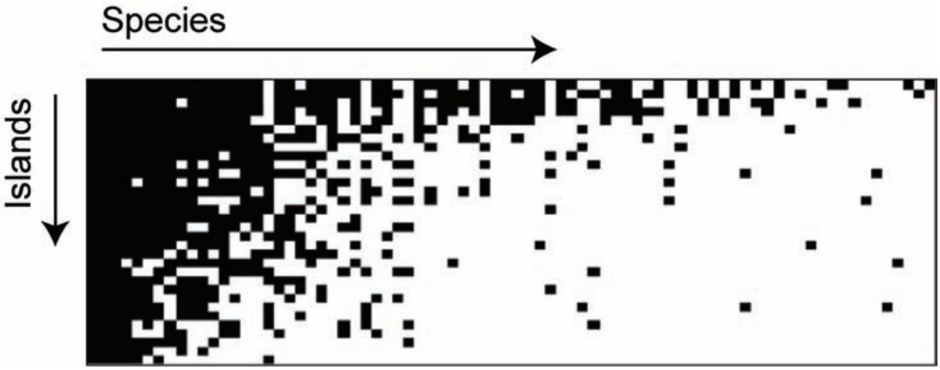


Figure 12 The packed matrix of species occurrence for all the western Mediterranean islands, where the species (columns) are ordered from the most frequent (left) to the least frequent (right) and the islands (rows) are ordered from the richest (top) to the poorest (bottom).

otherwise would not be expected to occur based only on its dispersal capabilities. A weak nested pattern indicates that island communities possess individual stories and protecting only the richest areas is not enough to preserve all the regional diversity since even the smallest community, such as the one on I3 in fig.11, may host fundamental fractions of diversity.

Indices of faunistic dissimilarity between communities

In order to evaluate how much two communities differ, one has to measure their degree of diversification by using specific indices. A diversity index is a numerical measurement of various aspects of the dissimilarities occurring in species composition.

Many such indices have been described in the last century and some of them are still widely used today because of their simple formulation and desirable algebraic properties (Koleff et al. 2003, Baselga 2010, Anderson 2011). Jaccard (1901) and Sørensen (1948) described two very similar and monotonically correlated indices that measure the overall degree of dissimilarity among communities, using the fraction of unshared species between two communities over the total diversity occurring there:

eq. 3 $$J = \frac{b+c}{a+b+c}$$

eq. 4 $$Sor = \frac{b+c}{2a+b+c}$$

where a is the number of shared species between two sites, and b and c are the species unique to the first and the second site, respectively.

Simpson (1960) provided a slightly modified index that measures the dissimilarity between communities based only on the species responsible for turnover:

eq. 5 $$Sim = \frac{\min(b,c)}{a+\min(b,c)}$$

where min (b,c) represents the minimum number of species unique to the first or to the second site. There has been much debate over which index is appropriate in specific contexts and it is important to have a clear understanding of what each of them measures (Morris et al. 2014). The methodological discussion of the meaning and appropriateness of community dissimilarity indices to expose different processes has been recently renewed by Baselga (2010). He stated that the overall diversity between two communities (measured by the Sørensen index) can be partitioned into two additive components: one derived only from differences due to species replacement (the Simpson index) and the other index, newly introduced, obtained from differences in nestedness (fig. 13). This study represented a turning point for community ecology, becoming one of the most cited articles in this field of study in the last five years (more than 250 citations recorded in Web Of Science). The possibility to separate the effects of phenomena occurring in ecological time, mostly producing contrasts in the nestedness component, from the phenomena occurring over historical time and producing species replacement, allowed the recognition of fundamental processes involved in the formation and maintenance of the community structure. In particular, the use of turnover and/or nestedness indices allowed to: delineate zoogeographic regions at global and local scales (Kraft & Jetz 2010, Svenning et al. 2011, Holt et al. 2013, Dapporto et al. 2013, 2015), recognize climatic, historical and neutral processes as drivers for regional diversity (Keil et al. 2012, Svenning et al. 2011, Dobrovolski et al. 2012), identify climate change fingerprints in diversity loss and faunal homogenization (Davey et al. 2013, Carvalheiro et al. 2013).

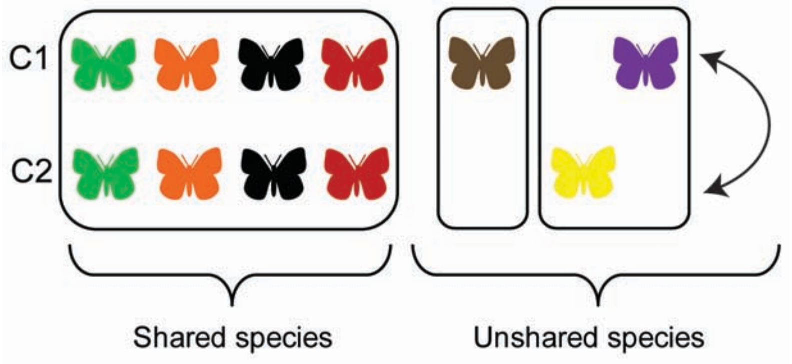


Figure 13 A graphical explanation of the possibility to partition the nestedness and turnover components of faunal dissimilarity between two communities, C1 and C2, using an additive approach. The four shared species do not contribute to the overall dissimilarity, which is only encompassed by the three unshared species on the right (brown, yellow and purple). The brown species is generating difference due to nestedness, while the yellow and purple species replace each other, thus producing the turnover component.

DNA barcoding for comparative phylogeography

Mitochondrial DNA (mtDNA) has been used in molecular studies for over four decades, proving to be a very effective marker to assess relationships among species that diverged relatively recently (Brown et al. 1979). MtDNA also helps accelerating the discovery and description of global biodiversity, particularly important in a time when about 80% of the species on earth are still undescribed and species extinction might outpace their identification and description (Pimm et al. 1995, 2014, Carlton et al. 1999, Barnosky et al. 2011, Bucklin et al. 2011).

The DNA barcoding technique was first proposed by Hebert et al. (2003a, 2003b) as "*a reliable, cost-effective and accessible solution to the current problem of species identification*". This method uses a short, standardized and highly variable DNA sequence acting as a species tag or 'barcode' for the species-level identification of specimens (Hebert et al. 2003a, 2003b). DNA barcoding in many cases allows for easier discrimination among similar species regardless of life stage (Godfray & Knapp 2004, Smith et al. 2006, Dincă et al. 2011a, 2011b, 2015) and thus accelerates their identification and description (Scheffers et al. 2012). In the animal kingdom a mtDNA fragment of aprox. 650 bp belonging to the cytochrome c oxidase subunit I (COI) represents the barcoding region (Hebert et al. 2003a, Savolainen et al. 2005). Another advantage associated with DNA barcoding was the establishment of a comprehensive reference library of barcode sequences for known species, where precise identifications can be made by comparing the DNA sequence obtained from any sample to a database that includes all the species from a given region (Hebert et al. 2003a, 2003b) (fig. 14). The Barcode of Life Data System (BOLD) is the world's largest repository that stores, analyses and publishes DNA barcode records, complemented by collateral data, such as specimen images, GPS coordinates and collecting information (Ratnasingham & Hebert 2007). Currently in BOLD there are over 500 000 species and 5 million specimens barcoded, representing about 5 to 10% of all the multicellular species on our planet.

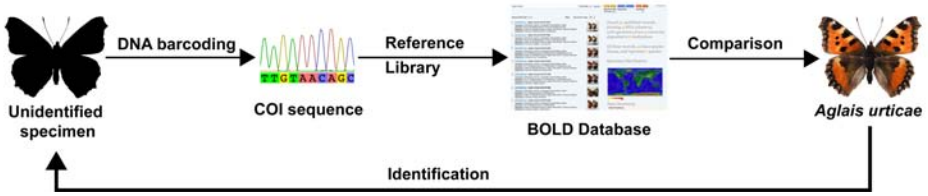


Figure 14 DNA barcoding aids a fast and reliable identification of unknown specimens by sequencing its mtDNA and comparing the sequence with other prior identified species existent in a reference library.

The effectiveness of DNA barcoding has been tested for various animal groups and different geographic regions: birds (e.g. Hebert et al. 2004a, Kerr et al. 2007), Lepidoptera (e.g. Hebert et al. 2004b, Janzen et al. 2005, Hajbabaei et al. 2006, Dincă et al. 2011c, 2015), fishes (e.g. Ward et al. 2005), mammals (e.g. Clare et al. 2007). The success rate for species identification varies between 70 % and 94%, strongly depending on the studied group and geographic coverage (Hebert et al. 2004a, Ward et al. 2005, Meyer & Paulay 2005, Vences et al. 2005, Hajibabaei et al. 2006, Elias et al. 2007, Wiemers & Fiedler 2007, Whitworth et al. 2007, Shearer & Coffroth 2008, Spooner 2009, Dincă et al. 2011c, 2015). For butterflies, DNA barcoding proved to be a very effective method since most barcoded specimens can be identified to a species or at least to a species-pair level (Lukhtanov et al. 2009, Dincă et al. 2011c, 2015, Wilson et al. 2013).

Barcodes are also used to reliably and rapidly identify a great variety of sample types (eggs, juvenile stages, fragments, semi-digested parts, etc.) (Jurado-Rivera et al. 2009), to distinguish males and females belonging to sexually dimorphic species (Janzen et al. 2005), to detect the presence of invasive species, and it has applications in forensics, food safety, or illegal animal traffic surveillance. Moreover, DNA barcoding is often used to detect possible cryptic species (Hebert et al. 2004b, Janzen et al. 2011, 2012, Chacon et al. 2013, Dincă et al. 2015).

However, DNA barcoding is still a debated technique (Mallet & Willmott 2003, Janzen 2004, Moritz & Cicero 2004, Will et al. 2005, Hebert & Gregory 2005, Smith 2005, Rubinoff 2006, DeSalle 2006, Hajibabaei et al. 2007, Collins & Cruickshank 2013, Kvist 2013), although most of the controversy stems from the mix-up of two different approaches simultaneously proposed: DNA-barcoding, which is mainly used for species identification (Hebert et al. 2003a) and DNA-taxonomy, which advocates for the use of DNA data to describe and delimit species (Tautz et al. 2003). Currently, the later approach is not accepted as a reliable means for species delimitation without being complemented by other methods (such as nuDNA sequencing, morphometry, karyology). Although DNA barcoding alone is not sufficient to verify and/or establish the specific status of taxa, it is a great tool to assess the variability of the mitochondrial genome, which typically displays similar phylogeographic patterns to the nuclear genome (Brower & deSalle 1998).

The controversy, but also the many studies that are using DNA barcoding led to a rapid increase in the number of publications as shown by the graphics in fig. 15.

DNA barcoding is widely used for phylogenetic inference and, although applied alone is not reliable to reconstruct deep phylogenetic relationships, it is useful to place unknown specimens into a broader taxonomic framework or to detect deeply diverged lineages within a known taxon that deserve more in depth studies.

Similarity among DNA sequences provides certain information about their relative divergence time, however it would be more reliable to obtain absolute divergence time values by applying what is known as a 'molecular clock' (Hillis et al. 1996, Li 1997, Page & Holmes 1998). This concept, first introduced in the 1960s (Zuckerkandl & Pauling 1962), assumed that molecular changes occur at a constant rate which allowed to date unknown divergence time among species just by comparing their DNA sequences.

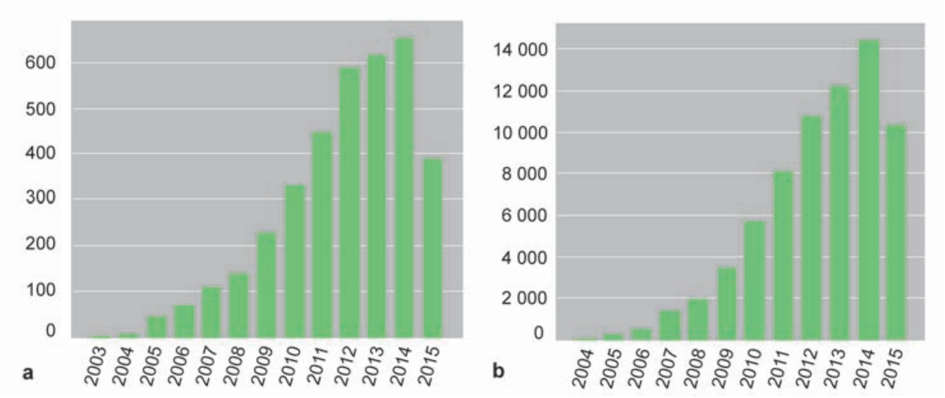


Figure 15 Publications (a) and citations (b) of articles dealing with DNA barcoding [downloaded from Web of Science on 6 September 2015]

However, the idea of using a strict molecular clock in phylogenies was criticised (Simpson 1964) and finally rejected (Kummar 2005). The subsequent development of relaxed molecular clock models that allow variation of evolutionary rates through time became an essential tool and set the basis for modern phylogenetics.

One of the key features of DNA barcoding for phylogeography is that it reflects the history of the maternal line (as the mitochondrial DNA is maternally inherited). Also callibration of the molecular clock is feasible through a wide array of estimated substitution rates for COI that have been published (e.g. Quek et al. 2004).

On the other hand, in addition to physical isolation, phylogeographical patterns may be affected by the history of infections by the endosymbiont *Wolbachia*. This bacterium is maternally inherited and can cause male killings or cytoplasmatic incompatibility. Knowing about the infection status and history of a species may however be interesting for conservation purposes or reintroduction programs, as the mixing of lineages with and without

Wolbachia may result in undesirable results and even negatively affect the populations.

Box 2 Reasons for using DNA barcoding

- It can identify fragments of organisms (e.g. useful for detecting undesirable material in processed foods);
- It can identify specimens at any stage of their life cycle (e.g. identifying eggs to species);
- It can distinguish among species that are morphologically very similar;
- It can accelerate the discovery of yet undescribed species;
- It adds a new value to museum collections.

Phylogeography: analysing distribution patterns of genetic lineages

Studies of dissimilarity indices for faunistic data can reveal many eco-evolutionary processes determining the composition of communities at the specific level (Svenning et al. 2011, Keil et al. 2012, Dobrovolski et al. 2012, Davey 2013, Dapporto et al. 2015). However, regional biodiversity is also driven by processes that are operating at the intraspecific level, such as local competition and adaptation, hybridization among lineages, appearance of reproductive barriers and dispersal or extirpation of populations (Ricklefs 2008, Marske et al. 2013, Mittelbach & Schemske 2015).

Genetic lineages evolve in time and space, and they disperse or remain isolated as a consequence of climatic or environmental changes, such as temporary connections between mainland and islands or the emergence of ecological barriers, such as glaciated areas (Hewitt 2001, Schmitt, 2007, Esselstyn & Brown 2009). Phylogeography, described as the bridge between intra- and interspecific patterns, is the discipline aiming to reconstruct the spatio-temporal evolution and dynamics of genetic lineages (Avise et al. 1987). This discipline focuses on relatively recent processes (usually Quaternary) leading to lineage diversification and to their interaction with a changing environment (Ricklefs 2008, Hickerson al. 2010, Marske et al. 2013). Ultimately, phylogeography compares the phylogenetic relationships among different lineages and their distributions in order to reconstruct their history and establish: i) where the populations remained isolated and evolved into different lineages, ii) where they dispersed and also iii) where they likely went extinct (Avise et al. 1987, Hickerson et al. 2010, Marske et al. 2013). In the last years, phylogeography has become one of the most integrative fields in biology, being rooted in population genetics, ecology, climatology, geospatial analyses and computer science (Hickerson et al. 2010). Molecular advances in recent decades allowed a better understanding of phylogeographic processes, which, in turn, facilitate the prediction of the future distributions of organisms and the development of conservation strategies (Thuiller et al. 2011, Ladle & Whittaker 2011, Barnagaud et al. 2014).

As mentioned in the previous chapter, mitochondrial DNA (mtDNA) has become in the last decades one of the most used markers for molecular assessments (Brown et al. 1979). The mtDNA is particularly suitable for phylogeography since it generally accumulates nucleotide mutations much faster than nuclear DNA (nuDNA), thus providing a highly sensitive tool for tracing the emergence and distribution of lineages in time (Avise 2009). Classic analytical procedures to display evolutionary relationships involve phylogenetic trees and the reconstruction of phylograms or haplotype networks, which represent genealogical relationships and the number of mutational steps among haplotypes (Templeton 1992, Posada & Crandall 2001). Projecting the occurrence of mutations in phylogenetic representations over space and comparing the observed patterns with paleogeographic reconstructions, allows us to understand the history of isolation, divergence and colonization routes of different genetic lineages (Avise 2000).

However, there are several phenomena that mtDNA alone cannot reveal: introgression, hybridization, incomplete lineage sorting, and most of the adaptive processes (Ebach & Holdrege 2005, Will et al. 2005, Prendini 2005, Cameron et al. 2006, Rubinoff et al. 2006, Currat et al. 2008, Excoffier et al. 2009, Toews et al. 2014). Therefore, other markers are needed to improve the power of molecular data to test phylogenetic and phylogeographic hypotheses, such as allozymes, nuDNA and phenotypic markers, mostly the ones linked to highly adaptive characters (mimicry, wing size and shape, in the case of butterflies) or to structures involved in sexual selection (for butterflies, genitalia and pheromones) (Schmitt 2007, Towes & Brelsford 2012, Dapporto et al. 2012, Hernández-Roldán et al. 2014, *Heliconius* Genome Consortium 2012). While genetic patterns reveal evolutionary relationships, hybridization processes and approximate divergence time, phenotypic traits represent the interface between genes and the environment and include key morphological traits for the interactions among individuals (revised in Habel et al. 2015). Another growing approach in phylogeography is to combine genetic data with ecological niche modelling (ENM), in order to understand if different genetic lineages live under different ecological conditions (usually climate) to which they adapted and that can be responsible for the maintenance of current distribution limits (Habel et al. 2010, Alvarado-Serrano & Knowles 2014). The most advanced phylogeographic assessments usually combine phylogenies for different markers, climatic and paleoclimatic modelling and paleogeographic reconstructions (for butterflies: Nazari et al. 2009, Dincă et al. 2011a, Hernández-Roldán et al. 2011, Sañudo-Restrepo et al. 2013, Kodandaramaiah et al. 2013, Zinetti et al. 2013, Toth et al. 2014).

Traditionally, Europe has been one of the most studied regions for the effects of paleoclimatic and paleogeographic events in shaping the genetic structure of the populations (Hewitt 1999, 2011). Butterflies make no exception and there are many studies underlying the importance of the supposed retraction and isolation of most species in the three southern European peninsulas and of the effect of postglacial expansion in determining the observed genetic structures (Schmitt 2007, Habel et al. 2010, 2011, Zinetti et al. 2013). The Mediterranean butterflies, in particular, have an astonishing intraspecific diversity that was built up during a complex history of paleogeographic and climatic events (Dennis et al. 1991, Schmitt 2007, Nazari et al. 2010, Dincă et al. 2011a). This is particularly evident in the western Mediterranean islands, where many lineages are endemic or show idiosyncratic distributions on both islands and mainland (Cesaroni et al. 1994, Dincă et al. 2011a, 2015, Sañudo-Restrepo et al. 2013, Zinetti et al. 2013).

However, the typical paradigms identifying the three main peninsulas as different glacial refugia (Hewitt 1999) rarely apply to butterflies, since the distribution patterns look much more complex and variegated (Dapporto et al. 2011b, Dincă et al. 2011a, Habel et al. 2011). There is strong evidence that other regions played a fundamental role in hosting butterfly populations during the glaciation cycles, prominently the Maghreb (Husemann et al. 2013), and the large islands: Corsica, Sardinia and Sicily (Dapporto 2010). After a series of studies using allozymes as molecular markers, Habel et al. 2005, 2010 described a fourth paradigm that provides evidence for high cohesiveness between the Maghreb, Sicily and the Italian Peninsula. On the other hand, Dapporto et al. (2012) found a different recurrent pattern in the shape of the male genitalia of several Satyrinae butterflies, showing that populations from the Italian Peninsula are highly similar to those from the Balkans, while the surrounding smaller (mostly Tuscan Archipelago) and larger islands (Sardinia, Corsica and Sicily) tend to host populations resembling those from Maghreb and Iberia. In general, molecular (mtDNA and nuDNA) and morphological markers concur for a strong separation between Europe and Maghreb, in large disagreement with the so-called butterfly paradigm (Weingartner et al. 2006, Nazari et al. 2010, Dincă et al. 2011a, Sañudo-Restrepo et al. 2013).

Some species are becoming models for phylogeography in Europe and the Mediterranean region and they have been studied independently for different markers by different authors (*Melanargia* spp. Nazari et al. 2010, Habel et al. 2011; *Polyommatus icarus-P. celina* Schmitt et al. 2003, Habel et al. 2010, Wiemers et al. 2010, Dincă et al. 2011a, Habel et al. 2010; *Maniola spp*. Thomson 1987, Dapporto et al. 2009, Dapporto 2010, Grill et al. 2006, Habel et al. 2010, Kreuzinger et al. 2015; *Leptidea spp*. Verovnik & Glogovcan 2007, Friberg & Wiklund 2009, Dincă et al. 2011b, 2013). In some cases the interest was due to the discovery of previously unknown cryptic entities displaying an idiosyncratic distribution, while in others it was due to the presence of genetic lineages showing clear morphologic variability, with the presence of intermediate individuals at the contact zone (Thomson 1987, Dapporto et al. 2009, Wiemers et al. 2010, Dincă et al. 2011a). Among these models, a few taxa have been used with particularly high frequency as models for the western Mediterranean area, including islands, and showed highly contrasting distribution even along very narrow sea straits: *Maniola jurtina*, *Polyommatus Icarus, P. celina*, *Aricia agestis, A. cramera*. All these cases were specifically studied in this thesis.

Maniola jurtina **(family Nymphalidae)**

The meadow brown butterfly *Maniola jurtina* (Linnaeus, 1758) forms conspicuous populations in the Mediterranean basin, including many islands (Schmitt et al. 2005b, Habel et al. 2009, Dapporto et al. 2009). Based on male genitalia and allozyme analyses there are two lineages that occur throughout Europe (Thomson 1987, 2011): a western Atlantic-Mediterranean lineage (*M. jurtina jurtina*) that occurs in the Maghreb, Spain, western France, Sicily and Sardinia; and an eastern Asian-Mediterranean lineage (*M. jurtina janira*) widespread from Asia to eastern and central Europe, including the Italian Peninsula (Thomson 1987, Dapporto et al. 2009, Dapporto & Bruschini 2012) (fig. 16).
There is also a contact zone between the two lineages extending from a few western Mediterranean islands (Corsica, Elba, Giglio, Capri) to the western Alps and the Benelux region (Thomson 1987, 2011, Schmitt et al. 2005b, Dapporto et al. 2009, Dapporto & Bruschini 2012).

Figure 16
Distribution of the two genetic lineages of *M. jurtina*. The orange band represents the areas where the two lineages overlap.

By modelling the morphological variation over the western Mediterranean area, Dapporto et al. (2009) and Dapporto & Bruschini (2012) suggested that the western morphotype was once widespread in the whole western Mediterranean, both on mainland and islands, including the Italian Peninsula. It is likely that *M. j. janira* expanded from eastern Europe and invaded the Italian Peninsula, replacing the original *M. j. jurtina* during one of the post-glacial periods in the last 7000 years (Dapporto et al. 2009, 2011b). But on several islands, such as Corsica, Elba, Giglio, Capri and Sicily, the supposed ancestral *M. j. jurtina* could have persisted even if narrow sea straits separates them from mainland (Dapporto et al. 2009). On the other hand, *M. j. janira* seems to be the ancestral morphotype in the Balkans, since it exclusively occurs on the islands surrounding the Balkan peninsula (Thomson 1987, Dapporto et al. 2012).

A comparative study of all the butterfly species in the Satyrinae subfamily revealed that this pattern of distribution of morphotypes is highly recurrent in the entire subfamily (Dapporto et al. 2012). This highly dynamic phylogeographic process could be at the basis of several observed contrasts between islands and mainland populations and can generate phenomena of paleoendemism, even on fragment islands where the evolution of new species is usually a rare phenomenon (Whittaker & Fernández-Palacios 2010).

Aricia agestis – A. cramera (family Lycaenidae)

Like *M. jurtina*, some *Aricia* species represent a constant presence over the western Mediterranean area. The taxonomy of the *Aricia* genus is still debated due to the occurrence of cryptic taxa that are regarded with different taxonomic status by different authors (Aagaard et al. 2002, Tolman & Lewington 2008,

Mallet et al. 2011, Tshikolovets 2011). A recent study argued that the western Mediterranean area hosts five species, of which two – *A. agestis* (Denis & Schiffermüller, 1775) and *A. cramera* (Eschscholtz, 1821) also occur on several Mediterranean islands and display a parapatric distribution with a narrow contact zone in north-eastern Spain (Catalonia) (Sañudo-Restrepo et al. 2013) (fig. 17). Interestingly, the two species were never found to co-exist on the same island, although some of the islands, like Corsica and Sardinia, are separated by sea straits of just a few kilometres (Sañudo-Restrepo et al. 2013).
Although only a few specimens were sequenced for each island, this distribution suggested the existence of a mutual exclusion pattern. Since these species have colonised many Mediterranean islands it looks improbable that the narrow sea strait between Sardinia and Corsica can represent an insurmountable barrier for their distribution. For this reason the two species are a perfect model to study the processes establishing and maintaining differences among communities. The distribution patterns are even more striking considering that Corsica and Sardinia were connected for most of the Pleistocene due to the lowering of the sea level (Lambeck & Purcell 2005).

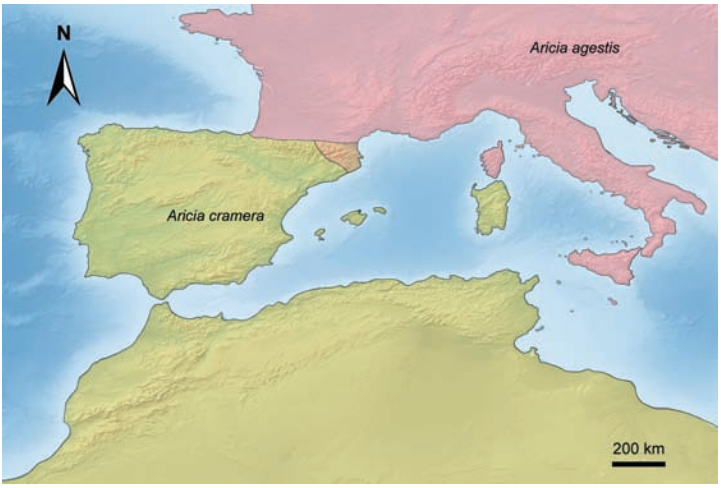


Figure 17
Distribution of the crypric species *A. agestis* and *A. cramera*. The orange band represents the contact area of the two taxa.

Polyommatus icarus – P. celina (family Lycaenidae)

The closely related species *Polyommatus icarus* (Rottemburg, 1775) and *P. celina* (Austaut, 1879) are also common in the western Mediterranean mainland and islands. Their taxonomic status was relatively recently resolved and it was shown that *P. celina* replaces *P. icarus* in North Africa (Wiemers 2003, Wiemers & Fiedler 2007, Vodolazhsky & Stradomsky 2008, Wiemers et al. 2010). Another study (Dincă et al. 2011a) proved that *P. celina* occurs not only in North Africa but also in Europe, including on several western Mediterranean islands, as well as large parts of Iberia, where it occurs in parapatry with *P. icarus* (fig. 18). The two species also exclude each other on islands, even across narrow sea straits like those separating Corsica from Sardinia, and Sicily from the Italian Peninsula (Dincă et al. 2011a). Moreover, in the Balearic Islands and Sardinia, on one hand, and in Sicily and Lipari on the other, strongly diverged lineages of *P. celina* occur, which are likely to represent relict populations produced by past range expansions and contractions (Dincă et al. 2011a). This group represents a great model system for examining patterns of mutual exclusion on western Mediterranean islands, including patterns occurring at intraspecific level.

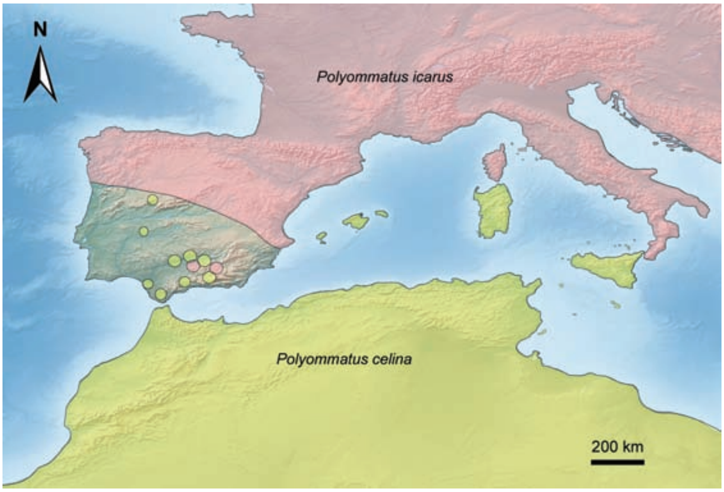


Figure 18 The distribution of *P. icarus* and *P. celina* in the western Mediterranean region.

1.3 The new era of island biogeography

For the last two centuries islands have provided some of the most important insights into the understanding and development of general ecological and evolutionary processes (Vitousek et al. 1995, Lomolino 2000, Fernández-Palacios et al. 2015, Warren et al. 2015). They have been rightfully considered as natural laboratories, relatively simplified models of our complex natural world that allow general theories to be developed and tested (Wittaker & Fernández-Palacios 2007).

Islands represented the starting point for Darwin and, independently, Wallace to decipher and formulate the theory of evolution (Darwin 1859, Wallace 1855, 1880), and decades later they inspired MacArthur and Wilson to develop the ETIB, one of the most important theories in the field of biogeography (MacArthur & Wilson 1963).

But the establishment of a solid framework for island research was a slow process. For more than a century after Darwin and Wallace introduced the theory of evolution by natural selection, eco-evolutionary, as well as biogeographic processes, were mostly descriptive. Decades later, in the 1960s, several major biogeographical concepts based on island research were formulated. Among them, the island theory of MacArthur and Wilson added the mathematical premises for understanding community ecology. However, for the next almost 40 years, the ETIB was one of the most debated theories in the field, mainly because, unlike what the theory states, most communities appear to be far from their equilibrium point (revised in Heaney, 2000, 2007, Lomolino 2000). Even though island studies provided a better understanding of their communities and of those occurring on fragmented mainland environments, in the last four decades an unbalanced amount of efforts have been devoted to testing the ETIB validity and generality rather than to the development of new conceptual frameworks (Fernández-Palacios et al. 2015a, Warren et al. 2015).

There are at least four shortcomings that prevented research on islands to reach its full potential and, based on Kueffer et al. 2014, they are the following:

i) Most studies rely on a few islands and taxonomic groups, which hampered researchers generalize their findings and formulate theories;
ii) There are few types of data to work with (e.g. species number on individual islands);
iii) Molecular, functional, community and paleoecological traits are usually not incorporated into island research. Combining such data would probably make possible the reconstruction of the evolution of islands population from the earliest phases of island colonization to the emergence of neo- and paleo-endemic species;
iv) There is a generalized lack of connection between research and applied conservation. As an example, conservation should evaluate the effects of invading species and lineages on autochthonous island populations and recent molecular advances could help in precisely identifying target species for preserving island uniqueness (cryptic endemics, diverging genetic lineages, single genes evolved for local adaptation).

But new theoretical and technological advances are changing the rules of the game and *"… islands are reclaiming a leading role in ecology and evolution, especially for synthetic studies at the intersection of macroecology, evolution, community ecology and applied ecology"* (Kueffer et al. 2014). We are now living in a new 'golden' era of island biogeography in which we have the means to re-assess fundamental issues in ecological and evolutionary theory (Fernández-Palacios et al. 2015a). Three major developments are the new trendsetters in this field:

i) An increase in data availability in fields such as geology, paleoecology, phylogenetics and community ecology.
   Geosciences are benefiting from new techniques that boost our knowledge of the oceans and seas: multi-beam scans of the ocean/sea floors, improved bathymetric maps, or revolutionary methodologies for dating island rocks and for reconstructing past climatic events, are just a few of the new developments (revised in Fernández-Palacios et al. 2015a). More precise descriptions of geological features and reconstructions of paleogeographic and

paleoecological settings are giving more insights into the effects of current and past events on the eco-evolutionary histories of island populations (Fernández-Palacios et al. 2015b). New studies in the field of paleoecology with great implications for conservation are providing increasing evidence of the high dynamics of island ecosystems before they were colonized by humans (Kueffer et al. 2014).

Phylogenetics, but also population genetics and genomics are advancing mostly due to less costly DNA sequencing and to a parallel increase in computational power.

And lastly, community ecology, macroecology and conservation biology are benefitting from the integration of data from all of the above-mentioned fields of research.

ii) An increase in studies that are covering more taxonomic groups and islands. Phylogeographic studies that analyse entire taxonomic groups and large areas have become more and more common mainly due to the increased availability of data (Carlquist et al. 2003, Caujapé-Castells 2013, Dincă et al. 2011b, 2015). Another positive point is that it is mandatory for researchers who use molecular data in their published articles to make them available in open public repositories (e.g. GenBank, BOLD). This does not only apply to genetic data since repositories for other kind of information exist (e.g. DRYAD, GBIF) and many journals are now explicitly requesting for the primary data to be made available as supplementary appendices. In this way other researchers can improve their datasets and compare patterns for different taxonomic groups and different areas at a much larger scale (e.g. Santons et al. 2010 for island communities).

iii) An increase in methodological and theoretical advances as a consequence of the first two types of development. The ever-increasing quantity and quality of data are challenging researchers to develop new methodologies for handling them. We now have the means to integrate large amounts of data: faunistic (new algorithms for large and small scale regionalization, Holt et al. 2013, Dapporto et al. 2015), climatic (climatic layers for species distribution modelling and software such as BIOMOD, Thuiller 2003 or MaxEnt, Phillips 2008), geologic-paleogeographic (e.g. explicit glacial-sensitive model, Fernández-Palacios et al. 2015b), genetic (e.g. genetic landscapes, Male et al. 2003) and phenotypic traits (e.g. projecting geometric morphometrics in a spatial context, Bookstein 1997, Dapporto et al. 2012).

This progress is also providing a great opportunity to study more in depth a series of particular and still poorly understood phenomena for the western Mediterranean butterflies. Some key aspects that could shed more light on eco-evolutionary processes mainly concern the relationship between island and mainland species and communities. The usage of different markers to explain spatial patterns of genetic and phenotypic characteristics, and the understanding of the mutual exclusive pattern of cryptic species and lineages, are just a few of the most burning issues.

Faunistic and genetic relationship between island and mainland species

Until recently, phylogeographic studies of island biotas were focused mostly on the islands of interest without reference to their continental counterparts. However, taking into account both island and mainland taxa could explain patterns of species assemblages and the establishment of island communities, detect how colonization events took place and where they originated.

The Mediterranean islands are continental and fragment islands (Fernández-Palacios 2007) since they are located relatively close to the mainland from which they originated. In this case, for highly dispersive taxa such as some species of butterflies, it can be postulated that a relatively constant exchange of individuals between islands and the nearest mainland can be maintained (Dennis et al. 2000, Dapporto & Dennis 2008). Indeed, in several studies, the effect of the closest mainland source in determining community and genetic structures in the Mediterranean explained most of the variance in island characteristics (e.g. species richness and occurrence, Dapporto & Cini 2007, Dapporto & Dennis 2008). In other studies however, the comparison between island and mainland populations showed unexpected contrasts and their distribution patterns led to the inference of complex and still undescribed colonization routes (Dincă et al. 2011, Dapporto et al. 2011).

Analysing both island and mainland taxa is necessary to explain patterns of species assemblages and the establishment of island communities, detect how, when and from where colonization events took place and ultimately understand if some island populations can be regarded as paleo- or neo-endemics, with profound implications in biological conservation.

In **Chapter I** of Results we investigate patterns of the butterfly beta-diversity in the western Mediterranean mainland and islands. We also test the power of various dissimilarity indices to explain different characteristics of insular and mainland communities and reconstruct the main dynamics determining their composition. In **Chapter V** of Results we use comparative analyses linking community and phylogenetic approaches for the butterfly fauna of the circum-Sicilian islands to understand the mechanisms determining and maintaining species assemblages on islands.

Combining different markers to explain patterns of genetic structure and distribution

Different markers (e.g. genetic, morphologic, climatic, etc.) provide different insights into the complex story of how species originated, evolved and expanded in search of suitable habitats. In most cases markers generally agree, but they can also show contradictory results.

Discordance among markers can be common in western Mediterranean butterflies (see Dapporto et al. 2011 for a review and a possible explanation). This could be a consequence of the particular geographic and climatological history of the study area, but also of the characteristics of the model group.

The Mediterranean region has provided rather stable habitats for most of the current temperate species during the strong Quaternary climatic oscillations (Hewitt 1999, Schmitt 2007). The structure of the Mediterranean area, formed by two separated continents and large islands (Africa, Europe, Corsica, Sardinia, Sicily), and the presence of three large peninsulas (Iberian, Italian, Balkan), facilitated the emergence of different lineages and morphotypes, as well as the evolution of local adaptations to particular climatic settings (Hewitt 1999, Schmitt 2007). Butterflies are known to be some of the fastest organisms to track suitable habitats, especially as a consequence of climate change (Devictor et al. 2012). Lineages and morphotypes emerged in allopatry during the long glacial periods can rapidly disperse in new territories when conditions are favourable and thus distributional shifts often occur in the case of butterflies. When various species get into secondary contact, introgression is likely to occur and in many cases it can be highly asymmetrical (Currat et al. 2008). Many mechanisms like differential adaptive value of different traits (The *Heliconius* Genome Consortium 2012, Toews et al. 2014), male or female biased dispersal (e.g. Kostro-Ambroziak & Rutkowsky 2011 for butterflies) and gene surfing (Lehe et al. 2012) can determine asymmetrical introgression and this can generate strong discrepancies, notably between mitochondrial and nuclear genes (Toews & Brelsford 2012).

Increasing availability of a large amount of data for different markers is facilitating comparative studies that can lead to the reconstruction of the complex patterns of contraction and expansion of butterflies during the geological history of the Mediterranean and ultimately to the understanding of the astonishing endemic diversity occurring in this region.

In **Chapter IV** of Results we provide a new method to directly compare the spatial dissimilarity patterns of two or more markers of any type, including genetic and morphologic, even if data available for each marker only partially overlap.

Understanding the patterns of mutual exclusion among cryptic species and lineages

If discrepancies among markers are mainly generated by introgression when lineages and species overlap over wide or narrow areas of secondary contact (Currat et al. 2008, Toews & Brelsford 2012), there is also growing evidence that genetic lineages and cryptic species are often displaying idiosyncratic distributions with unexpectedly narrow contact zones, mostly over sea straits (Dincă et al. 2011a, Waters et al. 2011, Sañudo-Restrepo et al. 2012). Recent molecular advances have fostered the discovery of cryptic species from different taxonomic groups and regions and have also led to the discovery that mutual exclusive patterns of distribution are particularly frequent in cryptic groups (Dincă et al. 2011a, Dincă et al. 2011b, Sañudo-Restrepo et al. 2012, Dapporto 2010). Many phenomena can concur in generating mutually exclusive patterns of distribution by limiting secondary contacts like founder effects followed by density dependent processes, reproductive interference, competitive exclusion and climatic-environmental preferences (revised in Waters et al. 2011, 2013). Mutually exclusive patterns are expected to generate regional diversity through the establishment and maintenance of species turnover, a fundamental component of beta-diversity. Disentangling such a large spectrum of potential determinants requires integrative multidisciplinary approaches that combine several types of markers: genetic, faunistic, climatic, etc.

In **Chapter II** of Results we analyse the patterns of distribution of the cryptic and non-cryptic fractions of the butterfly species occurring in the western Mediterranean islands and mainland, while in **Chapter III** we use genetic and presence data, haplotype networks and ecological niche modelling analyses for two pairs of cryptic species to specifically test for the factors that determine chequered patterns of distribution.

2. Objectives

2. Objectives

Butterflies are not just the ‘pretty faces’ of the insect realm, they have countless and diverse attributes that make them one of the most studied groups of invertebrates in the world. The amount and quality of data regarding their biology, ecology, taxonomy and evolution has currently reached an unprecedented level, which led to answering many questions that could not have been addressed before. However, as it has been shown in the Introduction, there are still many gaps that need to be filled and the research presented in this thesis focused on the unknown aspects of the butterfly fauna occurring in the western Mediterranean. The main aim was to have a better understanding of the processes that are governing butterfly diversity and distribution at both specific and intraspecific levels in one of the 35 biodiversity hotspots around the world.

- Although comprehensive museum sample collections and precise distribution maps exist for Mediterranean butterflies, there are hardly any collections designated to preserve material for high quality standard DNA analyses. Moreover, presence data for island communities are insufficient and highly dispersed. Given the scarcity of such data for the western Mediterranean butterflies, the first objective is to **gather a unique, extensive collection** consisting of faunistic data, DNA sequences and tissue samples for future research, complemented by photographs of the specimens (Results Chapter I, II, III, IV, V).

- Since a wide coverage of presence records for the western Mediterranean butterflies enables studies of distribution patterns of diversity, the second objective is to **investigate butterfly beta-diversity in the western Mediterranean mainland and islands**, recognize zoogeographic regions and test the power of different indices of community composition dissimilarity (Results Chapter I).

- Do cryptic species have particular properties? The recent increase of presence and genetic data allowed several studies to report unexpected levels of cryptic diversity for various taxa and geographic regions. However, it is not known whether cryptic species represent a homogeneous subset of biodiversity or they display differential characteristics. The third objective is to **evaluate the extent of cryptic species and analyse their patterns of distribution in the western Mediterranean islands and mainland** (Results Chapter II).

- It has been shown that several cryptic species of butterflies display parapatric distribution on mainland and they rarely, if ever, co-exist on islands. However, the mechanisms that are responsible for such patterns are still not well understood. The fourth objective is to **study potential determinants behind the observed spatial patterns of two pairs of cryptic species** by using a multidisciplinary approach (Results Chapter III).

- Different markers used for assessments of community traits can have complementary or contrasting results that uncover fundamental eco-evolutionary processes. However, a direct comparison among them is under some circumstances methodologically not possible. The fifth objective is to **develop and test new algorithms and procedures to compare spatial patterns of different markers even if they have different sample sizes** (Results Chapter IV).

- Because of the complexity of island systems, the mechanisms that determine and maintain their species composition are poorly understood. The sixth objective is to **examine the assembly rules that generated the observed communities on the circum-Sicilian islands and test for idiosyncratic processes leading to unique butterfly assemblages on island** (Results Chapter V).

2. Results

Chapter 1

Dapporto L., Fattorini S., Vodă R., Dincă V. & Vila R. 2014. Biogeography of western Mediterranean butterflies: combining turnover and nestedness components of faunal dissimilarity. *Journal of Biogeography,* **41**: 1639–1650.

Journal of Biogeography (*J. Biogeogr.*) (2014) **41**, 1639–1650

ORIGINAL ARTICLE

Biogeography of western Mediterranean butterflies: combining turnover and nestedness components of faunal dissimilarity

Leonardo Dapporto[1], Simone Fattorini[2]*, Raluca Vodă[3,4], Vlad Dincă[5,6] and Roger Vila[3]

[1]*Department of Biological and Medical Sciences, Centre for Ecology, Environment and Conservation, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK,* [2]*Departamento de Ciências Agrárias, Azorean Biodiversity Group (GBA, CITA-A) and Platform for Enhancing Ecological Research & Sustainability (PEERS), Universidade dos Açores, Rua Capitão João d'Ávila, Pico da Urze, Angra do Heroísmo, Terceira, Azores, Portugal,* [3]*Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37, Barcelona, Spain,* [4]*Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, Spain,* [5]*Department of Zoology, Stockholm University, Stockholm, Sweden,* [6]*Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario N1G 2W1, Canada*

*Correspondence: Simone Fattorini, Azorean Biodiversity Group (GBA, CITA-A) and Platform for Enhancing Ecological Research & Sustainability (PEERS), Departamento de Ciências Agrárias, Universidade dos Açores, Rua Capitão João d'Ávila, Pico da Urze, 9700-042, Angra do Heroísmo, Terceira, Azores, Portugal.
E-mail: simone_fattorini@virgilio.it

ABSTRACT

Aim Unpartitioned dissimilarity indices such as the Sørensen index (β_{sor}) tend to categorize areas according to species number. The use of turnover indices, such as the Simpson index (β_{simp}), may lead to the loss of important information represented by the nestedness component (β_{nest}). Recent studies have suggested the importance of integrating nestedness and turnover information. We evaluated this proposition by comparing biogeographical patterns obtained by unpartitioned (β_{sor}) and partitioned indices (β_{simp} and β_{nest}) on presence data of western Mediterranean butterflies.

Location Western Mediterranean.

Methods We assessed the regionalization of 81 mainland and island faunas according to partitioned and unpartitioned dissimilarity by using cluster analyses with the unweighted pair-group method using arithmetic averages (UPGMA) combined with non-metric multidimensional scaling (NMDS). We also carried out dissimilarity interpolation for β_{sor}, β_{simp}, β_{nest} and the $\beta_{\mathrm{nest}}/\beta_{\mathrm{sor}}$ ratio, to identify geographical patterns of variation in faunal dissimilarity.

Results When the unpartitioned β_{sor} index was used, the clustering of sites allowed a clear distinction between insular and mainland species assemblages. Most islands were grouped together, irrespective of their mainland source, because of the dominant effect of their shared low richness. β_{simp} was the most effective index for clustering islands with their respective mainland source. β_{simp} clustered mainland sites into broader regions than clusters obtained using β_{sor}. A comparison of regionalization and interpolation provided complementary information and revealed that, in different regions, the patterns highlighted by β_{sor} could largely be determined either by nestedness or turnover.

Main conclusions Partitioned and unpartitioned indices convey complementary information, and are able to reveal the influence of historical and ecological processes in structuring species assemblages. When the effect of nestedness is strong, the exclusive use of turnover indices can generate geographically coherent groupings, but can also result in the loss of important information. Indeed, various factors, such as colonization–extinction events, climatic parameters and the peninsular effect, may determine dissimilarity patterns expressed by the nestedness component.

Keywords

Beta diversity, butterflies, faunal dissimilarity, island biogeography, mainland regions, nestedness, regionalization, turnover, western Mediterranean.

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12315

INTRODUCTION

Dissimilarity indices, such as those described by Sørensen (1948) and Simpson (1960), have a long but contested history in biogeography (Shi, 1993; Baselga, 2010; Tuomisto, 2010; Almeida-Neto *et al.*, 2011). The widely used Sørensen index (β_{sor}) has been shown to comprise two additive components: (1) nestedness (β_{nest}) and (2) the spatial replacement (turnover) of species, corresponding to the Simpson index (β_{simp}) (Baselga, 2010). This insight has facilitated the identification of historical and ecological drivers of current similarities among species assemblages (Dobrovolski *et al.*, 2012; Fattorini & Baselga, 2012; Stuart *et al.*, 2012) and the development of broad-scale regionalizations (Kreft & Jetz, 2010; Holt *et al.*, 2013; Mouillot *et al.*, 2013). However, the meaning and appropriateness of their different uses remain subject to debate (Baselga, 2010; Tuomisto, 2010; Almeida-Neto *et al.*, 2011).

The relative importance of nestedness and turnover components in determining overall (dis)similarities among species assemblages varies as a result of different processes (Dobrovolski *et al.*, 2012; Mouillot *et al.*, 2013). Regarding island biogeography, a recursive and counterintuitive pattern occurs when islands are compared using β_{sor} or the Jaccard index (Jaccard, 1901), which is monotonically related to β_{sor}: the poorest islands tend to be grouped together, even if they belong to different archipelagos (Sfenthourakis, 1996; Dennis *et al.*, 2000; Gentile & Argano, 2005; Spengler *et al.*, 2011). Moreover, these indices tend to consider the species-poor islands as remarkably different from their neighbouring mainland sources (Dapporto & Cini, 2007; López-López *et al.*, 2008; Heiser & Schmitt, 2010). We refer to this pattern as the island aggregation rule. The aggregation rule may be a reflection of the impoverished and nested distribution of island biotas, and underlines the importance of understanding what the indices are actually measuring in each dataset before drawing conclusions.

The aggregation of areas according to their species richness represents a well-known phenomenon (Koleff *et al.*, 2003; Baselga, 2010) that has led to the general view that the ordered changes in richness among areas, expressed by the nestedness component, is noise that obscures biogeographical patterns, favouring instead the use of turnover indices (Baselga, 2010; Kreft & Jetz, 2010; Holt *et al.*, 2013; but see Mouillot *et al.*, 2013). From this perspective, turnover indices have become the preferred choice for expressing dissimilarity in species composition among species assemblages (e.g. Kreft & Jetz, 2010).

Despite the recent tendency to remove the signal produced by nestedness, nested patterns are considered to be widespread over most species assemblages as a result of a combination of ecological filtering and historical events (Ulrich *et al.*, 2009; Fattorini & Ulrich, 2012; Ulrich & Almeida-Neto, 2012). Dobrovolski *et al.* (2012) found that the variation in the proportion of dissimilarity attributed to nestedness (β_{nest}) on the overall unpartitioned index (β_{sor}) could explain colonization history, with a high β_{nest}/β_{sor} ratio underlying a predominance of extinction and recolonization events, and a low β_{nest}/β_{sor} ratio revealing historical fingerprints among more stable areas.

The different abilities of species to disperse and persist under various degrees of insularity may underpin a generalized trend for poor island species assemblages to be nested subsamples of species assemblages from larger islands or mainland regions (Ulrich *et al.*, 2009; Dennis *et al.*, 2012). As shown for coral reef fish, the nestedness component can group small assemblages not because they share a particularly high proportion of species, but because they share the absence of many species occurring in richer areas (Mouillot *et al.*, 2013). Thus a strong predominance of the nestedness component relative to turnover may explain the tendency of islands to aggregate.

We present a case study on the importance of pairing unpartitioned and partitioned components in order to better understand the biogeography of western Mediterranean butterflies that have a well-known distribution (e.g. Dennis & Schmitt, 2009). The aim of our study was to compare the pattern of each partitioned and unpartitioned component of faunal dissimilarity, and to dissect local evidence of historical and ecological phenomena (endemicity, relictuality, filtering and peninsular effects) (e.g. Dobrovolski *et al.*, 2012; Fattorini, 2013; Mouillot *et al.*, 2013) using two approaches. The first approach focused on detecting faunal regionalization based on overall dissimilarity matrices. Two recent studies (Kreft & Jetz, 2010; Holt *et al.*, 2013) used a combination of classification and ordination analyses [unweighted pair-group method using arithmetic averages (UPGMA) clustering and non-metric multidimensional scaling (NMDS)] for the identification of groups and visualization of global patterns. We applied a similar methodology to recognize coherent groups of sites in the studied region. The second approach highlighted changes in species composition among surrounding sites by projecting dissimilarity values among the nearest sites on a geographical map (Vandergast *et al.*, 2011; Keis *et al.*, 2013).

We show that unpartitioned and partitioned components provide a comprehensive representation of faunal affinity and its variation over space, thus facilitating identification of the main evolutionary processes, colonization routes and filtering mechanisms that eventually determine the observed species assemblages. We have also developed and made available new R functions to facilitate some of these analyses, because, as far as we know, there are no available scripts for this purpose.

MATERIALS AND METHODS

Sites, data and dissimilarity indices

Presence data of Lepidoptera species were gathered from several literature sources and from our own field studies carried out between 1999 and 2013 throughout the study region (see

Appendix S1 in Supporting Information). However, as for most arthropod groups, comprehensive and detailed distribution data were unavailable for many southern European areas and for most parts of the Maghreb. Therefore, we selected 81 sites with well-studied butterfly faunas comprising data belonging to 28 small islands (entire island area considered) and 53 sites of 25-km radius located in mainland regions (41) and on larger islands (Corsica, 3; Sardinia, 3; Sicily, 6) (see Appendix S1 and Fig. S1 in Appendix S2). To be selected, all sites were required to have in their species records at least four of five well-known migrant species (*Pieris brassicae*, *Pieris rapae*, *Colias crocea*, *Vanessa atalanta* and *Vanessa cardui*). These species are highly visible, occur throughout Europe during the summer season, and records for them can be used to establish a minimal surveying standard (Dapporto & Dennis, 2008). A total of 335 species had been recorded across the study sites. Nomenclature largely followed Kudrna *et al.* (2011).

To assess the importance of nestedness in determining the faunal structure of our dataset, we conducted nestedness analyses for island and mainland sites by calculating a nestedness measure based on overlap and decreasing fills (NODF) and temperature metrics with the NeD software (Strona *et al.*, in press). In addition to the absolute values of NODF and temperature, we also computed the respective standardized effect sizes as *z*-values based on 1000 simulated null matrices. Null matrices were constructed using the 'proportional row and column total' algorithm, which is considered to be the most biologically realistic (Bascompte *et al.*, 2003).

We considered the Sørensen index of overall (unpartitioned) dissimilarity (β_{sor}) and its two additive components: β_{simp}, which expresses species replacement among sites, and β_{nest}, which is a nestedness index (Baselga, 2010). The ratio between β_{nest} and β_{sor} can also provide information about the contributions of ecological and historical drivers to the observed diversity patterns (Dobrovolski *et al.*, 2012).

Combination of cluster analysis and multidimensional scaling

We visualized patterns of biogeographical regionalization by combining UPGMA cluster analyses and NMDS. UPGMA is considered to be the best method of producing trees for biogeographical analyses (Kreft & Jetz, 2010; Holt *et al.*, 2013). However, if there are tied values, the tree topology and bootstrap supports may be affected by the order of the sites in the original presence–absence matrix; thus we used the 'recluster' R package to reshuffle row order and create consensus trees (Dapporto *et al.*, 2013). We computed 10 levels of multiscale bootstrapping using from 1- to 10-fold the number of species in the original data frame. Subsequently, we verified which level showed the best separation into a high and low series of support for nodes, and used this categorization to indicate nodes with strong or weak support (Dapporto *et al.*, 2013; Appendix S3). To identify the best division of a tree into branches that defined separate zoogeographical groups of sites, we wrote an R function that allows a direct comparison of all the possible subdivisions of a given tree, providing a measure of their 'explained dissimilarity' (*sensu* Holt *et al.*, 2013) and illustrating cluster membership for each site (Appendix S3). As the best compromise, Holt *et al.* (2013) suggested using the minimal division explaining at least 90% of dissimilarity, and we adopted this threshold in our analyses.

As discussed in Kreft & Jetz (2010) and Holt *et al.* (2013), in addition to cluster analysis we used NMDS as a robust nonparametric procedure to project the dissimilarity matrices into a two-dimensional space. We attributed RGB (red, green, blue) colours to each site according to its position within the space defined by the first two NMDS axes. The axis with the highest variance was standardized between 0 and 1, and the second axis was rescaled according to the first one. Subsequently, the colours blue, green, yellow and red were assigned to the four corners. Finally, the contribution of each RGB colour to each site was calculated based on its position in the two-dimensional graph. We wrote a function for this analysis and have included it in the 'recluster' R package (Appendix S3). In order to maximize the similarity in colour pattern among indices without altering the relative distribution of points, we applied a procrustes analysis to the configuration of the three indices before projecting them into the RGB space (Appendix S3). The points belonging to each cluster identified by 90% of the explained dissimilarity were condensed into single points in the NMDS representation as centroids for clustered sites (Kreft & Jetz, 2010). The representative colours for such centroids were recomputed. Points with their centroid colour were plotted on the geographical map to show the overall biogeographical pattern.

Interpolation of dissimilarity values

Combining cluster analysis and NMDS provides patterns of regionalization, but grouping sites into larger units prevents the identification of particular areas of continuity and discontinuity. We therefore carried out an analysis to plot and interpolate on the map the distances between neighbouring sites for the three indices and for the β_{nest}/β_{sor} ratio. To do this, we combined the methods used by Vandergast *et al.* (2011) and Keis *et al.* (2013). We calculated a Delaunay triangulation among GPS coordinates for the study sites (Appendix S3). For each index, the dissimilarity value corresponding to the distance between the sites was attributed to the mid-point of each segment connecting two sites. Because segments have different lengths, and faunal dissimilarities tend to increase with distance following an asymptotic trend (Borcard *et al.*, 2011), we computed residuals of asymptotic regressions between dissimilarities and geographical distances using the R scripts provided by Keis *et al.* (2013). Finally, we interpolated the values of the mid-points with the inverse distance weighting algorithm provided by QGIS (http://www.

qgis.org/en/site/). The scripts for statistical analyses performed with R and the original data matrix are available in Appendices S1 and S3.

RESULTS

Nestedness pattern

Nestedness analyses revealed that both island and mainland assemblages were nested, but temperature was lower (and hence nestedness higher) for island sites (matrix size: 124 ×40, fill: 29%; NODF = 67.653, $z = 28.874$, $P < 0.001$; temperature = 12.820, $z = -20.119$, $P < 0.001$) than for mainland sites (matrix size: 319 × 41, fill: 37.9%; NODF = 59.972, $z = 24.631$, $P < 0.001$; temperature = 25.190, $z = -30.144$, $P < 0.001$). When sites were classified as mainland or island and then ordered according to their species richness, island assemblages were shown to be poorer than mainland assemblages, except for the Sicilian assemblages (Fig. S2 in Appendix S2).

Cluster analysis, NMDS ordination and RGB colour space

A consensus analysis revealed that β_{sor} UPGMA trees were not affected by row order because no nodes were collapsed (Fig. 1a, and Fig. S3 in Appendix S2). Support from multiscale bootstrap analysis increased with the number of species included in the resampled matrices (Fig. 1b). The 2× solution showed the highest power in separating nodes into two groups (Fig. 1a). The first resulting node identified two highly supported clusters: the first cluster grouped most island assemblages, thus confirming the aggregation rule, while the second cluster was composed of almost all the mainland assemblages plus the Sicilian assemblage (Fig. 1a). Because of this strong aggregation of sites into two main groups, we analysed the subdivision explaining at least 90% of the dissimilarity separately for the two subtrees. In the case of the island cluster, the sites were in turn aggregated into seven clusters (five of which had low support), mainly as a result of differences in richness. Although some spatial

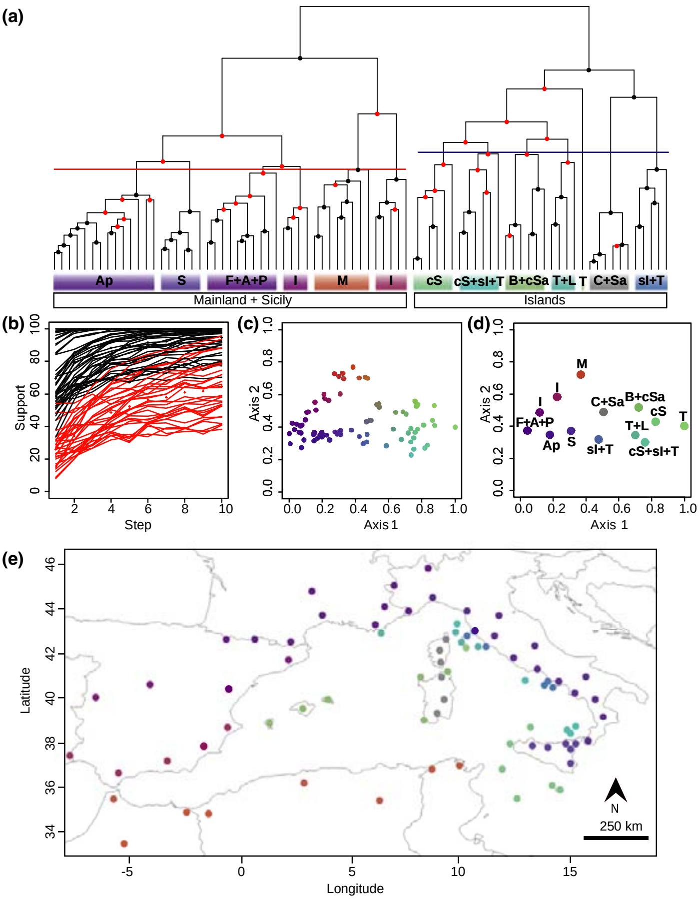


Figure 1 Combined analysis of the presence and absence of butterfly species in the western Mediterranean using UPGMA clustering and non-metric multidimensional scaling (NMDS) for the Sørensen index (β_{sor}). (a) The consensus UPGMA tree for β_{sor} with the selected thresholds, which resulted in two clusters: mainland sites + Sicily and the rest of islands grouped together. (b) Supports for nodes were obtained with multiscale bootstrap analysis and highly (black) and weakly (red) supported nodes were identified together with mean values. Highly and weakly supported nodes are indicated with circles in (a) by using the same black and red colours. (c) NMDS in RGB colour space for the sites, and (d) mean locations and colours for clusters according to the subdivision of the tree. (e) Finally, the colours obtained in (d) were attributed to all the sites in each cluster and the sites plotted on a map. The same tree with labels identifying each site can be found in Fig. S3 in Appendix S2. Abbreviations for the site groups are as follows: A, Alps; Ap, Apennines; F, France; P, Pyrenees; I, Iberia; M, Maghreb; Sa, Sardinia; C, Corsica; cSa, circum-Sardinian islands; S, Sicily; cS, circum-Sicilian; L, the Levant island; T, Tuscan islands; B, Balearics; sI, southern Italian.

structures could be seen with the identification of groups composed of circum-Sicilian and Sardo-Corsican sites, some inconsistencies occurred, for example the link between small circum-Sardinian islands (completely sourced by Sardinia) and the Balearics (Fig. 1a). In the mainland–Sicily subtree, the six identified clusters corresponded with well-recognized biogeographical units (Fig. 1a). A major aggregation identified a cluster comprising the Maghreb and southern Iberia; although this cluster was weakly supported by the bootstrap analysis, each of the individual areas was strongly supported. The rest formed another cluster that was in turn divided into two strongly supported groups of sites, corresponding to Sicily and the Apennines, and two weakly supported groups, corresponding to central Iberia and the Alps plus the Pyrenees (Fig. 1a).

NMDS with β_{sor} confirmed that island sites tended to cluster separately from the mainland sites, and the first axis was highly determined by a gradient in richness (Fig. 1c). The second axis was presumably characterized by the turnover because it separated the two broad areas of Sardinia–Corsica and the Maghreb, which have a high number of endemics. The mean colours for the 13 groups identified by hierarchical clustering were computed (Fig. 1d) and plotted on a map. Because of their low richness, all the islands (except Sicily) were separated from the mainland sites by the first NMDS axis. This resulted in similar greenish–cyan colours for the islands, contrasting with the red–purple–blue colours characterizing the mainland sites (Fig. 1e).

The β_{nest} UPGMA trees were also unaffected by row order because only a few terminal nodes were collapsed in the consensus analysis (Fig. 2a, and Fig. S4 in Appendix S2). The fifth step of the multiscale bootstrap analysis showed the highest power in separating nodes into highly and poorly supported categories (Fig. 2b). As expected, the pattern was highly dominated by species richness. The tree subdivision explaining at least 90% of the dissimilarity identified five clusters. The first cluster was highly supported and comprised only the poorest island of Montecristo, while all the other small islands were included in the second cluster, also highly supported. The third cluster, highly supported,

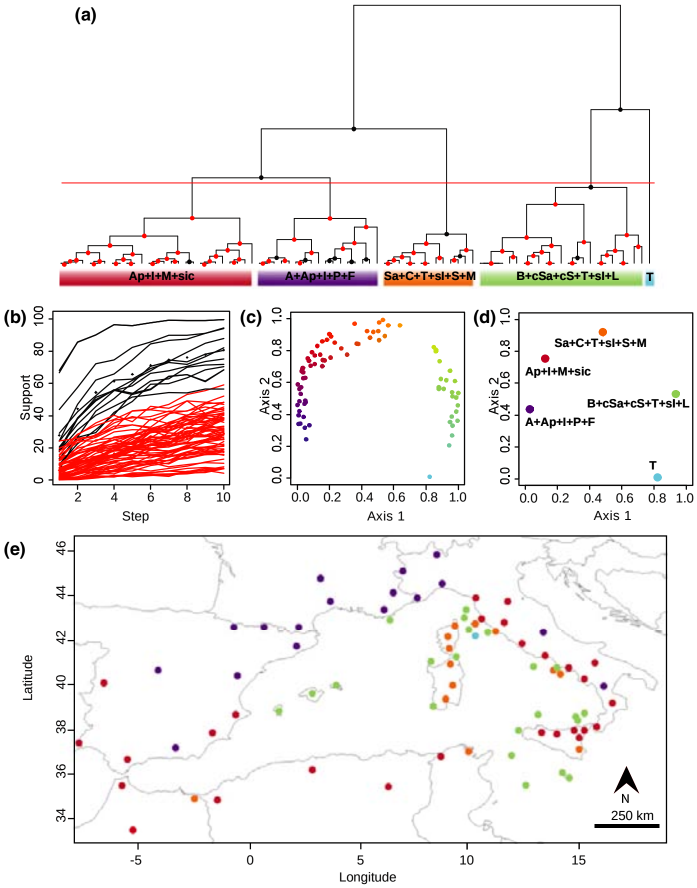


Figure 2 Combined analysis of the presence and absence of butterfly species in the western Mediterranean using UPGMA clustering and non-metric multidimensional scaling (NMDS) for the nestedness component (β_{nest}). (a) The consensus UPGMA tree for β_{nest} showing the thresholds selected to define the clusters, which explained 94% of the β_{nest} dissimilarity for mainland sites and divided the tree into five clusters. (b) Supports for nodes were obtained with multiscale bootstrap analysis and highly (black) and weakly (red) supported nodes were identified together with mean values. Highly and weakly supported nodes are indicated with circles in (a) by using the same black and red colours. (c) NMDS in RGB colour space for the sites, and (d) identification of the mean locations and colours for clusters according to the subdivision of the tree. (e) Finally, the colours obtained in (d) were attributed to all the sites of each cluster and the sites plotted on a map. The same tree with identifying labels for each site can be found in Fig. S4 in Appendix S2. Abbreviations for the site groups are as follows: A, Alps; Ap, Apennines; F, France; P, Pyrenees; I, Iberia; M, Maghreb; Sa, Sardinia; C, Corsica; cSa, circum-Sardinian islands; S, Sicily; cS, circum-Sicilian; L, the Levant; T, Tuscan islands; B, Balearics; sI, southern Italian.

included the largest islands but also excluded five out of six sites on Sicily and included two sites from the Maghreb. The fourth cluster was weakly supported and grouped most of the sites from the Maghreb, Sicily, southern Iberia and the Apennines, while most of the richest Alpine and Pyrenean sites were included in the fifth, poorly supported, cluster (Fig. 2a). The NMDS pattern revealed a rather semi-circular structure with an order almost completely determined by richness (Fig. 2c). When the mean colours for the five groups identified by hierarchical clustering were plotted on a map (Fig. 2d), the contrast between islands and mainland sites was less clear. This was also suggested by the UPGMA topology and by some sites in the Maghreb and Sicily that showed the same colour as Sardo-Corsican sites (Fig. 2e).

The β_{simp} trees were highly sensitive to row order because of the large occurrence of zero values. Accordingly, several nodes were collapsed in the consensus tree (Fig. 3a, and Fig. S5 in Appendix S2). Support for some nodes increased substantially with multiscale bootstrapping (Fig. 3b). This was an expected result for β_{simp}, where a few key species provide a disproportionate contribution to the index values. The step using 10× the original number of species revealed the best separation into highly and poorly supported nodes (Fig. 3b). The threshold for a minimum number of clusters explaining at least 90% of the dissimilarity identified eight groups (Fig. 3a). In contrast with the results obtained with β_{sor} and β_{nest}, small islands clustered with mainland sites in a highly coherent pattern (Fig. 3a). The Balearics clustered with southern Iberia in a cluster that also comprised Pantelleria. The Levant island, Tuscan islands near the Italian peninsula and other circum-Italian islands clustered with the Pyrenees–Alps–Apennines group. Lampedusa clustered with the Maghreb, and the circum-Sicilian islands were grouped with Sicily. The NMDS confirmed a polarization between the Alps–Pyrenees and the Maghreb along the first axis, whereas differences among the Sardo-Corsican islands and the Maghreb encompassed most of the variation along the second axis (Fig. 3c). Central and southern Iberian sites were located at an intermediate position between the Alps–Pyrenees–Apennines and the Maghreb. When RGB colours for mean

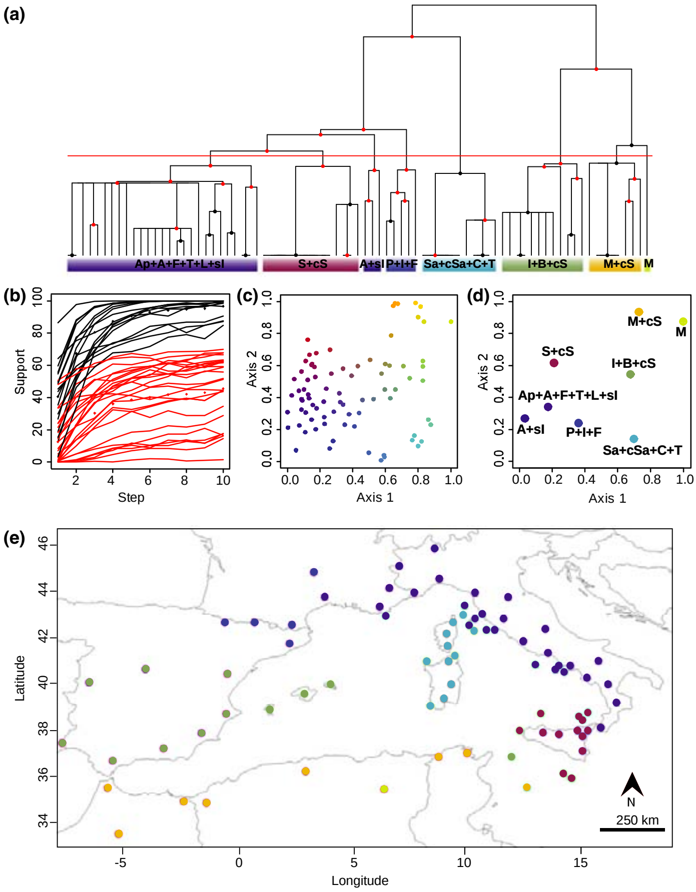


Figure 3 Combined analysis of the presence and absence of butterfly species in the western Mediterranean using UPGMA clustering and non-metric multidimensional scaling (NMDS) for the Simpson index (β_{simp}). (a) The consensus UPGMA tree for β_{simp} showing the thresholds selected to define the clusters, which explained 97% of the β_{simp} dissimilarity for mainland sites and divided the tree into eight clusters. (b) Supports for nodes were obtained with multiscale bootstrap analysis and highly (black) and weakly (red) supported nodes were identified together with mean values. Highly and weakly supported nodes are indicated with circles in (a) by using the same black and red colours. (c) NMDS in RGB colour space for the sites, (d) and identification of the mean locations and colours for clusters according to the subdivision of the tree. (e) Finally, the colours obtained in (d) were attributed to all the sites of each cluster and the sites plotted on a map. The same tree with labels identifying each site can be found in Fig. S5 in Appendix S2. Abbreviations for the site groups are as follows: A, Alps; Ap, Apennines; F, France; P, Pyrenees; I, Iberia; M, Maghreb; Sa, Sardinia; C, Corsica; cSa, circum-Sardinian islands; S, Sicily; cS, circum-Sicilian; L, the Levant; T, Tuscan islands; B, Balearics; sI, southern Italian.

cluster membership (Fig. 3d) were plotted on the map, a colour gradient from the Maghreb to Iberia, Sicily, the Pyrenees, the Apennines and the Alps was evident (Fig. 3e).

Interpolation of dissimilarity values

For both β_{sor} and β_{simp}, the asymptotic regression found a significant series of parameters to describe relationships between geographical distances and faunal dissimilarity values, but it failed to do so with β_{nest} and β_{nest}/β_{sor}. The absence of a spatial trend for these indices was probably because of their high dependence on richness, which was mostly explained by island attributes (area and isolation) rather than by geographical location. We then used residuals for β_{sor} and β_{simp} and crude dissimilarity values for β_{nest} and β_{nest}/β_{sor}. Projection of dissimilarity values over the Delaunay triangulation (Fig. S6) and their interpolation for β_{sor} (Fig. 4a) revealed a belt of high dissimilarity values between mainland and island sites but, as suggested by the cluster analysis, Sicily did not show such a strong contrast with respect to Italy. The Maghreb showed a strong contrast with respect to Sicily and a moderate dissimilarity compared with southern Iberia. Plotting partitioned components separately revealed that the main β_{sor} pattern separating islands from mainland sites had different meanings in different regions. Dissimilarity patterns separating the Balearics from Iberia and the Maghreb, the Tuscan islands from Corsica, and the circum-Sicilian islands from Sicily, respectively, appeared to be mostly generated by the nestedness component (Fig. 4b). Conversely, the β_{sor} dissimilarity around the Sardo-Corsican block with respect to all sites except the Tuscan archipelago, as well as that separating the Maghreb from southern Europe, was mainly the result of the turnover component (Fig. 4c). As expected, mainland sites showed much weaker dissimilarity discontinuities. The β_{nest}/β_{sor} index had low values for mainland sites, indicating a strong contribution of turnover in determining variations in β_{sor} values among mainland sites. The β_{nest}/β_{sor} ratio was much higher between islands and mainland sites where the nestedness component prevailed. There were, however, important exceptions for some sea areas: (1) the sea belt around Sardinia and Corsica, apart from the area of the Tuscan islands; (2) the sea channel separating Sicily from the Maghreb; and (3) the Gibraltar strait. The Balearics revealed a very interesting pattern, showing almost no effect of turnover with respect to both Iberia and the Maghreb (Fig. 4d).

DISCUSSION

We have shown that richness effects on unpartitioned β_{sor} vary largely according to the geographical characteristics of the areas (e.g. mainland versus islands). Because of the highly nested pattern of our dataset, differences in richness were expressed almost completely by the nestedness component. In Table 1 we summarize some examples of ecological, geographical/historical and other complex factors determining nestedness and turnover components. All these factors may contribute to the overall dissimilarity patterns and, in theory, one may decide to consider all of them simultaneously by using unpartitioned indices (Mouillot *et al.*, 2013). However, this will mix the different contributions of nestedness and turnover within the overall pattern. It is well known that island and mainland assemblages may be influenced by different factors. For example, sea straits, interrupting the habitat continuum and applying a strong differential filtering to dispersal, determine strong nestedness contrasts in island–mainland pairs (Stuart *et al.*, 2012). This is particularly evident where island populations are small and/or maintain some gene flow with the neighbouring mainland, so that the emergence of turnover through the evolution of new taxa is hampered (Whittaker & Fernández-Palacios, 2007). The western Mediterranean islands are mostly of this type and, therefore, apart from the Sardo-Corsican group, a strong influence of the nestedness component on the island–mainland dissimilarity patterns can be seen in the study area (Fig. 4d). However, ecological filtering is not the only phenomenon contributing to the identification of islands with the β_{sor} dissimilarity. Nestedness alone tended to group most of the medium and large islands with mainland sites. This suggests that western Mediterranean island faunas cannot be characterized simply by a shared and ordered impoverishment of their species assemblages.

The aggregation rule and its value in distinguishing island dynamics

Any a priori dichotomy distinguishing islands and mainland is subjective and simplistic. Some islands (usually the largest ones) show a faunal composition that is very similar to that of their neighbouring mainland, while some mainland patches of particular habitats (e.g. mountain peaks and wetlands) may have insular characteristics (Whittaker & Fernández-Palacios, 2007). We found that β_{sor} grouped together, with strong support, all island sites except for Sicily, thus confirming the tendency of this index to produce island aggregation (Fig. 1). However, a shared faunal impoverishment on islands, as expressed by the β_{nest} component, is not sufficient to explain the aggregation rule. The dissimilarity interpolation demonstrated that the belt of high β_{sor} values around islands can be generated locally by high values for either the β_{nest} or β_{simp} component. The turnover component in western Mediterranean butterflies has a complex origin. Whereas endemism tends to be generally low and mostly concentrated in the Sardo-Corsican area, there are several species that have higher incidence on islands than on the nearest mainland (e.g. *Pontia daplidice*, *Polyommatus celina*, *Aricia cramera*, *Argynnis pandora*, *Coenonympha lyllus* and several *Hipparchia* taxa), probably representing relict occurrences from past wider distributions (Dapporto & Dennis, 2008; Dincă *et al.*, 2011; Dapporto *et al.*, 2012; Sañudo-Restrepo *et al.*, 2013). As a result, these distributions produce an aggregation of islands in the β_{sor} analyses, reducing their

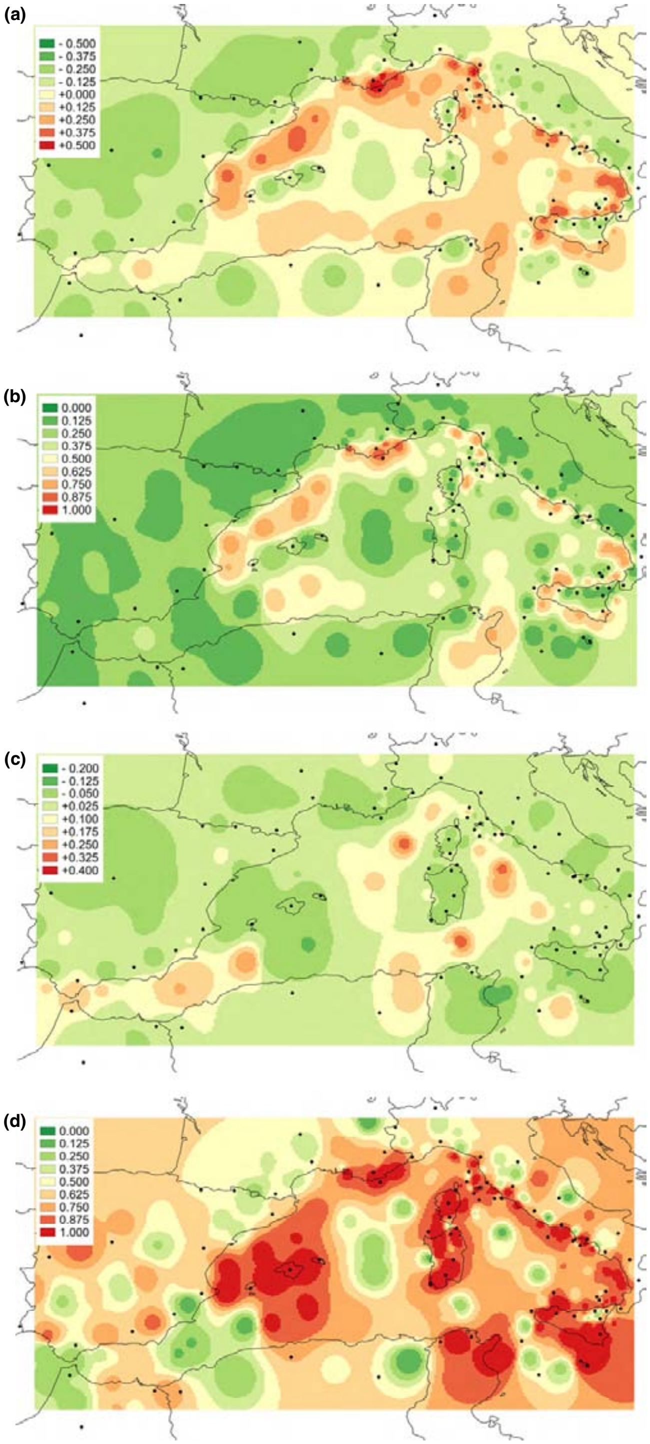


Figure 4 Inverse distance weighting interpolations of dissimilarity for the (a) Sørensen index (β_{sor}), (b) nestedness index (β_{nest}), (c) Simpson index (β_{simp}) and (d) β_{nest}/β_{sor} ratio. A red colour indicates a high dissimilarity for β_{sor}, β_{simp} and β_{nest}. β_{nest}/β_{sor} values higher than 0.5 (red) indicate the predominance of β_{nest} in determining the overall β_{sor} dissimilarity, while values below 0.5 (green) indicate a primacy of β_{simp}. Values for β_{sor} and β_{simp} represent residuals from a significant asymptotic dissimilarity–distance regression. No significant relationship was found for β_{nest} and the $\beta_{nest}/\ \beta_{sor}$ ratio; the interpolations were carried out on untransformed data.

Table 1 Factors responsible for the nested/richness and turnover components of dissimilarity.

Type of factors	Main factors affecting nestedness and richness	Main factors affecting turnover
Ecological	Area; isolation; environmental variability	Differences in species' ecological requirements; factors assisting dispersal (e.g. wind and sea current directions, ecological corridors)
Historical/ geographical	Presence of refugial areas, barriers and other geomorphological features regulating species numbers (e.g. via species–area relationships, faunal exchanges)	Historical dynamics of faunas (e.g. post-glacial colonization); distance to sources; evolution of endemics (e.g. via allopatric speciation)
Complex	Richness and composition of the nearest sources (in turn determined by both historical and ecological factors)	Interactions among species at their distribution borders limiting dispersal and stabilizing occupied areas (e.g. via competition)

similarity with respect to the mainland source sites. We can thus confirm that, in spite of the common belief that dissimilarity patterns of unpartitioned indices have minor biogeographical value (e.g. Baselga, 2010; Kreft & Jetz, 2010; Holt *et al.*, 2013), the basal partition for β_{sor} may identify which assemblages are constrained by ecological factors (such as isolation and limited resources) and characterized by common historical factors (that determined the occurrence of endemic and relict species). As a corollary, the aggregation rule revealed which mainland sites showed insular dynamics (such as Monte Argentario, a fossil island recently connected to the Italian peninsula by three stretches of sand) and which insular sites showed typical mainland assemblages (such as Sicily, the largest Mediterranean island, separated from the Italian peninsula by a narrow strait of 3 km and characterized by a high environmental heterogeneity that can offer suitable habitats to most mainland species).

Nestedness and turnover patterns within mainland areas

The recognition of typical mainland and island units was only the first piece of information provided by β_{sor}. The β_{sor} tree identified six clusters that almost perfectly corresponded with broad geographical mainland sites: the Maghreb, central Iberia, southern Iberia, the Alps and Pyrenees, the Apennines, and Sicily. At a local scale, the turnover component among mainland areas was much more influential than nestedness, as revealed by the low β_{nest}/β_{sor} patterns in the interpolation analyses (Fig. 4d). Nevertheless, nestedness largely influenced the regional pattern. This was revealed by the polarization of areas along the first NMDS axis, clearly reflecting a richness gradient (Fig. 2c). Such inconsistencies between regionalization and interpolation results can be explained by the geographical setting of the study area. Because of the absence of strong barriers between mainland sites, species replacement was fairly equally distributed over all the continental regions, with only two main areas of turnover corresponding to important geographical discontinuities: the sea channel separating the Maghreb from Sicily, and the Gibraltar strait, separating the Maghreb from Iberia.

The regionalization pattern obtained for the Italian region seemed to be largely the result of nestedness, and the β_{simp} index did not reveal a differentiation among most mainland sites from the Pyrenees to Sicily (as shown, for example, by a cluster including two Alpine areas and the Tyrrhenian island of Ponza). Although Sicily was identified as a distinct mainland site for both β_{sor} and β_{simp}, only two endemic butterfly taxa occur on this island (*Hipparchia blachieri* and *Melanargia pherusa*); two other species occurring on Sicily (*Spialia orbifer* and *Anthocharis damone*) have probably originated from the eastern Mediterranean because they are virtually absent from the western European mainland. Conversely, there are 14 species occurring at the southernmost tip of the Italian peninsula that are absent from Sicily (Appendix S1). Similarly, dozens of species occurring in the Alps and the Pyrenees do not occur in southern Italy, whereas the opposite rarely applies. There is therefore evidence that the clusters highlighted by β_{sor} from the Pyrenees to Sicily mostly describe an ordered gradient of impoverishment in a sort of peninsular effect or latitudinal gradient, as has been shown previously for the Iberian Lepidoptera (Martin & Gurrea, 1990).

Nevertheless, marked peninsular effects and nested patterns are not sufficient to explain the biogeography of butterflies from southern and central Iberia. Indeed, in both β_{sor} and β_{simp} NMDS analyses, these areas appeared to be clearly intermediate between Europe and the Maghreb, and the UPGMA clustered southern Iberia with the Maghreb for both indices (although the southern Iberia–Maghreb group received low support, as expected for a transitional area). This indicates that the latitudinal and/or peninsular effects are overshadowed by other phenomena. In central and southern Iberia there is a high frequency of endemic and relict species, with the presence of 15 taxa that do not occur elsewhere in the study region (*Euchloe bazae*, *Lycaena bleusei*, *Cupido carswelli*, *Kretania hesperica*, *Agriades zullichi*, *Polyommatus fabressei*, *Polyommatus violetae*, *Polyommatus fulgens*, *Polyommatus golgus*, *Polyommatus nivescens*, *Pseudophilotes panoptes*, *Aricia morronensis*, *Pseudochazara williamsi*, *Erebia hispania* and *Erebia zapateri*). On the other hand, only six species occurring in the Sicily–Apennines region do not occur in the rest of the study area (*Spialia orbifer*, *Euchloe ausonia*, *Anthocharis damone*, *Melanargia pherusa*, *Hipparchia blachieri* and *Hipparchia neapolitana*). These differences may be explained by the physical geography of the two peninsulas. The Iberian and Italian peninsulas may seem to be similar because of the existence of large mountain chains connecting them to the rest of the European mainland (the Pyrenees and the Alps) and the presence of high peaks at their southern Mediterranean margins (Sierra Nevada and Mount Etna). However, the Apennines constitute a mountainous corridor from the Alps to Sicily, hence facilitating gene flow among butterfly populations along the Italian peninsula, whereas in the Iberian Peninsula the mountain chains are generally longitudinally oriented, thus generating large gaps that favour genetic isolation.

The final and perhaps most important factor determining the observed pattern for central and southern Iberia is the influence of North Africa, which is much stronger than for the Apennines–Sicily. Central and southern Iberia host marginal populations of at least 18 species that are widespread in the Maghreb but do not occur in the Alps and the Pyrenees, whereas Sicily only hosts two, *Melitaea aetherie* and *Polyommatus celina*, the former also occurring in the southern Apennines.

Nestedness and turnover patterns for islands

The island–mainland pairs obtained with β_{simp} did not show a particularly high dissimilarity, and a basal dichotomy separating islands and mainland sites did not occur. In fact, β_{simp} is specifically designed to represent the proportion of species in the poorest areas that do not occur in the richest areas. Thus, for any dissimilarity to occur, it is necessary for the poorest area in the pair (usually the island in island–mainland pairs) to have at least one species not occurring in the richest one. In this respect, even a single species on a small island can determine a high value of turnover dissimilarity. Island endemics are by definition involved in turnover phenomena, but in the western Mediterranean they are mostly concentrated on the large and isolated islands of Sardinia and Corsica, which have, in turn, acted as a source of endemics for their satellite islands. When the dominant effect of nestedness was removed by using β_{simp}, small islands showed biogeographical patterns consistent with their location and recent geological history. All islands that had had contact with neighbouring areas during the Pleistocene clustered with their source mainland (Elba, Pianosa, Argentario, Giglio, Capri and Ischia with the Apennines; Malta, Gozo and Levanzo with Sicily; Lampedusa with the Maghreb). The Balearics were shown to be more similar to southern Iberia than to northern Iberia. The Tuscan islands of Montecristo and Capraia were confirmed as hosting typically Corsican faunas, although they were also close to Elba and Pianosa, which are grouped with the Apennines.

Nevertheless, regionalization revealed by turnover alone cannot fully reconstruct the colonization dynamics of islands. The interpolation analyses revealed that the main pattern of diversity obtained by β_{sor}, i.e. the separation of islands from mainland areas, had different explanations for different areas. A comparison of Fig. 4a–c reveals that the β_{sor} belt around the Sardo-Corsican area is the result of turnover phenomena, with the exception of a connection between Corsica and Tuscany through the Tuscan archipelago determined by nestedness. Conversely, the turnover UPGMA created a clear separation between the Sardo-Corsican area versus most Tuscan islands and the Tuscan mainland (Fig. 3e). However, a closer inspection of the species distributions revealed that the discontinuity revealed by the β_{simp} analysis was not fully justified. In fact, there are four Sardo-Corsican endemics extending to the Tuscan islands (*Plebejus bellieri*, *Lasiommata paramegaera*, *Hipparchia aristaeus* and *Hipparchia neomiris*) and one (*Coenonympha corinna*) also extending to the Tuscan mainland coast. There are also several mainland species occurring with decreasing frequency from the Tuscan islands to Sardinia. Accordingly, the interpolation analysis did not reveal any strong discontinuity in turnover dissimilarity, because the passage between typical Sardo-Corsican and Italian faunas is gradual, probably as a result of a bidirectional filtering effect generated by stepping-stone phenomena (Dapporto & Cini, 2007). Similar phenomena seem to have occurred on the Balearics, which showed no contrasting patterns of turnover with respect to both Iberia and the Maghreb, although these mainland sites themselves differ. In the case of the Balearics, the observed pattern is mainly the result of the presence of only a few species that are responsible for turnover. In the Balearics there are species (*Polyommatus celina* and *Coenonympha lyllus*) that also occur in southern Iberia and the Maghreb but are absent from the nearest areas of northern Iberia included in the study, such as Montseny and Teruel. Only one species (*Leptidea sinapis*) occurring on the Balearics and Spain is absent in the Maghreb, while *Gegenes pumilio* does occur on the Balearics but not in mainland Spain nor in any of the studied sites of the Maghreb. This is an indication that stepping-stone and relictuality phenomena are also operating on these islands.

CONCLUSIONS

There is a renewed debate about the meaning of partitioned and non-partitioned dissimilarity indices (Baselga, 2010; Tuomisto, 2010; Ulrich & Almeida-Neto, 2012). We have shown that such indices reveal complementary patterns and that, taken together, they can explain comprehensively dissimilarity patterns of insular and mainland sites. β_{sor} is particularly efficient in recognizing areas shaped by insular or mainland dynamics and in providing fundamental biogeographical information when applied to mainland areas (e.g. the existence of latitudinal and peninsular effects), as a result of a balanced contribution of the factors reported in Table 1 to assemblage composition. In contrast, this index was largely ineffective in recognizing regionalization between mainland sources and islands. Indeed, when islands were compared with mainland sites, unpartitioned indices tended to be dominated by the nestedness pattern produced by variations in island area and isolation. On the other hand, β_{simp} was effective in recovering island–mainland relationships but resulted in a low signal for mainland dissimilarities, thus limiting the distinction among recognized regions (Dennis & Schmitt, 2009).

Given the impossibility of identifying a single index as a panacea to retrieve comprehensive biogeographical patterns simultaneously for islands and mainland sites, we suggest comparing and combining results obtained using both unpartitioned and partitioned (nestedness and turnover) indices by means of regionalization and dissimilarity interpolation. It should also be noted that the discussion for different indices must follow different rules. While β_{sor} dissimilarity is generated by any species differing between two sites, the β_{simp} values are produced by a reduced subset of taxa occurring in the poorest site and being absent in the largest site. Thus, while β_{sor} can be discussed by taking into consideration faunas as a whole, in the case of β_{simp} even the effect of a single species may be of importance.

ACKNOWLEDGEMENTS

We acknowledge Andrés Baselga and Werner Ulrich for discussing with S.F. various theoretical topics involved in this paper. We are also very grateful to the editor José María Fernández-Palacios and three anonymous referees for their very constructive comments. Financial support was provided by the Spanish Ministerio de Ciencia e Innovación project CGL2010-21226/BOS and by the project ‘Definizione dello status di conservazione delle falene e della malacofauna terrestre del’Arcipelago Toscano’. V.D. was funded by a postdoctoral fellowship from the Wenner-Gren Foundation, Sweden, and R. Vodă by a FPU predoctoral fellowship (AP-2010-5409) from the Spanish Ministry.

REFERENCES

Almeida-Neto, M., Frensel, D.M.B. & Ulrich, W. (2011) Rethinking the relationship between nestedness and beta diversity: a comment on Baselga. *Global Ecology and Biogeography*, **21**, 772–777.

Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003) The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences USA*, **100**, 9383–9387.

Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **19**, 134–143.

Borcard, D., Gillet, F. & Legendre, P. (2011) *Numerical ecology with R*. Springer, New York.

Dapporto, L. & Cini, A. (2007) Faunal patterns in Tuscan archipelago butterflies: the dominant influence is recent geography not paleogeography. *European Journal of Entomology*, **104**, 497–503.

Dapporto, L. & Dennis, R.L.H. (2008) Species richness, rarity and endemicity on Italian offshore islands: complementary signals from island-focused and species-focused analyses. *Journal of Biogeography*, **35**, 664–674.

Dapporto, L., Bruschini, C., Dincă, V., Vila, R. & Dennis, R.L.H. (2012) Identifying zones of phenetic compression in West Mediterranean butterflies (Satyrinae): refugia, invasion and hybridization. *Diversity and Distributions*, **18**, 1066–1076.

Dapporto, L., Ramazzotti, M., Fattorini, S., Vila, R. & Dennis, R.L.H. (2013) recluster: an unbiased clustering procedure for beta-diversity turnover. *Ecography*, **36**, 1070–1075.

Dennis, R.L.H. & Schmitt, T. (2009) Faunal structures, phylogeography and historical inference. *Ecology of butterflies in Europe* (ed. by J. Settele, T. Shreeve, M. Konvièka and H. Van Dick), pp. 250–280. Cambridge University Press, Cambridge, UK.

Dennis, R.L.H., Shreeve, T.G., Olivier, A. & Coutsis, J.G. (2000) Contemporary geography dominates butterfly diversity gradients within the Aegean archipelago (Lepidoptera: Papilionoidea, Hesperioidea). *Journal of Biogeography*, **27**, 1365–1383.

Dennis, R.L.H., Hardy, P.B. & Dapporto, L. (2012) Nestedness in island faunas: novel insights into island biogeography through butterfly community profiles of colonization ability and migration capacity. *Journal of Biogeography*, **39**, 1412–1426.

Dincă, V., Dapporto, L. & Vila, R. (2011) A combined genetic–morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and *Polyommatus celina* Common Blue butterflies. *Molecular Ecology*, **20**, 3921–3935.

Dobrovolski, R., Melo, A.S., Cassemiro, F.A.S. & Diniz-Filho, J.A.F. (2012) Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **21**, 191–197.

Fattorini, S. (2013) Tenebrionid beetle distributional patterns in Italy: multiple colonisation trajectories in a biogeographical crossroad. *Insect Conservation and Diversity*, **7**, 144–160.

Fattorini, S. & Baselga, A. (2012) Species richness and turnover patterns in European tenebrionid beetles. *Insect Conservation and Diversity*, **5**, 331–345.

Fattorini, S. & Ulrich, W. (2012) Spatial distributions of European Tenebrionidae point to multiple postglacial colonization trajectories. *Biological Journal of the Linnean Society*, **105**, 318–329.

Gentile, G. & Argano, R. (2005) Island biogeography of the Mediterranean sea: the species–area relationship for terrestrial isopods. *Journal of Biogeography*, **32**, 1715–1726.

Heiser, M. & Schmitt, T. (2010) Do different dispersal capacities influence the biogeography of the western Palearctic dragonflies (Odonata)? *Biological Journal of the Linnean Society*, **99**, 177–195.

Holt, B.G., Lessard, J.-P., Borregaard, M.K., Fritz, S.A., Araújo, M.B., Dimitrov, D., Fabre, P.-H., Graham, C.H., Graves, G.R., Jønsson, K.A., Nogués-Bravo, D., Wang, Z., Whittaker, R.J., Fjeldså, J. & Rahbek, C. (2013) An update of Wallace's zoogeographic regions of the world. *Science*, **339**, 74–78.

Jaccard, P. (1901) Distribtion de la flore alpine dans le Bassin des Dranses et dans quelques regions voisines. *Bulletin de la Société Vaudoise des Sciences Naturelles*, **37**, 241–272.

Keis, M., Remm, J., Ho, S.Y.W., Davison, J., Tammeleht, E., Tumanov, I.L., Saveljev, A.P., Männil, P., Kojola, I., Abramov, A.V., Margus, T. & Saarma, U. (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. *Journal of Biogeography*, **40**, 915–927.

Koleff, P., Gaston, K.J. & Lennon, J.J. (2003) Measuring beta diversity for presence–absence data. *Journal of Animal Ecology*, **72**, 367–382.

Kreft, H. & Jetz, W. (2010) A framework for delineating biogeographic regions based on species distributions. *Journal of Biogeography*, **37**, 2029–2053.

Kudrna, O., Harpke, A., Lux, K., Pennersdorfer, J., Schweiger, O., Settele, J. & Wiemers, M. (2011) *Distribution atlas of butterflies in Europe*. Gesellschaft für Schmetterlingsschutz, Halle.

López-López, P., Benavent-Corai, J. & García-Ripollés, C. (2008) Geographical assemblages of European raptors and owls. *Acta Oecologica*, **34**, 252–257.

Martin, J. & Gurrea, P. (1990) The peninsula effect in Iberian butterflies. *Journal of Biogeography*, **17**, 85–96.

Mouillot, D., De Bortoli, J., Leprieur, F., Parravicini, V., Kulbicki, M. & Bellwood, D.R. (2013) The challenge of delineating biogeographical regions: nestedness matters for Indo-Pacific coral reef fishes. *Journal of Biogeography*, **40**, 2228–2237.

Sañudo-Restrepo, C.P., Dincă, V., Talavera, G. & Vila, R. (2013) Biogeography and systematics of *Aricia* butterflies (Lepidoptera, Lycaenidae). *Molecular Phylogenetics and Evolution*, **66**, 369–79.

Sfenthourakis, S. (1996) A biogeographic analysis of terrestrial isopods (Isopoda, Oniscidea) from the central Aegean islands. *Journal of Biogeography*, **23**, 687–698.

Shi, G.R. (1993) Multivariate data analysis in palaeoecology and palaeobiogeography: a review. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **105**, 199–234.

Simpson, G.G. (1960) Notes on the measurement of faunal resemblance. *American Journal of Science*, **258-A**, 300–311.

Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab*, **5**, 1–34.

Spengler, A., Hartmann, P., Buchori, D. & Schulze, C.H. (2011) How island size and isolation affect bee and wasp ensembles on small tropical islands: a case study from Kepulauan Seribu, Indonesia. *Journal of Biogeography*, **38**, 247–258.

Strona, G., Galli, P., Seveso, D., Montano, S. & Fattorini, S. (in press) Nestedness for Dummies (NeD): a user friendly web interface for exploratory nestedness analysis. *Journal of Statistical Software*. Software available at: http://ecosoft.alwaysdata.net/.

Stuart, Y.E., Losos, J.B. & Algar, A.C. (2012) The island–mainland species turnover relationship. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 4071–4077.

Tuomisto, H. (2010) A diversity of beta diversities: straightening up a concept gone awry. 2. Quantifying beta diversity and related phenomena. *Ecography*, **33**, 23–45.

Ulrich, W. & Almeida-Neto, M. (2012) On the meanings of nestedness: back to basics. *Ecography*, **35**, 1–7.

Ulrich, W., Almeida-Neto, M. & Gotelli, N.J. (2009) A consumer's guide to nestedness analysis. *Oikos*, **118**, 3–17.

Vandergast, A.G., Perry, W.M., Lugo, R.V. & Hathaway, S.A. (2011) Genetic Landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. *Molecular Ecology Resources*, **11**, 158–161.

Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography: ecology, evolution, and conservation*, 2nd edn. Oxford University Press, Oxford.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Presence data matrix for the butterflies from the 81 island and mainland sites in the western Mediterranean.

Appendix S2 Additional figures (Figures S1–S6).

Appendix S3 R scripts used to perform the analyses.

BIOSKETCH

Members of the research team are actively engaged in insect biogeography, systematics and conservation, with specific interest in unravelling the historical and present-day factors responsible for species distributions across mainland Europe and Mediterranean islands.

Author contributions: L.D., S.F. and R.Vila conceived the ideas; L.D., R.Vila, R.Vodă and V.D. collected the data; L.D, S.F, R.Vodă and V.D. analysed the data; L.D. led the writing; all authors contributed in the form of discussions and suggestions, and approved the final manuscript.

Editor: José María Fernández-Palacios

Chapter 2

Vodă R., Dapporto L., Dincă V. & Vila R. 2014. Cryptic matters: overlooked species generate most butterfly beta-diversity. *Ecography*, **38**: 405–409.

Ecography 38: 405–409, 2015
doi: 10.1111/ecog.00762

Subject Editor: Jeremy T. Kerr. Editor-in-Chief: Carsten Rahbek. Accepted 10 September 2014

Forum

Cryptic matters: overlooked species generate most butterfly beta-diversity

Raluca Vodă, Leonardo Dapporto, Vlad Dincă and Roger Vila

R. Vodă, V. Dincă and R. Vila (roger.vila@csic.es), Inst. de Biologia Evolutiva (CSIC-Univ. Pompeu Fabra), Passeig Marítim de la Barceloneta 37, ES-08003, Barcelona, Spain. R. Vodă also at: Dept de Genètica i Microbiologia, Univ. Autònoma de Barcelona, ES-08193, Bellaterra, Spain. VD also at: Dept of Zoology, Stockholm Univ., SE-106 91, Stockholm, Sweden, and Biodiversity Inst. of Ontario, Univ. of Guelph, Guelph, ON N1G 2W1, Canada. – L. Dapporto, Dept of Biological and Medical Sciences, Oxford Brookes Univ., Headington, Oxford, OX3 0BP, UK.

The cryptic fraction of biodiversity is composed of morphologically similar species that are or have been overlooked by scientists. Although current research is increasingly documenting new cases, cryptic species are frequently ignored in large-scale studies and monitoring programs, either because they have not yet been discovered, or because of the practical difficulties involved in differentiating them. However, it is unknown if this could represent a bias extending beyond the number of missed species. By analyzing the butterfly fauna of the west Mediterranean (335 species), we defined cryptic species based on the current consensus of the scientific community, compared their properties to other congeneric species and investigated the consequences of their inclusion/exclusion in beta-diversity analyses. We show that, as defined, the cryptic fraction of butterfly diversity represents about 25% of the west Mediterranean fauna and is overwhelmingly composed by groups of species that are not sympatric. Our results show that co-occurrence among cryptic species is significantly lower than among congeneric non-cryptic species. Accordingly, albeit the frequency of cryptic species is homogenously distributed over the study area, their distribution pattern accounts for most beta-diversity turnover over sea (from 50 to 100%). Beta-diversity turnover, a direct measure of the frequency of species replacement from site to site, is recognized as a fundamental parameter in ecology and is widely used to detect biogeographic patterns. These findings represent a change of paradigm in showing that cryptic diversity comprises original qualitative aspects in addition to merely quantitative ones. This highlights the importance of differentiating cryptic species for various research fields and opens the door to the study of further potential particularities of cryptic diversity.

Recent molecular studies unveiled a substantial fraction of biodiversity by recognizing a plethora of cryptic species in numerous taxonomic groups and across biogeographical regions (Pfenninger and Schwenk 2007). Cryptic species (Steyskal 1972) are those that have been erroneously treated as conspecific because of high morphological resemblance, and sometimes also behavioral or ecological similarity (Bickford et al. 2007). Thus, the concept of cryptic species is anthropocentric by definition because it is based on human perception of species. Irrespective of the evolutionary significance of delimiting such a group of species, it is justified to ask about the practical consequences of their discovery or omission. To date, no specific studies directly compared the properties of cryptic and non-cryptic diversity, perhaps due to the difficulty of unambiguously identifying cryptic species and their precise distributions for a sufficiently large area and taxonomic group. Consequently, researchers frequently simplify their datasets by not distinguishing cryptic species (Dennis 1997, Isaac et al. 2004, Pfenninger and Schwenk 2007, Descimon and Mallet 2009, Devictor et al. 2012).

Due to their similar natural history and appearance, cryptic species are predicted to compete for resources and to reproductively interfere between them (Pigot and Tobias 2013). Results for a series of cryptic butterfly species revealed only restricted contact zones on mainland and chequered distributions over narrow (3–10 km) sea straits (Dincă et al. 2011, Dapporto et al. 2012, Sañudo-Restrepo et al. 2013). If chequered distributions are overrepresented among cryptic species, they could encompass a disproportionally high fraction of beta-diversity turnover, a direct measure of the frequency of species replacement from site to site. Spatial turnover is recognized as a fundamental parameter in ecology and is increasingly used to detect biogeographic patterns (Holt et al. 2013).

To assess the effects of ignoring cryptic diversity, we studied the biogeographical properties of cryptic species uncovered since 1970, after the cornerstone work of Higgins and Riley (1970) that have been generally accepted by the scientific community. In particular, we explored the prevalence of chequered distribution patterns among groups of cryptic taxa in west Mediterranean butterflies, which represent one of the best-studied invertebrate groups, especially regarding their taxonomy and distribution. We also assessed if cryptic taxa account for an unbalanced fraction of the beta-diversity turnover and concluded that, although they are frequently excluded from large-scale surveys and studies, their contribution may be far more important than the mere percentage of species they represent.

Methods

Presence data and co-occurrence analyses

We analyzed butterfly presence data from 82 areas comprising 27 small islands and 55 sites of 25 km radius located on mainland (43) and on large islands (12) (Supplementary material Dataset 1). An objective identification of species is a crucial step for comparative studies at continental level, especially on a diverse group such as butterflies. Although European butterflies have been among the most studied invertebrates in the last centuries, recognition of species is not always straightforward and it has been based on different concepts for various taxa. In some cases species are readily recognizable based on their external morphology, while in other cases, in depth morphological and molecular analyses, as well as the assessment of reproductive barriers, have been needed (reviewed by Descimon and Mallet 2009). For these reasons, we used a taxonomic species concept (reviewed by Mayden 1997) recognizing species as 'whatever a competent taxonomist chooses to call a species' (Wilkins 2009). We applied this criterion and compared past taxonomy with the current consensus of the scientific community in the following way.

1) We produced a list of species based on the widely accepted checklist of the Fauna Europaea Project (Karsholt and Nieukerken 2013, <www.faunaeur.org>), funded by the European Commission (EVR1-1999-2001). We complemented this list with the north African species recognized in the recent synthesis works that include this region (Tennent 1996, Tarrier and Delacre 2008, Tshikolovets 2011). We also added four species (*Zerynthia cassandra*, *Polyommatus celina*, *Coenonympha lyllus*, *Coenonympha darwiniana*) and eliminated one (*Polyommatus galloi*) based on recent molecular studies, generally combined with morphometrics, mate choice experiments and/or ecological data (see Supplementary material Appendix 1 for references). These changes have not yet been introduced in Fauna Europaea but will most likely be in the next version.

2) Within the list of species obtained, we considered as cryptic those pairs or groups of species that have been considered as a single species in at least one of the most widely used synthesis works regarding Europe and/or north Africa published since the reference work of Higgins and Riley (1970) (see Supplementary material Appendix 1 for discussions on each case). The references provided should be regarded as examples and usually cover general works that likely represent widely accepted trends in European butterfly taxonomy. In specific cases where relevant studies were available, these have been cited as well.

Co-occurrence between pairs of species has been computed as the fraction of shared areas by using the Sørensen index QS = 2C/(A + B), where C is the number of areas where the two species co-occur and A and B are the number of areas where the first and second species are present, respectively. We computed the QS between pairs of cryptic species and between pairs of non-cryptic species belonging to the same genus. To test for significant differences in QS values we performed the Mann–Whitney U test with 1000 Monte Carlo permutations.

Beta-diversity turnover

To evaluate the Contribution of Cryptic Species to the beta-diversity Turnover (CCST), we calculated the Simpson index among the 82 study areas. Subsequently, for each presence submatrix composed of each cryptic group, we computed a matrix (sm) for areas by using the Simpson dissimilarity denominator. Finally, we added all sm and divided this sum matrix by the total Simpson matrix (also containing the contribution of non-cryptic species). Following the previous description, CCST can be computed by the formula:

$$CCST = \frac{\sum_{sm=1}^{36}\left(A_{sm} + \min\left(B_{sm}, C_{sm}\right)\right)}{A - \min(B, C)}$$

where A_{sm} is the number of shared species and B_{sm} and C_{sm} are the number of non-shared species among areas for each of the 36 cryptic groups considered separately (sm). A is the overall number of shared species and B and C the overall number of non-shared species among the same pairs of the 82 areas. With the 'deldir' R package we calculated a Delaunay triangulation among areas using their center GPS coordinates and plotted the CCST index on a map. The CCST among pairs of areas were attributed to the midpoints connecting those areas and were interpolated through Inverse-Distance-Weighting using Spatial Analyst (<www.esri.com>).

To evaluate the importance of cryptic species in determining the overall turnover pattern, we calculated the mean value of the ratio between the Simpson and the Sørensen dissimilarity matrices (βsimp/βsor) among areas. This ratio is considered to be the best measure describing the contribution of turnover to the overall beta-diversity pattern (Dobrovolski et al. 2012). We then removed all cryptic taxa and recalculated the same mean value. Finally, we created 10 000 matrices by removing a random number of species equal to the number of cryptic ones and calculated their mean βsimp/βsor. We tested if the matrix lacking all cryptic taxa shows a significantly lower ratio than the random ones (if at least 95% of the random matrices showed a βsimp/βsor higher than the matrix without all cryptic species).

Results

Co-occurrence analyses for cryptic and non-cryptic diversity fractions

In total, 86 species out of 335 (25.67%) are involved in 36 cryptic groups. In most groups (20), cryptic species have been separated only on the basis of differences in morphology (mostly male genitalia), while in six groups species were

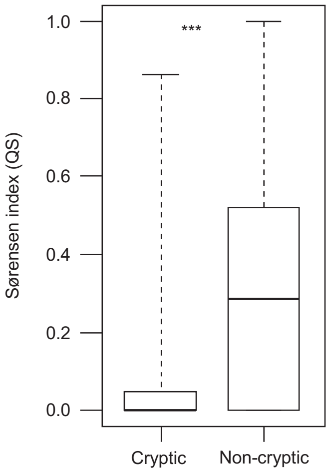


Figure 1. Boxplot indicating median values of the Sørensen index (QS) for pairs or groups of cryptic species (Cryptic) and congeneric pairs of non-cryptic species (Non-cryptic). Cryptic species share a significantly smaller fraction of their distributions compared with the non-cryptic ones.

recognized based exclusively on molecular evidence (DNA, allozyme or number of chromosomes) because no relevant differences in morphology have been detected. In nine groups morphological and molecular evidence have been coupled, and in one group (*Leptidea*) morphological and molecular analysis have been combined with mating experiments (Supplementary material Appendix 1). The Mann–Whitney test ($U = 15319.5$, Monte Carlo $p < 0.001$) revealed that cryptic species pairs have significantly lower QS than non-cryptic pairs of species (Fig. 1), which means that they share a smaller fraction of their distributions.

Contribution of non-sympatric cryptic species to the overall beta-diversity pattern

Although the fraction of cryptic diversity was relatively constant among the 82 study areas (23.02 ± 0.04, Fig. 2a), the fraction of beta-diversity turnover due to replacements between cryptic species showed pronounced variation over the west Mediterranean (Fig. 2b). As a general rule, the CCST was higher than 50% only over sea areas, with especially high values of more than 80% between Sardinia-Corsica-Italian mainland, Sicily-Italian mainland, Sicily-north Africa and Balearics-Catalonia-France. By contrast, the CCST was lower (about 10%) between sites separated by land (Fig. 2b). The complete presence matrix revealed a mean βsimp/βsor value of 0.450 (Fig. 2c). When all the 86 cryptic species have been removed the mean value decreased to 0.346, while the mean value for cryptic species alone was 0.666. None of the 10 000 matrices generated by removing 86 random species showed a βsimp/βsor mean value lower than the matrix lacking the 86 cryptic species ($p < 0.001$, Fig. 2c).

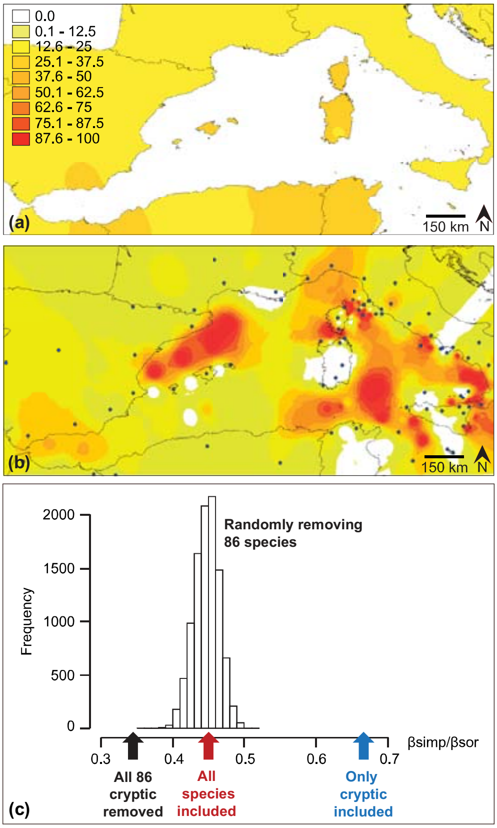


Figure 2. (a) Distribution map (equirectangular projection) of the percentage that cryptic butterfly species represent of the total diversity. (b) Landscape analysis for the Contribution of Cryptic Species to the beta-diversity Turnover (CCST). (c) Mean βsimp/βsor value of the complete matrix (red arrow) compared with the value obtained by removing all 86 cryptic species (black arrow) and by using only cryptic species (cyan arrow). The histogram represents frequencies of βsimp/βsor mean values for 10 000 matrices generated by removing 86 random species.

Discussion

Due to their morphological similarity and complex taxonomy, cryptic groups are frequently ignored or treated as a unit in large-scale studies and monitoring schemes. In this study we objectively defined cryptic species and showed that co-occurrence within cryptic groups over the west Mediterranean is much lower than among congeneric non-cryptic species. This result could be explained if we assume that cryptic groups are composed by entities that recently evolved in allopatry and maintained their distribution pattern simply because they did not have enough time to disperse. However, many cryptic taxa in our study are not sister species and show substantial genetic divergence, while other morphologically differentiated species are their sister taxa; in most such cases a chequered distribution still occurs (Dincă et al. 2011, Sañudo-Restrepo et al. 2013).

The mechanisms maintaining chequered distributions across specific narrow sea straits are still unknown, but a convincing hypothesis can be formulated. Recent studies have shown that butterflies can experience range expansions/contractions even over relatively short time periods (Dapporto and Bruschini 2012, Dapporto et al. 2012). It is thus likely that, in addition to temporal and physical constraints, the dispersal of most cryptic species has been hampered by the presence of other members of the same group in the recipient areas (Waters et al. 2013), a phenomenon recently demonstrated in birds (Pigot and Tobias 2013). Presumably, no strong precopulatory barriers exist between many cryptic butterfly species (Wickman 2009) and hybrids between related taxa often display reduced fitness (Descimon and Mallet 2009). In the absence of specific mate recognition mechanisms, contact areas may act as population sinks unlikely to enlarge, although they are predicted to shift until areas allowing low dispersal and low population densities are reached (Dasmahapatra et al. 2010). A concentration of boundaries between cryptic species over sea straits perfectly matches this hypothesis. Noteworthy, the fraction of cryptic diversity is homogeneous across the study region and almost all cryptic taxa that we analyzed occur on both mainland and islands. For this reason, the striking contribution to island divergence is not due to an unbalanced occurrence of island endemics among cryptic taxa. These mechanisms are generally applicable to most organisms, and thus it is probable that the phenomenon here discovered for butterflies holds for other taxa as well, albeit the exact impact on the overall patterns may vary.

We conclude that unveiling the cryptic fraction of biodiversity is of primary importance for a comprehensive description of biogeographical patterns. The finding that cryptic groups of species present characteristic geographical distributions highlights the need to study further potential particularities they may display, as well as to assess their implications in multiple research fields. We showed that ignoring cryptic species results in losing a significant fraction of spatially structured taxonomic variation, which hampers the opportunity to reveal phylogeographic suture zones (Moritz et al. 2009) and to investigate historical and ecological processes determining distribution range dynamics (Marske et al. 2013). Such phenomena underlie fundamental processes like dispersal, extinction, environmental filtering, competition, sexual interference, hybrid depression, and 'founder takes all' dynamics (Marske et al. 2013, Waters et al. 2013). Moreover, it is likely that cryptic species respond differently to climatic conditions (Pearman et al. 2010) and geographical shifts in their contact zones correlated with climate change have been documented (Mercader and Scriber 2008, Mallet et al. 2011). This suggests that studies quantifying the effects of global climatic changes on organisms may be missing part of the signal stemming from cryptic biodiversity. The same could apply to life history and other ecological preferences that may be differentially expressed in morphologically similar species. Morphological differentiation may well be the result of diverging selective pressures, which could generally lack in these groups of species. Different evolutionary and speciation factors and processes may result in varying degrees of morphological differentiation (*Heliconius* genome consortium 2012), and thus phylogenetic reconstruction and macroecological inference could be affected due to underrepresented cryptic diversity.

Case studies typically rely on well-assessed taxonomy, but macroecology and comparative studies are usually based on data that encompass a great number of species and are more susceptible of missing cryptic taxa. The costs and benefits of taking into consideration cryptic species have been rarely discussed due to a reductionist view of their importance limited to alpha-taxonomy, i.e. species numbers (Isaac et al. 2004). From this perspective, our study not only provides novel insights into biogeography, but it also serves as a heads-up about the importance of including the cryptic diversity in ecological, biodiversity and biogeographic studies to avoid biased results.

Acknowledgements – Funding was provided by the Spanish Ministerio de Ciencia e Innovación (Project CGL2010-21226/BOS and a predoctoral fellowship to R. Vodă, ref. AP-2010-5409), by the project 'Barcoding Italian butterflies' and by a postdoctoral fellowship from the Wenner-Gren Foundation Sweden and a EU Marie Curie International Outgoing Fellowship for Career Development (project no. 625997) to VD.

References

Bickford, D. et al. 2007. Cryptic species as a window on diversity and conservation. – Trends Ecol. Evol. 22: 148–155.

Dapporto, L. and Bruschini, C. 2012. Invading a refugium: post glacial replacement of the ancestral lineage of a Nymphalid butterfly in the west Mediterranean. – Org. Divers. Evol. 12: 39–49.

Dapporto, L. et al. 2012. Identifying zones of phenetic compression in west Mediterranean butterflies (Satyrinae): refugia, invasion and hybridization. – Divers. Distrib. 18: 1066–1076.

Dasmahapatra, K. K. et al. 2010. The anatomy of a 'suture zone' in Amazonian butterflies: a coalescent-based test for vicariant geographic divergence and speciation. – Mol. Ecol. 19: 4283–4301.

Dennis, R. L. H. 1997. An inflated conservation load for European butterflies: increases in rarity and endemism accompany increases in species richness. – J. Insect Conserv. 1: 43–62.

Descimon, H. and Mallet, J. 2009. Bad species. – In: Settele, J. et al. (eds), Ecology of butterflies in Europe. Cambridge Univ. Press, pp. 219–249.

Devictor, V. et al. 2012. Differences in the climatic debts of birds and butterflies at a continental scale. – Nat. Clim. Change 2: 121–124.

Dincă, V. et al. 2011. A combined genetic-morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and *Polyommatus celina* common blue butterflies. – Mol. Ecol. 20: 3921–3935.

Dobrovolski, R. et al. 2012. Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. – Global Ecol. Biogeogr. 21: 191–197.

Heliconius Genome Consortium 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptation among species. – Nature 487: 94–98.

Higgins, L. G. and Riley, N. D. 1970. A field guide to the butterflies of Britain and Europe. – Houghton Mifflin, 1st American ed.

Holt, B. et al. 2013. An update of Wallace's zoogeographic regions of the World. – Science 339: 74–78.

Isaac, N. J. B. et al. 2004. Taxonomic inflation: its influence on macroecology and conservation. – Trends Ecol. Evol. 19: 464–469.

Karsholt, O. and Nieukerken, E. J. van. 2013. Lepidoptera, Papilionoidea. – Fauna Europaea ver. 1.0, <www.faunaeur.org>

Mallet, J. et al. 2011. Hybridisation and climate change: brown argus butterflies in Britain (*Polyommatus* subgenus *Aricia*). – Insect Conserv. Divers. 4: 192–199.

Marske, K. A. et al. 2013. Phylogeography: spanning the ecology–evolution continuum. – Ecography 36: 1169–1181.

Mayden, R. L. 1997. A hierarchy of species concepts: the denouncement in the saga of the species problem – In: Claridge, M. F. et al. (eds), Species: the units of biodiversity. Chapman and Hall, pp. 381–424.

Mercader, R. J. and Scriber, J. M. 2008. Asymmetrical thermal constraints on the parapatric species boundaries of two widespread generalist butterflies. – Ecol. Entomol. 33: 537–545.

Moritz, C. et al. 2009. Identification and dynamics of a cryptic suture zone in tropical rainforest. – Proc. R. Soc. B 276: 1235–1244.

Pearman, P. B. et al. 2010. Within-taxon niche structure: niche conservatism, divergence and predicted effects of climate change. – Ecography 33: 990–1003.

Pfenninger, M. and Schwenk, K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. – BMC Evol. Biol. 7.

Supplementary material (Appendix ECOG-00762 at <www.ecography.org/readers/appendix>). Appendix 1 and Dataset 1.

Pigot, A. L. and Tobias, J. A. 2013. Species interactions constrain geographic range expansion over evolutionary time. – Ecol. Lett. 16: 330–338.

Sañudo-Restrepo, C. P. et al. 2013. Biogeography and systematics of *Aricia* butterflies (Lepidoptera, Lycaenidae). – Mol. Phylogenet. Evol. 66: 369–379

Steyskal, G. C. 1972. The meaning of the term “sibling species”. – Syst. Zool. 21: 239.

Tarrier, M. R. and Delacre, J. 2008. Les papillons de jour du Maroc. – Biotope, Mèze.

Tennent, J. 1996. The butterflies of Morocco, Algeria and Tunisia. – Gem Publishing, Walling for.

Tshikolovets, V. V. 2011. Butterflies of Europe and the Mediterranean area – Tshikolovets Publications.

Waters, J. M. et al. 2013 Founder takes all: density dependent processes structure biodiversity. – Trends Ecol. Evol. 28: 78–85.

Wickman, P.-O. 2009. Mating behaviour in butterflies. – In: Settele, J. et al. (eds), Ecology of butterflies in Europe. Cambridge Univ. Press, pp. 17–28.

Wilkins, J. S. 2009. Species: a history of the idea. – Univ. of California Press.

Chapter 3

Vodă R., Dapporto L., Dincă V. & Vila R. 2014. Why Do Cryptic Species Tend Not to Co-Occur? A Case Study on Two Cryptic Pairs of Butterflies. *PLoSONE*, **10**: e0117802. doi:10.1371/journal.pone.0117802.

PLOS | ONE

RESEARCH ARTICLE

Why Do Cryptic Species Tend Not to Co-Occur? A Case Study on Two Cryptic Pairs of Butterflies

Raluca Vodă[1,2], Leonardo Dapporto[3], Vlad Dincă[1,4], Roger Vila[1]*

1 Butterfly Diversity and Evolution Lab, Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Spain, **2** Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, Spain, **3** Department of Biological and Medical Sciences, Oxford Brookes University, Headington, Oxford, United Kingdom, **4** Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada

* roger.vila@csic.es

Abstract

As cryptic diversity is being discovered, mostly thanks to advances in molecular techniques, it is becoming evident that many of these taxa display parapatric distributions in mainland and that they rarely coexist on islands. Genetic landscapes, haplotype networks and ecological niche modeling analyses were performed for two pairs of non-sister cryptic butterfly species, *Aricia agestis*-A. *cramera* and *Polyommatus icarus—P. celina* (Lycaenidae), to specifically assess non-coexistence on western Mediterranean islands, and to test potential causes producing such chequered distribution patterns. We show that the morphologically and ecologically equivalent pairs of species do not coexist on any of the studied islands, although nearly all islands are colonized by one of them. According to our models, the cryptic pairs displayed marked climatic preferences and 'precipitation during the driest quarter' was recovered as the most important climatic determinant. However, neither dispersal capacity, nor climatic or ecological factors fully explain the observed distributions across particular sea straits, and the existence of species interactions resulting in mutual exclusion is suggested as a necessary hypothesis. Given that the studied species are habitat generalists, feeding on virtually unlimited resources, we propose that reproductive interference, together with climatic preferences, sustain density-dependent mechanisms like "founder takes all" and impede coexistence on islands. Chequered distributions among cryptic taxa, both sister and non-sister, are common in butterflies, suggesting that the phenomenon revealed here could be important in determining biodiversity patterns.

OPEN ACCESS

Citation: Vodă R, Dapporto L, Dincă V, Vila R (2015) Why Do Cryptic Species Tend Not to Co-Occur? A Case Study on Two Cryptic Pairs of Butterflies. PLoS ONE 10(2): e0117802. doi:10.1371/journal.pone.0117802

Academic Editor: Carlos López-Vaamonde, Institut National de la Recherche Agronomique (INRA), FRANCE

Received: February 24, 2014

Accepted: December 31, 2014

Published: February 18, 2015

Funding: Funding for this research was provided by the Spanish Ministerio de Economía y Competitividad (Projects CGL2010-21226/BOS and CGL2013-48277-P and a predoctoral fellowship to R. Vodă, AP-2010-5409), by the European Commission through a Marie Curie International Outgoing Fellowship for Career Development to V. Dincă (project no. 625997), and by the projects "Barcoding Italian Butterflies" and "Barcoding Butterflies of the Tuscan Archipelago National Park." The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Introduction

Adaptation to different abiotic elements (e.g. climate, geomorphology and soil) and biotic resources determines the fundamental niche of a species. The occupancy of this theoretically suitable space is in turn constrained by other abiotic and biotic factors (such as geographic barriers and species interactions) that shape the realized niche of a species [1–4]. In the last decade,

Competing Interests: The authors have declared that no competing interests exist.

considerable progress has been made to provide researchers with essential data (climatic, environmental and distributional, e.g. GBIF http://www.gbif.org/, BIOCLIM http://www.worldclim.org/bioclim [5], Socioeconomic Data And Applications Center (SEDAC) http://sedac.ciesin.columbia.edu/) and methodological tools to unravel the relative importance of biotic and abiotic factors in determining the observed distribution of species [6–9]. Among all concurrent effects, a direct evaluation of biotic interactions such as competition and reproductive interference requires particularly complex and long-term studies [10–13]. For this reason, in most cases, the importance of species interactions is indirectly evaluated through comparative studies or by testing the power of models that, in addition to climatic and environmental variables, include biotic variables potentially correlated with the supposed interaction, such as the presence of the presumably competing species [9,14–18].

A recent study on passerine birds examined a series of sister species and revealed that rates of secondary sympatry are positively associated with both phylogenetic and morphological distance between species. The authors suggested that competition between ecologically similar species limits their possibilities to occur in sympatry [17]. The use of a large number of sister species allowed a comprehensive comparative analysis, but eliminated the possibility to examine the nature of the interactions between non-sister species that can also be morphologically and ecologically similar, despite being phylogenetically relatively distant. A recent comparative study of the entire butterfly fauna of the western Mediterranean revealed that complexes of cryptic taxa show lower co-occurrence than other congeneric non-cryptic species [19]. Among these, several examples of non-sister cryptic taxa with chequered distributions have been reported. Interestingly, such ecologically and morphologically similar species tend to be parapatric on mainland and apparently many display chequered distributions on islands, even over narrow sea straits [20–25]. Such a pattern could represent a signal of interspecific interactions [17,26,27].

Here, we indirectly test for the existence of mutual exclusion, using as a model two pairs of non-sister cryptic butterfly species: *Polyommatus icarus*-*P. celina*, and *Aricia agestis*-*A. cramera*. The genetic structure and evolutionary relationships of these species have been recently documented [21,22]. *Polyommatus icarus* and *P. celina* are parapatric, habitat generalist species, are found over a broad altitudinal range, and feed on a wide array of host plants (Fig. 1A). *Polyommatus celina* has been only recently recognized as a distinct species, and it can be reliably distinguished from *P. Icarus* based on both nuclear and mitochondrial genetic markers, while morphological traits only show minor trends of variation[21,28]. However, *P. icarus* is phylogenetically closer to *P. eros*, a genetically and morphologically differentiated mountain taxon that feeds on a more restricted number of host plants [21,28]. Similarly, *A. agestis* is a habitat generalist and ubiquitous Palaearctic species that is phylogenetically closer to the boreo-montane *A. artaxerxes* and the montane *A. montensis*. The sister of this clade is *A. cramera*, which occurs in the south-western Mediterranean and is almost identical to *A. agestis* in ecology and external morphology, but it is differentiated genetically and in the male genital morphology[22] (Fig. 1B). The recognition of these taxa as species or subspecies has been debated [29,30], but what it is important in our case is that the phylogenetically divergent pairs are parapatric and show higher morphological similarity than the sympatric and phylogenetic closer taxa. Moreover, the parapatric pairs are habitat generalists [31] that occur from sea level to high mountains, both in highly anthropic and in non-managed areas, and their varied host plants represent a virtually continuous and unlimited resource over space. Thus, it is unlikely that their distributions are constrained by habitat quality and host plant presence [10].

We sequenced a large number of specimens for all four cryptic species from the entire western Mediterranean region, with a special effort on the currently known contact areas (Iberia, the Sicily channel, the Messina channel and the strait of Bonifacio). We combined phylogeny,

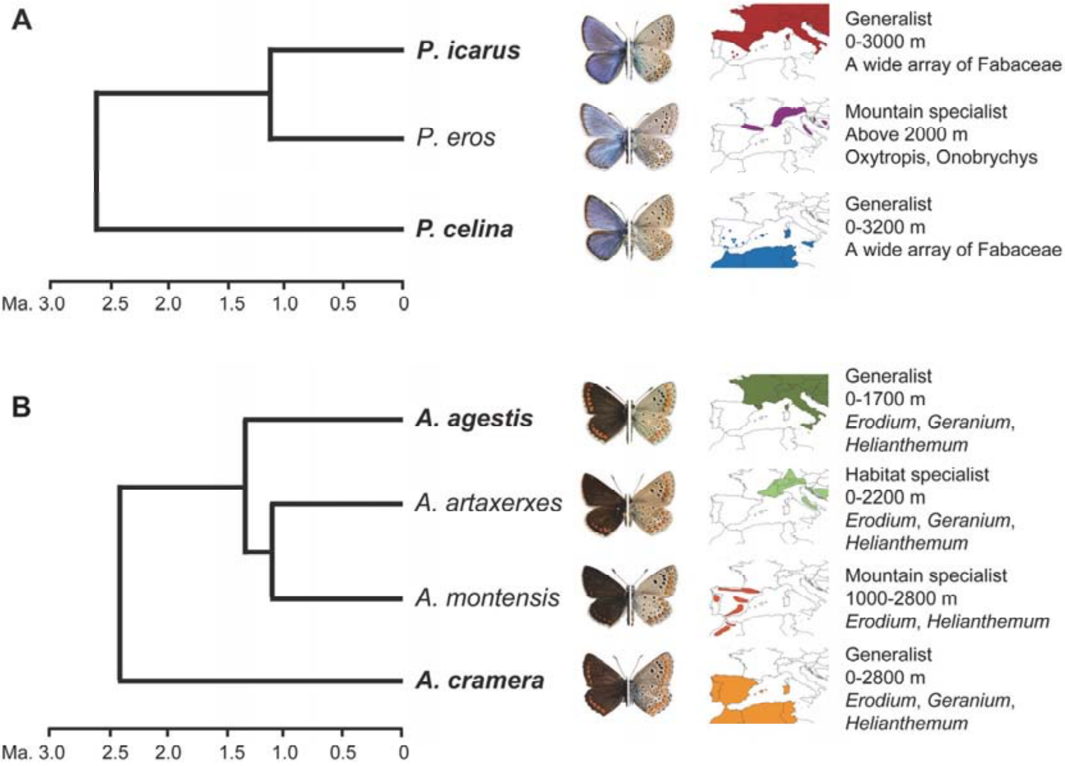


Fig 1. Model species. A. The evolutionary relationships, external appearance, western Mediterranean distribution and ecological preferences for *Polyommatus icarus*, *P*. *eros* and *P*. *celina*. *Polyommatus icarus* and *P*. *celina* are almost identical ecologically and morphologically, but *P*. *icarus* is phylogenetically closer to *P*. *eros*, a morphologically differentiated mountain species that feeds on a more restricted number of host plants. **B**. The evolutionary relationships, external appearance, western Mediterranean distribution and ecological preferences for *Aricia agestis*, *A*. *artaxerxes*, *A*. *montensis* and *A*. *cramera*. *Aricia agestis* is a habitat generalist phylogenetically the closest to a pair of specialist species:*Aricia artaxerxes* (boreo-montane) and *A*. *montensis* (montane). *Aricia cramera* is phylogenetically more distant to *A*. *agestis*, although both species are almost identical ecologically and morphologically.

doi:10.1371/journal.pone.0117802.g001

distribution modeling, genetic landscapes, and the evaluation of dispersal capabilities for each taxon and showed that the biogeographical patterns consistent with mutual exclusion for these cryptic species cannot be fully explained either by climatic preferences or by limited dispersal capabilities.

Methods

Ethics Statement

No specific permits were required for the studied areas in France, Italy and North Africa (Morocco, Algeria and Tunisia), because the locations were not privately owned or protected in any way, and the field studies did not involve endangered or protected species. All necessary permits for the field studies in Spain were obtained from the competent public bodies: environmental agencies of the Comunidades Autónomas, the Natural Parks of Sierra Nevada, Montseny and Picos de Europa, and Reserva Natural Illes Columbretes.

Molecular data and genetic landscape

We analyzed the cytochrome *c* oxidase I(COI) mitochondrial marker for 325 *P. icarus-P. celina* and 262 *A. agestis-A. cramera* specimens from the western Mediterranean mainland and most of the islands in this region (S1 Table). Total genomic DNA was extracted using Chelex 100 resin, 100–200 mesh, sodium form (Biorad), under the following protocol: one leg was removed and introduced into 100 μL of Chelex 10% and 5 μL of Proteinase K (20 mg/mL) were added. The samples were incubated overnight at 55°C and were subsequently incubated at 100°C for 15 minutes. Samples were then centrifuged for 10 seconds at 3.000 rpm. A 658-bp fragment at the 5' end of the mitochondrial gene (COI) was amplified by polymerase chain reaction using the primers LepF1 (5´-ATTCAACCAATCATAAAGATATTGG-3´) and LepR1 (5´-TAAACTTCTGGATGTCCAAAAAATCA-3´) [32]. Double-stranded DNA was amplified through polymerase chain reactions (PCR) in 25-μL volumes containing: 14.4 μL autoclaved Milli-Q water, 5 μL 5x buffer, 2 μL 25 mM MgCl2, 0.5 μL 10 mM dNTPs, 0.5 μL of each primer (10 μM), 0.1 μL Taq DNA Polymerase (Promega, 5U/ μL) and 2 μL of extracted DNA. The typical thermal cycling profile was: first denaturation at 92°C for 60 s, followed by five cycles of 92°C for 15 s, 49°C for 45 s and 62°C for 150 s, and then by 35 cycles of 92°C for 15 s, 52°C for 45 s and 62°C for 150 s and a final extension at 62°C for 420 s. PCR products were purified and sequenced by Macrogen Inc. Sequences were edited and aligned using GENEIOUS PRO 6.0.5 created by Biomatters (http://www.geneious.com/). A part of the sequences generated by this study have been obtained at the Biodiversity Institute of Ontario, Canada. In this case a glass fibre protocol [33] was employed to extract DNA and PCR and DNA sequencing were carried out following standard DNA barcoding procedures for Lepidoptera [34]. All new sequences have been deposited in GenBank (accession numbers KM459029—KM459439, and KP052710). Neighbour-Joining (NJ) phylogenetic trees for identification purposes were obtained using MEGA 5.05 [35], with 100 bootstrap pseudo-replicates to assess the robustness of the tree clades.

To create genetic landscapes for each species pair, a matrix of p-distances and a table of GPS coordinates (decimal degrees format) for the corresponding samples were imported in R 3.0.2. When more than one sample belonged to the same location, mean genetic distances have been computed. Using the "deldir" package, we calculated a Delaunay triangulation among the GPS coordinates of the collection sites (S1 Fig.). Since the segments composing the triangles have different lengths and genetic distances tend to increase with geographic distance following different trends according to organisms and scale [36–38], we computed a series of preliminary regressions (linear with original values, linear with log-transformed values, and asymptotic regression) and selected the one showing the highest fit (R^2). Subsequently, we computed residuals of the selected regressions between dissimilarities and geographical distances. For all pairs of areas connected in the Delaunay triangulations, we attributed the residual p-distance calculated between those areas to the midpoint of each segment [39]. The residual values and midpoint locations were imported in QGIS 2.0.1. (www.qgis.org), and the values interpolated using the inverse distance weighting algorithm to generate a visual representation of the spatial distribution of genetic divergence [39].

Haplotype networks and evaluation of dispersal constraints

Haplotype networks for each species were inferred with the program TCS 1.21 from subsets consisting exclusively of sequences without ambiguities (S1 Table). On the basis of the relationships highlighted by these networks, it has been possible to identify a minimum number of sea crossing events necessary to produce the observed pattern of distributions. In practice, by assuming that genetic convergence is much less probable than dispersal, we scored a cross-sea dispersal event when a haplotype was shared between two areas separated by a sea strait. We also scored a dispersal event if two haplotypes occurring in areas separated by sea straits were directly linked in the haplotype network. When calculating the sea crossing capacity of the species, we used the-50m isobath, representing an approximate mean sea depth between glacial maxima and interglacial periods. Some of the investigated islands were connected to mainland areas according to this isobath (Elba, Pianosa and Ischia to Italy, Levant to France, Levanzo to Sicily) and hence no dispersal was scored in these cases. We conservatively excluded particularly long dispersal events when the genetic pattern could also be explained by extinction or non-detection of some haplotypes. As an additional source of information, we used island groups that are outside the studied area, but where some of the four studied species occur (eastern Mediterranean for *P. icarus* and *A. agestis*, Canary Islands for *P. celina* and *A. cramera*), and reconstructed the minimum number and shortest over-sea dispersal events required for these species to achieve the observed distributional pattern. Finally, we plotted the frequency of the observed dispersal lengths in order to verify if the unobserved events determining the chequered distributions lie within the inferred dispersal capability of each species.

Ecological niche modeling

In order to test if the observed chequered distributions of the species pairs can be explained by climatic factors, we performed ecological niche modeling for each of the four species. Based on molecular results, reliable observations, and data from literature, we gathered over 7000 presence points for these species in the study area. For areas close to the contact zones we exclusively used presence points based on molecular results. In order to eliminate clustered occurrences due to unbalanced sampling, we filtered the datasets by randomly selecting points at a minimum distance of 0.5 degrees using the function gridSample of the R package “dismo”. The selection of 0.5 degrees has been made as a visual best compromise between maintaining the largest possible number of occurrence data and a good level of homogeneity. After filtering, we obtained 599 records for *P. icarus*, 128 for *P. celina*, 444 for *A. agestis* and 322 for *A. cramera* (S1 Dataset). We downloaded the 19 climatic layers from WorldClim (http://www.worldclim.org/, [5]) at a resolution of 30 arc-seconds. We cropped the layers to include only the western Mediterranean area and downscaled their resolution to 0.1 degrees of latitude and longitude to match the resolution of occurrence data. Climatic variables tend to be highly correlated mainly when regions with similar climate are analyzed. Collinearity does not largely affect the prediction of occurrence [40], but it may bias the estimate of the relative importance of predictors [6]. We evaluated the correlation among variables by randomly selecting 10,000 points of the WorldClim variables and analyzed the Pearson correlation between all pairs of variables. By inspecting the correlation matrix, we selected the maximum possible number of variables showing a Pearson correlation coefficient lower than 0.8 [41,42]. Eight out of 19 variables were retained: mean temperature diurnal range (°C), isothermality (°C), temperature seasonality (coefficient of variation in %), mean temperature of wettest quarter (°C), mean temperature of driest quarter (°C), precipitation of wettest period (mm), precipitation of driest quarter (mm) and precipitation of coldest quarter (mm). We then performed the ecological niche modeling analyses using the maximum entropy algorithm MaxEnt v3.3.3k. When using the Auto features, MaxEnt [6] generates five classes of predictor variables (linear, quadratic, product, threshold and hinge) often resulting in data over-fitting [43], and thus producing non-conservative results. According to recent reviews we performed two series of models. In the first series we only used hinge features (H models), which produce similar results to the generalized additive models (GAM)[44]. In the second series we used hinge, quadratic and product features (HQP models) because hinge features tend to be redundant with linear and threshold ones [44]. The distribution of each species was modeled using 100 replicates, subsample and random seed, with 25% of the presence data to test the model and 75% to train the model. All other settings were left by default. To plot the maps for the predicted species occurrences we considered as potentially suitable areas, those cells showing logistic values higher than the maximum training sensitivity plus specificity logistic threshold [45]. In order to test the importance of species interactions, we performed another set of species distribution modeling analyses in which we used the presence of the corresponding cryptic species as an additional categorical layer. The new layers were generated by alpha-convex hulls based on the presence data of each species. In order to obtain the best compromise between local convexity and the actual distribution of species we used the function alhull of the R package "alphahull" with the alpha value set to ten. Each 0.1 degree cell of the layers was categorized as 1 if internal to the convex hull (species occurring in the area of the cell) and as 0 if external (species not recorded to occur in the area of the cell). The convex hull layer of each species was then included as a categorical variable in a new MaxEnt analysis of its corresponding cryptic species.

In order to evaluate the relative importance of highly correlated predictors in the models, we inspected both the percent contribution of each variable to the model and the Jackknife output produced by MaxEnt. Because the Jackknife regularized training gain for models performed without one of the variables are expected uninformative in case of high collinearity, we used instead the training gain of the models, which tested the performance of each variable individually [41]. The least important variable returned by the climatic models was removed from the interaction analyses in order to use the same number of variables in both series of analyses.

Results

Locating hotspots of mutual exclusion: genetic landscape

The phylogenetic trees based on COI sequences (S2 Fig.) resulted in clearly differentiated clades for each species in accordance with previous studies [21,22]. Among the regressions used to correct the genetic distances according to geographic distances, the relationship of original p-distances against log-transformed geographic distances showed the best fit. As expected for the reduced geographic scale involved in the Delaunay triangulation, the asymptotic regression failed in finding any significant solution. We thus computed and interpolated the residuals between p-distance and log-geographic distance. The resulting genetic divergence landscapes and the assignment of individuals to a species based on their position in the phylogenetic tree indicated, for both cryptic pairs, parapatry on mainland with a contact zone in the Iberian Peninsula, and not a single case of coexistence on islands (Fig. 2). For *Polyommatus*, the strongest genetic divergence with respect to geographic distance corresponded to abrupt changes of distributions across the narrow Messina and Bonifacio straits, along the Tyrrhenian Sea between Italy and Sardinia, between the Balearics and Iberia and between northern and southern Iberia. In southern Iberia the picture became more complex due to the existence of isolated populations of *P. Icarus* in particular mountaintops, such as Sierra Nevada and Sierra de La Sagra. For *Aricia*, the highest divergence emerged across the Bonifacio strait, along the Tyrrhenian Sea, between North Africa and Sicily, between the Balearics and France, in Catalonia and along the Pyrenees. As a result of intraspecific divergence, minor differentiation was also found between Corsica and the Tuscan Archipelago and between Sicily and the Italian Peninsula. In summary, for both species pairs the most pronounced genetic differences were located over sea areas, confirming that sea straits have strong power in the formation and maintenance of non-sympatric distributions.

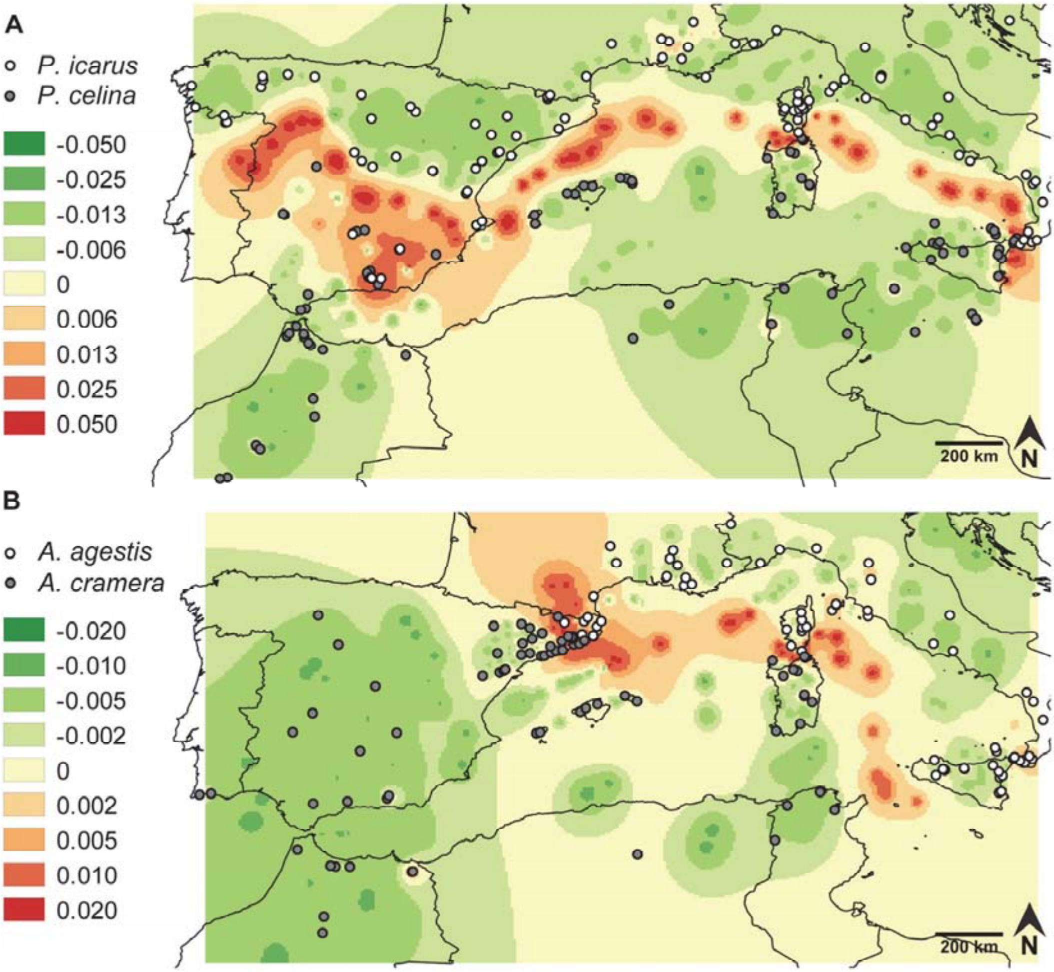


Fig 2. Genetic landscapes obtained for the sequenced specimens of the two pairs of cryptic species. The color gradient represents residuals of COI genetic p-distances. **A**. *Polyommatus icarus* and *P*. *celina*. **B**. *Aricia agestis* and *A*. *cramera*.

doi:10.1371/journal.pone.0117802.g002

Evaluation of dispersal constraints based on haplotype networks

Haplotype networks of the studied species showed different degrees of complexity, but in all cases intraspecific divergences were small enough so that the biogeographical history could be explained by Quaternary environmental changes and by the dispersal capabilities of the butterflies, instead of vicariance caused by older geological events. Regardless of the hypothetical location of the ancestor, the most parsimonious series of dispersal events that could have produced the observed patterns are highlighted in Fig. 3, S3 and S4 Figs. *Aricia cramera* showed a simple pattern with a widespread haplotype and some satellite ones connected by a single substitution (Fig. 3). Among dispersal events that could have produced the observed distributions (Fig. 3 and S3a Fig.), we retained that the pattern showed by haplotype hcr5,

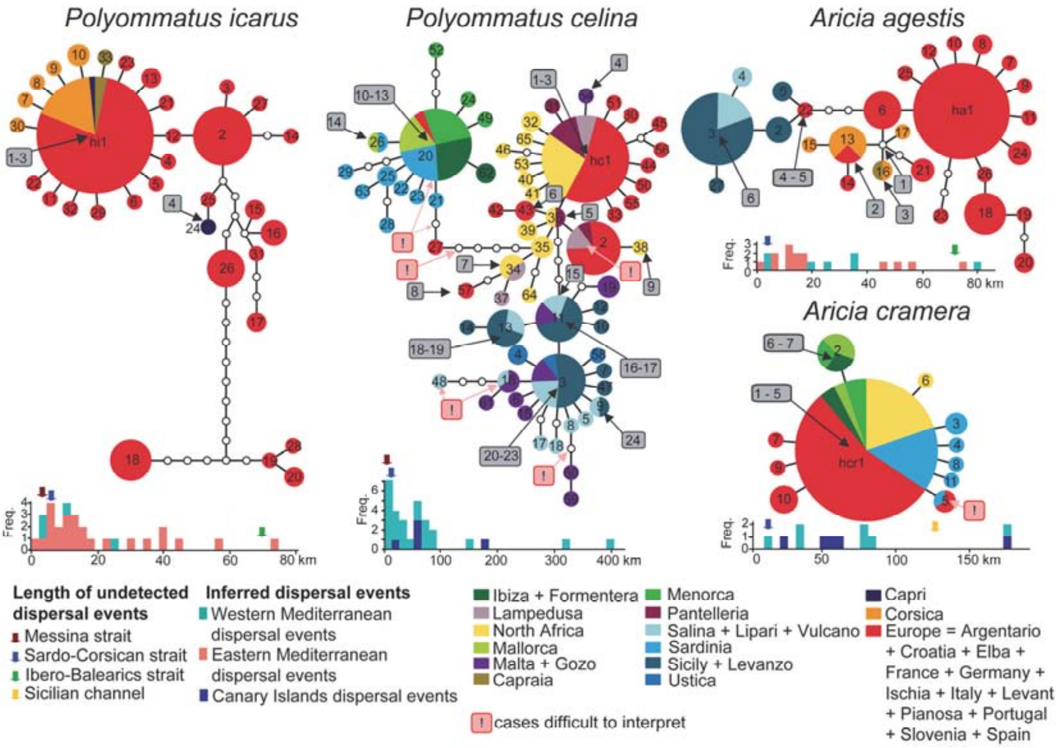


Fig 3. Haplotype networks and evaluation of dispersal events over the sea. Inferred over-sea dispersal events are numbered in grey rectangles, and ambiguous cases that were excluded from the analyses are highlighted in pink squares. The lengths of the dispersal events for the study area have been measured according to the-50m isobath and are displayed in histograms (cyan blocks), together with minimal dispersal events based on the distribution of the same species in the eastern Mediterranean (red blocks) and Canary Islands (blue blocks). The width of the key sea-straits Sardinia-Corsica, Italy-Sicily, Tunisia-Sicily and Ibero-Balearic are indicated by blue, red and green arrows, respectively.

doi:10.1371/journal.pone.0117802.g003

indicating dispersal from Spain to Sardinia, can be more parsimoniously explained by convergence or by the extinction or undetected presence of this haplotype in the Balearics. In order to obtain a conservative measurement of the frequencies of dispersal lengths, we excluded this ambiguous event. A minimum of five steps was required to attain the occurrence pattern of *A. cramera* over the Canary Islands (S3b Fig.). The dispersal length frequency, measured between the-50m isobaths, showed that the unrecorded Sardinia-Corsica dispersal event, potentially allowing *A. cramera* to mix with *A. agestis* populations in Corsica, would have been among the shortest dispersal events detected for this species. Conversely, the length of a hypothetical dispersal between Tunisia and Sicily is showed to be much less frequent (Fig. 3).

Aricia agestis showed a slightly more complex pattern. Corsica has been presumably colonized only once by an undetected haplotype that has drifted on both Corsica and mainland and successively reached Elba Island (connected to Italy in the-50 m isobath) and Capraia from Corsica (Fig. 3). Sicily has been apparently colonized in two events followed by differentiation and successive colonization of the Aeolian Islands. We thus scored six dispersal events (S3c Fig.), to which another between Sicily and Malta should be added because *A. agestis* was present in Malta at least until 30 years ago [46]. The examination of the distribution in the eastern Mediterranean, where only *A. agestis* is to be found, showed that a minimum of 14 dispersal events occurred in that region (S3d Fig.). Again, the Corsica-Sardinia dispersal would be among the shortest events recorded for this species, thus demonstrating that the dispersal capabilities have not been the limiting factor for the colonization of Sardinia by *A. agestis*.

According to previous results [21], *P. celina* showed three main lineages associated with i) Spain-Maghreb-southern Sicilian islands, ii) Balearics-Sardinia-Spain, iii) Sicily, circum-Sicilian islands. Spain and Malta harbour specimens belonging to two different lineages, which indicates that they can coexist in sympatry (S2 Fig.). According to such a complex pattern we scored a minimum of 24 overseas dispersal events (Fig. 3, S4a Fig.). We did not consider seven ambiguous events for the same reasons discussed for *A. cramera*. Four other events can be recognized for the Canary Islands (S4b Fig.). The comparison of the length of these events in the-50 m isobath with those required to colonize Italy from Sicily and Corsica from Sardinia revealed that such unrecorded dispersal events would have been among the shortest ones performed by this species.

Two well-separated lineages could be identified for *P. icarus*: one widely distributed over the whole Palaearctic region and one ranging from Sierra Nevada to France and Crete[21]. One Palaearctic haplotype found on Corsica, Capraia and Capri also occurs on mainland. Therefore, Corsica and Capraia seem to have been recently colonized and only once, while Capri has been colonized at least two times (Fig. 3). No evidence for other overseas dispersal events in the-50m isobath could be detected (Fig. 3, S4c Fig.). In the eastern Mediterranean at least 25 events(S4d Fig.) were necessary to justify the distribution on islands and most of them are longer than those required to reach Sicily from the Italian Peninsula and Sardinia from Corsica.

In summary, a number of successful colonization events were inferred from the haplotype networks and distributions on islands for the studied species. The lengths of the estimated sea-crossing dispersal events were frequently much longer than the distances required to cross the sea barriers where the most striking genetic divergences have been detected, especially for the narrow Messina and Bonifacio straits, just 3 and 5 kilometers long when considering the-50 m isobath (Fig. 3).

Evaluation of climatic and interaction constraints: species distribution modeling

All the climatic models showed a good fit and with very similar results between H and HQP models (AUC for H models: *P. icarus* = 0.758; *P. celina* = 0.901; *A. agestis* = 0.791; *A. cramera* = 0.833; AUC for HQP models: *P. icarus* = 0.758; *P. celina* = 0.904; *A. agestis* = 0.792; *A. cramera* = 0.835; Fig. 4 and S5 Fig. for map projections). The climatic variable showing the highest percentage of contribution to the model in both the H and HQP analyses was precipitation of the driest quarter. This variable also showed the highest regularized gain when tested alone in the Jackknife evaluation for all species except *A. agestis*, for which the highest contribution was given by mean temperature diurnal range. By projecting the values of precipitation of the driest quarter on a map and comparing it with the plots representing the response of logistic occurrences of the four species (Fig. 5), it is clear that values of 50–100 mm of precipitation determine a threshold that is highly correlated with the distribution of the four species (Fig. 2).

According to the climatic models, only *P. icarus* was predicted to occur in areas that agreed with the actual distribution of the species, also between the two sides of the narrow sea straits of Messina and Bonifacio channels (Fig. 4A and S5a Fig.). For the other three species MaxEnt highlighted as climatically suitable large areas where they do not actually occur. These areas largely corresponded to the unoccupied sides of the sea straits identified as areas with the strongest genetic contrasts (Sicily-Calabria, Sardinia-Corsica, North Africa-Sicily, Iberia-Balearics;

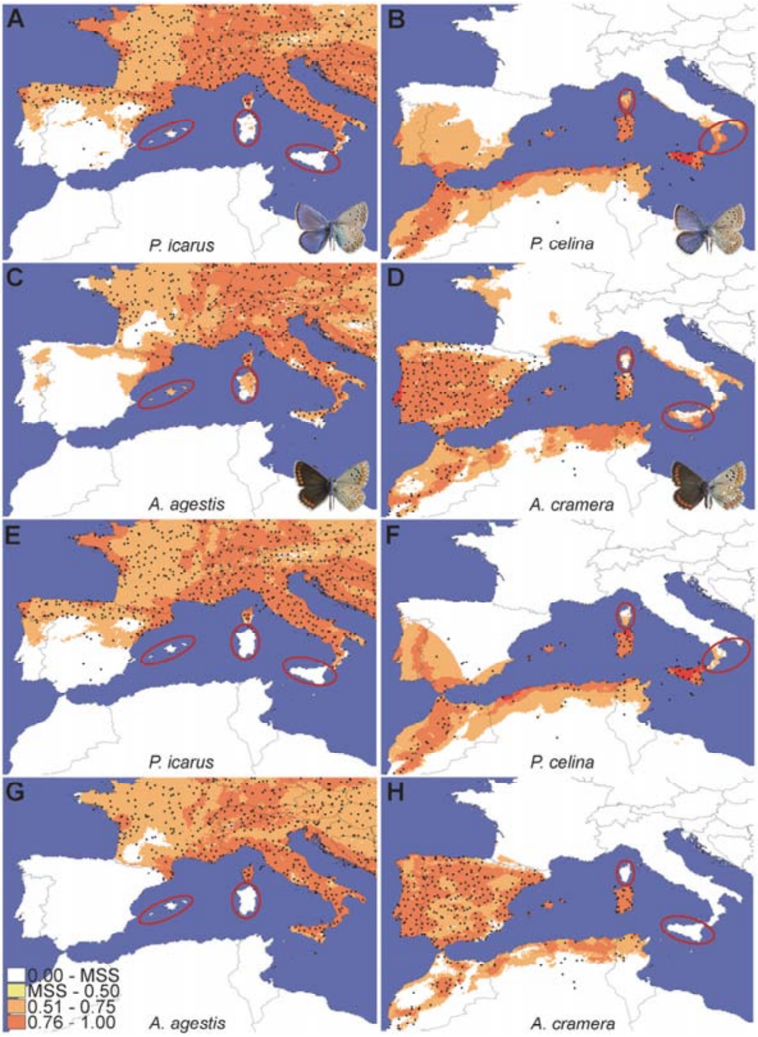


Fig 4. Projection over the study area of the logistic values obtained by MaxEnt analyses (H model) for the four species. The color gradient indicates the logistic probabilities of species occurrence, increasing from white (very low probability) to red (very high probability). The maximum value of the lowest class (white) is represented by the maximum training sensitivity plus specificity logistic threshold returned by MaxEnt. The key areas where the species are predicted to occur but are not present despite geographical proximity are highlighted.

doi:10.1371/journal.pone.0117802.g004

compare Fig. 2 and Fig. 4A–D and S5a–d Fig.), thus indicating that climatic constraints cannot entirely explain the chequered distributions of three of the studied species.

The second series of maximum entropy models, in which we included the convex hull representing the presence area of the corresponding cryptic species as a categorical predictive layer, revealed a strong influence of the presence of the peer cryptic species in determining the

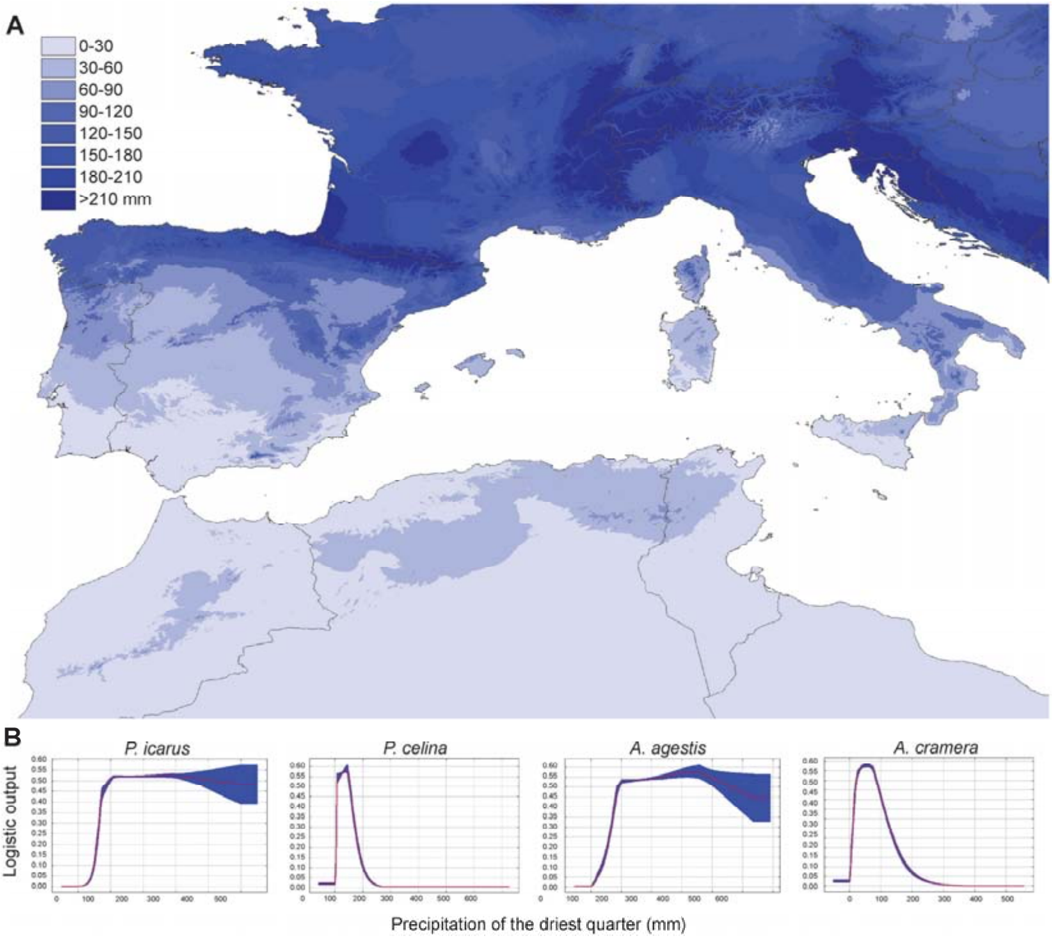


Fig 5. Projection of the climatic variable *precipitation in the driest quarter* and logistic responses for each species. A. The color gradient on the map indicates the precipitations (in mm) over the study area. **B**. The logistic responses of the precipitation in the driest quarter tested alone for the four studied species (*P. icarus*, *P. celina*, *A. agestis*, *A. cramera* from left to right). The response showed a clear and recurrent threshold between 50 and 100 mm of precipitation.

doi:10.1371/journal.pone.0117802.g005

observed pattern. The models revealed high AUC values also in this case (AUC for H models: *P. icarus* = 0.754; *P. celina* = 0.902; *A. agestis* = 0.799; *A. cramera* = 0.867); AUC for HQP models: *P. icarus* = 0.755; *P. celina* = 0.904; *A. agestis* = 0.805; *A. cramera* = 0.904). The presence of the related cryptic species was indicated as the variable showing the highest percentage contribution to the model for all species except *P. icarus*, for which it is ranked as second after precipitation of the driest quarter. Conversely, in the Jackknife evaluation of the regularized gain, the presence of the related cryptic species was the most important variable only for *A. cramera*. In these analyses, logistic predictions for presence/absence along sea straits showed a much higher correspondence with the observed patterns than those obtained only with climatic variables (Fig. 4E–H and S5e–h Fig.).

Discussion

This study shows that the strongest genetic contrasts between the two pairs of cryptic species occur over short distances, which prominently correspond to the Bonifacio and Messina straits. On the basis of the estimated dispersal capabilities, we show that each species are able to cross much longer sea barriers than Bonifacio and Messina. The two pairs of cryptic species revealed strong climatic preferences but these preferences alone could not fully explain the observed distribution pattern. In fact, on the basis of distribution modeling, both cryptic pairs were expected to coexist at least on one side of these relatively narrow sea straits. Conversely, only one of the two sibling species was documented on each side (Figs. 3 and 4), although dozens of individuals from the potential contact zones were examined (S1 Table). Previous papers suggested potentially chequered distribution patterns for several butterfly taxa in the Mediterranean region [20–22,47,48] and a recent study showed that the cryptic butterfly diversity in this region is overwhelmingly composed by groups of species that are not sympatric [19]. Since geographic isolation is considered to be the main driver for speciation, the tendency for sister species to show allopatric distributions is to be expected. However, the complete segregation of species and lineages over narrow sea straits that we observed is intriguing. Several hypotheses can be proposed to explain this pattern.

Dispersal capabilities and ecological constraints

A direct comparison of dispersal capabilities against the observed barriers potentially maintaining the vicariance patterns failed to explain the observed distribution. This is not surprising since recent studies have shown that butterflies can experience range expansions/contractions even over relatively short periods of time [20,49–51] and can rapidly adjust their distribution to track suitable environments [49]. The absence of specific resources, such as host plants, is unlikely to be an explanation for the observed mutual exclusion, because the four studied species are habitat and trophic generalists, occurring in a wide array of environments, from anthropic to mountain areas, and feeding on a variety of similar, and even identical, ubiquitous plant resources [52–54]. Since the host plants for the more recently discovered species *P. celina* were not well documented, we provide a table of our field observations (S2 Table) showing that it feeds on at least four widespread genera of Fabaceae. Species distribution modeling revealed that the two pairs of species experience different climate settings in the areas where they occur and that a common threshold of about 50–100 mm of precipitation in the driest quarter (summer in the Mediterranean) is highly correlated with all the observed distributions. Warm and dry conditions in the Mediterranean are well known to affect the life history of many butterfly species which, in many cases, emerge at the beginning of summer and estivate to delay reproduction in colder and wetter conditions [55]. There is growing evidence that interspecific variation in mitochondrial genes can determine a different respiration metabolism [56]. The strong differences in COI among the morphologically similar studied species may reveal to be functional for surviving in different climate settings and to be directly involved in maintaining the observed chequered patterns. However, it is difficult to evaluate if the observed climatic threshold has a real causal effect in determining the spatial separation among the species or if it only coincides with three main areas of phylogeographic breaks (Iberia-Maghreb, Sardinia-Corsica, Maghreb-Sicily-Calabria). Nevertheless, when included in the model, the presence of the corresponding cryptic pair had a strong influence in the models and explained most of the observed discordances between climatic predictions and the actual distribution of the four species. This suggests that the examined cryptic pairs have been influenced by the presence of other members of the same group in the recipient areas, apparently as much as by temporal and physical constraints.

Competition

Studies focused on birds demonstrated a correlation between differences in functional beak morphology and the rate of secondary sympatry, which suggests that competition instead of sexual interference was the main determinant of chequered distributions [17,18]. However, we specifically selected two pairs of species with high genetic divergence, thus with long potential time for secondary contacts, and a high degree of generalism, which renders competition for resources rather unlikely [10]. The evidence for a highly nested structure in butterfly communities, assembled in an order that reflects well their degree of specialization, provides empirical evidence that generalist butterflies do not tend to exclude each other [57]. Indeed, most of the widespread species of butterflies are habitat generalists that have the tendency to largely co-occur [58]. Thus, the two pairs of species studied here show idiosyncratic distributions beyond the general hypothesis stating that sister species are primarily allopatric and maintained their distribution pattern due to a short dispersal and evolutionary time, but competition for resources does not seem to be a suitable alternative explanation either.

Reproductive interference

It should be noted that the observation of strict mutual exclusion among these pairs of species only applies to islands, because on restricted mainland areas these cryptic species are known to display contact zones [21,22] where occasional potential hybrids are found. Accordingly, distribution modeling showed that the two pairs experience different climatic conditions but also that suitable areas largely overlap. Presumably, no strong precopulatory barriers exist between these species, as is the case in many butterflies [59,60], but hybrids between closely related species often display reduced fitness [30,61]. In the absence of precise recognition mechanisms, contact areas can be seen as population sinks, unlikely to enlarge given the cost for the neighboring populations [26]. Moreover, hybrid zones are predicted to shift until areas allowing low dispersal and low population densities are reached [26,62]. A concentration of boundaries between cryptic species over sea straits located over areas separating the climatic preferences of the different taxa, highly matches this hypothesis [19]. Thus, we hypothesize that the observed mutual exclusion, at least for the two pairs of cryptic species studied here, could be mainly due to a combination of climatic preference reinforced by reproductive interference between species [63].

Interestingly, our data on *P. celina* show that different lineages coexist in Malta (Sicilian and North African lineages) and Spain (European and Sardo-Corsican lineages), suggesting that multiple successful colonization events over notably long distances can take place, but only when propagules involve the same species as recipient populations (a case in which reproductive interference would not exist). In fact, these observations suggest that the niche carrying capacities in these islands are not saturated and allow the establishment of additional incoming lineages.

Conclusion

There is comparative evidence in literature that interactions between sister species constrain their range expansion [17,19]. Based on a completely different approach, our results point to the conclusion that a combination of climatic preferences and biotic interactions limit geographic range overlap and reject models limited purely by dispersal constraints. We show that this phenomenon can be more acute in particular non-sister cryptic than in non-cryptic sister taxa, which agrees with the positive correlation found between morphological distance and secondary contact [17], but not with that of phylogenetic distance versus secondary contact, since phylogenetically nearer species (*P. icarus*-*P. eros* and *A. agestis*-*A.artaxerxes/montensis*) are largely sympatric. Our data support productive interference plus climatic preference hypothesis over ecological competition, but the mechanisms contributing to the realized distributions may vary depending on the taxonomic group.

Climatic preference and density-dependent processes have been supposed to be key factors in determining the evolution of species, their mutual exclusion on oceanic islands and in generating and maintaining the phylogeographic structures of many species in Europe [21,22,47,48]. Indeed, density-dependent processes, mostly at the leading edge of colonization events, can generate striking geographic contrasts in the distribution of genes and species. We hypothesize that a well-established population of a taxon in a recipient area can strongly interfere with the dispersing individuals belonging to the other cryptic taxon, and thus maintain the geographic patterns established over sea straits by a "founder takes all" mechanism [64]. With increasing knowledge of cryptic Mediterranean butterflies there is evidence for recurrent separation of different species/lineages over the Bonifacio and Messina channels (e.g. *Spialia orbifer/sertorius*, *Melanargia arge/pherusa* [65], *Coenonympha pamphilus/lyllus* [66], *Lysandra coridon* group [67], *Pararge aegeria* lineages [20]). This pattern is reinforced by the observation that among all western Mediterranean butterfly species, cryptic taxa tend to establish contact zones over the same sea straits [19]. Notably, the Messina and Bonifacio straits, but also the area between Iberia and Maghreb, show a similar abrupt change in the quantity of precipitations during the driest quarter. Precipitation is an important factor influencing butterfly survival in the Mediterranean, since it determines availability of key resources for adults and larvae (water, nectar sources, host plants)[55]. Thus, the phenomenon studied here for two pairs of species may have a prominent impact in determining the overall patterns of butterfly diversity in the region.

The mechanisms here discussed are theoretically applicable to most organisms, and a direct implication of our results is the necessity to consider interactions among cryptic entities when aiming at documenting diversity and its dynamics, including the effects of global changes [49,68,69]. Unfortunately, including the cryptic fraction of biodiversity is not straightforward because its complex recognition requires extensive morphological and/or genetic assessments, and these species are frequently amalgamated in wide-scale surveys [30,49,70].

Supporting Information

S1 Dataset. GPS coordinates for the specimens used in the ecological niche modeling.
(XLSX)

S1 Fig. Delaunay triangulation obtained for the sequenced specimens of *Polyommatus* spp. and *Aricia* spp.
(TIF)

S2 Fig. Neighbour-Joining trees based on COI sequences for *Polyommatus* spp. and for *Aricia* spp.
(TIF)

S3 Fig. Minimum dispersal events recorded for *Aricia* spp.
(TIF)

S4 Fig. Minimum dispersal events recorded for *Polyommatus* spp.
(TIF)

S5 Fig. Projection over the study area of the logistic values obtained by MaxEnt analyses for the four species, with the HQP models.
(TIF)

S1 Table. List of specimens used in the molecular analysis.
(XLSX)

S2 Table. List of host plant records for *Polyommatus celina*.
(XLSX)

Acknowledgments

We are grateful to all the people and institutions that provided presence records and/or samples: B. Acosta, Z. Bálint, G. Barech, D. Carreras, J. Dantart, S. Cuvelier, M. Đurić, S. Estradé, B. Fontaine, O. Garcia, J. A. García-Alamá, F. Gil-T., E. García-Barros, F. González, J. Hernández-Roldán, M. A. Ibáñez-Orrico, J. Jubany, M. Khaldi, E. Maravalhas, X. Merit, S. Montagud, M. Munguira, J. Olivares, J. Pérez-López, K. Rebbas, J. Requejo, P. Sammut, J. M. Sesma, F. Slamka, C. Stefanescu, M. R. Tarrier, M. A. Travassos, S. Viader and Juan Carlos Vicente. Special thanks to H. Romo and G. Talavera for their help with the ecological niche modeling analyses.

Author Contributions

Conceived and designed the experiments: ROV LD. Performed the experiments: RAV VD. Analyzed the data: ROV RAV LD VD. Contributed reagents/materials/analysis tools: ROV RAV LD VD. Wrote the paper: ROV RAV LD VD.

References

1. Meier ES, Kienast F, Pearman PB, Svenning JC, Thuiller W, et al. (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. Ecography 33: 1038–1048.
2. Hutchinson GE (1957) Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology 22: 415–427.
3. Pianka ER, Huey RB, Lawlor LP (1979) Niche segregation in desert lizards. In: Horn DJ, Stairs GR, Mitchell RD, editors. Analysis of Ecological Systems Columbus: Ohio State Univ. Press. pp. 67–116
4. Soberon J (2007) Grinnellian and Eltonian niches and geographic distributions of species. Ecology Letters 10: 1115–1123. PMID: 17850335
5. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965–1978.
6. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231–259.
7. Thuiller W (2003) BIOMOD—optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biology 9: 1353–1362.
8. Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters 8: 993–1009.
9. Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, et al. (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews 88: 15–30. doi: 10.1111/j.1469-185X.2012.00235.x PMID: 22686347
10. Friberg M, Leimar O, Wiklund C (2013) Heterospecific courtship, minority effects and niche separation between cryptic butterfly species. J Evol Biol 26: 971–979. doi: 10.1111/jeb.12106 PMID: 23480828
11. Ord TJ, King L, Young AR (2011) Contrasting Theory with the Empirical Data of Species Recognition. Evolution 65: 2572–2591. doi: 10.1111/j.1558-5646.2011.01319.x PMID: 21884057
12. Kishi S, Nishida T, Tsubaki Y (2009) Reproductive interference determines persistence and exclusion in species interactions. Journal of Animal Ecology 78: 1043–1049. doi: 10.1111/j.1365-2656.2009.01560.x PMID: 19457018
13. Zhou ZS, Chen ZP, Xu ZF (2010) Niches and interspecific competitive relationships of the parasitoids, *Microplitis prodeniae* and *Campoletis chlorideae*, of the Oriental leafworm moth, *Spodoptera litura*, in tobacco. Journal of Insect Science 10: 10. doi: 10.1673/031.010.1001 PMID: 20575741

14. Wardle DA, Barker GM, Bonner KI, Nicholson KS (1998) Can comparative approaches based on plant ecophysiological traits predict the nature of biotic interactions and individual plant species effects in ecosystems? Journal of Ecology 86: 405–420.

15. Keddy PA (1992) A Pragmatic Approach to Functional Ecology. Functional Ecology 6: 621–626.

16. Heikkinen RK, Luoto M, Virkkala R, Pearson RG, Korber JH (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. Global Ecology and Biogeography 16: 754–763.

17. Pigot AL, Tobias JA (2013) Species interactions constrain geographic range expansion over evolutionary time. Ecology Letters 16: 330–338. doi: 10.1111/ele.12043 PMID: 23231353

18. Tobias JA, Cornwallis CK, Derryberry EP, Claramunt S, Brumfield RT, et al. (2013) Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. Nature 506: 359–363. doi: 10.1038/nature12874 PMID: 24362572

19. Vodă R, Dapporto L, Dincă V, Vila R (in press) Cryptic matters: Overlooked buttery beta-diversity. Ecography. DOI: 10.1111/ecog.00762. PMID: 25076812

20. Dapporto L, Bruschini C, Dincă V, Vila R, Dennis RLH (2012) Identifying zones of phenetic compression in West Mediterranean butterflies (Satyrinae): refugia, invasion and hybridization. Diversity and Distributions 18: 1066–1076.

21. Dincă V, Dapporto L, Vila R (2011) A combined genetic-morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and *Polyommatus celina* Common Blue butterflies. Molecular Ecology 20: 3921–3935. doi: 10.1111/j.1365-294X.2011.05223.x PMID: 21883579

22. Sañudo-Restrepo CP, Dincă V, Talavera G, Vila R (2013) Biogeography and systematics of *Aricia* butterflies (Lepidoptera, Lycaenidae). Molecular Phylogenetics and Evolution 66: 369–379. doi: 10.1016/j.ympev.2012.10.010 PMID: 23099146

23. Dincă V, Wiklund C, Lukhtanov VA, Kodandaramaiah U, Noren K, et al. (2013) Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. J Evol Biol 26: 2095–2106. doi: 10.1111/jeb.12211 PMID: 23909947

24. McBride CS, van Velzen R, Larsen TB (2009) Allopatric origin of cryptic butterfly species that were discovered feeding on distinct host plants in sympatry. Molecular Ecology 18: 3639–3651. doi: 10.1111/j.1365-294X.2009.04309.x PMID: 19674305

25. Dapporto L (2010) Speciation in Mediterranean refugia and post-glacial expansion of *Zerynthia polyxena* (Lepidoptera, Papilionidae). Journal of Zoological Systematics and Evolutionary Research 48: 229–237.

26. Dasmahapatra KK, Lamas G, Simpson F, Mallet J (2010) The anatomy of a 'suture zone' in Amazonian butterflies: a coalescent-based test for vicariant geographic divergence and speciation. Molecular Ecology 19: 4283–4301. doi: 10.1111/j.1365-294X.2010.04802.x PMID: 20819158

27. Porter AH (2009) Ecological genetics and gvolutionary ecology in hybrid zones. In: Settele J, Shreeve TG, Konvicka M, Van Dyck H, editors. Ecology of Butterflies in Europe. Cambridge: Cambridge University Press. pp. 296–312.

28. Wiemers M, Stradomsky BV, Vodolazhsky DI (2010) A molecular phylogeny of *Polyommatus* s. str. and *Plebicula* based on mitochondrial COI and nuclear ITS2 sequences (Lepidoptera: Lycaenidae). European Journal of Entomology 107: 325–336.

29. FaunaEuropea (2013) http://wwwfaunaeurorg/.

30. Descimon H, Mallet J (2009) Bad species. In: Settele J, Shreeve TG, Konvicka M, Van Dyck H, editors. Ecology of Butterflies in Europe. Cambridge: Cambridge University Press. pp. 219–249.

31. Dapporto L, Dennis RLH (2013) The generalist-specialist continuum: Testing predictions for distribution and trends in British butterflies. Biological Conservation 157: 229–236.

32. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America 101: 14812–14817. PMID: 15465915

33. Ivanova NV, Dewaard JR, Hebert PDN (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes 6: 998–1002.

34. deWaard JR, Ivanova NV, Hajibabaei M, Hebert PDN (2008) Assembling DNA barcodes: analytical protocols. In: Martin C, editor. Methods in molecular biology: environmental genetics. Totowa, NJ: Humana Press. pp. 275–293.

35. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596–1599. PMID: 17488738

36. Borcard D, Gillet Fo, Legendre P (2011) Numerical ecology with R. New York: Springer. xi, 306 p. p.

37. Keis M, Remm J, Ho SYW, Davison J, Tammeleht E, et al. (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. Journal of Biogeography 40: 915–927.

38. Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, et al. (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. Proceedings of the National Academy of Sciences of the United States of America 102: 15942–15947. PMID: 16243969

39. Vandergast AG, Perry WM, Lugo RV, Hathaway SA (2011) Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. Molecular Ecology Resources 11: 158–161. doi: 10.1111/j.1755-0998.2010.02904.x PMID: 21429115

40. Harrell FE (2001) Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer. xxii, 568 p. p.

41. Blach-Overgaard A, Svenning JC, Dransfield J, Greve M, Balslev H (2010) Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. Ecography 33: 380–391.

42. Santos H, Juste J, Ibanez C, Palmeirim JM, Godinho R, et al. (2014) Influences of ecology and biogeography on shaping the distributions of cryptic species: three bat tales in Iberia. Biological Journal of the Linnean Society 112: 150–162.

43. Merow C, Smith MJ, Silander JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36: 1058–1069.

44. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, et al. (2011) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17: 43–57.

45. Liu CR, White M, Newell G (2013) Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of Biogeography 40: 778–789.

46. Sammut P (1984) A systematic and synonymic list of the Lepidoptera of the Maltese Islands. Neue Entomologische Nachrichten 13: 1–124.

47. Dincă V, Lukhtanov VA, Talavera G, Vila R (2011) Unexpected layers of cryptic diversity in wood white *Leptidea* butterflies. Nature Communications 2: 324. doi: 10.1038/ncomms1329 PMID: 21610727

48. Nazari V, Ten Hagen W, Bozano GC (2010) Molecular systematics and phylogeny of the 'Marbled Whites' (Lepidoptera: Nymphalidae, Satyrinae, *Melanargia* Meigen). Systematic Entomology 35: 132–147.

49. Devictor V, van Swaay C, Brereton T, Brotons L, Chamberlain D, et al. (2012) Differences in the climatic debts of birds and butterflies at a continental scale. Nature Climate Change 2: 121–124.

50. Dapporto L, Bruschini C (2012) Invading a refugium: post glacial replacement of the ancestral lineage of a Nymphalid butterfly in the West Mediterranean. Organisms Diversity & Evolution 12: 39–49. doi: 10.1016/j.ijfoodmicro.2015.01.008 PMID: 25626150

51. Habel JC, Lens L, Rodder D, Schmitt T (2011) From Africa to Europe and back: refugia and range shifts cause high genetic differentiation in the Marbled White butterfly *Melanargia galathea*. Bmc Evolutionary Biology 11: 215. doi: 10.1186/1471-2148-11-215 PMID: 21777453

52. Tennent J (1996) The Butterflies of Morocco, Algeria and Tunisia. Wallingford: Gem Publishing Company.

53. Tarrier MR, Delacre J (2008) Les Papillons de jour du Moroc. Mèze: Biotope.

54. Dennis RLH (2010) A resource-based habitat view for conservation: butterflies in the British landscape. Chichester, UK; Hoboken, NJ: Wiley-Blackwell. xii, 406 p. p.

55. Grill A, Cerny A, Fiedler K (2013) Hot summers, long life: egg laying strategies of Maniola butterflies are affected by geographic provenance rather than adult diet. Contributions to Zoology 82: 27–36.

56. Toews DPL, Mandic M, Richards JG, Irwin DE (2014) Migration, Mitochondria, and the Yellow-Rumped Warbler. Evolution 68: 241–255. doi: 10.1111/evo.12260 PMID: 24102562

57. Ulrich W, Gotelli NJ (2007) Disentangling community patterns of nestedness and species co-occurrence. Oikos 116: 2053–2061.

58. Dennis RLH, Hardy PB, Dapporto L (2012) Nestedness in island faunas: novel insights into island biogeography through butterfly community profiles of colonization ability and migration capacity. Journal of Biogeography 39: 1412–1426.

59. Friberg M, Vongvanich N, Borg-Karlson AK, Kemp DJ, Merilaita S, et al. (2008) Female mate choice determines reproductive isolation between sympatric butterflies. Behavioral Ecology and Sociobiology 62: 873–886.

60. Wickman P-O (2009) Mating behaviour in butterflies. In: Settele J, Shreeve TG, Konvicka M, Van Dyck H, editors. Ecology of Butterflies in Europe. Cambridge: Cambridge University Press. pp. 17–28.

61. Biermann H, Eitschberger U (1996) Bemerkungen zu *Lasiommata megera* (Linnaeus, 1767) und *Lasiommata paramegaera* (Hübner, [1824]) (Lepidoptera Saryridae). Atalanta 27: 253–319.

62. Mallet J (2010) Shift happens! Shifting balance and the evolution of diversity in warning colour and mimicry. Ecological Entomology 35: 90–104.

63. Groning J, Hochkirch A (2008) Reproductive interference between animal species. Quarterly Review of Biology 83: 257–282. PMID: 18792662

64. Waters JM, Fraser CI, Hewitt GM (2013) Founder takes all: density-dependent processes structure biodversity. Trends in Ecology & Evolution 28: 78–85. doi: 10.1016/j.cub.2014.11.067 PMID: 25619764

65. Kudrna O, Harpke A, Lux K, Pennerstorfer J, Schweiger O, et al. (2011) Distribution atlas of butterflies in Europe. Halle: Gesellschaft fuer Schmetterlingsschutz e.V. 1–576 p.

66. Boillat H (2002) *Coenonympha lyllus* Esper, 1805, spec. rev. Une nouvelle approche taxinomique du complexe *pamphilus*. Alexanor 22: 243–309.

67. Talavera G, Lukhtanov VA, Rieppel L, Pierce NE, Vila R (2013) In the shadow of phylogenetic uncertainty: The recent diversification of *Lysandra* butterflies through chromosomal change. Molecular Phylogenetics and Evolution 69: 469–478. doi: 10.1016/j.ympev.2013.08.004 PMID: 23954756

68. Hill JK, Thomas CD, Fox R, Telfer MG, Willis SG, et al. (2002) Responses of butterflies to twentieth century climate warming: implications for future ranges. Proceedings of the Royal Society B-Biological Sciences 269: 2163–2171. PMID: 12396492

69. Parmesan C, Ryrholm N, Stefanescu C, Hill JK, Thomas CD, et al. (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. Nature 399: 579–583.

70. Isaac NJB, Mallet J, Mace GM (2004) Taxonomic inflation: its influence on macroecology and conservation. Trends in Ecology & Evolution 19: 464–469. doi: 10.1016/j.cub.2014.11.067 PMID: 25619764

Chapter 4

Dapporto L., Vodă R., Dincă V. & Vila R. 2014. Comparing population patterns for genetic and morphologicalmarkers with uneven sample sizes. An example for the butterfly *Maniola jurtina*. *Methods in Ecology and Evolution*, **5**: 834–843.

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2014, **5**, 834–843 doi: 10.1111/2041-210X.12220

Comparing population patterns for genetic and morphological markers with uneven sample sizes. An example for the butterfly *Maniola jurtina*

Leonardo Dapporto[1]*,†, Raluca Vodă[2,3],†, Vlad Dincă[2,4,5] and Roger Vila[2]

[1] *Department of Biological and Medical Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, UK;* [2] *Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37, Barcelona 08003, Spain;* [3] *Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain;* [4] *Department of Zoology, Stockholm University, Stockholm, S-106 91, Sweden; and* [5] *Biodiversity Institute of Ontario, University of Guelph, Guelph, N1G 2W1 ON, Canada*

Summary

1. Integrating genetic and/or phenotypic traits at population level is considered a fundamental approach in the study of evolutionary processes, systematics, biogeography and conservation. But combining the two types of data remain a complex task, mostly due to the high, and sometimes different, sample sizes required for reliable assessments of community traits. Data availability has been increasing in recent years, thanks to online resources, but it is uncommon that different types of markers are available for any given specimen.

2. We provide new R functions aimed at directly correlating traits at population level, even if data sets only overlap partially. The new functions are based on a modified Procrustes algorithm that minimizes differences between bidimensional ordinations of two different markers, based on a subsample of specimens for which both characters are known. To test the new functions, we used a molecular and morphological data set comprising Mediterranean specimens of the butterfly *Maniola jurtina*.

3. By using this method, we have been able to maximize similarities between genotypic and phenotypic configurations obtained after principal coordinate analysis for the model species and evaluated their degree of correlation at both individual and population level. The new recluster.procrustes function retained the information of the relative importance of different morphological variables in determining the observed ordinations and preserved it in the transformed configurations. This allowed calculating the best combination of morphological variables mirroring genetic relationships among specimens and populations. Finally, it was possible to analyse the modality and variance of the phenotypic characters correlated with the genetic structure among populations.

4. The genetic and phenotypic markers displayed high overall correlation in the study area except in the contact zone, where discrepancies for particular populations were detected. Interestingly, such discrepancies were spatially structured, with southern populations displaying typical western morphotype and eastern haplotypes, while the opposite occurred in the northern populations. The methodology here described can be applied to any number and type of traits for which bidimensional configurations can be obtained, and opens new possibilities for data mining and for meta-analyses combining existing data sets in biogeography, systematics and ecology.

Key-words: Cytochrome c oxidase subunit 1, geometric morphometrics, Lepidoptera, Mediterranean, *recluster* R package, recluster.procrustes function

Introduction

Integrating different types of data has become a usual procedure in the study of speciation, biogeography and conservation ecology (Pergams & Lacy 2008; Dincă, Dapporto & Vila 2011; Mila *et al.* 2011; Toews *et al.* 2014; Derryberry *et al.* 2014). There is growing evidence that different types of markers can display contrasting spatial patterns (Dapporto *et al.* 2011; Toews & Brelsford 2012; Pavlova *et al.* 2013; Toews *et al.* 2014), and such discrepancies are hypothesized to encompass important evolutionary and population processes like local adaptation, character displacement and sex-biased asymmetries in hybridization processes (Toews & Brelsford 2012). Biogeographical, ecological and evolutionary studies typically combine different DNA sequences (usually nuclear and mitochondrial genes) or include both genetic and morphological traits to compare taxa or populations. While genetic patterns can reveal evolutionary relationships, hybridization processes and approximate divergence time, phenotypic traits represent

*Correspondence author. E-mail: leondap@gmail.com

†These authors equally contributed to this article.

the interface between genes and the environment and include key morphological traits for the interactions among individuals.

Nevertheless, comparing genetic and morphological patterns among populations in a spatial context is methodologically challenging because genetic and morphological analyses produce different raw data (e.g. DNA sequences vs. linear or shape measurements) and are usually examined with different techniques (Claude 2008; Paradis 2012). Moreover, while DNA sequencing follows highly standardized procedures producing directly comparable results, morphological assessments are affected by stochastic factors such as sample preparation, measurement errors and by a multitude of environmental events determining the observed individual's phenotype. As a consequence, the morphological assessment of populations usually requires a higher number of samples compared with genetic analyses. Yet, most of the existing methods for direct comparisons between different types of markers only allow correlations on specimens analysed for both characters. Importantly, the difficulty of directly comparing data sets that do not fully overlap (i.e. with different numbers of samples assessed for each marker) prevents a full use of data available in public repositories (e.g. GenBank, BOLD systems), where typically only one of the markers is available for any given specimen.

Even when an adequate sample size is gathered, the analyses for the spatial distribution of morphological and genetic markers among populations are often carried out independently (Pergams & Lacy 2008; Dincă, Dapporto & Vila 2011; Seraphim *et al.* 2013; Toews *et al.* 2014). Direct comparisons are complicated by the different approaches used to retrieve genetic and morphological patterns. Genetic patterns are usually represented by phylogenetic trees or haplotype networks, which are postulated to mirror the evolutionary relationships among taxa (Paradis 2012). However, although phylogenetic trees can be conveniently used to order taxa, they can be inefficient to sort hybrid individuals or to analyse the overall genetic structure of mixed populations (Kalinowski 2009). In these cases, less constrained representations avoiding bi- or polytomic ordinations should be preferred (Kalinowski 2009). Such continuous patterns for genetic data can be obtained by ordination methods, like multidimensional scaling, principal coordinates analysis, phylogenetic PCA (pPCA), etc., which reduce the dimensionality of any dissimilarity matrix obtained by different markers such as DNA sequences, microsatellites or allozymes (Pritchard, Stephens & Donnelly 2000; Paradis 2012). Interestingly, similar algorithms are the standard methods used to visualize patterns of variation in continuous morphological traits (Claude 2008). Such parallel approaches provide the basis for a direct comparison of genetic and morphological patterns. Direct correlations are usually tested through the overall correlation coefficients between original distance matrices or configurations obtained, for example, by means of Mantel tests or protest analyses (Renvoise *et al.* 2012). These analyses evaluate the degree of correlation between matrices and calculate the associated *P* value. However, they do not allow for: (i) the identification of particularly discrepant individuals or populations, (ii) the analysis of unimodal or bimodal distributions in phenotypic traits as associated with genetic ones and (iii) the representation of the variation of these characteristics in the geographical space.

Such analyses have been carried out only in a few studies that mainly compared single components (usually belonging to PCA analyses) of genetic and morphological variation between them and with environmental determinants (Gompert *et al.* 2010).

In this article, we developed new functions for the *recluster* R package that facilitate in-depth comparisons of bidimensional configurations (genotypic and phenotypic markers), even when only a limited series of specimens is analysed for the two traits. The main new function recluster.procrustes is a modified Procrustes algorithm which allows indicating the subset of specimens that overlap for the two data sets. The function minimizes the differences between two configurations by applying a classical Procrustes on individuals for which both traits have been assessed, and subsequently applies the same transformation to the specimens that were not analysed for both markers. Importantly, the new functions allow entering the coordinates of the variables determining the observed pattern (as it is usually obtained by PCA or MDS), and also applies the same transformation to them to maintain the information regarding their contribution.

To show the characteristics of the analysis and to provide a practical example, we tested the new method on two data sets for the butterfly *Maniola jurtina* (Nymphalidae) covering the west Mediterranean area and eastern Europe. We examined if the COI mitochondrial gene and the morphological data reveal with precision the location and nature of the contact zone between the two main lineages described for this species. Discrepancies between allozyme and morphological patterns have been described in *M. jurtina* (Dapporto *et al.* 2011), suggesting that this species is a suitable model for our analysis. We show how the proposed analysis can identify and model correspondences and discrepancies in a spatial framework. Due to the increasing availability of similar data for a large number of organisms, such a theoretical and practical assessment can represent a useful model and resource for future studies.

Methods

SPECIES, MARKERS AND BACKGROUND

Maniola jurtina

The meadow brown butterfly *Maniola jurtina* (Linnaeus, 1758) is a species that has often been targeted for phylogeography and speciation studies. It forms conspicuous populations over Europe and the Mediterranean basin including many islands. Two lineages, identified on the basis of male genitalia and allozyme analysis, occur in Europe: a western Atlantic-Mediterranean lineage (*M. j. jurtina*) in the Maghreb, Spain, western France, Sicily and Sardinia and an eastern-Mediterranean-Asian lineage (*M. j. janira*) widespread from Asia to eastern and central Europe, including the Italian Peninsula (Schmitt, Rober & Seitz 2005; Dapporto *et al.* 2009; Thomson 2011). On the basis of allozyme data, the two lineages seem to have diverged during the late Pleistocene

(Schmitt, Rober & Seitz 2005) and may have experienced a series of range contraction/expansion cycles during the following glacial–interglacial periods (Schmitt, Rober & Seitz 2005; Dapporto *et al.* 2011; Thomson 2011). Morphological and allozyme data also revealed the presence of a contact zone extending from a few western Mediterranean islands (Corsica, Elba, Giglio and Capri) to the western Alps and the Benelux region (Schmitt, Rober & Seitz 2005; Dapporto *et al.* 2011; Thomson 2011; but see also Habel, Dieker & Schmitt 2009). Modelling of continuous morphological variation and discrepancies between morphological and allozyme data have been used to infer the phylogeography of this species over the western Mediterranean area, and led to the formulation of a new post-glacial colonization paradigm over Europe (Dapporto *et al.* 2009, 2011; Habel, Dieker & Schmitt 2009; Dapporto & Bruschini 2012). Despite the considerable number of studies dealing with patterns in morphology and allozymes for this species, no comprehensive genetic assessments have been published and no direct evidence exists for a correspondence between morphology and DNA. We identified 39 geographic areas, comprising 20 insular and 19 mainland areas (Table 1). All areas included at least two specimens analysed for both markers (see Appendix S2 for details).

Table 1. Studied areas and their abbreviations

Number	Area	Abbreviation
1	Eastern Europe	E_Eur
2	Dolomites	Dolom
3	Central-northern Italy	N_Ita
4	Argentario	Arg
5	Giglio Island	Giglio
6	Pianosa Island	Pian
7	Elba Island	Elba
8	Piombino	Piomb
9	Northern Corsica	N_Cor
10	Southern Corsica	S_Cor
11	Northern Sardinia	N_Sar
12	Central Sardinia	C_Sar
13	Southern Sardinia	S_Sar
14	Ischia Island	Ischia
15	Capri Island	Capri
16	Sorrento Peninsula	Sorren
17	Southern Italian Peninsula	S_Ita
18	Aspromonte	Aspr
19	Eastern Sicily	E_Sic
20	Vulcano Island	Vulc
21	Lipari Island	Lipari
22	Western Sicily	W_Sic
23	Gozo Island	Gozo
24	Tunisia	Tun
25	Algeria	Alg
26	Morocco	Mor
27	Southern Iberia	S_Ibe
28	Central Iberia	C_Ibe
29	Northern Iberia	N_Ibe
30	North-eastern Iberia	NE_Ibe
31	Ibiza Island	Ibiza
32	Mallorca Island	Mall
33	Menorca Island	Men
34	Southern France	S_Fra
35	Levant Island	Levant
36	North-western Alps	NW_Alps
37	South-western Alps	SW_Alps
38	Liguria	Liguria
39	Switzerland	Suisse

Genetic markers

We analysed the COI mitochondrial gene because it is a widely used marker in systematics and phylogeography (Avise 2009) and extensive libraries of publicly available sequences exist (e.g. GenBank and BOLD systems). A total of 218 *M. jurtina* COI sequences were used in the analysis, of which 45 were obtained from publicly available data in BOLD and 173 have been sequenced specifically for this study (GenBank accession numbers KM020807–KM020882, KJ994239–KJ994253 and KM033847–KM033941; see Appendix S2 for sampling localities). Well-assessed protocols described in Appendix S2 were used for DNA extraction and sequencing. A dissimilarity p-distance matrix among COI sequences was obtained with default settings in MEGA 5.05. The ordination of genetic markers has been obtained by projecting the dissimilarity matrix into a two-dimensional configuration through a principal coordinate analysis (PCoA). A neighbour-joining tree was constructed in MEGA 5.05 using p-distance and assessing node supports by 100 bootstrap pseudoreplicates. One sequence of *Hyponephele lupina* was used as outgroup.

Morphological markers

We selected the shape of male genitalia as morphological marker. We analysed 616 male *M. jurtina* specimens by means of geometric morphometrics (Bookstein 1991), a method that produces relative warps (PCs) representing continuous variables of shape variations. Variables from different structures of the genitalia can be combined in successive ordination analyses to reveal the overall patterns of variation among specimens and the shape variation associated with such patterns. Geometric morphometrics is a powerful method for analysing morphological species traits in a wide range of organisms (Viscosi & Cardini 2011; Madeira *et al.* 2012; Zelditch, Swiderski & Sheets 2012). It has been successfully used as a quantitative method to distinguish western and eastern morphotypes, as well as intermediate (presumably hybrid) individuals and populations of *M. jurtina* (Dapporto *et al.* 2009; Dapporto & Bruschini 2012). In this species, two structures (valva and brachium, Appendix S2) are mostly involved in differentiating morphotypes (Dapporto *et al.* 2009; Thomson 2011). Moreover, the continuous variation of these markers has been used to model the distribution of morphotypes (Dapporto *et al.* 2009; Dapporto & Bruschini 2012). We thus conducted geometric morphometrics of brachium and valva following the methodology employed by Dapporto *et al.* (2012), which is specifically described in Appendix S2.

Similarly to what has been carried out for the genetic markers, we used relative warps obtained for the brachium and valva analysis to perform a PCoA. The contribution of the variables to the configuration has been obtained by computing the weighted average scores of variables by using the wascores *vegan* function. We retained the two most important variables for each of the two axes used.

THE NEW RECLUSTER.PROCRUSTES FUNCTION

The *recluster* package was originally created to solve a bias in the application of cluster analysis to turnover dissimilarity matrices (Dapporto *et al.* 2013). Subsequently, the package has been improved to allow the analysis of diversity patterns in a spatial context. This version of the *recluster* package enables coupling tree-based results with two-dimensional representations in RGB colour space, so that the nearest points in the bivariate configuration are represented with similar colours. This approach is becoming increasingly used in biogeography studies (Kreft & Jetz 2010; Holt *et al.* 2013).

The core of the proposed analysis is an algorithm maximizing similarities among bivariate configurations (in this case based on genotypic and phenotypic patterns). Such configurations can be obtained through any method for phenotypic traits (a crude combination of two measurements, the two main relative warps of a geometric morphometrics output, scatterplots of PCA, PCoA, non-metric multidimensional scaling (MDS), etc.). For genotypic data, bidimensional configurations can be obtained, for example, by multidimensional scaling or PCoA of any genetic distance matrix and by pPCA.

Some of these analyses (e.g. PCA, PCoA and MDS) not only order the specimens, but also provide the contribution of the variables to the obtained patterns (Legendre & Legendre 1998). This is much more common for morphological data where patterns are usually established by ordination of a multidimensional character matrix.

Two similar bidimensional configurations can appear highly different due to relative rotation, orientation and scaling (Appendix S1). The classic tools used to maximize similarity between configurations is the Procrustes analysis, which scales, flips and rotates a configuration to maximize its similarity to another one (Mardia, Kent & Bibby 1979). The *vegan* R package provides a Procrustes function to compute this analysis. However, it has two characteristics representing problematic limitations for our purposes:

1 The Procrustes *vegan* function only works when the specimens occurring in the two configurations are exactly the same.

2 The *vegan* Procrustes cannot maintain the correspondence between the rotated and rescaled configurations for specimens and the ordination of the variables.

The Procrustes analysis provides a rigid transformation of the second configuration. In theory, the second matrix can be thus transformed to maximally fit the first one on the basis of a subsample of corresponding specimens. Then, the same solid transformation can be applied to remnant (not shared) specimens, as well as to the coordinates of variables in the configuration to maintain the information about their contribution in the observed pattern. For this purpose, we created the recluster.procrustes function by modifying the Procrustes function of *vegan*. The recluster.procrustes function allows the user to indicate the number of common specimens, which must be listed first and in the same order in the two matrices. Moreover, it is possible to include other matrices containing variable contributions (coordinates) for each marker. An in-depth explanation of the recluster.procrustes function is given in Appendix S1.

Analyses

ORIENTING THE FIRST AXIS OF THE GENETIC CONFIGURATION ACCORDING TO A PHYLOGENETIC TREE

Using the recluster.group.col function, we computed the mean positions in the genetic PCoA configuration of specimens belonging to each of the two main clades in the phylogenetic tree (Fig. 1a). In the resulting matrix, all the specimens belonging to the same clade were collapsed to their barycentre in the PCoA configuration (Appendix S1). Subsequently, we created two functions: recluster.line and recluster.rotate. The first identifies the line connecting the most distant points in a configuration and computes its intercept and angular coefficient; the second rotates the points of a configuration to a new configuration where a line identified by its intercept and its angular coefficient is rotated to become horizontal. In practice, by using these functions in series, the specimens belonging to the two most diverging clades are polarized along the *x*-axis by maintaining the original configuration among individuals.

The first Procrustes analysis re-aligned the PCoA configuration obtained by genetic distances with the configuration of polarized clades. In such a way, the phylogenetic tree information is retained in the subsequent analyses as a horizontal ordination along a main genetic *x*-axis.

MINIMIZING DISSIMILARITIES BETWEEN CONFIGURATIONS AND COMPUTING POPULATION MEANS

We performed a second recluster.procrustes transformation of the morphological PCoA configuration based on the previously obtained

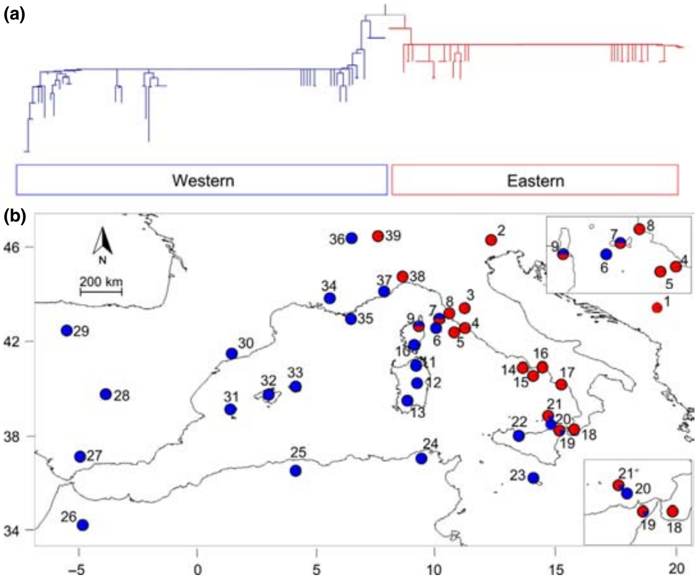


Fig. 1. (a) Structure of the neighbour-joining tree based on 218 COI sequences for *Maniola jurtina* confirming the existence of two main western and eastern clades (see Appendix S2 for a detailed representation). (b) Map of the study area indicating the proportion of individuals belonging to the western (blue) or eastern (red) COI lineages for sampled populations. Enlarged representations for the critical areas of Tuscan and Aeolian Islands are provided. Numbers correspond to areas listed in Table 1.

genetic one. Thus, the initial information fitting the phylogenetic variation along the *x*-axis was directly transferred to the resulting morphological configuration. The degree and significance of the correlation between the configuration of genetic and morphological data on the subset of shared specimens was evaluated by using the *vegan* protest function.

To study the genetic and morphological patterns at population level, we computed barycentre of specimens belonging to the same area in the two configurations. The protest function was applied to mean population coordinates to verify concordance between genetic and morphological configurations.

ASSESSING DISSIMILARITY LOSS AFTER COMPUTING POPULATION MEANS AND TESTING ITS SIGNIFICANCE

Because the two genetic PCoA axes were oriented to maximize differences among the most diverging phylogenetic clades along the first axis, we analysed results separately on the first and second axes. As a first step, we evaluated the amount of configuration variance lost along both axes by grouping specimens according to their population barycentre. If the distribution of specimens in the configuration is randomly scattered among populations, almost all population barycentres are expected to attain a rather central position with respect to the original individual points, which would result in a small mean distance between barycentres. The new recluster.test.dist function produces a ratio between the mean squared pairwise distance for all individuals and the mean squared pairwise distance for population barycentres (Bookstein 1991). This ratio is calculated for the overall configuration and for the two axes separately. Moreover, this function provides a test for the significance of the variation preserved by population means. To do this, 1000 matrices were obtained by randomly sampling the original vector defining population membership for each specimen. Then, we computed the frequency of mean squared distance ratios in random configurations higher than the observed ratio.

PROJECTING THE ROTATED CONFIGURATIONS IN RGB COLOUR SPACE AND COMPUTING POPULATION MEANS

We projected the two configurations together in the RGB space. For this purpose, the axis with the highest variance was standardized between 0 and 1 and the variance along the second axis rescaled according to the first one. Subsequently, the colours blue, green, yellow and red were assigned to the four corners. Finally, the contribution of each RGB colour to each site has been calculated on the basis of its position in the two-dimensional graph. This method is receiving an increasing interest in biogeography where it is employed to visually depict similarity patterns among elements (Kreft & Jetz 2010; Holt *et al.* 2013). We projected the colours for populations on a map, separately for genetic and morphological configurations.

ANALYSING MODALITY FOR AXIS VALUES IN POPULATIONS, REGRESSING THE GENETIC AND MORPHOLOGICAL POPULATION VALUES AND DEPICTING SPATIAL DISTRIBUTION OF RESIDUALS

We applied Hartigan's dip test of unimodality (Hartigan & Hartigan 1985) by using the R package *diptest* to verify whether the distribution along the axes significantly deviates from the unimodal distribution in all populations. Variance has been computed to verify whether populations hosting intermediate individuals show higher variance in shape. For this purpose, we made quadratic regressions for populations using mean values in morphological axes as predictors and their variance in the same axes as dependent variables. Finally, we regressed the mean values of morphological axes against the mean values of genetic axes and computed the residuals. In this case, we used major axis regression (MA) since it is specifically designed for minimizing errors for both variables. Thus, the existence of a dependent and independent variable is not presumed and the residuals reflect discrepancies in both genotypic and phenotypic traits (Claude 2008). Finally, we plotted these residuals on a map. To obtain a more conservative representation, we set to zero the residuals within the interval of standard deviation of absolute residual values.

Results

THE GENETIC MARKER

The 218 COI sequences obtained for *M. jurtina* represented 55 different haplotypes. The neighbour-joining phylogenetic tree supported the existence of two main clades largely matching the distribution of western and eastern populations previously recognized based on morphological and allozyme data (Dapporto *et al.* 2011; Thomson 2011) (Fig. 1a). In some areas located at contact zones, the two genetic lineages were found to coexist (Fig. 1b). A PCoA based on genetic distances confirmed a polarization between eastern and western lineages together with a higher differentiation among the haplotypes belonging to the western group (Fig. 2a, Appendix S1). The mean position in the PCoA configuration for the specimens belonging to each of the two clades was computed and the two points rotated to be aligned with the *x*-axis (Appendix S1). Subsequently, a second Procrustes aligned the original genetic configuration with this phylogenetic configuration of clades (see Appendix S1 for a step-by-step guide and for plots of intermediate configurations).

THE MORPHOLOGICAL MARKERS

We analysed 616 male specimens and obtained 12 and 42 relative warps from the brachium and valva analysis, respectively. A PCoA analysis based on these 53 variables revealed that most of the variance along both PCoA axes is explained by variation in the first PC of the valva and the brachium (as expected, since they explain most of the variance in their respective analyses, Appendix S1).

MINIMIZING DISSIMILARITIES BETWEEN CONFIGURATIONS

A series of 169 specimens was analysed for both genetic and morphological markers. A preliminary protest analysis on the rotated genetic and morphological PCoA configurations revealed that they are highly correlated (correlation coefficient 0·612, $P < 0{\cdot}001$). It should be noted that a higher correlation coefficient would be difficult to obtain due to the strict qualitative nature of COI. This resulted in high residuals for morphologically intermediate (potentially

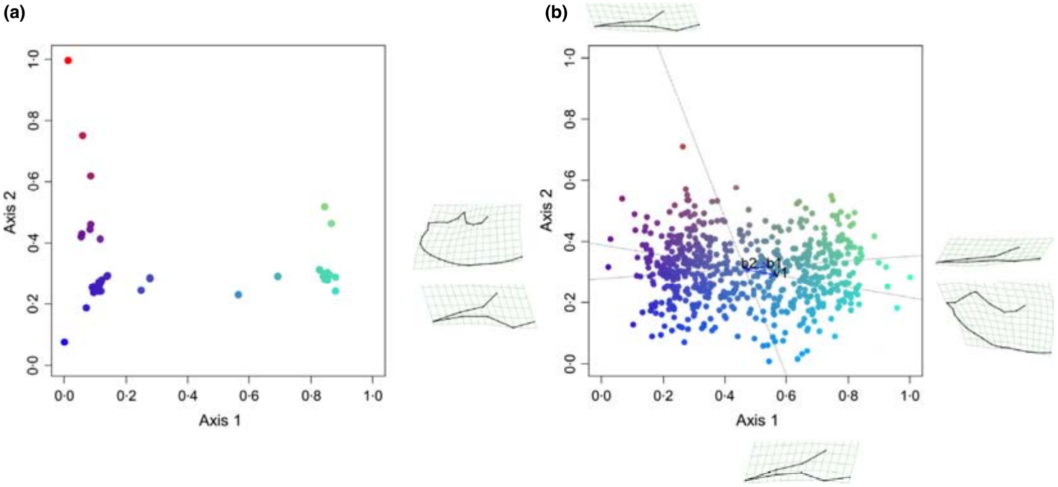


Fig. 2. (a) Principal coordinate analysis (PCoA) of genetic distance for sequenced specimens projected in the RGB colour space and rotated to maximize concordance with the existence of two phylogenetic clades over the *x*-axis (left, western lineage; right, eastern lineage). Specimens with identical haplotypes visually overlap and appear as a single dot. (b) PCoA of morphological data projected in the RGB colour space. Each point represents an individual with studied morphology. The arrows indicate the importance (length) and direction of the three main shape principal components (first PC for valva, v1, and first and second PCs for brachium, b1 and b2) in determining the configurations. The thin plate splines represent shape variation along the direction of main variables.

hybrid) specimens, which cannot have a corresponding intermediate genetic characteristic (see Appendix S1 for a graph). Subsequently, the morphological configuration with its variable coordinates was aligned with the phylogenetically rotated configuration of genetic distances by using the 169 shared specimens.

After computing the barycentre for populations, the correlation between genetic and morphological configurations resulted to be higher than for individuals. This was actually expected since, due to the presence of populations with both COI lineages, intermediate values for the genetic configuration are possible at population level (protest: correlation coefficient 0·797, $P = 0{\cdot}001$).

ASSESSING DISSIMILARITY LOSS AFTER COMPUTING POPULATION MEANS AND TESTING ITS SIGNIFICANCE

After grouping individuals according to population means, most of the original variability of specimen configuration was maintained for genetic data (85·04% *x*-axis and 61·23% *y*-axis). Both axes maintained a significantly higher divergence than in random configurations ($P < 0{\cdot}001$ and $P = 0{\cdot}007$, as it can be also verified by visual comparisons of Figs 2a and 3a). Conversely, the grouped morphological configuration (Fig. 2b) only maintained a larger than random amount of variation in the first *x*-axis (61·79%, $P < 0{\cdot}001$). In the *y*-axis, the maintained variance was only 14·15% with $P = 0{\cdot}079$. The *y*-axis in Fig. 3b is much more flattened than in Fig 2b, suggesting that most of the variation in genitalia is linked to the *x*-axis, oriented according to the western–eastern clade membership.

PROJECTING THE ROTATED CONFIGURATIONS IN RGB COLOUR SPACE AND COMPUTING POPULATION MEANS

The transformed genetic and morphological configurations have been projected in the RGB colour space. For the morphological data, the points have been plotted together with the standardized contribution of shape variables. Maintaining the contribution of variables enabled the inspection of the main pattern of shape variation, as carried out by thin plate spline in geometric morphometrics (Bookstein 1991). The alignment with the genetic axis also allowed us to clearly recognize the direction of morphological variation with respect to genetic variation (Fig. 2a,b).

Mean population values have been projected in the RGB space (Fig. 3) and the colours obtained for each population have been plotted in the geographic space (Fig. 4). There is high concordance between genetic and morphological data over most of the study area, but some exceptions exist. They are mostly due to the existence of endemic lineages in Mallorca and the Aeolian islands and to strong discordances located from the Maritime Alps to the Messina strait, along the Tyrrhenian coast of the Italian Peninsula.

ANALYSING MODALITY FOR AXIS VALUES IN POPULATIONS, REGRESSING THE GENETIC AND MORPHOLOGICAL POPULATION VALUES AND DEPICTING THE SPATIAL DISTRIBUTION OF RESIDUALS

In the first morphological axis, no population showed a distribution that significantly differed from unimodal (Hartigan's dip test of unimodality $P > 0{\cdot}050$ in each case, see Appendix

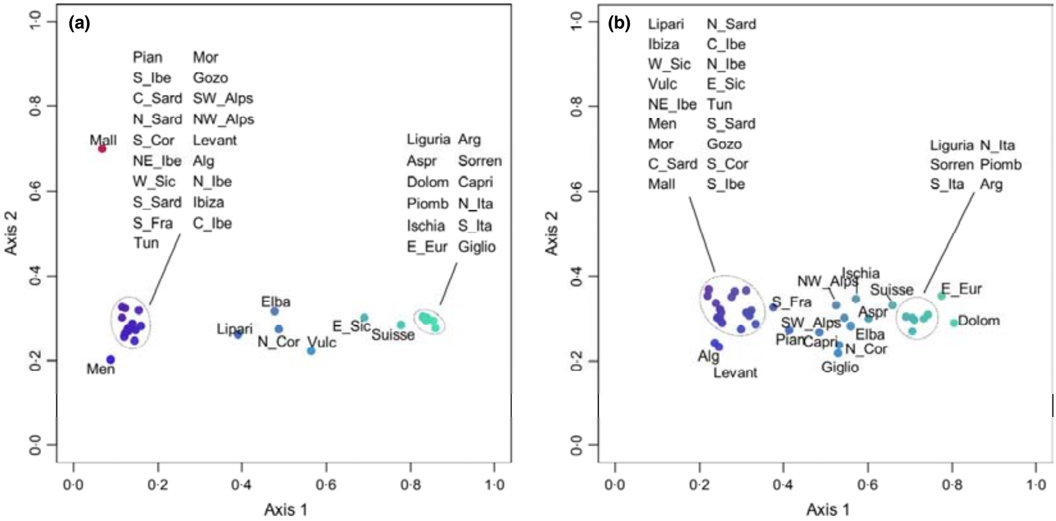


Fig. 3. Principal coordinate analysis (PCoA) projected in the RGB colour space representing mean population configurations for (a) genetic and (b) morphological data. See Table 1 for population abbreviations.

S1 for histograms). Variance in the x-axis of morphological configuration did not reveal higher values for intermediate populations, indicating that supposedly hybrid populations are not more morphologically variable (Appendix S1 for figure). Both quadratic coefficients revealed non-significant ($P > 0{\cdot}050$) effects on variance (Appendix S1).

A linear MA regression revealed a highly significant correlation between genetic and morphological x-axis values among populations (slope = 0·587, elevation = 0·194, $P < 0{\cdot}001$, $R^2 = 0{\cdot}732$, Fig. 5a). Nevertheless, the scatterplot revealed that several populations show large discrepancies (Fig. 5a). The values for residuals exceeding the standard deviation limits have been plotted on a map by using an appropriate colour scale for the dots (Fig. 5b). This allowed highlighting particularly discordant populations: populations showing more eastern genitalia with respect to their western genetic affiliation (north-western Alps, south-western Alps and, to a lesser extent, Pianosa) and vice versa (eastern Sicily, Vulcano, Capri and, to a lesser extent, Lipari, Giglio and Ischia). The sign of the residuals showed a highly coherent spatial pattern, with populations displaying more eastern morphology than expected based on genetic data being located in the northern Mediterranean, between the western Alps, Corsica and the Tuscan islands, while populations showing more western morphology than expected were located along the southern Tyrrhenian coast (Fig. 5b).

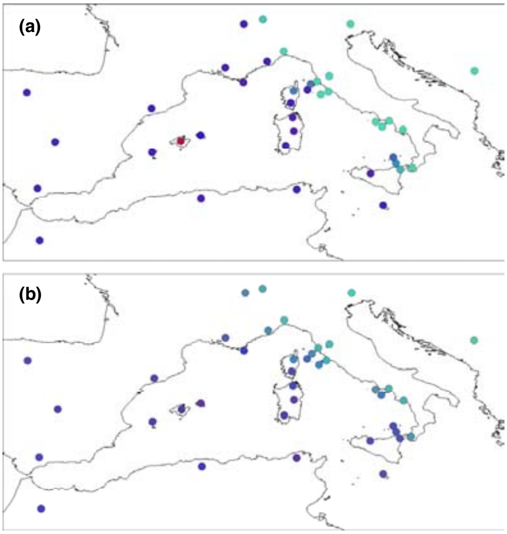


Fig. 4. Geographical location of populations with RGB colours as displayed in Fig. 3. (a) COI; (b) genitalia morphology.

Discussion

The method implemented by using the new *recluster* functions allowed a detailed assessment of the degree of co-variation between a genetic marker and a continuous morphological marker at population level. As a novelty, this method enabled the alignment and thus the simultaneous analysis of specimens for which data on either one or both traits was available. The possibility to use such heterogeneous data sets makes it possible, for example, to optimize the balance between the costs of DNA sequencing and the necessity to examine large numbers of specimens for morphological analyses. Undoubtedly, the new recluster.consensus function will also facilitate the use of public databanks and open the door to new data mining strategies. In particular, population studies such as the one performed in the model data set are becoming a powerful tool for biogeography and conservation biology and are facilitating the

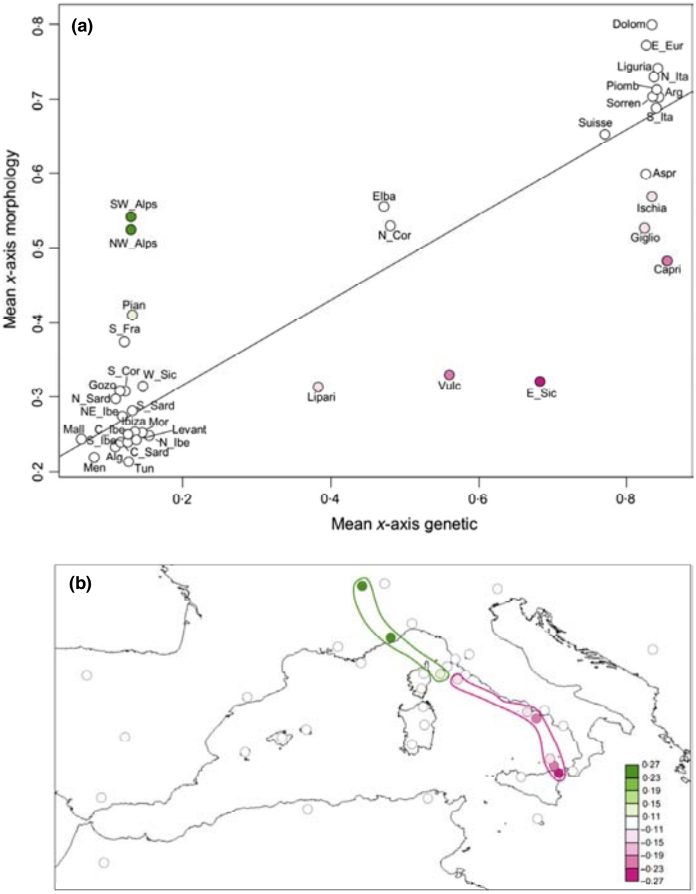


Fig. 5. (a) Major axis regression between genetic and morphological mean configurations for populations. Positive and negative residuals exceeding the standard deviation have been coloured with a green and purple scale, respectively. Values within the ± standard deviation interval are displayed in white. The positive residual for the completely eastern Dolomites population (Dolom) is represented in white. (b) Map showing the localization of large residuals over the study area.

recognition of speciation events, dynamics of taxa distribution and micro-evolutionary phenomena driven by various forces, including climate change (Schlick-Steiner *et al.* 2007; Dincă, Dapporto & Vila 2011; Dapporto *et al.* 2012; Renvoise *et al.* 2012; Toews & Brelsford 2012). Unlike previous assessments, the new procedure does not require the selection of morphological variables or components to be correlated with genetic patterns. Indeed, it has to be noted that the first morphological *x*-axis, to which most of our results belong, is profoundly different from a single morphological variable or PCA component since it is determined by a known contribution of several variables from two different genitalic structures. In most cases, several morphological components are correlated with the genetic signal and their contribution can vary. The first morphological axis we examined represents the best combination of such components to align the overall bidimensional morphological data to a genetic configuration obtained by combining phylogenetic relationships and overall genetic distance patterns.

The analysis allowed describing and evaluating the correlation between COI and morphological markers in *M. jurtina*. Most importantly, the possibility to align morphological and genetic data and to obtain mean values for populations allowed further examinations such as the modality and variance of the morphological traits as associated with genetic variation. Morphological traits revealed to be unimodal in all populations, thus excluding the possibility that the two lineages can cohabit while maintaining different shapes of genitalia, as found, for example, for the butterfly *Zerynthia cassandra* in the same geographic area (Zinetti *et al.* 2013). Populations at contact areas did not display higher variance than the rest, indicating that the mixing of the two genetic lineages produces morphologically homogeneous intermediate populations, a result that fully supports the subspecific status of the two taxa studied. However, the analysis revealed that some populations belong to a single genetic clade while showing intermediate morphology (like in the western Alps or in Capri). Some populations were also characterized by high residuals between morphological and genetic traits, like those in the western Alps and in eastern Sicily. The Vulcano population is especially interesting because, while it demonstrated a rather typical western morphology, it displayed an endemic genetic lineage (also detected in the neighbouring island of Lipari) that was sister to the rest of the western clade. Indeed, in the mean genetic PCoA configuration, the Vulcano population was placed at a rather intermediate position between the eastern and western clades, and this was not a consequence of hosting genetically mixed populations as was the case of northern Corsica, Elba, eastern Sicily and Lipari (Fig. 3a). This result suggests a rather old origin for the Vulcano–Lipari endemic genetic lineage, probably closely following the original split between the eastern and western clades. Mallorca hosts another example of an endemic genetic lineage, which was well placed within the western clade but showed signs of substantial drift, as indicated by the genetic PCoA y-axis (Fig. 3a). This lineage has been found at higher altitudes in the Serra de Tramuntana, while in lowland areas, the typical western lineage exists. In this apparently chaotic situation, a highly ordered spatial pattern of discordance was revealed, with discrepancies being located along the contact zone between the French Alps and the Messina strait. The sign of the residuals also clustered well in the two northern and southern halves of the contact zone. Several forces may determine such a pattern, for example, different selective pressures on the two latitudinal areas, differential introgression or dispersal between males and females (Descimon & Mallet 2009; Dasmahapatra *et al.* 2010; Gompert *et al.* 2010; Habel *et al.* 2011; Mallet, Wynne & Thomas 2011; Renvoise *et al.* 2012; Toews & Brelsford 2012; Mende & Hundsdoerfer 2013; Zinetti *et al.* 2013; Toews *et al.* 2014). What is important to point out is that a detailed study of the correlation between genetic and morphological markers within a geographic framework revealed that phylogeographic histories and ongoing dynamics are more multifaceted than hypothesized. The algorithms described in this article can be used with any kind of traits for which bidimensional configurations can be obtained, not necessarily a genetic and a morphological one. Moreover, more than two traits can be studied by using a series of Procrustes analyses and the possibility to handle missing data makes the new functions suitable for meta-analyses. The COI sequences and genitalia morphology of *M. jurtina* have provided an adequate model for this methodological study, and it is our hope that this procedure will be useful to many other organisms and characters.

Acknowledgements

We thank colleagues who provided samples used in this study. Financial support was provided by the Spanish Ministerio de Ciencia e Innovación project CGL2010-21226/BOS, by the project "Barcoding Italian Butterflies and by the project “Definizione dello status di conservazione delle falene e della malacofauna terrestre dell'Arcipelago Toscano”. Vlad Dincă was funded by a post-doctoral fellowship from the Wenner-Gren Foundation, Sweden and by a EU Marie Curie International Outgoing Fellowship for Career Development (project no. 625997). Raluca Vodă is funded by a FPU predoctoral fellowship (AP-2010-5409) from the Spanish Ministerio de Ciencia e Innovación. This study was performed in collaboration with the Parco Nazionale dell'Arcipelago Toscano.

Data accessibility

DNA sequences: GenBank accession numbers

KM020807–KM020882

KJ994239–KJ994253KM033847–KM033941

Morphological data: online supporting information

Geographic coordinates for specimens: online supporting information

R scripts: online supporting information and CRAN (http://cran.at.r-project.org/web/packages/recluster/index.html)

References

Avise, J.C. (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3–15.

Bookstein, F.L. (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge UK; New York.

Claude, J. (2008) *Morphometrics with R*. Springer, New York, London.

Dapporto, L. & Bruschini, C. (2012) Invading a refugium: post glacial replacement of the ancestral lineage of a Nymphalid butterfly in the West Mediterranean. *Organisms Diversity & Evolution*, **12**, 39–49.

Dapporto, L., Bruschini, C., Baracchi, D., Cini, A., Gayubo, S.F., Gonzalez, J.A. & Dennis, R.L.H. (2009) Phylogeography and counter-intuitive inferences in island biogeography: evidence from morphometric markers in the mobile butterfly *Maniola jurtina* (Linnaeus) (Lepidoptera, Nymphalidae). *Biological Journal of the Linnean Society*, **98**, 677–692.

Dapporto, L., Habel, J.C., Dennis, R.L.H. & Schmitt, T. (2011) The biogeography of the western Mediterranean: elucidating contradictory distribution patterns of differentiation in *Maniola jurtina* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, **103**, 571–577.

Dapporto, L., Bruschini, C., Dinca, V., Vila, R. & Dennis, R.L.H. (2012) Identifying zones of phenetic compression in West Mediterranean butterflies (Satyrinae): refugia, invasion and hybridization. *Diversity and Distributions*, **18**, 1066–1076.

Dapporto, L., Ramazzotti, M., Fattorini, S., Talavera, G., Vila, R. & Dennis, R.L.H. (2013) recluster: an unbiased clustering procedure for beta-diversity turnover. *Ecography*, **36**, 1070–1075.

Dasmahapatra, K.K., Lamas, G., Simpson, F. & Mallet, J. (2010) The anatomy of a ‘suture zone’ in Amazonian butterflies: a coalescent-based test for vicariant geographic divergence and speciation. *Molecular Ecology*, **19**, 4283–4301.

Derryberry, E.P., Derryberry, G.E., Maley, J.M. & Brumfield, R.T. (2014) hzar: hybrid zone analysis using an R software package. *Molecular Ecology Resources* **14**, 652–663.

Descimon, H. & Mallet, J. (2009) Bad species. *Ecology of Butterflies in Europe* (eds J. Settele, T.G. Shreeve, M. Konvicka & H. Van Dyck), pp. 219–249. Cambridge University Press, Cambridge.

Dincă, V., Dapporto, L. & Vila, R. (2011) A combined genetic-morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and Polyommatus celina Common Blue butterflies. *Molecular Ecology*, **20**, 3921–3935.

Gompert, Z., Lucas, L.K., Fordyce, J.A., Forister, M.L. & Nice, C.C. (2010) Secondary contact between *Lycaeides idas* and L-melissa in the Rocky Mountains: extensive admixture and a patchy hybrid zone. *Molecular Ecology*, **19**, 3171–3192.

Habel, J.C., Dieker, P. & Schmitt, T. (2009) Biogeographical connections between the Maghreb and the Mediterranean peninsulas of southern Europe. *Biological Journal of the Linnean Society*, **98**, 693–703.

Habel, J.C., Lens, L., Rödder, D. & Schmitt, T. (2011) From Africa to Europe and back: refugia and range shifts cause high genetic differentiation in the Marbled White butterfly *Melanargia galathea*. *Bmc Evolutionary Biology*, **11**, 215.

Hartigan, J.A. & Hartigan, P.M. (1985) The dip test of unimodality. *Annals of Statistics*, **13**, 70–84.

Holt, B., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araujo, M.B., Dimitrov, D. *et al.* (2013) An Update of Wallace's Zoogeographic Regions of the World. *Science*, **339**, 74–78.

Kalinowski, S.T. (2009) How well do evolutionary trees describe genetic relationships among populations? *Heredity*, **102**, 506–513.

Kreft, H. & Jetz, W. (2010) A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, **37**, 2029–2053.

Legendre, P. & Legendre, L. (1998) *Numerical Ecology*, 2nd English edn. Elsevier, Amsterdam; New York.

Madeira, C., Alves, M.J., Mesquita, N., Silva, S.E. & Paula, J. (2012) Tracing geographical patterns of population differentiation in a widespread mangrove gastropod: genetic and geometric morphometrics surveys along the eastern African coast. *Biological Journal of the Linnean Society*, **107**, 647–663.

Mallet, J., Wynne, I.R. & Thomas, C.D. (2011) Hybridisation and climate change: brown argus butterflies in Britain (Polyommatus subgenus Aricia). *Insect Conservation and Diversity*, **4**, 192–199.

Mardia, K.V., Kent, J.T. & Bibby, J.M. (1979) *Multivariate Analysis*. Academic Press, London; New York.

Mende, M.B. & Hundsdoerfer, A.K. (2013) Mitochondrial lineage sorting in action – historical biogeography of the *Hyles euphorbiae* complex (Sphingidae, Lepidoptera) in Italy. *Bmc Evolutionary Biology*, **13**, 83.

Mila, B., Toews, D.P.L., Smith, T.B. & Wayne, R.K. (2011) A cryptic contact zone between divergent mitochondrial DNA lineages in southwestern North America supports past introgressive hybridization in the yellow-rumped warbler complex (Aves: Dendroica coronata). *Biological Journal of the Linnean Society*, **103**, 696–706.

Paradis, E. (2012) *Analysis of Phylogenetics and Evolution with R*, 2nd edn. Springer, New York.

Pavlova, A., Amos, J.N., Joseph, L., Loynes, K., Austin, J.J., Keogh, J.S., Stone, G.N., Nicholls, J.A. & Sunnucks, P. (2013) Perched at the mito-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution*, **67**, 3412–3428.

Pergams, O.R.W. & Lacy, R.C. (2008) Rapid morphological and genetic change in Chicago-area Peromyscus. *Molecular Ecology*, **17**, 450–463.

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

Renvoise, E., Montuire, S., Richard, Y., Quere, J.P., Gerber, S., Cucchi, T., Chateau-Smith, C. & Tougard, C. (2012) Microevolutionary relationships between phylogeographical history, climate change and morphological variability in the common vole (*Microtus arvalis*) across France. *Journal of Biogeography*, **39**, 698–712.

Schlick-Steiner, B.C., Seifert, B., Stauffer, C., Christian, E., Crozier, R.H. & Steiner, F.M. (2007) Without morphology, cryptic species stay in taxonomic crypsis following discovery. *Trends in Ecology & Evolution*, **22**, 391–392.

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

Seraphim, N., Marín, M.A., Freitas, A.V.L. & Silva-Brandão, K.L. (2013) Morphological and molecular marker contributions to disentangling the cryptic *Hermeuptychia hermes species* complex (Nymphalidae: Satyrinae: Euptychiina). *Molecular Ecology Resources*, **14**, 39–49.

Thomson, G. (2011) *The Meadow Brown Butterflies*. Waterbeck, Scotland.

Toews, D.P.L. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Toews, D.P.L., Mandic, M., Richards, J.G. & Irwin, D.E. (2014) Migration, mitochondria, and the Yellow-Rumped Warbler. *Evolution*, **68**, 241–255.

Viscosi, V. & Cardini, A. (2011) Leaf Morphology, Taxonomy and Geometric Morphometrics: a Simplified Protocol for Beginners. *PLoS ONE*, **6**, e25630.

Zelditch, M., Swiderski, D.L. & Sheets, H.D. (2012) *Geometric Morphometrics for Biologists: A Primer*, 2nd edn. Elsevier Academic Press.

Zinetti, F., Dapporto, L., Vovlas, A., Chelazzi, G., Bonelli, S., Balletto, E. & Ciofi, C. (2013) When the rule becomes the exception. No evidence of gene flow between two Zerynthia cryptic butterflies suggests the emergence of a new model group. *PLoS ONE*, **8**, e65746.

Received 2 April 2014; accepted 10 June 2014
Handling Editor: Louise Johnson

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. R scripts (AppendixS1.doc) and data (geno.txt, morpho.txt, names.txt, new functions.R and gendist.csv) used to perform the analyses.

Appendix S2. Supplementary methods.

Chapter 5

Vodă R., Dapporto L., Dincă V., Shreeve T., Khaldi M., Barech G., Rebbas K., Sammut P., Scalercio S., Hebert P.D.N., Vila R. What makes island communities unique? Understanding and protecting butterfly diversity at the interface between Africa and Europe.

What makes island communities unique? Understanding and protecting butterfly diversity at the interface between Africa and Europe

Raluca Vodă[a,b]*, Leonardo Dapporto[c,1]*, Vlad Dincă [a,d], Tim Shreeve[c], Mourad Khaldi[e], Ghania Barech[f], Khellaf Rebbas[f], Paul Sammut[g], Stefano Scalercio[h], Paul D.N. Hebert[d], Roger Vila[a]

[a]Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37, 08003, Barcelona, Spain
[b]Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain
[c]Department of Biological and Medical Sciences, Oxford Brookes University, Headington, Oxford, OX3 0BP, UK
[d]Biodiversity Institute of Ontario, University of Guelph, Guelph, N1G 2W1, Ontario, Canada
[e] Département d'Agronomie, Université Mohamed Boudiaf de M'sila, 28000 M'sila, Algeria
[f] Département des sciences de la nature et de la vie, Université Mohamed Boudiaf de M'sila, 28000 M'sila, Algeria
[g]137, "Fawkner/2" Dingli Road, Rabat RBT 9023, Malta
[h]Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Unità di Ricerca per la Selvicoltura in Ambiente Mediterraneo, c.da Li Rocchi, I-87036 Rende (CS), Italy

*Equal contribution

[1]To whom correspondence should be addressed: Leonardo Dapporto, e-mail: leondap@gmail.com

Keywords: community assembly, conservation biogeography, Lepidoptera, Pleistocene, relict species, western Mediterranean

ABSTRACT

Islands possess varied histories, eco-geographical settings, and experience stochastic events. Because of this complexity, the mechanisms determining and maintaining insular species assemblages are poorly understood. We used a comparative approach linking community and phylogeographic analyses for the entire butterfly fauna of the circum-Sicilian islands, a key intercontinental region of the Mediterranean. This allowed to recognizing the processes determining the observed assemblages and highlighting the uniqueness of each community. Species richness was mainly influenced by contemporary factors, but Pleistocene connections also had a significant impact. Species frequency on mainland was a weaker predictor than dispersal tendency for their occurrence on islands. When all species were considered, nestedness and species turnover had similar contributions in determining faunistic dissimilarities with a ratio between nestedness and Sørensen indexes of 0.603. When divided into sub-matrices, the widespread species showed a predominance of the nestedness component (0.814) and the uncommon species a predominance of species replacement (0.271). Many populations from Europe and north Africa formed well-differentiated and mutually-excluding genetic groups, while island populations showed a high similarity with their closest mainland and with the areas they were connected to during glaciations. "Uncommon" species, tended to be genetically more diverse than "widespread" ones and played a key role in determining the uniqueness of the communities. Our results highlight the value of coupling community structure and phylogeography for reconstructing biogeographic processes and for taking evidence-based conservation decisions.

BACKGROUND

Island communities share several features: the number of species is lower than on similarly sized mainland areas (impoverishment), they host a disproportionate fraction of non-predatory and highly dispersive species (disharmony) and they are characterized by a high fraction of endemic elements [1]. In the last decades it has been largely recognized that these features are shaped by events of colonization and extinction but also by the long-term persistence of populations, resulting in relictuality and endemicity [1,2,3]. However, these processes depend on a large number of mechanisms related to both species and island characteristics, as well as to stochastic events [1,4]. As a result, no island populations and communities are identical and they have been regarded as "individual" [5]. Hence, recognizing the main drivers behind the formation of each community and developing specific conservation plans require in-depth comparative analyses linking community and phylogenetic approaches [6,7,8].

An important property of island communities is that their structure tends to be nested, with some species occurring on most islands, and others occurring on fewer islands, usually the largest and least isolated [9,10]. The existence of nested patterns has important implications for conservation; when communities are highly nested, conservation decisions are simplified and the usual strategy is to concentrate efforts on the most diverse communities [1]. After a series of studies suggesting that nested structures universally occur within different taxa and island systems [11], the recent use of strict null models revealed that significantly nested patterns are less common than previously hypothesized [12]. There is also a debate [6,13] as to whether species occurrence in island communities is mostly determined by their frequency in neighbouring source areas, as postulated by neutral theories [2,14], or whether competition and interactions between species traits and island characteristics play a primary role [3,10,15].

Genetic variation can also provide fundamental information to understand the historical and ecological dynamics that shaped island populations and communities [6,16]. Past connections or long-term isolation are well-known determinants of the genetic structure of island populations, but they exert variable effects on different species, with highly migratory species expected to show the least geographic variation due to frequent gene flow [16,17,18]. However, the assumption that species lacking genetic structure are the most vagile and widespread has rarely been tested [6,16,19]. Following these premises, it is not surprising that, although numerous studies exposed the role of complex phenomena in producing faunistic [20,21] and genetic community structures for selected insular taxa [19,22], comprehensive studies aiming to disentangle large arrays of ecological-historical and deterministic-stochastic factors over large taxonomic groups are uncommon [8,16,17].

In this study, we examined butterfly communities occurring on the circum-Sicilian islands and compared them with populations from southern Italy, Sicily and north Africa. The Sicilian Channel represents a biogeographic crossroad with high species richness and contrasting biodiversity [23,24,25]. These islands have different geological histories and locations with respect to the two main faunistic sources (southern Europe and north Africa) and possess different environmental settings. European butterflies constitute an excellent model system since their taxonomy and distribution are well known. However, the sparse information about their genetic structure has impeded the link between community composition and phylogeographic patterns. In this study we: i) model island richness for butterflies over the entire western Mediterranean region, ii) examine the pattern of nestedness in the study area, iii) analyse the relative importance of species dispersal tendency and frequency at source in determining their frequency on islands, iv) assess the phylogeographic patterns of 27 species and test for a correlation between dispersal tendency and regional genetic variation, v) document an overall phylogeographic pattern for north Africa, circum-Sicilian islands, Sicily, and the Italian Peninsula, and vi) categorize the species based on their island occupancy and genetic structure. This integrative approach aids the recognition of the multiple processes generating species assemblages for an entire and diverse superfamily and provides the information needed to prioritize conservation decisions in a key biogeographical contact zone.

METHODS

Study area and data collection

We analysed the butterfly faunas of 11 circum-Sicilan islands (Lampedusa, Levanzo, Linosa, Lipari, Maltese islands, Marettimo, Pantelleria, Salina, Stromboli, Ustica, Vulcano), Sicily itself and nearby mainland locations in southern Italy, Tunisia and Algeria. Presence data were gathered from several post 1980 literature sources and from field surveys carried out by the authors between 1999 and 2015. Specimens used for genetic analyses were collected outside protected areas and were deposited at the Institute of Evolutionary Biology (CSIC-UPF), Barcelona, Spain.

Determinants for island species richness

To evaluate the influence of different factors on island richness on the islands, we used Generalized Linear Models (GLM) with Akaike Information Criterion (AIC). We also assessed the relative importance of potentially correlated variables using hierarchical partitioning of variance, employing the *lmg* metric implemented in the “relaimpo” R Package. To place the richness of the study islands in a broader framework, we employed data on the butterfly fauna from all the western Mediterranean islands [24] (excluding Sicily, Sardinia and Corsica because their much larger size would provide an unbalanced contribution). Islands were included in the study if they harboured at least four out of five migrant species (*Pieris brassicae*, *Pieris rapae*, *Colias croceus*, *Vanessa atalanta* and *Vanessa cardui*). These species are conspicuous and widespread throughout Europe and north Africa, and their presence establishes a minimal surveying standard [24].

In the GLMs, species richness was modelled against the following biotic, geographic and climatic predictors: i) mean annual temperatures (MT), ii) annual precipitation (AP), iii) island area (IA), iv) maximum elevation of island (EL), v) isolation (IS), vi) source richness (SR) and vii) the occurrence of the Pleistocene connections (PC, factor variable). MT and AP were obtained by computing the mean values of the cells corresponding to islands in Bio1 and Bio12 layers from Bioclim (http://www.worldclim.org). For each island the faunistic source was identified as the nearest 50×50 km area, either mainland (southern Europe or north Africa), large island (Sicily, Sardinia, Corsica), or an island at least ten times larger than the target one. We calculated SR as the number of species reported from the source area and IS as the minimum distance between the target island and this source. To linearize the relationships, we log transformed richness, area, isolation and source richness. We computed a first GLM only using contemporary variables and a second GLM in which we included PC. We compared the AIC, explanatory power and residuals of the selected variables between the two models.

Community structure

We estimated the degree of nestedness of butterfly communities in circum-Sicilian islands by applying the widely accepted NODF metric [9]. Recent reviews have shown that the detected degree of nestedness depends heavily on the selected null model due to their different tendency to preserve features of the original matrix [12]. Therefore, to assess the significance of the observed NODF, we used the NeD program [26] and computed the z-values using 999 null matrices, built by applying a series of different null models (equiprobable rows and columns, EE; proportional rows and columns total, CE; and fixed rows and columns, FF) [12].

Prior studies have divided species into “core” and “satellite” species [27] in which “core” species are those occurring on more than half of the islands, and “satellite” those present on fewer islands. We used the same approach and divided the 27 species occurring on all the studied islands into “widespread” (those occurring on six or more of the 11 islands) and "uncommon" (those found on no more than five islands). We also analysed the nested structure of the two subsets separately.

The overall faunistic dissimilarities among areas can be partitioned between a component generated by nestedness and one by species replacement (turnover) [28]. We calculated the relative contribution of nestedness and turnover as the mean value of the ratio between the faunistic dissimilarities obtained by nestedness and Sørensen indexes (see Appendix S1 and S3 in Supporting Information) [29]. This ratio has been computed for all species and for uncommon and widespread species, separately.

We tested for the effect of species’ dispersal tendency and their frequency at source on their occupancy on islands. Measurements of the dispersal tendency in butterflies are complex and in previous studies have been mostly based on the agreement between subjective evaluations by experts and less commonly by using objective species traits. Here we combined the indexes provided by four papers [30, 31, 32, 33] by standardizing their range between 0 and 1 and by computing, for each species, an average dispersal tendency based on the available measurements. For *Danaus chrysippus* and *Zizeeria karsandra* no evaluation has been made and they have been excluded from the analysis. To estimate the frequency at source we counted the number of cells of 0.1×0.1 degrees between 34° and 40° latitude and 7° and 18° longitude, in which each of the 31 investigated species has been reported. As standardized sources we used “CKmap2000”, an online checklist of the Italian butterfly fauna [34], and the database of the Butterfly Diversity and Evolution Lab at The Institute of Evolutionary Biology (CSIC-UPF), Spain. The frequency of each species on islands was regressed against dispersal tendency and frequency at source and their relative importance tested by hierarchical partitioning of variance.

Genetic analyses and identification of study units

Using standard sequencing procedures (Appendix S1) we obtained cytochrome *c* oxidase subunit I (COI) sequences for 888 specimens (Appendix S2) from the study islands and from five surrounding areas: mainland Italy (Calabria), eastern Sicily (>14 ° longitude), western Sicily (<14 ° longitude), Tunisia and Algeria. We considered only the islands for which we had available more than 80% of the reported fauna. This led to the exclusion of Linosa and left ten islands and a total of 15 areas to be analysed.

The butterfly species currently recognized by taxonomists show different levels of intraspecific genetic divergence and they can have an unbalanced contribution to the overall biogeographic pattern [35, 36]. To reduce this bias, we identified as units the groups of individuals having minimum COI p-distances to the nearest neighbour of less than 3%, a measure that, in Lepidoptera, was reported to separate more than 90% of recognized sister species [37]. By applying this threshold we identified 27 units (termed species) closely matching the taxonomy proposed by Fauna Europaea (www.faunaeur.org). The list of species is provided in Appendix S1.

Overall phylogenetic pattern

To examine patterns of genetic variation for each species in the study area we constructed haplotype networks for all species with TCS Network in PopART (http://popart.otago.ac.nz). We also calculated the genetic p-distances among all sequenced specimens for each species, and obtained Specimens Dissimilarity Matrices (SDM, n=27). For each SDM we calculated the mean genetic distance among specimens from the same study area and produced the Area Dissimilarity Matrices (ADM, n=27). The mean values of the corresponding cells with available distances between all 27 ADM were used to produce a final Mean Inter-Deme Genetic Variation Matrix (GVM), representing the degree of mtDNA differentiation among areas for all species. A Principal Coordinate Analysis (PCoA) was applied to the GVM to obtain the overall genetic pattern for the 15 areas. Subsequently, we aligned this configuration with the geographic location of the areas by using the "procrustes" analysis from the "vegan" R package (Appendix S3 for R scripts).

To visualize the pattern of similarity among islands over geographic space, we projected the final configuration of genetic distances among areas in the RGB space [24,38] using the R package “recluster”. The colour resemblance of the resulting dots is directly proportional to the genetic similarity among the communities. These dots were then placed in their corresponding geographic positions on a bathymetric map, generated with the GEBCO Grid display software, where we outlined the -100m depth contour as a reliable reconstruction of land during the last glacial maximum (LGM) [20].

Islands located between two genetically contrasting sources, can show intermediate communities if: i) islands host species characterized by genetic variation and half individuals belong to each source, or ii) islands host only a reduced fraction of species characterized by genetic diversification, and therefore look equally similar to both sources. To visualize the pattern of genetic similarity and to compare the degree of genetic variation among areas we computed the Mean Inter-Deme Variation (IDV) value for each species, as the mean of all the values in the ADM of each species. We then plotted the relationship between the PCoA configuration and the mean IDV of the species from each area in 3D. If the first hypothesis is met, we expect that intermediate islands will have similar mean IDV to other islands, while if the second is true we expect that intermediate islands will have lower mean IDV values.

Analysis of spatial genetic variation

We tested for the correlation between IDV and both dispersal tendency and island occupancy (number of studied islands occupied by each species) using the non-parametric Spearman correlation rank test. We also square root transformed the IDV values to improve normality and divided the bivariate relationship between occupancy and IDV into four quadrants, and considered the species either as ‘widespread’ or ‘uncommon’ and the genetic variation as low (< half) or high (>half) of the maximum square root transformed IDV. The number of species falling into each quadrant was scored. Subsequently, 999 null matrices were constructed by attributing to each species random values of occupancy and IDV. For each square, we assessed the frequency for which the observed number of species was lower than in random configurations. Values < 0.050 were considered as significantly empty quadrants, and values > 0.950 as significantly full quadrants.

RESULTS

Determinants for island species richness

For the GLM calculated on the basis of ecological factors, a model with four variables (IA, IS, SR and EL) best fitted the data, with isolation explaining the largest part of variance (Table 1). The Maltese islands, Lampedusa, Lipari and Vulcano were richer in species than expected (Fig. S1a), while Pantelleria, Linosa, Stromboli and Marettimo had fewer species than expected. In the second GLM (including the occurrence of PC as a variable) a six variable model best fitted the data (Table S3), adding MT and PC to the previous model. In order to make this model directly comparable with the first one, we excluded MT (results including these variables are shown in Table S3). This model increased the explained variance to 74.65% with PC having the lowest *lmg* value (Table 1). With this model, richness on the Maltese islands was almost perfectly predicted, whilst Lampedusa still had more species than expected and Levanzo, compared to the perfect fit in the previous model, had a negative residual (Fig S1b).

Table 1 Results for the AIC-based stepwise Generalized Linear Model predicting species richness on islands based on: IA = island area; EL = maximum elevation of island; IS = island isolation from the nearest source; SR = butterfly richness of the nearest source; PC = occurrence of Pleistocene connection. Mean annual temperature and annual precipitation did not enter the model. Abbreviations: Est = estimated parameter; SE = Standard Error; t = t value; P = P value; lmg = percentage of explained variance attributed by hierarchical partition of variance.

	Est.	S.E.	t	P	lmg%
MT					
AP					
IA	0.074	0.038	1.923	0.066	14.8
EL	0.190	0.068	2.796	0.010	12.7
IS	-0.111	0.047	-2.346	0.027	15.0
SR	0.597	0.126	4.740	<0.001	19.8
PC	0.167	0.069	2.413	0.023	12.3
					74.65%

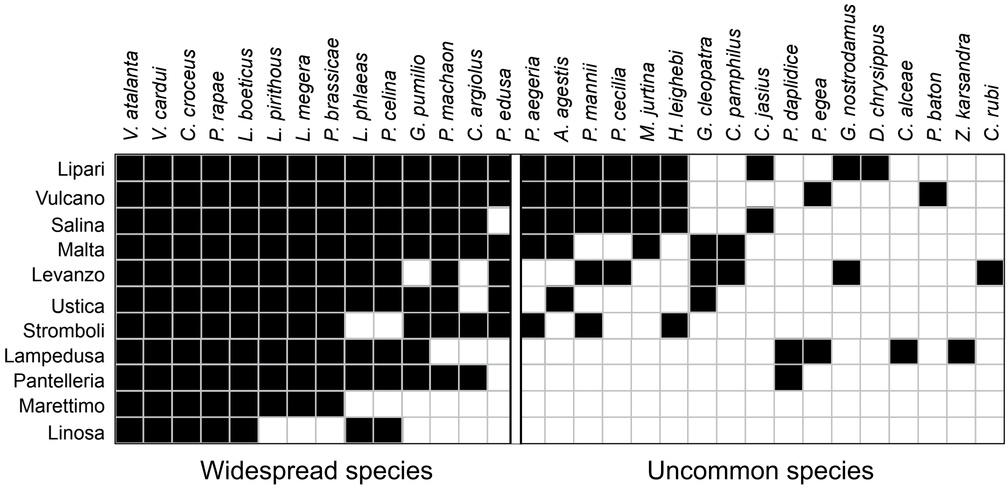
The richness residuals of the circum-Sicilian islands that were connected with the mainland during the Pleistocene (Table 2), suggest that the Maltese islands have particularly high residuals in the model without the PC, compared to the full model. This is the first evidence supporting an important contribution of this historical event in assembling the current butterfly fauna of the Maltese islands.

Table 2 The residuals scored in the Generalized Linear Models by each island having had contact with mainland areas during the LGM, obtained either by removing (Res. 1) or including (Res. 2) the Pleistocene connection as a categorical predictor. The difference between the two values can serve as an estimate for the number of species that colonized the islands during the Last Glacial Maximum and survived until present. The Maltese islands had the highest difference in residuals between the two models.

Island	Res. 1	Res. 2	Diff.
Maltese	3.5	-1.2	4.6
Levanzo	-0.1	-2.6	2.6
Lampedusa	4.6	2.2	2.4

Community structure

Various null models provided different evidence for the existence of a significantly nested pattern. The observed NODF of the entire packed matrix (Fig. 1) was significantly higher than the mean NODF obtained with the EE and CE null models (Table 3). However, the mean NODF for the FF null model was significantly higher than the observed value, revealing an anti-nested pattern (Table 3). After dividing the matrix into two sub-matrices, widespread species showed a nested pattern for the EE and CE null models, while uncommon species only showed a significantly nested pattern for the non-restrictive EE null model and a significantly anti-nested pattern with the FF model (Table 3). In particular, species occurring on less than four islands showed a highly chequered distribution (Fig. 1).



	V. atalanta	V. cardui	C. croceus	P. rapae	L. boeticus	L. pirithous	L. megera	P. brassicae	L. phlaeas	P. celina	G. pumilio	P. machaon	C. argiolus	P. edusa	P. aegeria	A. agestis	P. mannii	P. cecilia	M. jurtina	H. leighebi	G. cleopatra	C. pamphilus	C. jasius	P. daplidice	P. egea	G. nostrodamus	D. chrysippus	C. alceae	P. baton	Z. karsandra	C. rubi
Lipari	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■			■	■				
Vulcano	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■					■				■		
Salina	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■			■								
Malta	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■		■	■									
Levanzo	■	■	■	■	■	■	■	■	■	■		■		■			■	■			■	■				■					■
Ustica	■	■	■	■	■	■	■	■	■	■	■	■		■		■					■										
Stromboli	■	■	■	■	■	■	■	■			■	■	■	■	■		■			■											
Lampedusa	■	■	■	■	■	■	■	■	■	■	■													■	■			■		■	
Pantelleria	■	■	■	■	■	■	■	■	■	■	■	■	■											■							
Marettimo	■	■	■	■	■	■	■	■																							
Linosa	■	■	■	■	■				■	■																					

Figure 1 The packed matrix minimizing the NODF metric for the butterfly species present on the islands examined in this study. The rows represent the studied islands and the columns the species occurring in each island. Widespread species showed a much more nested pattern than uncommon ones.

Accordingly, on the basis of all the species, nestedness and turnover had similar contributions in determining faunistic dissimilarities with a ratio between nestedness and Sorensen indexes of 0.603. When divided into sub-matrices, the widespread species showed a predominance of the nestedness component (0.814) and the uncommon predominance of species replacement (0.271).

A multiple regression showed that species frequency on islands was significantly associated with both frequency in the study area (n=29, Est.=0.027, t=2.76, P=0.010, Fig. 2a, S2) and with dispersal tendency (n=29, Est.=6.038, t=3.04, P<0.001, Fig. 2b) but partitioning of variance revealed that dispersal tendency explained a higher fraction of variation than frequency at source (0.318 for dispersal tendency and 0.187 for frequency at source).

Genetic pattern

Some species showed very little genetic diversification with only one or a few closely related haplotypes (e.g. *C. croceus*, *Gonepteryx cleopatra*, *Pseudophilotes baton*, Appendix S4). Other highly mobile species were more variable, but the different haplotypes did not show any geographic pattern (e.g. *P. rapae* and *V. cardui*, Fig. 3).

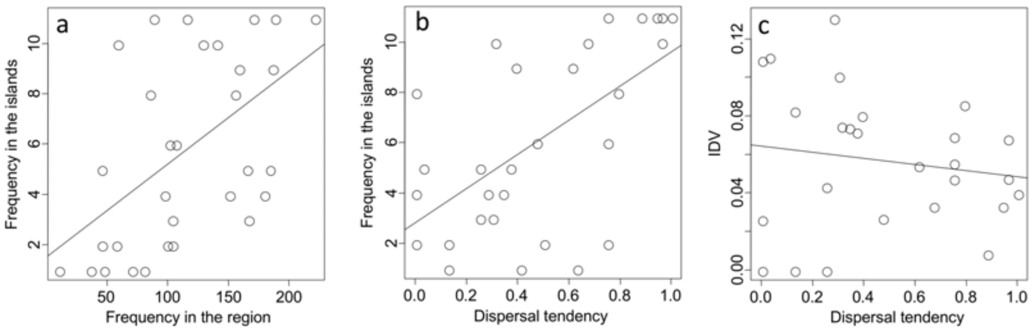


Figure 2 (a) The correlation between species occurrence on islands and their frequency at source in cells of 0.1×0.1 latitude-longitude degrees and **(b)** the correlation between species occurrence on islands and their dispersal tendency; the size of the circles is proportional with the number of species having identical values. **(c)** The correlation between IDV and dispersal tendency is non-significant.

Table 3 Results for the ratio between the nestedness beta diversity index (βnest) and the unpartitioned Sorensen index (βsor) and observed NODF for the nestedness analyses, using all species, only uncommon species and only widespread species. Mean NODF and Z values are provided for different series of 999 matrices generated by different null models (EE, equiprobable rows and columns; CE, proportional rows and columns total; and FF, fixed rows and columns). Asterisks represent associated P values.

	βnest/βsor	**NODF**	**Null Model**	**Mean NODF**	**Z**
All species	0.603	75.613	EE	51.673	9.664***
			CE	60.560	4.739***
			FF	77.549	-2.698**
Uncommon species	0.271	39.491	EE	32.451	1.649*
			CE	37.739	0.275
			FF	43.557	-2.016*
Widespread species	0.814	82.013	EE	71.095	2.378*
			CE	74.553	1.989*
			FF	82.206	-1.196

Several species were genetically differentiated between north Africa and Europe (examples provided in Fig. 3, the remaining species are available in Appendix S4). Lineages of some species like *Papilio machaon*, *Lycaena phlaeas*, *Lasiommata megera* (Fig. 3), *Carcharodus alceae, Aricia agestis* and *Polyommatus celina* (Appendix S4), showed patterns of mutual exclusion, even between the two sides of the narrow Messina strait. Other species like *Hipparchia* spp., *Maniola jurtina* and *Pyronia cecilia* had distinct lineages that in some areas occur in sympatry (Fig. 3).

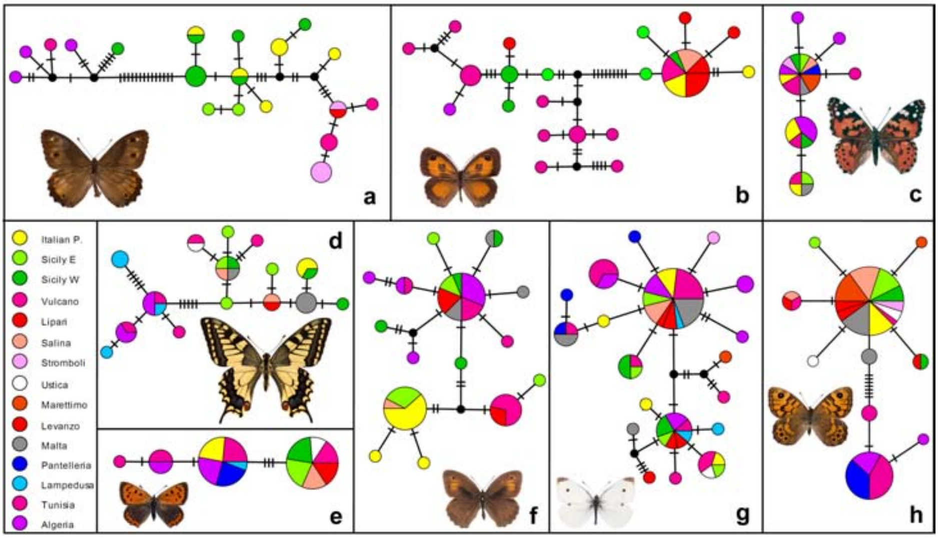


Figure 3 COI haplotype networks for eight of the 27 investigated species (for the rest see Appendix S4): species with genetic variation but no spatial segregation (*V. cardui* c, *P. rapae* g); species with spatial differentiation but no mutually exclusive pattern among lineages (*Hipparchia spp.* a, *P. cecilia* b*, M. jurtina* f); species showing evidence of mutual exclusion (*P. machaon* d, *L. phlaeas* e, *L. megera* h). Some of these species also have endemic lineages.

As for the characteristics of island communities, Lampedusa and Pantelleria usually hosted north African lineages, while the rest of the islands had lineages from Sicily and the Italian Peninsula. Accordingly, the PCoA based on the mean phylogeographic variation (GVM) showed that Europe and north Africa were genetically well-differentiated. Levanzo, Ustica, the Maltese and Aeolian islands were genetically similar to Sicily-Italy, while Marettimo, Pantelleria and Lampedusa had a more intermediate position between Europe and North Africa (Fig. 4a,b, Appendix S1). When the mean IDV of the species occupying each area was plotted on the z-axis, islands with intermediate geographic positions showed that they contained species with low genetic variation (Fig. 4a).

Correlations between species occurrence and genetic variation

IDV was not significantly correlated either with dispersal tendency (Spearman rho=-0.102, P=0.612, Fig 3c) or with species frequency on islands (Spearman rho=-0.178, P=0.375). When all islands are considered, no sector was significantly full or empty (Fig. 5a). But, when examined individually, the faunas of Lipari and Salina showed a significant negative correlation between IDV and frequency (Table S3). Moreover, the widespread and genetically undiversified species were always more common than expected in a random distribution and there was a deficit of uncommon, little diversified species on all islands except for Lampedusa and Vulcano (Table S3 and Fig. 5). The Aeolian Islands, together with the Maltese Islands and Levanzo did not show a significant deficit for uncommon and diversified species.

DISCUSSION

Island communities

The integration of faunistic and phylogeographic information ultimately allowed the recognition of the main processes shaping island communities in a key biogeographic region. The evaluation of the contribution of singular factors in characterizing island diversity facilitated the identification of species that deserve conservation efforts on individual islands. Species entering the butterfly communities and their genetic structure were determined by neutral (frequency at source) and non-neutral (dispersal tendency, mutual exclusion) mechanisms and by unique combinations of island traits like isolation, area and paleogeography. One of the main findings of this study, stemming from the community and genetic data, is that even in a relatively restricted area, island populations have very different histories and are subjected to different recolonization probabilities following local extinctions. Similar comprehensive examinations of island

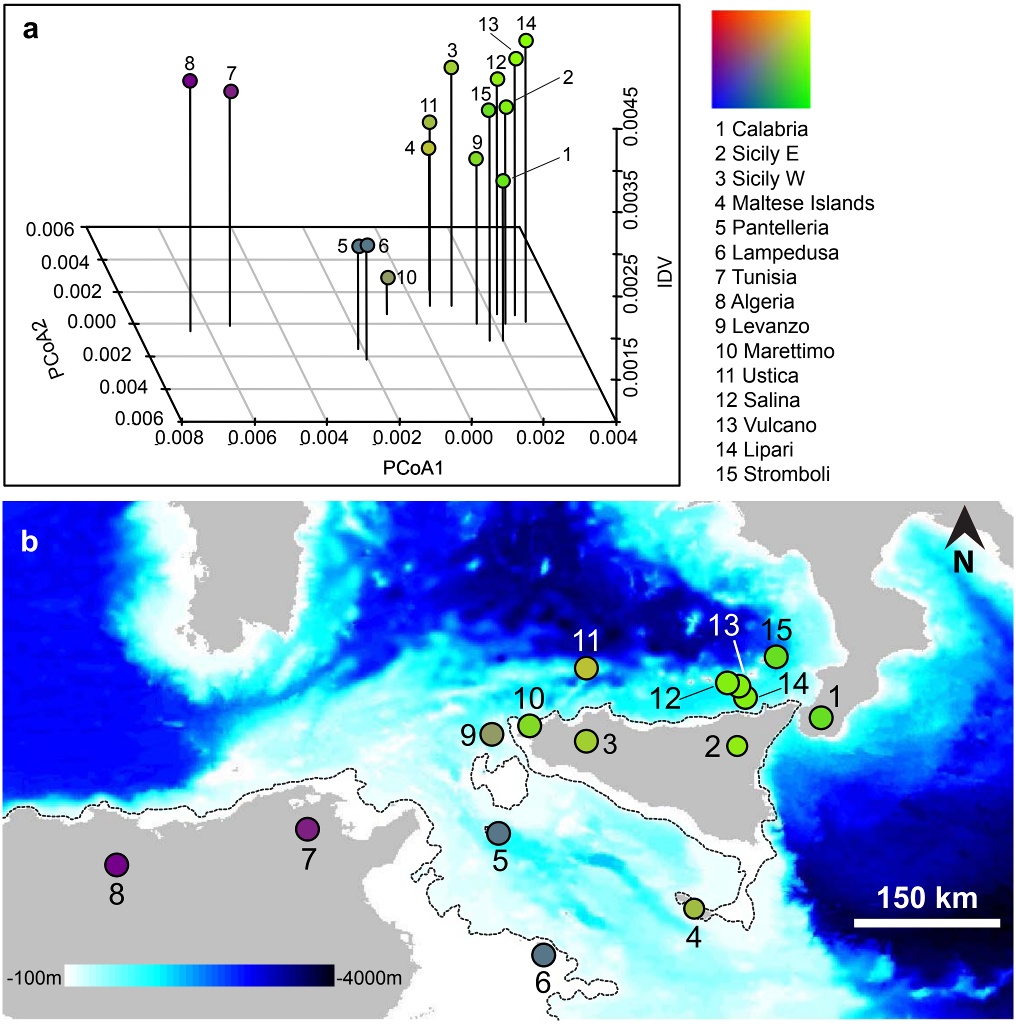


Figure 4 (a) Principal Coordinates Analysis (PCoA) showing the pattern of similarity among islands based on the mean genetic distances among the 27 studied species for the 15 areas (Axis 1 and 2). The vertical axis represents the mean inter-deme genetic variation (IDV) of the species on each island; the top right coloured square is the RBG configuration used for attributing colours to the areas. **(b)** The coloured dots in the bi-dimensional representation of PCoA have been plotted on a bathymetric map (equirectangular projection) showing the -100m isobaths (white), which represent a reliable reconstruction of seashores during the last glacial maximum. The area codes for the numbers in (a) are provided in (b).

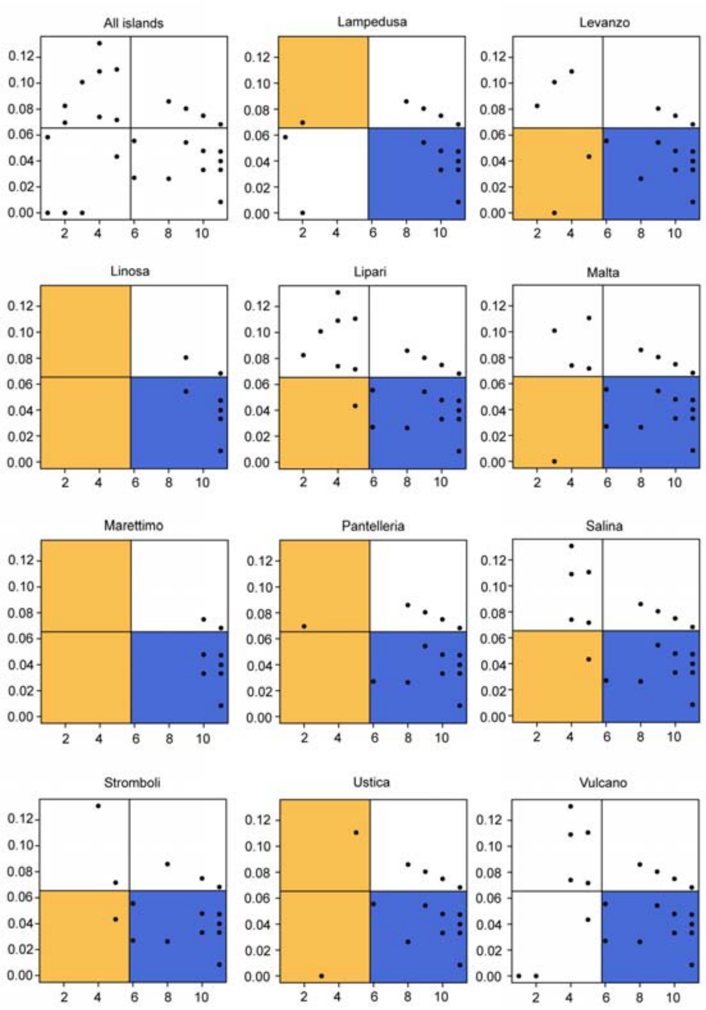


Figure 5 Correlations between species occurrences (x-axis) and inter-deme genetic variation (IDV) (y-axis) for the species occurring on: **(a)** all the islands, **(b)** the Maltese islands, **(c)** Pantelleria, **(d)** Lampedusa, **(e)** Levanzo, **(f)** Marettimo, **(g)** Ustica, **(h)** Vulcano, **(i)** Lipari, **(j)** Salina, **(k)** Stromboli and **(l)** Linosa. The colours indicate if a certain square is significantly filled (blue), significantly empty (orange) or showing no significant differences compared to the values obtained by chance (white). All islands are significantly enriched in widespread species with low genetic variation, and most are impoverished in uncommon taxa.

communities are scarce, especially due to the difficulty of obtaining high quality data at both faunistic and genetic levels for large groups of organisms [8,22,39,40].

Patterns of diversity and their determinants

In contrast to very isolated oceanic islands, the Western Mediterranean islands maintained relatively strong biotic relationships with the neighbouring mainland, especially for organisms with high dispersal ability, such as butterflies [24,41]. Accordingly, our analyses of faunistic diversity and intraspecific genetic variability showed that contemporary island-source dynamics, filtered by the main ecological characteristics of islands (predominantly area and isolation), are the main determinants structuring butterfly communities in the circum-Sicilian islands. Although a direct comparison between dispersal tendency and genetic variation did not result in a significant correlation (Fig. 2c), about half of the species in these communities are taxa with high dispersal ability (*P. rapae, P. brassicae, C. croceus, Leptotes pirithous, Lampides boeticus, V. atalanta* and *V. cardui*) [42] and/or with low genetic diversification (Fig. 5). Island populations of these species likely include both local residents and recent immigrants [43]. Besides these widespread species, island communities include uncommon species, most of which showed considerable genetic variation over the study area. Out of 31 species, 9 occur only on one or two islands and 17 on less than half of them. Although their overall number is high, they represent a relatively small part of each island's diversity (on average 27.8%). These uncommon species had a non-significant nested pattern and generated inter-island faunistic dissimilarity mostly because of turnover. Frequency at source had a contribution in determining the occurrence of species on islands, but contrary to what was expected by a purely neutral hypothesis, dispersal tendency played a major role and many sedentary species were uncommon on islands despite their ubiquitous distribution in the entire study area (*M. jurtina, C. pamphilus, P. aegeria, A. agestis-A. cramera, P. cecilia*).

The non-nested pattern of uncommon species reveals that basic rules of ecological filtering and turnover dynamics are not sufficient to explain occupancy patterns and a larger array of determinants is involved, including: i) dispersal from local source populations (e.g. *P. baton* to Aeolian Islands, widespread north African species to Lampedusa), ii) interaction with sister species and lineages (as revealed by mutual exclusion, see below), iii) human impact and iv) Pleistocene connections (having probably enriched at least the Maltese fauna). Theory predicts that relict species from Pleistocene connections should be sought in the group of uncommon and sedentary species with relatively high genetic variation and similarity between formerly connected areas. Marked differences between the two continents in butterfly diversity are well known [24,25,35,44] and individual islands showed different degrees of similarity to the two sources, clearly based on their relative distance and paleogeographic connections. The richness analysis attributed to Malta a set of species that colonized it from Sicily during the LGM and in fact all the uncommon species occurring there and showing regional genetic diversification have COI sequences that are identical to populations in Sicily.

Another strong pattern revealed by the molecular analyses is the pervasive occurrence of mutual exclusion. This is evident not only for several pairs of cryptic species (*Pontia edusa-P. daplidice, P. icarus-P. celina, A. agestis-A. cramera, Coenonympha pamphilus-C. lyllus*) [44] but also at the intraspecific level. Many of the geographically clustered lineages displaying over 1% intraspecific genetic divergence (*C. alceae, P. machaon, L. phlaeas, A. agestis, L. megera, P. aegeria*) show mutually exclusive patterns not only across the wide channel separating Sicily from north Africa but also across the 3 km Messina strait between Sicily and Calabria. Mutually exclusive distributions challenge the hypothesis of neutrality among species and lineages, but the relative importance of possible mechanisms is still unknown. A combination of reproductive interference, reduced dispersal, density-dependent phenomena and differences in climatic niches is probably at the basis of the observed patterns [44].

Conservation of communities as individuals

All the islands investigated host a significantly high fraction of dispersive species having widespread distributions in the study area and displaying low levels of genetic differentiation. Their continuous presence on islands is expected because any local extinction event would soon be followed by recolonization involving a population genetically similar to the original one. This is obvious for those species that reach Europe every spring but are unable to overwinter there [43], but here we show that many others display a similar distributional and genetic signature. Any action to preserve local island populations of these widespread species should be cautiously considered since it could represent a loss of economic resources. On the contrary, uncommon species should be carefully monitored and conservation efforts immediately directed towards declining populations [45]. Their non-nested distribution patterns, with a prevalence of unpredictable species replacements between islands, and their regionally structured genetic diversity, are strong indications that their island populations have unique histories [1,6,12]. Moreover, uncommon species generate the faunistic identity and the particular genetic structure of each island community, thus providing an unbalanced contribution to diversity [35].

Only three endemic lineages were detected in the study islands. We found that the supposed Aeolian endemic *H. leighebi* likely represents only an endemic lineage of *H. semele* based on mitochondrial DNA, while a COI haplotype of *M. jurtina,* common on Vulcano and Lipari, was not found elsewhere [46] (Fig. 3, Appendix 4). Moreover, Lampedusa hosts a diverging lineage of the north African clade of *P. machaon.* If population declines are detected for these entities, they should be considered as priority taxa for conservation in the islands where they occur. Priority for conservation should be also given to populations that apparently represent glacial relicts and that, in the current ecological settings, would unlikely recolonize a specific island. In fact, biodiversity loss on the studied islands seems to primarily affect these species (see Appendix S4 for detailed information), possibly because they lack metapopulation dynamics that maintain population numbers (rescue effect) and are affected by inbreeding depression. The most striking example is that of the Maltese islands, where *L. phlaeas* and *A. agestis* are believed to be extinct, while *M. jurtina, C. pamphilus* and *P. aegeria* are rapidly declining due to human impacts (Paul Sammut personal observations). These populations require immediate conservation actions to preserve the uniqueness of the Maltese butterfly fauna.

Levanzo and Lampedusa also had connections with Sicily and Tunisia, respectively, during the LGM. Due to the proximity to Sicily, most butterflies on Levanzo are probably part of a metapopulation and no positive richness residuals were found for this island (Fig. 1a). An exception could be *C. pamphilus*, which was represented on Levanzo by distinct haplotypes. In Lampedusa, all the uncommon species are highly dispersive and widely distributed in north Africa but not in Europe, probably because of climatic restrictions or species interactions leading to mutual exclusion [35,44].

However, the divergent haplotype of *P. machaon* should be considered as a first rank candidate species for local conservation actions.

Mutual exclusion should also be taken into consideration for conservation decisions. Density dependent and founder-takes-all mechanisms [47,48] suggest that the populations existing in an area/island represent the main barrier to the colonization by other lineages, which would change the original genetic structure, probably established after a series of unrepeatable historical events [23]. The fauna of Pantelleria is mostly composed of widespread and undiversified species, but three of them (*P. celina, L. phlaeas, L. megera*) have different lineages in Sicily and Tunisia. Interestingly, all three species are represented on Pantelleria by typical north African populations, although the island was closer to Sicily during the lowered sea levels of the last glaciation. The similarity to Tunisia is the most important characteristic of this island's butterfly fauna, but its cause is unknown. The extinction of these populations may result in recolonization from Sicily, thus erasing a key biogeographic signal on Pantelleria.

Documenting the uniqueness of single islands and of butterfly populations requires diverse analyses. Our approach provides an example of how a series of strict and comprehensive analyses on a wide area and large taxonomic group can test rarely assessed biogeographic principles, while providing the baseline information essential for developing conservation strategies that maximize biodiversity at both the species and genetic levels.

Acknowledgements

This research was supported by the Spanish Ministerio de Economía y Competitividad (Project CGL2013-48277-P and a predoctoral fellowship to R. Vodă, AP- 2010-5409) and by the projects "Barcoding Italian Butterflies" and "Barcoding Butterflies of the Tuscan Archipelago National Park." V. Dincă was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme (project no. 625997).We thank Axel Hausmann, Anthony Seguna and Michael Seizmair for providing samples.

REFERENCES

1. Whittaker RJ, Fernández-Palacios JM. 2007 Island biogeography: ecology, evolution, and conservation. Oxford; New York: Oxford University Press.
2. MacArthur RH, Wilson EO. 1967 The theory of island biogeography. Monographs in Population Biology, 1.
3. Lomolino MV. 2000 A species-based theory of insular zoogeography. *Global Ecology and Biogeography* **9**, 39-58.
4. Schoener TW. 2009 The MacArthur-Wilson Equilibrium Model: What it said and how it was tested. pp 52-87 in J. B. Losos and R. E. Ricklefs (Eds.). The Theory of Island Biogeography Revisited. Princeton University Press, Princeton NJ.
5. Cronk Q. 1999 Islands as individuals. *Trends in Ecology & Evolution* **14**, 81.
6. Cavender-Bares J, Kozak KH, Fine PV, Kembel SW. 2009 The merging of community ecology and phylogenetic biology. *Ecology Letters* **12**, 693-715.
7. Fernández-Palacios JM, Kueffer C, Drake D. 2015 A new golden era in island biogeography. *Frontiers of Biogeography* **7**: 14-20.
8. bis Kueffer C, Drake DR, Fernández-Palacios JM 2014 Island biology: looking towards the future. *Biology letters*, **10**: 20140719.
9. Ulrich W, Almeida M, Gotelli NJ. 2009 A consumer's guide to nestedness analysis. *Oikos* **118**, 3-17.
10. Dennis RLH, Hardy PB, Dapporto L. 2012 Nestedness in island faunas: novel insights into island biogeography through butterfly community profiles of colonization ability and migration capacity. *Journal of Biogeography* **39**: 1412-1426.
11. Watling JI, Donnelly MA. 2006 Fragments as islands: a synthesis of faunal responses to habitat patchiness. *Conservation Biology* **20**, 1016-1025.
12. Matthews TJ, Cottee-Jones HEW, Whittaker RJ. 2015 Quantifying and interpreting nestedness in habitat islands: a synthetic analysis of multiple datasets. *Diversity and Distributions* **21**: 392-404.
13. Clark JS. 2012 The coherence problem with the Unified Neutral Theory of Biodiversity. *Trends in Ecology & Evolution* **27**, 198-202.
14. Hubbell SP. 2010 Neutral Theory and the Theory of Island Biogeography, pp. 264-292. in J. B. Losos and R. E. Ricklefs (Eds.). The Theory of Island Biogeography Revisited. Princeton University Press, Princeton NJ.

15. Diamond JM. 1975 Assembly of species communities. In: Ecology and Evolution of Communities (eds Cody, M.L. & Diamond, J.M.). Belknap Press of Harvard University Press, Cambridge, pp. 342-444.
16. Dawson MN, Hays CG, Grosberg RK, Raimondi PT. 2014 Dispersal potential and population genetic structure in the marine intertidal of the eastern North Pacific. *Ecological Monographs* **84**, 435-456.
17. Burney CW, Brumfield RT. 2009 Ecology predicts levels of genetic differentiation in Neotropical birds. *The American Naturalist* **174**, 358-368.
18. Carnicer J, Stefanescu C, Vila R, Dincă V, Font X, Peñuelas J 2013 A unified framework for diversity gradients: the adaptive trait continuum. *Global Ecol Biogeogr* **22**, 6-18.
19. Papadopoulou A, Anastasiou I, Spagopoulou F, Stalimerou M, Terzopoulou S, Legakis A, Vogler AP. 2011 Testing the species–genetic diversity correlation in the Aegean Archipelago: Toward a haplotype-based macroecology? *The American Naturalist* **178**, 241-255.
20. Fattorini S. 2011 Biogeography of tenebrionid beetles (Coleoptera: Tenebrionidae) in the circum-Sicilian islands (Italy, Sicily): Multiple biogeographical patterns require multiple explanations. *European Journal of Entomology* **108**, 659-672.
21. Luna-Jorquera G, Fernández CE, Rivadeneira MM. 2012 Determinants of the diversity of plants, birds and mammals of coastal islands of the Humboldt current systems: implications for conservation. *Biodiversity and Conservation* **21**, 13-32.
22. Papadopoulou A, Anastasiou I, Keskin B, Vogler AP. 2009 Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference. *Molecular Ecology* **18**, 2503-2517.
23. Dapporto L, Bruschini C, Dincă V, Vila R, Dennis RLH. 2012 Identifying zones of phenetic compression in West Mediterranean butterflies (Satyrinae): refugia, invasion and hybridization. *Diversity and Distributions* **18**, 1066-1076.
24. Dapporto L, Fattorini S, Vodă R, Dincă V, Vila R. 2014 Biogeography of western Mediterranean butterflies: combining turnover and nestedness components of faunal dissimilarity. *Journal of Biogeography* **41**, 1639-1650.
25. Husemann M, Schmitt T, Zachos FE, Ulrich W, Habel JC. 2014 Palaearctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaearctic biota. *Journal of Biogeography* **41**, 81-94.

26. Strona G, Galli P, Seveso D, Montano S, Fattorini S. 2014 Nestedness for Dummies (NeD): a user friendly web interface for exploratory nestedness analysis. *Journal of Statistical Software* **59**, 1-9.
27. Ulrich W, Zalewski M. 2006 Abundance and co-occurrence patterns of core and satellite species of ground beetles on small lake islands. *Oikos* **114**, 338-348.
28. Baselga A. 2010 Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* **19**, 134-143.
29. Dobrovolski R, Melo AS, Cassemiro FA, Diniz-Filho JAF. 2012 Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* **21**, 191-197.
30. Balletto E, Kudrna O. 1985 Some aspects of the conservation of butterflies in Italy, with recommendations for a future strategy. *Boll Soc Ent Ital* **11**7, 39-59.
31. Komonen A, Grapputo A, Kaitala V, Kotiaho JS, Päivinen, J. 2004 The role of niche breadth, resource availability and range position on the life history of butterflies. *Oikos*, **105**, 41-54.
32. Heikkinen RK, Luoto M, Leikola N, Pöyry J, Settele J, Kudrna O, Marmion M, Fronkek S, Thuiller W 2010 Assessing the vulnerability of European butterflies to climate change using multiple criteria. *Biodiv Conserv*, **19**, 695-723.
33. Dennis RL 2012 A resource-based habitat view for conservation: butterflies in the British landscape. John Wiley & Sons.
34. Balletto E, Bonelli S, Cassulo L. 2005 Insecta Lepidoptera Papilionoidea. In: S. Ruffo & F. Stoch (Eds)- Checklist e distribuzione della fauna italiana. 10.000 specie terrestri e delle acque interne. *Mem. Mus. civico St. nat. Verona* **16**, 259-263.
35. Vodă R, Dapporto L, Dincă V, Vila R. 2015 Cryptic matters: overlooked species generate most butterfly beta-diversity. *Ecography* **38**, 405-409.
36. Dincă V, Montagud S, Talavera G, Hernández-Roldán J, Munguira ML, García-Barros E, Hebert PDN, Vila R. 2015 DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. *Scientific Reports* **5**:12395. doi:10.1038/srep12395.
37. Hebert PDN, Ratnasingham S, de Waard JR. 2003 Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(Suppl 1), S96-S99.

38. Kreft H, Jetz W. 2010 A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography* **37**, 2029-2053.

39. Helmus MR, Savage K, Diebel MW, Maxted JT, Ives AR. 2007 Separating the determinants of phylogenetic community structure. *Ecology Letters* **10**, 917-925.

40. Pyron RA, Burbrink FT. 2014 Ecological and evolutionary determinants of species richness and phylogenetic diversity for island snakes. *Global Ecology and Biogeography* **23**, 848-856.

41. Dennis RLH, Shreeve TG, Olivier A, Coutsis JG. 2000 Contemporary geography dominates butterfly diversity gradients within the Aegean archipelago (Lepidoptera: Papilionoidea, Hesperioidea). *Journal of Biogeography* **27**, 1365-1383.

42. Stefanescu C, Peñuelas J, Filella I. 2003 Effects of climatic change on the phenology of butterflies in the northwest Mediterranean Basin. *Global Change Biology* **9**, 1494-1506.

43. Stefanescu C, Páramo F, Åkesson S, Alarcón M, Ávila A, Brereton T, Carnicer J, Cassar LF, Fox R, Heliölä J, et al. 2013 Multi-generational long-distance migration of insects: studying the painted lady butterfly in the Western Palaearctic. *Ecography* **36**, 474-486.

44. Vodă R, Dapporto L, Dincă V, Vila R. 2015 Why do cryptic species tend not to co-occur? A case study on two cryptic pairs of butterflies. *PLoS ONE* **10**(2): e0117802. doi:10.1371/journal. pone.0117802.

45. Hunter ML, Hutchinson A. 1994 The virtues and shortcomings of parochialism: conserving species that are locally rare, but globally common. *Conservation Biology* **8**, 1163-1165.

46. Dapporto L, Vodă R, Dincă V, Vila R. 2014 Comparing population patterns for genetic and morphological markers with uneven sample sizes. An example for the butterfly *Maniola jurtina*. *Methods in Ecology and Evolution* **5**, 834-843.

47. Waters JM. 2011 Competitive exclusion: phylogeography's 'elephant in the room'? *Molecular Ecology* **20**, 4388-4394.

48. Waters JM., Fraser CI, Hewitt GM. 2013 Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution* **28**, 78-85.

Supporting Information:

Appendix S1 Supplementary methods and results

Appendix S2 List of samples sequenced for this study

Appendix S3 R scripts used in the study

Appendix S4 Detailed information about the species and the islands analysed

4. General Discussion

4. General discussion

As shown in chapter 1.3 of the Introduction we are currently experiencing a new 'golden era', not only in island biogeography, but in biological sciences in general. Based on an impressive amount of knowledge gathered in various research fields in the last centuries, the recent technological and methodological advances are taking biogeography and evolutionary biology to a new level, where researchers can test and validate complex eco-evolutionary hypotheses that were previously unimaginable (Fernández-Palacios et al. 2015a).

The data accumulated and the results of this PhD thesis are aiding this 'transition' from studies restricted taxonomically and geographically towards comprehensive macroecological approaches and analyses of large datasets in this field of research.

A prerequisite for the analytical work carried out and one of the main contributions of the studies here presented was **creating a unique sample collection and DNA sequence dataset** that represents a major product. Indeed, a comprehensive butterfly DNA and tissues collection for the western Mediterranean region was assembled and the occurrence data for island and mainland communities improved. During the four years that this PhD was developed, more than 50 field expeditions have been organised to five countries (Spain, France, Italy, Morocco, Tunisia) and 34 islands (Mallorca, Menorca, Ibiza, Levant, Gorgona, Elba, Pianosa, Capraia, Montecristo, Giglio, Giannutri, Capri, Ischia, Ventotene, Ponza, Sicily, Vulcano, Stromboli, Lipari, Salina, Alicudi, Levanzo, Marettimo, Ustica, Linosa, Corsica, Sardinia, Santa Maria, Spargi, San Pietro, La Maddalena, Caprera, Pantelleria) throughout the western Mediterranean. These field trips, carried out each spring-summer for the entire duration of the seasons (April-October), required considerable effort largely because several areas were often difficult to sample: high mountains, remote islands, countries with unstable political/social conditions. More than 10000 specimens were gathered corresponding to at least 240 species, which represent almost 72% of the entire western Mediterranean butterfly fauna. Together with the samples previously collected by the group from 2006 to 2011, the collection is the best existing for this region for butterfly DNA studies (fig. 19). This collection is deposited at the Institute of Evolutionary Biology (CSIC-UPF) in Barcelona, Spain, where the research has been done. One of the main purposes of this vast collection was to assemble a library of DNA sequences for as many species possible from the study area and with a good geographic coverage. Almost 6000 COI sequences belonging to the Western Mediterranean Butterflies (WMB) project, which our team manages, are currently deposited in BOLD, one of the largest repositories for DNA sequences in the world. BOLD imposes strict quality requirements for entries that gain the status of DNA barcodes: besides the DNA sequences, a series of collateral data are required, which include precise collection data (including GPS coordinates) and photographs of all the barcoded specimens. Moreover, the electropherograms based on which the final DNA barcode is obtained are also available. Other 850 specimens were sequenced by Macrogen and the DNA sequences are available in GenBank.

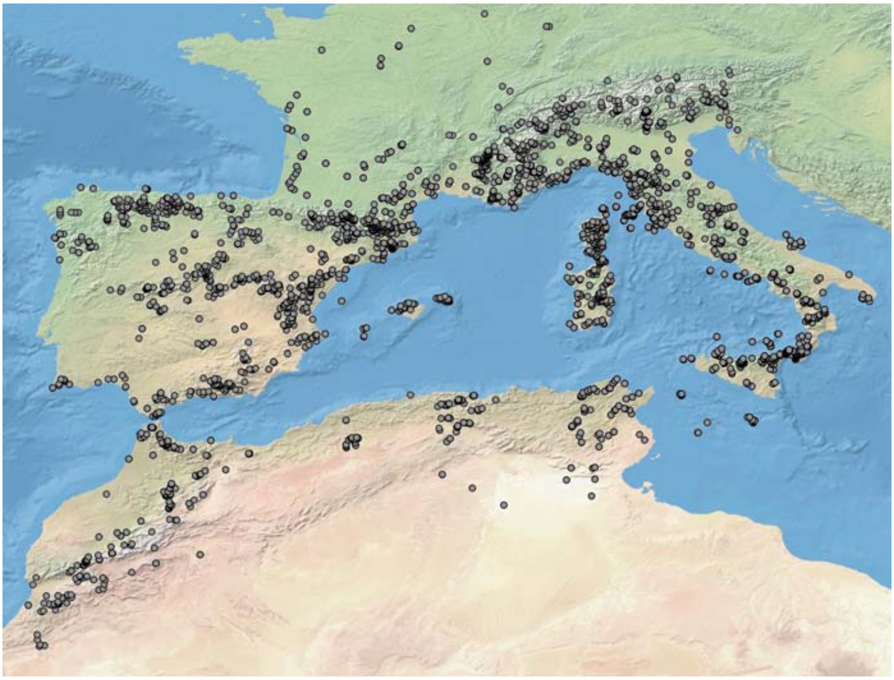


Figure 19 The western Mediterranean localities sampled for butterfly specimens between 2006-2015.

It is important to highlight the usefulness of the data produced as an original contribution of this PhD. Currently in Europe (and probably in the world) there are no other similar datasets available for an entire taxonomic group with such a wide distribution and with a comparable spatial resolution. This great amount of species presence data and DNA sequences provided the basis for the studies that have already been published, but they also represent a valuable source for future research.

Advancing towards obtaining a comprehensive map of genetic diversity for the entire western Mediterranean butterfly fauna that includes the temporal dimension in addition to the spatial ones is only possible on the basis of such a comprehensive database. Therefore, this collection represents a snapshot of the present Mediterranean diversity that could be compared with the future butterfly composition in order to trace changes at faunistic or genetic level. Moreover, this valuable resource will also open the door to large-scale comparative phylogeography at an unprecedented scale and resolution for a speciose taxonomic group. This will helps us enormously to obtain unprecedented insights into several key questions in fields such as wide-scale biogeography, climate change effects on biodiversity, genetics of extinction and will likely result in generalizations of patterns and concepts that are now only barely grasped.

Based on the highly improved dataset of presence records for the western Mediterranean butterflies, in the **first Chapter** of Results we investigated patterns of the butterfly beta-diversity in the western Mediterranean mainland and islands. The meaning of partitioned ($ß_{simp}$ and $ß_{nest}$) and non-partitioned ($ß_{sor}$) dissimilarity indices as measures of beta-diversity is still a highly debated issue (Shi 1993, Baselga 2010, Tuomisto 2010, Almeida-Neto et al. 2011). We showed that such indices reveal complementary patterns and that, taken together, they can comprehensively explain dissimilarity patterns of insular and mainland communities and reconstruct the main dynamics determining their composition (glacial interglacial pulses, past land-bridge connections, ongoing mainland-island faunal exchange). $ß_{sor}$ provided the first fundamental information for island biogeography about which communities were shaped by insular dynamics, since on the basis of this index it was possible to clearly identify the two clusters formed by island and mainland communities. Indeed, when islands were compared with mainland sites, the unpartitioned index was

dominated by the nestedness pattern, produced by the well-known impoverishment process occurring in island communities (Wittaker & Fernández-Palacios 2007, Dapporto & Dennis 2009) mostly determined by variations in island area and isolation. As a consequence, this index was largely ineffective in recognizing regionalization patterns between mainland sources and islands. Mainland communities are much less variable than islands in their richness and, when applied to mainland areas, the $ß_{sor}$ index provided fundamental biogeographical information (e.g. the location of different evolutionary regions and the existence of latitudinal and peninsular effects). On the other hand, $ß_{simp}$ was effective in recovering island–mainland relationships but resulted in a low signal for mainland dissimilarities, thus limiting the distinction among recognized regions. We suggest that there is no single index able to retrieve comprehensive biogeographical patterns simultaneously for islands and mainland and we propose comparing and combining results obtained using both unpartitioned and partitioned (nestedness and turnover) indices by means of regionalization and dissimilarity interpolation as the best solution.
The results of this study shed light on the effectiveness of different measures of beta-diversity to highlight specific eco-evolutionary processes that determine community assemblages, with important implications for zooregionalization and biogeography in general.

In **Chapter II** of Results we analysed the patterns of distribution for the cryptic and non-cryptic fractions of the butterfly species occurring in the western Mediterranean islands and mainland. For the first time, we objectively defined cryptic species and produced a list of cryptic groups for all the butterflies in the study region. As defined, the cryptic fraction of butterfly diversity represents about 25% of the western Mediterranean fauna. Based on this list and on the newly updated database of presence records we showed that co-occurrence within cryptic groups is much lower than among congeneric non-cryptic species. This result could be explained if we assume that cryptic groups are composed by entities that recently evolved in allopatry and maintained their distribution pattern simply because they did not have enough time to disperse (Pigot & Tobias 2013). However, many cryptic taxa in our study are not sister species and show substantial genetic divergence, while other morphologically differentiated species are their sister taxa (Dincă et al. 2011a, Sañudo-Restrepo et al. 2013). The mechanisms maintaining chequered distributions, in particular across specific narrow sea straits, are still unknown, but a combination of density dependent, adaptive and interactive processes could represent the most probable explanation. It is likely that, in addition to temporal and physical constraints that hamper dispersal, most cryptic species have been impeded to colonize new areas by the presence of other members of the same group in the recipient areas (Waters et al. 2013), a phenomenon that was also demonstrated in birds (Pigot & Tobias 2013).

These findings represent a change of paradigm in showing that cryptic diversity comprises idiosyncratic qualitative aspects in addition to merely quantitative ones and highlight the importance of differentiating cryptic species for various research fields.

This study showed that the chequered patterns of distribution are a general phenomenon for cryptic butterfly taxa in the western Mediterranean as could also be the case for other organisms. But which are the mechanisms that generate and maintain them?

In **Chapter III** of Results we addressed this question and investigated potential causes producing such chequered distribution patterns by using a multidisciplinary approach. We performed genetic landscape, haplotype networks and ecological niche modelling analyses for two pairs of cryptic species, *Aricia agestis* - *A. cramera* and *Polyommatus icarus* - *P. celina* (Lycaenidae) to specifically assess non-coexistence on western Mediterranean islands and test potential causes for their distribution. We showed that the strongest genetic contrasts between both cryptic pairs occurred over short distances, which corresponded to the Bonifacio and Messina straits, even though on the basis of the estimated dispersal capabilities each species was able to cross much longer sea barriers than Bonifacio and Messina straits. The two pairs of cryptic species revealed marked climatic preferences and in particular the climatic variable measuring 'precipitation during the driest quarter' was the most influential determinant for the observed boundaries between the pairs of cryptic taxa. However, the climate alone could not fully explain the observed distribution pattern. In fact, on the basis of distribution modelling, both cryptic pairs were expected to coexist at least on one side of these relatively narrow sea straits. Conversely, only one of the two sibling species was documented on each side, although dozens of individuals from the potential contact zones were examined. We show that this phenomenon can be more acute in particular non-sister cryptic than in non-cryptic sister taxa, which agrees with a positive correlation between morphological distance and tendency for secondary contact found in birds by Pigot & Tobias (2013). In fact, while the non-sister cryptic taxa in our analysis largely exclude each others (*P. celina - P. icarus*, *A. agestis - A. cramera*) phylogenetically nearer species showing diverging morphology and/or ecology (*P. icarus - P. eros* and *A. agestis - A. artaxerxes/montensis*) are often sympatric.

The generalist habitus of the species showing mutually exclusive patterns supports the hypothesis that a combination between reproductive interference and climatic preference is more likely to produce the observed distributional pattern than ecological competition, but the mechanisms contributing to the realized distributions may vary depending on the taxonomic group.

While in the previous chapter we highlighted the general importance of this phenomenon at a large scale, here we addressed for the first time the mechanisms that are behind it and provided answers. We show that this is a multifaceted phenomenon that cannot be explained by simple hypotheses and we pinpoint some of the key players, including species interactions, which are usually forgotten in ecology given the difficulty to assess their importance.

In **Chapter IV** of Results we provide a new method to directly compare different types of markers (e.g. genetic, morphologic) even if data available for each marker only partially overlap and discuss the biogeographic implications of the observed concordances and discrepancies. This method is specifically designed to compare the spatial patterns of the markers and to pinpoint the geographical areas where discordance occurs. This method is applicable to a wide array of markers and taxa. Here we took as an example the widespread butterfly *Maniola jurtina*.

This novel method, implemented by introducing new functions in the 'recluster' R package, allowed a detailed assessment of the degree of co-variation between the pattern revealed by a genetic marker and that expressed by a continuous morphological marker at population level. For the first time, this method enabled the alignment and thus the simultaneous analysis of specimens for which data on either one or both traits was available. The possibility to use such heterogeneous data sets makes it possible, for example, to optimize the balance between the costs of DNA sequencing and the necessity to examine large numbers of specimens for morphological analyses. Undoubtedly, the new 'recluster.consensus' function will also facilitate the use of public databanks and facilitate new data mining strategies. The algorithms described in this study can be used with any kind of traits for which dissimilarity matrices and bidimensional configurations can be obtained.

Population studies such as the one performed in this model dataset are becoming a powerful tool for biogeography and conservation biology and are facilitating the recognition of speciation events, dynamics of taxa distribution and micro-evolutionary phenomena driven by various forces, including climate change (Schlick-Steiner et al. 2007, Dincă et al. 2011a, Dapporto et al. 2012, Renvoise et al. 2012, Toews & Brelsford 2012).

In **Chapter V** of Results we investigated the mechanisms determining and maintaining the butterfly community that occurs on the circum-Sicilian islands, at the interface between two continents – Africa and Europe. Analyses of the faunistic diversity and intraspecific genetic variability allowed the recognition of the main processes shaping the island communities of this key biogeographic region. We showed that contemporary island-source dynamics, filtered by the main ecological characteristics of islands (predominantly area and isolation), are the main determinants structuring butterfly communities in the circum-Sicilian islands. Their communities but also their genetic structure were determined by neutral (frequency at source) and non-neutral (dispersal tendency) mechanisms (with a clear predominance of the latter) and by unique combinations of island traits like isolation, area and paleogeography. The main phylogeographic pattern we identified was a strong diversification between Africa and Europe, but also a strong tendency of islands to host identical lineages to those occurring in areas to which they were connected during the glaciations, supporting the importance of past land-bridges in determining the observed diversity. A fundamental finding provided by this study was that the mutually exclusive patterns of distribution, widespread among cryptic taxa on the basis of occurrence data (see Chapter II) are also pervasive at the intraspecific level. Almost all the genetic lineages with more than 1% divergence revealed mutual exclusive patterns of distribution. Another important finding of this study, stemming from the community and genetic data, was that even in a relatively restricted area, island populations have very different histories and are subjected to different recolonization probabilities following local extinctions. A similar evaluation of the contribution of singular factors in characterizing island diversity facilitated the identification of populations that deserve immediate conservation efforts on individual islands. Similar comprehensive examinations of island communities are scarce, especially due to the difficulty of obtaining high quality data at both faunistic and genetic levels for large groups of organisms, therefore this study represents one of the first attempts to explain the main processes governing the island communities in the western Mediterranean.

The results of this PhD thesis contribute to a better understanding of the butterfly diversity in the western Mediterranean and they also provide a framework for future studies that investigate not only butterflies, but other organisms as well.

The work performed during this PhD brings novel data for research (a comprehensive collection of specimens, presence data and DNA sequences), original tools for comparative phylogeography (improved algorithms, R functions), descriptive results (zoogeographic maps, phylogeographic patterns) and conceptual contributions (mutual exclusion, unique properties of cryptic biodiversity, islands as individuals, processes that determine island communities, the phylogeographic value of populations for conservation prioritization).

5. Conclusions

5. Conclusions

The following conclusions can be drawn from the studies included in this thesis:

- We produced a unique sample collection that includes more than 10000 specimens, which together with specimens already existing in the collection at IBE, account for a virtually complete representation of the existing butterfly species in the western Mediterranean.

- We produced a DNA dataset of almost 7000 COI sequences, complemented by GPS coordinates and photographs of the corresponding specimens. Most of these sequences and the collateral data are deposited in BOLD, one of the largest repository for DNA sequences.

- There are no other similar datasets available for an entire taxonomic group with such a wide distribution and with a comparable spatial resolution. This great amount of data provided the basis for the studies of this thesis, but they also represent a valuable resource for future research.

- We provide new evidence for the effectiveness of different measures of beta diversity and demonstrate that no index by itself is able to retrieve comprehensive biogeographical patterns simultaneously for islands and mainland. We propose to compare and combine results obtained by using both unpartitioned and partitioned (nestedness and turnover) to obtain unbiased results.

- We objectively defined cryptic species and produced a list of cryptic groups for all the butterflies in the western Mediterranean. As defined, the cryptic fraction of butterfly diversity represents about 25% of the western Mediterranean fauna and has a homogenous distribution.

- Chequered patterns of distribution within cryptic groups are much more frequent than among congeneric non-cryptic species and they preponderantly occur between non-sister species that show substantial genetic divergence. Therefore, cryptic diversity comprises idiosyncratic qualitative aspects in addition to merely quantitative ones, which highlights the importance of differentiating cryptic species for various research fields.

- The strongest genetic contrasts between cryptic species usually occur between islands located at short distances, which in the western Mediterranean mainly corresponded to the Bonifacio and Messina sea straits.

- Dispersal capacity, ecological preferences and climate alone do not fully explain the chequered distribution patterns of the two pairs of cryptic species, *Aricia agestis* - *A. cramera* and *Polyommatus icarus* - *P. celina*. Most probably, a combination of reproductive interference and climatic preferences produces the observed distributional pattern.

- We provide a new method to directly compare different types of markers (e.g. genetic, morphologic) even if data available for each marker only partially overlap. This method is specifically designed to compare the spatial patterns of the various markers and to pinpoint the geographical areas where discordance occurs. A great attribute is that it is applicable to a wide array of markers and taxa.

- We reveal the main factors that are shaping island communities in a key biogeographical region. Contemporary island-source dynamics, filtered by the main ecological characteristics of islands (predominantly area and isolation), are the main determinants structuring butterfly communities in the circum-Sicilian islands. Their communities and also their genetic structure were determined by neutral (frequency at source) and non-neutral (dispersal tendency) mechanisms and by unique combinations of island traits like isolation, area and paleogeography.

- We identified a strong genetic diversification between the north African and European butterfly populations, and also a tendency of islands to host identical lineages to those occurring in areas to which they were connected during the glaciations, supporting the minor but still significant importance of past land-bridges in determining the observed diversity.

- Another fundamental finding was that the chequered patterns of distribution, widespread among cryptic taxa on the basis of occurrence data, are also pervasive at the intraspecific level. Almost all the genetic lineages with more than 1% divergence revealed chequered distributions in the circum-Sicilian islands.

- We show that island populations have very different histories and are subjected to different recolonization probabilities following local extinctions. These findings have great implications for focusing conservation efforts on specific islands.

- Butterfly species occurring on circum-Sicilian islands can be divided in two main groups (widespread and uncommon) with different frequencies, dispersal capacity, average intraspecific genetic diversity and structure, contribution in creating contrasting island communities, and importance for conservation.

- We introduce the phylogeographic value of a population as a factor that should be taken into account for conservation together with other factors like population trends or endemicity at the species level.

References

Aagaard K., Hindar K., Pullin A.S., James C.H., Hammarstedt O., Balstad T. & Hanssen O. 2002. Phylogenetic relationships in brown argus butterflies (Lepidoptera: Lycaenidae: *Aricia*) from north-western Europe. *Biol. J. of the Linnean Soc.*, **75**: 27–37.

Almeida-Neto M., Frensel D.M.B. & Ulrich W. 2011. Rethinking the relationship between nestedness and beta diversity: a comment on Baselga. *Glob. Ecol. and Biogeo.*, **21**: 772–777.

Als D.T., Vila R., Kandul N.P., Nash D.R., Yen S-H., HsuY-F., Mignault A.A., Boomsma J.J. & Pierce N.E. 2004. The evolution of alternative parasitic life histories in large blue butterflies. *Nature*, **432**: 386–390.

Alvarado-Serrano D.F. & Knowles L.L. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molec. Ecol. Res.*, **14**: 233-248.

Anderson M. C., Hain C. R., Wardlow B. D., Pimstein A., Mecikalski J. R. & Kustas W. P. 2011. Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. *J. of Climate*, **24**: 2025-2044.

Anthes N., Fartmann T., Hermann G. & Kaule G. 2003. Combining larval habitat quality and metapopulation structure – the key for successful management of pre-alpine *Euphydryas aurinia* colonies. *J. of Insect Conserv.*, **7**: 175–183.

Arrhenius O. 1921. Species and area. *J. of Ecol.*, **9**: 95–99.

Asher J., Warren M., Fox R., Harding P., Jeffcoate G. & Jeffcoate S. 2001. The Millenium Atlas of Butterflies in Britain and Ireland. Oxford University Press.

Atmar W. & Patterson B.D. 1993. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, **96**: 373–382.

Avise J.C. 2009. Phylogeography: retrospect and prospect. *J. of Biogeo.*, **36**: 3–15.

Avise J.C., Arnold J., Ball Jr R.M., Bermingham E., Lamb T., Neigel J.E., Reeb C.A. & Saunders N.C. 1987. Intraspecific phylogeography: the mitocondrial DNA bridge between population genetics and systematics. *Ann. Rev. of*

Ecol. and System., **18**: 489-522.

Back W., Olivares J. & Leestmans R. 2005. Une nouvelle sous-espèce d' *Euchloe bazae bazae* (Fabiano, 1993) d' Aragón, dans le N.-E. de l' Espagne: iberae ssp. Nova (Lepidoptera: Pieridae). *Linneana Belgica,* **20**: 67-72.

Baguette M., Clobert J. & Schtickzelle N. 2011. Metapopulation dynamics of the bog fritillary butterfly: experimental changes in habitat quality induced negative density-dependent dispersal. *Ecography,* **34**: 170-176.

Barnagaud J.Y., Kissling W.D., Sandel B., Eiserhardt W.L., Sekercioglu C.H. & Enquist B.J. et al. 2014. Ecological traits influence the phylogenetic structure of bird species co-occurrences worldwide. *Ecol. Lett.*, **17**: 811–820.

Barnosky A.D., Matzke N., Tomiya S., Wogan G.O.U, Swartz B., Quental T.B., Marshall C., McGuire J.L., Lindsey E.L., Maguire K.C., Mersey B. & Ferrer E.A. 2011. Has the Earth's sixth mass extinction already arrived? *Nature,* **471**: 51-57.

Barnosky A.D., Matzke N., Tomiya S., Wogan G.O.U. & Swartz B. et al. 1999. Has the Earth's sixth mass extinction already arrived? *Nature,* **471**: 51-57.

Baselga A. 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecol. Biogeo.*, **19**: 134–143.

Begon M., Townsend C.R. & Harper J.L. 2006. Ecology From Individuals to Ecosystems. Blackwell Publishing.

Belyea L.R. & Lancaster J. 1999. Assembly Rules within a Contingent Ecology. *Oikos,* **86**: 402-416.

Beneš J., Konvička M., Dvořák J., Fric Z., Havelda Z., Pavlíčo A., Vrabec V. & Weidenhoffer Z. (eds.) 2002. Motýli Česk republiky: Rozšíření a ochrana I, II / Butterflies of the Czech Republic: Distribution and conservation I, II. SOM, Praha.

Benson W.W. 1972. Natural Selection for Miillerian Mimicry in *Heliconius erato* in Costa Rica. *Science,* **176**: 936-939.

Bickford D., Lohman D.J., Sodhi N.S., Ng P.K.L., Meier R., Winker K., Ingram K.K. & Das I. 2006. Cryptic species as a window on diversity and conservation. *Trends in Ecol. and Evol.*, **22**: 148-155.

Blanc P.L. 2002. The opening of the Plio-Quaternary Gibraltar strait: assessing the size of a cataclysm. *Geodyn. Acta*, **15**: 303-317.

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

Bookstein F.L. 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medic. image anal.*, **1**: 225-243.

Brakefield P.M. & French V. 1999. Butterfly wings: the evolution of development of colour patterns. *BioEssays*, **21**: 391-401.

Brakefield P.M. 1998. The evolution-development interface and advances with the eyespot patterns of Bicyclus butterflies. *Heredity*, **80**: 265–272.

Brakefield P.M., Beldade P. & Zwaan B.J. 2009. The African Butterfly *Bicyclus anynana*: a Model for Evolutionary Genetics and Evolutionary Developmental Biology. *Cold Spring Harb. Protoc.*: pdb.em0122.

Braudel F. 1985. La Méditerranée. L'Espace et l'Historire. Flammarion, Paris.

Briggs J.C. 1987. Anthropical distribution and evolution in the indo-West Pacific Ocean. *System. Zool.*, **36**: 237-247.

Briggs, J.C. 2003. Marine centres of origin as evolutionary engines. *J. of Biogeo.*, **30**: 1–18.

Brower A.V.Z. & DeSalle R. 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of *wingless* as a source of characters for phylogenetic inference. *Insect Molec. Biol.*, **7**: 73–82.

Brown J.H. & Kodric-Brown A. 1977. Turnover rates in insular biogeography: effect of immigration and extinction. *Ecology,* **58**: 445-449.

Brown K.S. Jr. 1991. Conservation of neotropical environments: insects as indicators, pp. 349–404. In: Collins, N. & Thomas, J. (eds.). Conservation of Insects and their Environments. Academic Press, London.

Brown W.M., George M. JR. & Wilson A.C. 1979. Rapid evolution of animal mitochondrial DNA. *PNAS*, **76**: 1967-1971.

Brunzel S. 2002. Experimental Density-Related Emigration in the Cranberry Fritillary *Boloria aquilonaris*. *J. of Insect Behav.*, **15**: 739–750.

Bucklin A., Steinke D. & Blanco-Bercial L. 2011. DNA Barcoding of Marine Metazoa. *Annu. Rev. Mar. Sci.*, **3**: 471–508.

Caley J.M., Fisher R. & Mengersen K. 2014. Global species richness estimates have not converged. *Trends in Ecol. and Evol.*, **29**: 187-188.

Calvo M., Alda F., Oliverio M., Templado J. & Machordom A. 2015. Surviving the Messinian Salinity Crisis? Divergence patterns in the genus Dendropoma (Gastropoda: Vermetidae) in the Mediterranean Sea. *Molec. Phylogenet. Evol.*, **91**: 17-26.

Cameron S., Rubinoff D. & Will K. 2006. Who Will Actually Use DNA barcoding and What Will It Cost? *Syst. Biol.*, **55**: 844–847.

Carbayo F. & Marques A.C. 2011. The costs of describing the entire animal kingdom. *Trends in Ecol. and Evol.*, **26**: 154-155.

Carlquist S., Baldwin B.G. & Carr, G.D. (eds.). 2003. Tarweeds & silverswords: evolution of the Madiinae (Asteraceae). Missouri Botanical Garden Press, St. Louis.

Carlton J.T., Geller J.B., Reaka-Kudla M.L. & Norse E.A. 1999. Historical extinctions in the sea. *Annu. Rev. Ecol. Syst*, **30**: 515–38.

Carlton J.T., Geller J.B., Reaka-Kudla M.L. & Norse E.A. 1999. Historical Extinctions in the Sea. *Annu. Rev. Ecol. Syst.*, **30**: 515–38.

Carvalheiro L.G., Kunin W.E., Keil P., Aguirre-Gutiérrez J. & Ellis W.N. 2013. Species richness declines and biotic homogenisation have slowed down for NW-European pollinators and plants. *Ecol. Lett.*, **16**: 870–878.

Casevitz-Weulersse J. 1992. Analyse biogéographique de la myrmécofaune Corse et comparaison avec celle des regions voisines. *Comp. Rend. des Séan. de la Soc. de Biogéo.*, **68**: 105-129.

Caujapé-Castells J. 2013. Jesters, red queens, boomerangs and surfers: a molecular outlook on the diversity of the Canarian endemic flora, pp. 284–324. In: Bramwell D. & Caujapé-Castells J. (eds.). The biology of island floras. Cambridge University Press, Cambridge.

CEC 2008. North American Monarch Conservation Plan. Commission for Environmental Cooperation, Montreal.

Cesaroni D., Lucarelli M., Allori P., Russo F. & Sbordoni V. 1994. Patterns of evolution and multidimensional systematics in graylings (Lepidoptera, *Hipparchia*). *Biol. J. of the Linnean Soc,* **52**: 101–111.

Chacon I.A., Janzen D.H., Hallwachs W., Sullivan J.B. & Hajibabaei M. 2013. Cryptic species within cryptic moths: new species of *Dunama* Schaus (Notodontidae, Nystaleinae) in Costa Rica. *Zookeys*, **264**: 11–45.

Chapman A.D. 2006. Numbers of Living Species in Australia and the World. Australian Biological Resources Study, Canbera.

Chapman A.D. 2009. Numbers of Living Species in Australia and the World. Australian Biological Resources Study, Canbera.

Cianchi R., Ungaro A., Marini M. & Bullini L. 2003. Differential patterns of hybridization and introgression between the swallowtails *Papilio machaon* and *P. Hospiton* from Sardinia and Corsica islands (Lepidoptera, Papilionidae). *Molec. Ecol.*, **12**: 1461–1471.

Clare E.L., Lim B.K., Engstrom M.D., Eger J.L. & Hebert P.D.N. 2007. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. *Molec. Ecol. Notes*, **7**: 184–190.

Collins R.A. & Cruickshank. 2013. The seven deadly sins of DNA barcoding. *Molec. Ecol. Res.*, **13**: 969-975.

Cosentino D., Buchwaldt R., Sampalmieri G., Iadanza A., Cipollari P., Schildgen T.F., Hinnov L.A., Ramezani J. & Bowring S.A. 2013. Refining the Mediterranean “Messinian gap” with high-precision U-Pb zircón geochronology, central and northern Italy. *Geology*, **41**: 323-326.

Costello M.J. 2015. Biodiversity: the known, unknown and rates of extinction. *Curr. Biol.*, **25**: 368-371.

Costello M.J., May R.M. & Stork N.E. 2013. Can we name Earth’s species before they go extinct? *Science*, **339**: 413-416.

Coward M. & Dietrich D. 1989. Alpine tectonics – an overview. *Geological Society, London, Special Publications*, **45**: 1-29.

Coyne J.A. & Orr H.A. 2004. Speciation. Sinauer Associates, Sunderland.

Cronk Q.C.B. 1992. Relict floras of Atlantic islands: paterns assessed. *Biol. J. Linnean Soc.*, **45**: 91-103.

Cronk Q.C.B. 1999. Islands as individuals. *Trends in Ecol. and Evol.*, 14: 81.

Currat M., Ruedi M., Petit R.J. & Excoffier L. 2008. The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**: 1908–1920.

Cuttelod A., García N., Malak D.A., Temple H. & Katariya V. 2008. The Mediterranean: A Biodiversity Hotspot Under Threat. In: Vié J.-C., Hilton-Taylor C. & Stuart S.N. (eds). The 2008 Review of The IUCN Red List of Threatened Species. IUCN Gland, Switzerland.

Dapporto L. & Bruschini C. 2012. Invading a refugium: post glacial replacement of the ancestral lineage of a Nymphalid butterfly in the West Mediterranean. *Org. Divers. Evol.*, **12**: 39-49.

Dapporto L. & Cini A. 2007. Faunal patterns in Tuscan archipelago butterflies: the dominant influence is recent geography not paleogeography. European *J. of Entomol.*, **104**: 497–503.

Dapporto L. & Dennis R.L.H. 2008. Island size is not the only consideration. Ranking priorities for the conservation of butterflies on Italian offshore islands. *J. Insect Conserv.*, **12**: 237-249.

Dapporto L. & Dennis R.L.H. 2009. Conservation biogeography of large Mediterranean islands. Butterfly impoverishment, conservation priorities and inferences for an ecological “island paradigm”. *Ecography*, **32**: 169-179.

Dapporto L. & Dennis RLH. 2008. Species’ richness, rarity and endemicity of Italian offshore islands: complementary signals from island–focused and species–focused analyses. *J. of Biogeo.*, **35**: 664–674.

Dapporto L. 2009. Speciation in Mediterranean refugia and post-glacial expansion of *Zerynthia polyxena* (Lepidoptera, Papilionidae). *J. Of Zool. System. And Evol. Res.*, **48**: 229-237.

Dapporto L. 2010. Satyrinae butterflies from Sardinia and Corsica show a kaleidoscopic intraspecific biogeography (Lepidoptera, Nymphlidae). *Biol. J. of the Linnean Soc.*, **100**: 195–212.

Dapporto L., Bruschini C., Baracchi D., Cini A., Gayubo S.F., Gonzàlez J.A. & Dennis R.L.H. 2009. Phylogeography and counter-intuitive inferences in island biogeography: evidence from morphometric markers in the mobile butterfly *Maniola jurtina* (Linnaeus) (Lepidoptera, Nymphalidae). *Biol. J. of the Linnean Soc.*, **98**: 677–692.

Dapporto L., Bruschini C., Dincă V., Vila R. & Dennis R.L.H. 2012. Identifying zones of phenetic compression in West Mediterranean butterflies (Satyrinae): refugia, invasion and hybridization. *Div. and Distrib.*, **18**: 1066–1076.

Dapporto L., Dennis R.L.H., Ciolli G., Fox R. & Shreeve T.G. 2015. A new procedure for extrapolating turnover regionalization at mid-small spatial scales, tested on British butterflies. *Methods in Ecol. And Evol.*, doi: 10.1111/2041-210X.12415.

Dapporto L., Habel J.C., Dennis R.L.H. & Schmitt T. 2011b. The biogeography of the western Mediterranean: elucidating contradictory distribution patterns of differentiation in *Maniola jurtina* (Lepidoptera, Nymphalidae). *Biol J Linn Soc.,* **103**: 571–577.

Dapporto L., Ramazzotti M., Fattorini S., Talavera G., Vila R. & Dennis R.L.H. 2013. recluster: an unbiased clustering procedure for beta-diversity turnover. *Ecography*, **36**: 1070-1075.

Dapporto L., Schmitt T., Vila R., Scalercio S., Biermann H., Dincă V., Gayubo S.F., González J.A., Lo Cascio P. & Dennis R.L.H. 2011a. Phylogenetic island disequilibrium: evidence for ongoing long-term population dynamics in two Mediterranean butterflies. *J. of Biogeo.*, **38**: 854–867.

Darwin C. 1859. On the origin of species by means of natural selection. J. Murray, London.

Davey J. W., Cezard T., Fuentes-Utrilla P., Eland C., Gharbi K. & Blaxter M.L. 2013. Special features of RAD Sequencing data: implications for genotyping. *Molec. Ecol.*, **22**: 3151-3164.

Davis M.B. 1976. Pleistocene biogeography of temperate deciduous forests. *Geoscience Canada*, **13**: 13-26.

De Jong H. 1998. In search of historical biogeographic patterns in the western Mediterranean terrestrial fauna. *Biol. J. of the Linnean Soc.*, **65**: 99- 164.

de Jong, R. 1972. Systematics and geographic history of hte genus *Pyrgus* in the Palaearctic region (Lepidoptera, Hesperiidae). *Tijdsch. voor Entomol.*, **155**: 1-121.

de Lesse H. 1969. Les nombres de chromosomes dans le groupe de *Lysandra coridon* (Lep., Lycaenidae). *Ann. Soc. Ent. Fr.,* **5**: 469–522.

DeLong M. & Martinson B.(eds.). 2012. Color and design. Berg, London.

Dennis R. L. 2012. A resource-based habitat view for conservation: butterflies in the British landscape. John Wiley & Sons.

Dennis R. L., Shreeve T. G., Olivier A., & Coutsis J. G. 2000. Contemporary geography dominates butterfly diversity gradients within the Aegean archipelago (Lepidoptera: Papilionoidea, Hesperioidea). *J. of Biogeo*., **27**: 1365-1383.

Dennis R.L.H. & Shreeve T.G. 1997. Diversity of butterfly species on British islands: ecological influences underlying the roles of area, isolation and faunal source. *Biol. J. Linnean Soc.,* **60**: 257–275.

Dennis R.L.H. 2009. Ecology of butterflies in Europe–where are we now and where to go? In: Settele, J., Shreeve, T., Konvička, M., & Dyck, H. V. (eds). Ecology of butterflies in Europe. Cambridge University Press.

Dennis R.L.H., Dapporto L., Shreeve T.G., John E., Coutsis J.G., Kudrna O., Saarinen K., Ryrholm N. & Williams W.R. 2008. Butterflies of European islands: the implications of the geography and ecology of rarity and endemicity for conservation. *J. Insect Conserv.,* **12**: 205-236.

Dennis R.L.H., Shreeve T.G. & Williams W.R. 1995. Taxonomic differentiation in species richness gradients among European butterflies (Papilionoidea, Hesperioidea): contribution of macroevolutionary dynamics. *Ecography,* **18**: 27-40.

Dennis R.L.H., Williams W.R & Shreeve T.G. 1991. A multivariate approach to the determination of faunal structures among European butterfly species (Lepidoptera: Rhopalocera). *Biol. J. of the Linnean Soc.,* **101**: 1-49.

Dennis R.L.H., Williams W.R & Shreeve T.G. 1998. Faunal structures among European butterflies: evolutionary implications of bias for geography, endemisms and taxonomic affiliation. *Ecography*, **21**: 181-203.

DeSalle R. 2006. Species Discovery versus Species Identification in DNA Barcoding Efforts: Response to Rubinoff. *Conserv. Biol*., **20**: 1545–1547.

Descimon H. & Mallet J. 2009. Bad species, pp. 219–249. In: Settele J., Shreeve J., Konvička M. & Van Dyck H. (eds.). Ecology of Butterflies in Europe. Cambridge University Press.

Devictor V., van Swaay C., Brereton T., Brotons L., Chamberlain D., Heliölä J., Herrando S., Julliard R., Kuussaari M., Lindström Å., Reif J., Roy D.B., Schweiger O., Settele J., Stefanescu C., Van Strien A., Van Turnhout C., Vermouzek Z., WallisDeVries M., Wynhoff I. & Jiguet F. 2012. Differences in the climatic debts of birds and butterflies at a continental scale. *Nat. Clim. Change,* **2**: 121–124.

Dewey J.F., Helman M.L., Turco E., Hutton D.H.W. & Knott S.D. 1989. Kinematics of the western Mediterranean, pp. 265-283. In: Coward M.P., Dietrich D. & Park R.G. (eds.). *Alpine Tectonics*, Geological Society Special Publication, **45**.

Diamond J.M. 1975. Assembly of species communities. In: Cody M.L. & Diamond J.M. (eds.). Ecology and Evolution of Communities. Belknap Press of Harvard University Press, Cambridge.

Dincă V., Dapporto & Vila R. 2011a. A combined genetic morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and *Polyommatus celina* Common Blue butterflies. *Molec. Ecol.*, **20**: 3921–3935.

Dincă V., Lukhtanov V. A., Talavera G. & Vila, R. 2011b. Unexpected layers of cryptic diversity in wood white Leptidea butterflies. *Nat. Commun.,* **2***:* 324.

Dincă V., Montagud S., Talavera G., Hernández-Roldán J., Munguira M.L., García-Barros E., Hebert P.D.N. & Vila R. 2015. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. *Scientific Reports*, **5**: 12395. doi:10.1038/srep12395.

Dincă V., Runquist M., Nilsson M. & Vila R. 2013. Dispersal, fragmentation and isolation shape the phylogeography of the European lineages of *Polyommatus* (*Agrodiaetus*) *ripartii* (Lepidoptera: Lycaenidae). *Biol. J. of the Linnean Soc.*, **109**: 817–829.

Dincă V., Wiklund C., Lukhtanov V.A., Kodandaramaiah U., Noren K., Dapporto L., Wahlberg N., Vila R. & Frieberg M. 2013. Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. *J. Evol. Biol.*, **26**: 2095–2106.

Dincă V., Zakharov E.V., Hebert P.D.N. & Vila R. 2011c. Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. *Proc. R. Soc. B,* **278**: 347-355.

Dirzo R. & Raven P.H. 2003. Global state of biodiversity and loss. *Annu. Rev. Environ. Resour*, **28**: 137-167.

Dobrovolski R., Melo A.S., Cassemiro F.A.S. & Diniz-Filho J.A.F. 2012. Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Global Ecol. Biogeogr.*, **21**: 191–197.

Duggen S., Hoernle K., van den Bogaard P., Rüpke L. & Morgan J.P. 2003. Deep roots of the Messinian salinity crisis. *Nature*, **422**: 602-606.

Ebach M.C. & Holdrege C. 2005. DNA barcoding is no substitute for taxonomy. *Nature*, **434**: 697.

Elias M., Hill R.I., Willmott K.R., Dasmahapatra K.K., Brower A.V.Z., Mallet J. & Jiggins C.D. 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies. *Proc. R. Soc. B,* **274**: 2881–2889.

Eliasson C.U., Ryrholm N., Holmer M., Jilg K. & Gärdenfors U. 2005. Nationalnyckeln till Sveriges flora och fauna. Fjärilar: Dagfjärilar. Hesperiidae – Nymphalidae. ArtDatabanken, SLU, Uppsala.

Elmes G.W. & Thomas J.A. 1992. Complexity of species conservation in managed habitats: interaction between *Maculinea* butterflies and their ant hosts. *Biodiv. Conserv.*, **1**: 155–169.

Emerson C. 2002. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molec. Ecol.*, **11**: 951–966.

Erwin T.L. 1982. Tropical forests: their richness in Coleoptera and other Arthropod species. *The Coleop. Bull.*, **36**: 74–75.

Espeland M., Hall J., DeVries P., Lees D., Cornwall M., Yu-Feng Hsu, Li-Wei Wu, Campbell D., Talavera G., Vila R., Salzman S., Ruehr S., Lohman D. & Pierce N. 2015. Ancient Neotropical origin and recent recolonisation: Phylogeny, biogeography and diversification of the Riodinidae (Lepidoptera: Papilionoidea). *Molec. Phylog. and Evol.,* **93**: 296–306.

Esselstyn J.A. & Brown R.M. 2009. The role of repeated sea-level fluctuations in the generation of shrew (Soricidae: Crocidura) diversity in the Philippine Archipelago. *Molec. Phylog. and Evol.*, **53**: 171–181.

Excoffier L., Hofer T. & Foll M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**: 285–298.

Faccenna C., Funiciello F., Giardini D. & Lucente P. 2001. Episodic back-arc extension during restricted mantle convection in the Central Mediterranean. *Earth and Planetary Science Letters,* **187**: 105-116.

Fagua G., Amarillo A. & Andrade-C. M.G. 1999. Las mariposas (Lepidoptera: Rhopalocera) como indicadores del grado de intervención en la cuenca del Río Pato (Caquetá, Colombia), pp. 285–315. In: Amat, G., Andrade-C, M.G., Fernández, F. (eds). Insectos de Colombia, vol. 2. Academia Colombiana de Ciencias Exactas Físicas y Naturales, Bogotá.

Fattorini S. 2002. Biogeography of the tenebrionid beetles (Coleoptera, Tenebrionidae) on the Aegean Islands (Greece). *J. Of Biogeo*., **29**: 49–67.

Fattorini S. 2009. Both Recent and Pleistocene geography determine animal distributional patterns in the Tuscan Archipelago. *J. of Zool.*, **277**: 291–301.

Fattorini S. 2011. Biogeography of tenebrionid beetles (Coleoptera: Tenebrionidae) in the circum-Sicilian islands (Italy, Sicily): Multiple biogeographical patterns require multiple explanations. *Eur. J. Entomol.*, **108**: 659–672.

Fernández-Palacios J.M. & Whittaker R.J. 2010. El ciclo de la isla, pp.18-27. In: Martín Esquivel J.L. (ed.). Atlas de la Biodiversidad de Canarias. Turquesa Ediciones, Santa Cruz de Tenerife.

Fernández-Palacios J.M., Kueffer C. & Drake D.R. 2015a. A new golden era in island biogeography. *Front. In Biogeo*., **7**: 14-20.

Fernández-Palacios J. M., Rijsdijk K. F., Norder S. J., Otto R., Nascimento L., Fernández-Lugo S., Tjørve E. & Whittaker R. J. 2015b. Towards a glacial-sensitive model of island biogeography. *Glob. Ecol. and Biogeo*, doi: 10.1111/geb.12320.

Fisher R., O'Leary R.A., Low-Choy S., Mengersen K., Knowlton N., Brainard R.E. & Caley J.M. 2015. Species richness on coral reefs and the pursuit of convergent global estimates. *Curr. Biol.*, **25**: 500-505.

Friberg M. & Wiklund C. 2009. Host plant preference and performance of the sibling species of butterflies Leptidea sinapis and Leptidea reali: a test of the trade-off hypothesis for food specialization. *Oecologia,* **159**: 127–137.

Fric Z., Wahlberg N., Pech P. & Zrzavý, J. 2007. Phylogeny and classification of the *Phengaris–Maculinea* clade (Lepidoptera: Lycaenidae): total evidence and phylogenetic species concepts. *System. Entomol.,* **32**: 558–567.

García-Barros E., Munguira M.L., Stefanescu C. & Vives Moreno A. 2013. Lepidoptera Papilionoidea. In: Ramos M.A. et al. (Eds.). Fauna Iberica, vol. 37. Museo Nacional de Ciencias Naturales. CSIC. Madrid.

Gaullier V., Loncke L., Vendeville B., Déverchère J. & Droz L. 2010. Interactions between salt tectonics and deep-seated tectonics. Part I: Examples from the western Mediterranean (Abstr.). In: International Conference SEPM –The Geological Society, "Salt tectonics, sediments and prospectivity", 20-22 January 2010, London.

Geiger W. (ed.) 1987. Les papillons de jour et leurs biotopes. Bâle, Ligue Suisse pour la Protection de la Nature.

Godfray C.H.J. 2002. Challenges for taxonomy. *Nature*, **417**: 17-19.

Godfray H.C.J. & Knapp S. 2004. Introduction to Theme Issue, "Taxonomy for the 21st Century". *Phil. Trans. R. Soc. Lond. B,* ***359**: 559-570.*

Goffart P., Baguette M., Dufrêne M., Mousson L., Nève G., Sawchik J., Weiserbs A. & Lebrun P. 2001. Gestion des milieux semi-naturels et restauration de populations menacées de papillons de jour. Travaux 25, Division de la nature et des forêts, Région Wallone.

Grill A., Gkiokia E. & Alvarez N. 2006. Evolutionary history and patterns of differentiation among European *Maniola* butterflies (Lepidoptera: Satyrinae). *Eur. J. Entomol.*, **103**: 613.

Grill A., Raijmann L.E.L., Van Ginkel W., Gkioka E., Menken S.B. 2007. Genetic differentiation and naturalhybridization between the Sardinian endemic *Maniola nurag* and the European *Maniola jurtina*. *J. of Evol. Biol.*, **20**: 1255–1270.

Groombridge B. & Jenkins M.D. 2002. World Atlas of Biodiversity. Prepared by the UNEP World Conservation Monitoring Centre, University of California Press, Berkeley.

Gutiérrez D., Vila R. & Wilson R.J. Asymmetric constraints on limits to species distributions influence consumer-host diversity over an environmental gradient, in prep.

Habel J.C., Brückmann S.V., Krauss J., Schwarzer J., Weig A., Husemann M. &

Steffan-Dewenter I. 2015. Fragmentation genetics of the grassland butterfly*Polyommatus coridon*: Stable genetic diversity or extinction debt? *Conserv. Genetics*, **16**: 549-558.

Habel J.C., Dieker P. & Schmitt T. 2009. Biogeographical connections between the Maghreb and the Mediterraenan peninsulas of southern Europe. *Biol. J. of the Linnean Soc.*, **98**: 693–703.

Habel J.C., Husemann M., Rödder D. & Schmitt T. 2011. Biogeographical dynamics of the Spanish Marbled White *Melanargia ines* (Lepidoptera: Satyridae) in the Western Mediterranean: does the Atlanto-Mediterranean refuge exist? *Biol. J. of the Linnean Soc.,* **104**: 828-837.

Habel J.C., Rödder D., Stefano S., Meyer M. & Schmitt T. 2010. Strong genetic cohesiveness between Italy and North Africa in four butterfly species. *Biol. J. of the Linnean Soc.*, **99**: 818–830.

Habel J.C., Schmitt T. & Müller P. 2005. The fourth paradigm pattern of postglacial range expansion of European terrestrial species: the phylogeography of the Marbled White butterfly (Satyrinae, Lepidoptera). *J. of Biogeo*, **32**: 1489–1497.

Hajibabaei M., Janzen D.H., Burns J.M., Hallwachs W. & Hebert P.D.N. 2006. DNA barcodes distinguish species of tropical Lepidoptera. *PNAS*, **103**: 968–971.

Hajibabaei M., Singer G.A.C., Hebert P.D.N. & Hickey D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, **23**: 167–172.

Hamilton A.J., Basset Y., Benke K.K., Grimbacher P.S., Miller S.E., Novotny V., Samuelson A.G., Stork N.E., Weiblen G.D. & Yen J.D.L. 2010. Quantifying Uncertainty in Estimation of Tropical Arthropod Species Richness. *The Am. Nat.*, **176**: 90-95.

Hamilton A.J., Novotny V., Waters E.K., Basset Y., Benke K.K., Grimbacher P.S., Miller S.E., Samuelson A.G., Weiblen G.D., Yen J.D.L. & Stork N.E. 2013. Estimating global arthropod species richness: refining probabilistic models using probability bounds analysis. *Oecologia*, **171**: 357–365.

Hammond P. 1992. Species Inventory, pp. 17-39. In: Brian Groombridge (ed.). Global biodiversity. Chapman & Hall, London.

Handy M.R., Schmid S.M., Bousquet R., Kissling E. & Bernoulli D. 2010. Reconciling plate-tectonic reconstructions of Alpine Tethys with the geological geophysical record of spreading and subduction in the Alps. *Earth Sci. Rev.*, **102**: 121–158.

Hanksi I. 1999. Metapopulation ecology. Oxford University Press, New York.

Hanksi I., Kuussaari M. & Nieminen M. 1994. Metapopulation structure and migration in the butterfly Melitaea cinxia. *Ecology*, **75**: 747–762.

Hanski I. 2010. The Theories of Island Biogeography and Metapopulation Dynamics, pp. 186-213. In: Losos J.B. & Ricklefs R.E. (eds.). The Theory of Island Biogeography Revisited. Princeton University Press. Princeton and Oxford.

Hausdorf B. & Henning C. 2005. The influences of recent geography, paleogeography and climate on the composition of the fauna of the central Aegean Islands. *Biol. J. Linnean. Soc.*, **84**: 785–795.

Hawkins B.A. 2004. Summer vegetation, deglaciation and the anomalous bird diversity gradient in eastern North America. *Global Ecol. Biogeogr.*, **13**: 321–325.

Hawksworth D.L. & Kalin-Arroyo M.T. 1995. Magnitude and Distribution of Biodiversity, pp. 107–191. In: Heywood V. (ed.) Global biodiversity assessment. Cambridge University Press, Cambridge.

Hays J.D., Imbrie J. & Shackleton N.J. 1976. Variations in the Earth's Orbit: Pacemaker of the Ice Ages. *Science*, **194**: 1121-1132.

Heaney L.R. 2000. Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. *Glob. Ecol. and Biogeo.*, **9**: 59-74.

Heaney L.R. 2007. Is a new paradigm emerging for oceanic island biogeography?. *J. of Biogeo.*, **34**: 753-757.

Hebert P.D.N. & Gregory R. 2005. The Promise of DNA Barcoding for Taxonomy. *Syst. Biol.*, **54**: 852–859.

Hebert P.D.N., Cywinska A., Ball S.L. & deWaard J.R. 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. B*, **270**: 313–321.

Hebert P.D.N., Penton E.H., Burns J.M., Janzen D.H. & Hallwachs W. 2004b. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *PNAS*, **101**: 14812-14817.

Hebert P.D.N., Ratnasingham S. & deWaard J.R. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 diverges among closely related species. *Proc. R. Soc. B* (Suppl.), **270**: 96–99.

Hebert P.D.N., Stoeckle M.Y., Zemlak T.S. & Francis C.M. 2004a. Identification of Birds through DNA Barcodes. *PLoS Biology,* **2**: 1657–1663.

Henriksen H.J. & Kreutzer I., 1982. The butterflies of Scandinavia in nature. Skandinavisk Bogforlag, Odense.

Hernández-Roldán J.L., Bofill R., Dapporto L., Munguira M.L. & Vila R. 2014. Morphological and chemical analysis of male scent organs in the butterfly genus *Pyrgus* (Lepidoptera: Hesperiidae). *Org. Divers. Evol.*, **14**: 269–278.

Hewitt G.M. 1993. Postglacial distribution and species substructure: lessons from pollen, insects and hybrid zones, pp.97-123. In: Lees D.R. & Edwards D. (eds.). Evolutionary patterns and processes. Linnean Society Symposium Series 14. Academic, London.

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

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

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

Hewitt G.M. 2001. Speciation, hybrid zones and phylogeography - seeing genes in space and time. *Mol Ecol.,* **10**: 537–549.

Hewitt G.M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Phil. Trans. R. Soc. Lond. B,* **359**: 183–195.

Hewitt G.M. 2011. Quaternary phylogeography: the roots of hybrid zones. *Genetica*, **139**: 617-38.

Hickerson M.J., Carstens B.C., Cavender-Bares J., Crandall K.A., Graham C.H., Johnson J.B., Rissler L., Victoriano P.F. & Yoder A.D. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molec. Phylo. and Evol.,* **54**: 291–301.

Higgins L.G. 1975. The Classification of European Butterflies. Collins, London.

Hillis D.M., Moritz C., Mable B.K. 1996 . Molecular Systematics. Sunderland, MA: Sinauer Associates.

Hodkinson I.D. & Casson D. 1991. A lesser predilection for bugs: Hemiptera (Insecta) diversity in tropical rain forests. *Biol. J. of the Linnean Soc.*, **43**: 101-109.

Holt B.G., Lessard J.-P., Borregaard M.K., Fritz S.A., Araújo M.B., Dimitrov D., Fabre P.-H., Graham C.H., Graves G.R., Jønsson K.A., Nogués Bravo D., Wang Z., Whittaker R.J., Fjeldsa, J. & Rahbek C. 2013. An update of Wallace's zoogeographic regions of the world. *Science*, **339**: 74–78.

Hortal J., Triantis K.A., Meiri S. Thébault & Sfenthourakis S. 2009. Island Species Richness Increases with Habitat Diversity. *The Am. Nat.*, **174**: 205-217.

Hsü K.J. 1971. Origin of the Alps and western Mediterranean. *Nature*, **233**: 44-48.

Hsü K.J., Ryan W.B.F. & Cita M.B. 1973. Late Miocene Desiccation of the Mediterranean. *Nature*, **242**: 240-244.

Hubbell S.P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press.

Hubbell S.P. 2009. Neutral Theory and the Theory of Island Biogeography, pp. 264-292. In: Losos J.B. & Ricklefs R.E. (eds.). The Theory of Island Biogeography Revisited. Princeton University Press, Princeton and Oxford.

Hula V., Konvicka M., Pavlicko A. & Fric Z. 2004. Marsh Fritillary (*Euphydryas aurinia*) in the Czech Republic: monitoring, metapopulation structure, and conservation of an endangered butterfly. *Entomol. Fennica,* **15**: 231–241.

Husemann M., Schmitt T., Zachos F.E., Ulrich W. & Habel J.C. 2014. Palaearctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaearctic biota. *J. of Biogeo.*, **41**: 81–94.

Janzen D. J. 2004. Now is the time. *Phil. Trans. R. Soc. Lond. B,* **359**: 731–732.

Janzen D.H., Hajibabaei M., Burns J.M., Hallwachs W., Remigio E. & Hebert, P.D.N. 2005. Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philos Trans R Soc Lond B,* **360**: 1835–1845.

Janzen D.H., Hallwachs W., Burns J.M., Hajibabaei M., Bertrand C. & Hebert P.D.N. 2011. Reading the Complex Skipper Butterfly Fauna of One Tropical Place. *PLoS ONE,* **6**: e19874.

Janzen DH, Hallwachs W, Harvey DJ, et al. 2012. What happens to the traditional taxonomy when a well-known tropical saturniid moth fauna is DNA barcoded? *Invert. System.*, **26**: 478–505.

Jolivet L. & C. Faccenna. 2000. Mediterranean extension and the Africa-Eurasia collision. *Tectonics*, **9**: 1095-1106.

Jurado-Rivera J., Vogler A.P., Reid C.A.M., Petitpierre E., Gómez-Zurita J. 2009. DNA barcoding insect-host plant associations. *Proc. R. Soc. B*, **276**: 639–648.

Kalmar A. & Currie D.J. 2006. A global model of island biogeography. *Glob. Ecol. And Biogeo.*, **15**: 72-81.

Kandul N.P., Lukhtanov V.A. & Pierce N.E. 2007. Karyotypic diversity and speciation in *Agrodiaetus* butterflies. *Evolution*, **61**: 546–559.

Karsholt O. & Nieukerken E.J. 2013. Lepidoptera, Papilionoidea. Fauna Europaea ver. 1.0, www.faunaeur.org.

Karsholt O. & Razowski, J. 1996. The Lepidoptera of Europe. A distributional checklist. Apollo Books, Stenstrup.

Keil P., Schweiger O., Kühn I., Kunin W.E., Kuussaari M., Settele J., Henle K., Brotons L., Pe'er G., Lengyel S., Moustakas A., Steinicke H., & Storch D. 2012. Patterns of beta diversity in Europe: The role of climate, land cover and distance across scales. *J. of Biogeo.*, **39**: 1473–1486.

Keller L.F. & Waller D.M. 2002. Inbreeding effects in wild populations. *Trends in Ecol. And Evol.*, **17**: 230-241.

Kellert S.R. 1993. Values and perceptions of invertebrates. *Conserv. Biol.,* **7**: 845-855.

Kerr K.C.R., Stoeckle M.Y., Dove C.J., Weigt L.A., Francis C.M. & Hebert P.D.N. 2007. Comprehensive DNA barcode coverage of North American birds. *Molec. Ecol. Notes,* **7**: 535–543.

King R., Proudfoot L. & Smith B. (eds.). 1997. The Mediterranean : environment and society. Hodder Arnold, London.

Kodandaramaiah U. & Wahlberg N. 2009. Phylogeny and biogeography of *Coenonympha* butterflies (Nymphalidae: Satyrinae) - patterns of colonization in the Holarctic. *Syst. Entomol.*, **34**: 315-323.

Kodandaramaiah U., Lindenfors P. & Tullberg B.S. 2013. Deflective and intimidating eyespots: a comparative study of eyespot size and position in *Junonia* butterflies. *Ecol Evol.*, **3**: 4518–4524.

Kolbert E. 2014. The sixth extinction. Henry Holt and Company, New York.

Koleff P., Gaston K.J. & Lennon J.J. 2003. Measuring beta diversity for presence-absence data. *J. of Anim. Ecol.*, **72**: 367–382.

Kostro-Ambroziak A. & Rutkowski R. 2011. Divergent patterns in the mitochondrial and nuclear diversity of the specialized butterfly *Plebejus argus* (Lepidoptera: Lycaenidae). *Eur. J. Entomol*, **108**: 537-545.

Kreft H. & Jetz W. 2010. A framework for delineating biogeographic regions based on species distributions. *J. of Biogeo*., **37**: 2029–2053.

Kreft H., Jetz W., Mutke J., Kier G. & Barthlott W. 2008. Global diversity of island floras from a macroecological perspective. *Ecol. Letters*, **11**: 116–127.

Kremen C., Colwell R.K., Erwin T.L., Murphy D.D., Noss R.F. & Sanjayan M.A. 1993. Terrestrial arthropod assemblage: Their use in conservation planning. *Conserv. Biol.*, **7**: 796–808.

Kreuzinger A. J., Fiedler K., Letsch H. & Grill A. 2015. Tracing the radiation of *Maniola* (Nymphalidae) butterflies: new insights from phylogeography hint at one single incompletely differentiated species complex. *Ecol. and Evol.*, **5**: 46-58.

Krijgsman W., Hilgen F.J., Raffi I. & Sierro F.J. 1999. Chronology, causes and progression of the Messinian salinity crisis. *Nature*, **400**: 652-654.

Kudrna O. 2002. The Distribution Atlas of European Butterflies. Oedippus 20.

Kudrna O., Harpke A., Lux C., Pennerstorfer J., Schweiger O., Settele J. & Wiemers M. 2011. Distribution Atlas of Butterflies in Europe. Gesellschaft für Schmetterlingschutz, Halle.

Kudrna O., Pennerstorfer J. & Lux K. 2015. Distribution Atlas of European Butterflies and Skippers. Peks, Schwanfeld.

Kueffer C., Drake D.R. & Fernández-Palacios J.M. 2014. Island biology: looking towards the future. *Biol. Lett.*, **10**: 20140719.

Kumar S. 2005. Molecular clocks: four decades of evolution. *Nature Reviews Genetics*, **6**: 654-662.

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. *Molec. Phylog. and Evol.*, **69**: 39–45.

Ladle R.J. & Wittaker R.J. 2011. Conservation Biogeography. Wiley-Blackwell

Lafranchis T. 2000. Les Papillons de jour de France, Belgique et Luxembourg et leurs chenilles. Biotope, Mèze.

Lambeck K. & Purcell A. 2005. Sea-level change in the Mediterranean Sea since the LGM: Model predictions for tectonically stable areas. *Quaternary Science Reviews,* **24**: 1969–1988.

Larsen T. B.. 1986. Tropical butterflies of the Mediterranean. *Nota Lepid.*, **9**: 63-77.

Lehe R., Hallatschek O. & Peliti L. 2012. The Rate of Beneficial Mutations Surfing on the Wave of a Range Expansion. *PLoS Comput. Biol.*, **8**: e1002447.

Leneveu J., Chichvarkhin A. & Wahlberg N. 2009. Varying rates of diversification in the genus *Melitaea* (Lepidoptera: Nymphalidae) during the past 20 million years. *Biol. J. Linnean Soc.,* **97**: 346–361.

Levins R. & McArthur R. 1969. An hypothesis to explain the incidence of monophagy. *Ecology*, **50**: 910-911.

Li W.H. 1997. Molecular Evolution. Sunderland, MA: Sinauer Associates.

Liénard M.A., Wang H-L., Lassance J-M. & Löfstedt C. 2014. Sex pheromone biosynthetic pathways are conserved between moths and the butterfly *Bicyclus anynana. Nat. communic,* **5:** doi:10.1038/ncomms4957.

Lofi J., Déverchère J., Gaullier V., Gillet H., Guennoc P., Gorini C., Loncke L., Maillard A., Sage F. & Thinon I. 2011. Seismic atlas of the Messinian salinity crisis markers in the offshore Mediterranean domain. CCGM & Mém. Soc. géol. de France.

Lofi J., Gorini C., Berne S., Clauzon G., Dos Reist T., Ryan W.B.F. & Steckler M.S. 2005. Erosional processes and paleo-environmental changes in the western gulf of Lions (SW France) during the Messinian salinity crisis. *Mar. Geol.*, **217**: 1-30.

Loget N. & Van Den Driessche J. 2006. On the origin of the strait of Gibraltar. *Sediment. Geol.*, **188-189**: 341-356.

Lomolino M.V. 1996. Investigating causality of nestedness of insular communities: selective immigrations or extinctions*? J. of Biogeo*., **23**: 699–703.

Lomolino M.V. 2000. A species-based theory of insular zoogeography. *Glob. Ecol. and Biogeo.*, **9**: 39-58.

Lonergan L. & White N. 1997. Origin of the Betic-Rif mountain belt. *Tectonics*, **16**: 504-522.

Lorković Z. 1941. Die Chromosomenzahlen in der Spermatogenese der Tagfalter. *Chromosoma*, **2**: 155–191.

Lorković Z. 1990. The butterfly chromosome and their application in systematics and phylogeny, pp. 332–396. In: Kudrna O. (ed.). Butterflies of Europe, Vol. 2. Aula-Verlag, Wiesbaden.

Lorković Z. 1993. *Leptidea reali* Reissinger, 1989 (= lorkovicii Real 1988), a new European species (Lepid., Pieridae). *Nat. Croatica*, **2**: 1–26.

Losos J.B. & Ricklefs R.E. (eds.). 2010. The Theory of Island Biogeography Revisited. Princeton University Press. Princeton and Oxford.

Lukhtanov A. V., Dincă V., Talavera G. & Vila R. 2011. Unprecedented within-species chromosome number cline in the Wood White butterfly *Leptidea sinapis* and its significance for karyotype evolution and speciation. *BMC Evol. Biol.,* **11**: 109

Lukhtanov V.A., Kandul N.P., Plotkin J., Dantchenko A.V., Haig D. & Pierce N.E. 2005. Reinforcement of pre-zygotic isolation and karyotype evolution in Agrodiaetus butterflies. *Nature*, **436**: 385–389.

Lukhtanov V.A., Sourakov A., Zakharov E.V. & Hebert P.D.N. 2009. DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. *Molec. Ecol. Res*., **9**: 1302-1310.

MacArthur R.H. & Wilson E.O. 1963 An equilibrium theory of insular zoogeography. *Evolution,* **17**: 373–387.

MacArthur R.H. & Wilson E.O. 1967 The theory of island biogeography. Princeton University Press.

Malinverno A. & Ryan W.B.F. 1986. Extension on the Tyrrhenian Sea and shortening in the Apennines as result of arc migration driven by sinking of the lithosphere. *Tectonics*, **5**: 227-245.

Mallet J., & K Willmott. 2003. Taxonomy: Renaissance or Tower of Babel? *Trends in Ecol. and Evol.,* **18**: 57–59.

Mallet J., Wynne I.R. & Thomas C.D. 2011. Hybridisation and climate change: brown argus butterflies in Britain (Polyommatus subgenus Aricia). *Insect Conserv. and Div.*, **4**: 192–199.

Manel S., Schwartz M.K., Luikart G., & Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecol. and Evol.*, **18**: 189-197.

Manzi V., Gennari R., Hilgen F., Krijgsman W., Lugli S., Roveri M. & Sierro F.J. 2013. Age refinement of the Messinian salinity crisis onset in the Mediterranean. *Terra Nova*, **25**: 315–322.

Maravalhas E. 2003. As Borboletas de Portugal. The Butterflies of Portugal. Apollo Books, Stenstrup.

Marske K.A., Rahbek C. & Nogués-Bravo D. 2013. Phylogeography: spanning the ecology-evolution continuum *Ecography*, **36**: 1169-1181.

Martin J., Gilles A. & Descimon H. 2003. Species Concepts and Sibling Species: The Case of *Leptidea sinapis* and *Leptidea reali*, pp. 459–476. In: Boggs, C.L., Watt, W.B., Ehrlich, P.R. (eds.). Butterflies. Ecology and Evolution Taking Flight. The University of Chicago Press, Chicago.

Martin S.H., Dasmahapatra K.K., Nadeau N.J., Salazar C., Walters J.R., Simpson F., Blaxter M., Manica A., Mallet J. & Jiggins C.D. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.*, **23**: 1817-1828.

Masini F., Petruso D., Bonfiglio L. & Mangano G. 2008. Origination and extinction patterns of Mammals in three central western Mediterranean islands in the Late Miocene to Quaternary. *Quatern. Intl.*, **182**: 63-79.

Mavárez J., Salazar C.A., Bermingham E., Salcedo C., Jiggins C.D. & Linares M. 2006. Speciation through hybridization in Hernánd butterflies. *Nature*: **441**: 868–871.

May R.M. 1988. How Many Species are There on Earth? *Science*, **241**: 1441-1449.

May R.M. 2002. The future of biological diversity in a crowded world. *Curr. Science*, **82**: 1325-1331.

May R.M., Lawton J.H. & Stork N.E. 1995. Assessing extinction rates, pp. 1-24. In: Lawton J.H. & May R.M. (eds.). Extinction Rates. Oxford University Press, Oxford.

McQuarrie N. & van Hinsbergen D.J.J. 2013. Retrodeforming the Arabia-Eurasia collision zone: Age of collision versus magnitude of continental subduction. *Geology*, **41**: 315-318.

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

Médail F. 2008. Mediterranean. In: S.E. Jorgensen (ed.). Encyclopedia of Ecology. Elsevier, Amsterdam.

Meyer C.P. & Paulay G. 2005. DNA Barcoding: Error Rates Based on Comprehensive Sampling. *PLoS Biology*, **3**: e422.

Mittelbach G.G. & Schemske D.W. 2015. Ecological and evolutionary perspectives on community assembly. *Trends in Ecol. and Evol.*, **30**: 241–247.

Mittelbach G.G. & Schemske D.W. 2015. Ecological and evolutionary perspectives on commmunity assembly. *Trends in Ecol. and Evol.*, **30**: 241-247.

Mittermeier R.A., Gil P.R., Hoffmann M., Pilgrim J., Brooks T., Mittermeier C.G., Lamoreux J. & da Fonseca G.A.B. 2004. Hotspots Revisited: Earth's Biologically Richest and Most Endangered Ecoregions. Mexico City (Mexico): CEMEX.

Mora C., Rollo A. & Tittensor D.P. 2013. Comment on "Can we name Earth's species before they go extinct?". *Science*, **341**: 237.

Mora C., Tittensor D.P., Adl S., Simpson A.G.B. & Worm B. 2011. How Many Species Are There on Earth and in the Ocean? *PLoS Biol*, **9**: e1001127.

Moreira F. Arianoutsou M., Leone V., Moya D., Lovreglio R., Delipetrou P. & De Las Heras J. 2012. Management of Threatened, High Conservation Value, Forest Hotspots Under Changing Fire Regimes, pp. 257-291. In : Moreira F. Arianoutsou M., Corona P. & De Las Heras J. (eds.). Post-Fire

Management and Restoration of Southern European Forests, Managing Forest Ecosystems 24, Springer, Netherlands.

Moritz C. & Cicero C. 2004. DNA Barcoding: Promise and Pitfalls. *PLoS Biol.*, **2**: e354.

Morris E.K., Caruso T., Buscot F., Morris E.K., Caruso T. et al. 2014. Choosing and using diversity indices: insights for ecological applications from the German Biodiversity Exploratories. *Ecol. and Evol.*, **4**: 3514–3524.

Mousson L., Neve G. & Baguette M. 1999. Metapopulation structure and conservation of the cranberry fritillary *Boloria aquilonaris* (Lepidoptera, Nymphalidae) in Belgium. *Biol. Conserv.*, **87**: 285–293.

Myers N. & Cowling R. 1999. Mediterranean Basin, pp. 254-267. In: Mittermeier R.A., Myers N. & Mittermeier C.G. (eds.). Hotspots – Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. Mexico City: CEMEX & Conservation International.

Nazari V. & Sperling F.A.H. 2007. Mitochondrial DNA divergence and phylogeography in western Palaearctic Parnassiinae (Lepidoptera: Papilionidae): How many species are there? *Insect Syst. Evol.*, **38**: 121–138.

Nazari V., Hagen W.T. & Bozano G.C. 2010. Molecular systematics and phylogeny of the 'Marbled Whites' (Lepidoptera: Nymphalidae, Satyrinae, *Melanargia* Meigen). *Syst. Entomol.*, **35**: 132-147.

Öztürk Ö. 2011. Pontus: antikçağ'dan günümüze Karadeniz'in etnik ve siyasi tarihi. Genesis Kitap, Ankara.

P. Jaccard. 1091. Distribution de la florine alpine dans la Bassin de Dranses et dans quelques regiones voisines. *Bull. de la Soc. Vaudoise des Scien. Nat.,* **37**: 241–272.

Page R.D.M. & Holmes E.C. 1998 . Molecular Evolution: a phylogenetic approach. Blackwell Science, Oxford.

Pamperis L.N. 2009. The Butterflies of Greece. Editions Pamperis, Athens.

Papadopoulou A. & Knowles L.L. 2015. Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. *Evolution*, **69**: 1501-1517.

Pares J.M., Freeman R. & Roca E., 1992. Neogene structural development in the Valencia Trough margins from palaeomagnetic data. *Tectonophysics*, **203**: 111-124.

Parmesan C., Ryrholm N., Stefanescu C., Hill J., Thomas C., Descimon H., Huntley B., Kaila L., Kullberg J. & Tammaru T. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, **399**: 579–583.

Patterson B.D. & Atmar W. 1986. Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biol. J. of the Linnean Soc.*, **28**: 65-82.

Petrakis P.V. & and Legakis A. 2005. Insect migration and dispersal with emphasis on Mediterranean ecosystems, pp. 85-126. In: Ashraf M. & Elewa T. (eds.). Insect migration and dispersal with emphasis on Mediterranean ecosystems. Springer, Berlin.

Pfiffner O.A. 1992. Alpine Orogeny, pp. 180-190. In: Blundell D., Freeman R. & Mueller S. (eds). A continent revealed: The European Geotraverse. European Science Foundation, Cambridge.

Phillips S. J. & Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**: 161-175.

Pigot A.L. & Etienne R.S. 2014. A new dynamic null model for phylogenetic community structure. *Ecol. Letters*, **18**: 153-163.

Pigot A.L. & Tobias J.A. 2013. Species interactions constrain geographic range expansion over evolutionary time. *Ecol. Lett.*, **16**: 330–338.

Pimm S.L. & Brooks T.M. 1997. The sixth extinction: how large, where, and when?, pp. 46-62. In: Raven P.H. (ed.). Nature and Human Society: The Quest for a Sustainable World. National Academy Press, Washington DC.

Pimm S.L., Jenkins C.N., Abell R., Brooks T.M., Gittleman J.L., Joppa L.N., Raven P.H., Roberts C.M. & Sexton J.O. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**: 1246752-1 - 1246752-10.

Pimm S.L., Russell G.J., Gittleman, J.L. & Brooks T.M. 1995. The future of biodiversity. *Science*, **269**: 347-350.

Pollard E. & Yates T. J. 1994. Monitoring butterflies for ecology and conservation: the British butterfly monitoring scheme. Springer Science & Business Media.

Por F.D., 2009. Tethys returns to the Mediterranean: success and limits of tropical re-colonization. *BioRisk*, **3**: 5–19.

Posada D. & Crandall K.A. 2001. Intraspecific phylogenetics: Trees grafting into networks. *Trends in Ecol. and Evol.*, **16**: 37-45.

Prendini L. 2005. Comment on 'Identifying spiders through DNA barcoding'. *Canadian J. of Zool.*, **83**: 498–504.

Provan J. & Bennett K.D. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecol. and Evol.*, **23**: 564-571.

Quek S-P., Davies S.J., Itino T. & Pierce N.E. 2004. Codiversification in an ant-plant mutualism: stem texture and the evolution of host use in Crematogaster (Formicidae: Myrmicinae) inhabitants of Macaranga (Euphorbiaceae). *Evolution*, **58**: 554–570.

Ratnasingham S. & Hebert P.D.N. 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molec. Ecol. Notes*, **7**: 355–364.

Raven P.H. 1964. Catastrophic selection and edaphic endemism. *The Amer. Natur.*, **98**: 336-338.

Raven P.H. 1983. The challenge of tropical biology. *Bull. Entomol. Soc. Am.*, **29**: 4-13.

Raven P.H., Roberts C.M. & Sexton J.O. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**: 1246752-1 - 1246752-10.

Rehault J.-P., Boillot, G. & Mauffret A. 1985. The Western Mediterranean Basin, pp. 101-129. In: Stanley D.J. & Wezel F.-C. (eds.). Geological evolution of the Mediterranean Basin. Springer, New York.

Renvoise E., Montuire S., Richard Y., Quere J.P., Gerber S., Cucchi T., Chateau-Smith C.& Tougard C. 2012. Microevolutionary relationships between phylogeographical history, climate change and morphological variability in the common vole (*Microtus arvalis*) across France. *J. of Biogeo.*, **39**: 698–712.

Ricklefs R.E. & Lovette I.J. 1999. The roles of island area *per se* and habitat diversity in the species–area relationships of four Lesser Antillean faunal groups. *J. of Anim. Ecol.*, **68**: 1142-1160.

Ricklefs R.E. 2008. Disintegration of the Ecological Community. *The Am. Nat.*, **172**: 741–750.

Ricou L.E., Dercourt J., Geyssant J., Grandjacquet C., Lepvrier C. & Biju-Duval B. 1986. Geological constraints on the Alpine evolution of the Mediterranean Tethys. *Tectonophysics*, **123**: 83-122.

Roe A., Weller S., Baixeras J., Brown JW., Cummings M.P., Davis D.R., Horak M., Kawahara A.Y., Mitter C., Parr C.S., Regier J.C., Rubinoff D., Simonsen T.J., Wahlberg N. & Zwick A. 2010. Evolutionary framework for Lepidoptera model systems, pp. 1–24. In: Goldsmith, M.R. & Marec, F. (eds.). Molecular biology and genetics of the Lepidoptera. CRC Press, Boca Raton.

Rögl F. & Steininger F.F. 1983. Vom Zerfall der Tethys zu Mediterran und Paratethys. Die neogene Palaogeographie und Palinspastik des zirkum-mediterranen Raumes. *Ann. Naturhist. Mus. Wien*, **85**: 135-163.

Rosenbaum G. & Lister G.S. 2004. Formation of arcuate orogenic belts in the western Mediterranean region, *Geol. Soc. Am. Special Paper*, **383**: 41–56.

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

Rosenbaum G., Lister G.S. & Duboz C. 2002b. Relative motions of Africa, Iberia and Europe during Alpine orogeny. *Tectonophysics*, **359**: 117-129.

Rosenzweig M.L. 2005. Avoiding Mass Extinction: Basic and Applied Challenges. *The Amer. Midland Natur.*, **153**: 195-208.

Rouchy J-M. & Caruso A. 2006. The Messinian salinity crisis in the Mediterranean basin: A reassessment of the data and an integrated scenario. *Sediment. Geol.*, **188-189**: 35-67.

Roveri M. & Manzi V. 2006. The Messinian salinity crisis: Looking for a new paradigm? *Palaeogeo., Palaeoclim., Palaeoecol.*, **238**: 386–398.

Roveri M., Lugli S., Manzi V., Gennari R. & Schreiber C.B. 2014. High-resolution strontium isotope stratigraphy of the Messinian deep Mediterranean basins: Implications for marginal to central basins correlation. *Marine Geology*, **349**: 113–125.

Roveri M., Manzi V., Lugli S., Schreiber B.C., Caruso A., Rouchy J.M., Iaccarino S.M., Gennari R. & Vitale F.P. 2006. Clastic vs. primary precipitated evaporites in the Messinian Sicilian basins. Preprint *Acta Naturalia de l'Ateneo-Parmense,* **42**: 125–199. RCMNS Interim Colloquium "The Messinian salinity crisis revisited II", Parma 7-9 September 2006.

Royden L.H. 1993. Evolution of retreating subduction boundaries formed during continental collision. *Tectonics*, **12**: 629-638.

Rozensweig M.L. 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Rubinoff D. 2006. Utility of Mitochondrial DNA Barcodes in Species Conservation. *Conserv. Biol.*, **20**: 1026-1033.

Rubinoff D., Cameron S. & Will K. 2006. A Genomic Perspective on the Shortcomings of Mitochondrial DNA for "Barcoding" Identification. *J. of Hered.*, **97**: 581–594.

Ryan W.B.F. & Cita M.B. 1978. The nature and distribution of Messinian erosional surfaces; indicators of a several-kilometer-deep Mediterranean in the Miocene. *Mar. Geol.*, **27**: 193-230.

Ryan W.B.F. 2011. Geodynamic response to a two-step model of the Messinian salinity crisis. *Bull. Soc. Géol. Fr.*: **182**: 73-78.

Sadori L., Bertini A., Combourieu-Nebout N., Kouli K., Mariotti Lippi M., Roberts N. & Mercuri A.M. 2013. Palynology and Mediterranean vegetation history. *Fl. Medit.*, **23**: 141-156.

Santos A., Whittaker R. J., Triantis K. A., Borges P. A., Jones O. R., Quicke D. L., & Hortal J. 2010. Are species–area relationships from entire archipelagos congruent with those of their constituent islands?. *Glob. Ecol. and Biogeo.*, **19**: 527-540.

Sañudo-Restrepo C.P., Dincă V., Talavera G. & Vila R. 2013. Biogeography and systematics of *Aricia* butterflies (Lepidoptera, Lycaenidae). *Molec. Phylo. and Evol.*, **66**: 369–379.

Savolainen V., Cowan R.S., Vogler A.P., Roderick G.K. & Lane R. 2005. Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Phil. Trans. R. Soc. Lond. B,* **360**: 1805–1811.

Savostin L.A., Sibuet J.C., Zonenshain L.P., Le Pichon X. and Roulet M.J. 1986. Kinematic evolution of Tethys belt from the Atlantic Ocean to the Pamirs since the Triassic. *Tectonophysics*, **123**: 1-35.

Scheffers B.R., Joppa L.N., Pimm S.L. & Laurance W. 2012. What we know and don't know about Earth's missing biodiversity. *Trends in Ecol. and Evol.*, **27**: 501-510.

Schettino A. & Turco E. 2006. Plate kinematics of the western Mediterranean region during the Oligocene and early Miocene. *Geophys. J. Int.*, **166**: 1398–1423.

Schlick-Steiner B.C., Seifert B., Stauffer C., Christian E., Crozier R.H. & Steiner F.M. 2007. Without morphology, cryptic species stay in taxonomic crypsis following discovery. *Trends in Ecol. and Evol.*, **22**: 391–392.

Schmitt T. & Seitz A. 2001. Allozyme variation in *Polyommatus coridon* (Lepidoptera: Lycaenidae): identification of ice-age refugia and reconstruction of post-glacial expansion. *J. of Biogeo.*, **28**: 1129–1136.

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

Schmitt T., Giessl A. & Seitz A. 2003. Did *Polyommatus icarus* (Lepidoptera: Lycaenidae) have distinct glacial refugia in southern Europe? Evidence from population genetics. *Biol. J. of the Linnean Soc.*, **80**: 529-538.

Schmitt T., Varga Z. & Seitz A. 2005a. Are *Polyommatus hispana* and *Polyommatus slovacus* bivoltine *Polyommatus coridon* (Lepidoptera: Lycaenidae)? The discriminatory value of genetics in the taxonomy. *Org. Divers. Evol.*, **5**: 297–307.

Schmitt, T., Röber, S. & Seitz, A. 2005b. Is the last glaciation the only relevant event for the present genetic population structure of the Meadow Brown butterfly *Maniola jurtina* (Lepidoptera: Nymphalidae)? *Biol. J. of the Linnean Soc.*, **85**: 419–431.

Settele J. & Kühn E. 2009. Insect Conservation. *Science*, **325**: 41–42.

Settele J., Feldmann R. & Reinhardt R. 2000. Die Tagfalter Deutschlands. Ulmer, Stuttgart.

Settele J., Kudrna O., Harpke A., Kühn I., van Swaay C., Verovnik R., Warren M., Wiemers M., Hanspach J., Hickler T., Kühn E., van Halder I., Veling K., Vliegenthart A., Wynhoff I. & Schweiger O. 2008. Climatic Risk Atlas of European Buterflies. Biorisk 1 (special issue), Pensoft Sofia-Moscow.

Settele J., Kühn E. & Thomas J.A. (eds.). 2005. Studies in the Ecology and Conservation of Butterflies in Europe 2. Pensoft, Sofia.

Settele J., Shreeve T., Konvička M. & Van Dyck H. (eds.) 2009. Ecology of Butterflies in Europe. Cambridge University Press, Cambridge.

Shearer T.L. & Coffroth M.A. 2008. Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Molec. Ecol. Res.*, **8**: 247–255.

Shi G.R. 1993. Multivariate data analysis in palaeoecology and palaeobiogeography: a review. *Palaeogeo., Palaeoclim., Palaeoecol.*, **105**: 199–234.

Šíchová J., Voleníková A., Dincă V., Nguyen P., Vila R., Sahara K. & Marec F. 2015. Dynamic karyotype evolution and unique sex determination systems in *Leptidea* wood white butterflies. *BMC Evol. Biol.,* **15**: 89.

Signor P.W. 1990. The geologic history of diversity. *Annu. Rev. Ecol. Syst.*, **21**: 509-539.

Simpson G.G. 1964 . Organisms and moleclues in evolution. *Science*, **146**: 1535-1538.

Smith A.M., Rodriguez J.J., Whitfield J.B., Deans A.R., Janzen D.H., Hallwachs W. & Hebert P.D.N. 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *PNAS*, **105**: 12359–12364.

Smith A.M., Woodley N.E., Janzen D.H., Hallwachs W. & Hebert, P.D.N. 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *PNAS*, **103**: 3657–3662.

Smith V. S. 2005. DNA barcoding: Perspectives from a "Partnerships for Enhancing Expertise in Taxonomy" (PEET) debate. *Syst. Biol.*, **54**: 841–844.

Sørensen, T. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter*, **5**: 1-34.

Spooner D.M. 2009. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *Am. J. of Botany*, **96**: 1177–1189.

Stefanescu C. 2000. El Butterfly Monitoring Scheme en Catalunya: los primeros cinco años. *Treballs de la Soc. Catal. de Lepidopt.*, **15**: 5–48.

Stefanescu C., Alarcón M. & Àvila A. 2007. Migration of the painted lady butterfly, *Vanessa cardui*, to north-eastern Spain is aided by African wind currents. *J. of Anim. Ecol.*, **76**: 888-898.

Stoltze M. 1996. Danske dagsommerfugle. Gyldendal, Copenhagen.

Stork N. & Gaston K.J. 1990. Counting species one by one. *New Sci*, **1729**: 43-7.

Stork N.E. 1993. How many species are there? *Biodiv. and Conserv.*, **2**: 215-232.

Stork N.E. 2010. Re-assessing current extinction rates. *Biodiv. Conserv.*, **19**: 357–371.

Stork N.E., McBroom J., Gely C. & Hamilton A.J. 2015. New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods. *PNAS*, **112**: 7519-7523.

Suc J.-P. 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature*, **307**: 429-432.

Suess E. 1893. Are great ocean depths permanent? *Natural Sciences*, **2**: 180-187.

Sundseth K. 2009. Natura 2000 in the Mediterranean Region. European Commission Environment Directorate General.

Svenning J.–C., Flojgaard C. & Baselga A. 2011. Climate, history and neutrality as drivers of mammal beta diversity in Europe: insights from multiscale deconstruction. *J. of Anim. Ecol.*, **80**: 393–402.

Taberlet P. & Cheddadi R. 2002. Quaternary Refugia and Persistence of Biodiversity. *Science*, **297**: 2009-2010.

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

Talavera G, Lukthanov VA, Pierce NE & Vila R. 2013. Establishing criteria for higher level taxonomic classification using molecular data: the systematics of *Polyommatus* blue butterflies (Lepidoptera,Lycaenidae). *Cladistics*, **29**: 166–192.

Tangley, L. 1997. How many species are there? *U.S. News and World Report*, **123**: 78–80.

Tarrier M. R. & Delacre J. 2008. Les papillons de jour du Maroc. Biotope, Mèze.

Tautz D., Arctander P., Minelli A.,Thomas R.H., Vogler A.P. 2003. A plea for DNA taxonomy. *Trends in Ecol. and Evol.*, **18**: 70–74.

Templeton A.R. 1990. The role of genetics in captive breeding and reintroduction for species conservation. *Endangered Species UPDATE*, **8**: 14-17.

Templeton A.R. 1992. Human origins and analysis of mitocondrial DNA sequences. *Science*, **255**: 737

Tennent J. 1996. The butterflies of Morocco, Algeria and Tunisia. Gem Publishing, Wallingford, Oxfordshire.

The *Heliconius* Genome Consortum. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, **487**: 94–98.

The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. Downloaded on 20 August 2015.

Thomas J. Matthews T.J., Cottee-Jonesm H.E.W. & Whittaker R.J. 2015. Quantifying and interpreting nestedness in habitat islands: a synthetic analysis of multiple datasets. *Div. and Distrib.*, **21**: 392–404.

Thomas J.A. & Settele J. 2004. Butterfly mimics of ants. *Nature*, **432**: 283–284.

Thomas J.A. 1994. The Ecology and Conservation of *Maculinea arion* and Other European Species of Butterfly, pp. 180–197. In: Pullin A.S. (ed.). Ecology and Conservation of Butterflies. Chapman and Hall, London.

Thomson G. 1987. Enzyme variation at morphological boundaries in Maniola and related genera (Lepidoptera: Nymphalidae: Satyrinae). Unpublished D. Phil. Thesis, University of Stirling, Scotland.

Thomson G. 2011. The Meadow Brown Butterflies.Waterbeck, Scotland.

Thornton I. 1997. Krakatau. The Destruction and Reassembly of an Island Ecosystem. Harvard University Press, Cambridge.

Thornton I.W.B. 2001. Colonization of an island volcano, Long Island, Papua New Guinea, and an emergent island, Motmot, in its caldera lake. I. General introduction. *J. of Biogeo*, **28**: 1299-1310.

Thuiller W. 2003. BIOMOD–optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. change biol.,*b: 1353-1362.

Thuiller W., Lavergne S., Roquet C., Boulangeat I., Lafourcade B. & Araújo M.B. 2011. Consequences of climate change on the Tree of Life in Europe. *Nature*, **470**: 531-534.

Toews D.P.L. & Brelsford A. 2012.The biogeography of mitochondrial and nuclear discordance in animals. *Molec. Ecol.*, **16**: 3907-30.

Toews D.P.L., Mandic M., Richards J.G. & Irwin D.E. 2014. Migration, mitochondria, and the yellow-rumped warbler. *Evolution*, **68**: 241–255.

Tolman T. & Lewington R. 1997. Collins Field Guide to the Butterflies of Britain and Europe. Collins, London.

Tomaselli R. 1977. The degradation of the Mediterranean maquis. *Ambio*, The Mediterranean: A Special Issue, **6**: 356-362.

Tóth J. P., Bereczki J., Varga Z., Rota J., Sramkó G. & Wahlberg N. 2014. Relationships within the *Melitaea phoebe* species group (Lepidoptera: Nymphalidae): new insights from molecular and morphometric information. *System. Entomol.,* **39**: 749-757.

Tricart P., Torelli L., Argnani A., Rekhiss F. & Zitellini N. 1994. Extensional collapse related to compressional uplift in the Alpine Chain off northern Tunisia (Central Mediterranean). *Tectonophysics*, **238**: 317-329.

Tshikolovets V. V. 2011. Butterflies of Europe and the Mediterranean área. Tshikolovets Publications, Pardubice.

Tuomisto H. 2010. A diversity of beta diversities: straightening up a concept gone awry. 2. Quantifying beta diversity and related phenomena. *Ecography*, **33**: 23–45.

Turner J.R.G. & Hawkins B.A. 2004. The global diversity gradient, pp. 171–190. In: Lomolino M. & Heaney L. (eds.). Frontiers of biogeography: new directions in the geography of nature. Mass: Sinauer Associates, Sunderland.

Ulrich W., Almeida M. & Gotelli N.J. 2009. A consumer's guide to nestedness analysis. *Oikos*, **118**: 3-17.

Urban M.C. 2015. Accelerating extinction risk from climate change. *Science*, **348**: 571-573.

Valdiosera C.E., García-Garitagoitia J.-L., Garcia N., Doadrio I., Thomas M.G., Hänni C., Arsuaga J.-L., Barnes I., Hofreiter M., Orlando L. & Götherström A. 2008. Surprising migration and population size dynamics in ancient Iberian brown bears (*Ursus arctos*). *PNAS*, **105**: 5123–5128.

Valentine J.W. 1970. How many marine invertebrate fossils? *J. Paleontol.*, **44**: 410-15.

Van Hinsbergen D.J.J., Vissers R.L.M. & Spakman W. 2014. Origin and consequences of western Mediterranean subduction, rollback, and slab segmentation. *Tectonics*, **33**: 393-419.

van Swaay C. A., Nowicki, P. Settele, J. & van Strien A. J. 2008. Butterfly monitoring in Europe: methods, applications and perspectives. *Biodiv. and Conserv.*, **17**: 3455-3469.

Van Swaay C., Cuttelod A., Collins S., Maes D., Lopez Munguira M., Šašić M., Settele J., Verovnik R., Verstrael T., Warren M., Wiemers M. & Wynhof I. 2010. European Red List of Butterfies. Publications Office of the European Union, Luxembourg.

Vane-Wright R.I. & Ackery P.R. (eds.). 1989. The Biology of Butterflies. Princeton University Press.

Vences M., Thomas M., Bonett R.M. & Vieites D.R. 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Phil. Trans. R. Soc. Lond. B,* **360**: 1859–1868.

Verovnik R. & Glogovcan P. 2007. Morphological and molecular evidence of a possible hybrid zone of *Leptidea sinapis* and *L. reali* (Lepidoptera: Pieridae). *Eur. J. Entomol.,* **104**: 667–674.

Vitousek P.M., Loope L.L. & Andersen H. (eds.). 1995. Islands: biological diversity and ecosystem function. Ecological studies 115. Springer-Verlag, Berlin.

Vodolazhsky D.I. & Stradomsky B.V. 2008. Phylogenetic analysis of subgenus *Polyommatus* (s. str.) Latreille, 1804 (Lepidoptera: Lycaenidae) based on mtDNA markers. *Part II. Kavk. Entomol. Bjul.* **4**: 237–242.

Vogiatzakis I.N., Mannion A.M. & Griffiths G.H. 2006. Mediterranean ecosystems: problems and tools for conservation. *Progr. in Phys. Geog.*, **30**: 175-200.

Wahlberg N. & Saccheri I. 2007. The effects of Pleistocene glaciations on the phylogeography of *Melitaea cinxia* (Lepidoptera: Nymphalidae). *Eur. J. of Entomol.,* **104**: 675-684.

Wahlberg N., Klemetti T. & Hanski I. 2002. Dynamic populations in a dynamic landscape: the metapopulation structure of the marsh fritillary butterfly. *Ecography*, **25**: 224–232.

Wake D.B & Vredenburg V.T. 2008. Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *PNAS,* **105**: 11466–11473.

Wallace A.R. 1855. On the law which has regulated the introduction of new species. *Ann. Magazine Nat. His.*, **16**: 184–196.

Wallace A.R. 1880. Island life. Macmillan, London.

Ward R.D., Zemlak T.S., Innes B.H., Last P.R. & Hebert P.D.N. 2005. DNA barcoding Australia's fish species. *Philos Trans R Soc Lond B,* **360**: 1847–1858.

Warren B.C.S. 1936. Monograph of the genus *Erebia* Dalman. British Museum of Natural History, London.

Warren B.H., Simberloff D., Ricklefs R.E., Aguil R. & Condamine F.L. et al. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecol. Letters*, **18**: 200–217.

Warren M.S. 1994. The UK status and suspected metapopulation structure of a threatened European butterfly, the marsh fritillary *Eurodryas aurinia*. *Biol. Conserv.*, **67**: 239–249.

Wasik B.R., Liew S.F., Lilien D.A., Dinwiddie A.J., Noh H., Cao H. & Monteiro A. 2014. Artificial selection for structural color on butterfly wings and comparison with natural evolution. *PNAS,* **111**: 12109-12114.

Waters J.M. 2011. Competitive exclusion: phylogeography's 'elephant in the room'? *Molec. Ecol.*, **20**: 4388–4394.

Waters J.M., Fraser C.I. & Hewitt G.M. 2013. Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecol. and Evol.*, **28**: 78-85.

Watson H.C. 1835. Remarks on the geographical distribution of British plants. Longman, Rees, Orme, Brown, Green, and Longman, London.

Watt W.B. & Boggs C.L. 2003. Synthesis: Butterflies as Model Systems in Ecology and Evolution - Present and Future, pp. 603–613. In: Boggs C.L., Watt W.B., Ehrlich P.R. (eds.). Butterflies. Ecology and Evolution Taking Flight. The University of Chicago Press, Chicago.

Weingartner E., Wahlberg N. & Nylin S. 2006. Speciation in Pararge (Satyrinae: Nymphalidae) butterflies - North Africa is the source of ancestral populations of all Pararge species. *System. Entomol.*, **31**: 621–632.

Whittaker R.J. & Fernández-Palacios J.M. 2007. Island Biogeography: Ecology, Evolution and Conservation. Oxford University Press, Oxford.

Whittaker R.J., Araújo M.B., Jepson P., Ladle R.J., Watson J.E.M. & Willis K.J. 2005. Conservation Biogeography: assessment and prospect. *Div. and Distrib.*, **11**: 3–23.

Whittaker R.J., Triantis K. & Ladle R.J. 2008. A general dynamic theory of oceanic island biogeography. *J. of Biogeo.*, **35**: 977–994.

Whitworth T. L., Dawson R. D., Magalon H. & Baudry E. 2007. DNA barcoding cannot reliably identify species of the blowfly genus Protocalliphora (Diptera: Calliphoridae). *Proc. R. Soc. B*, **274**: 1731–1739.

Wiemers M. & Fiedler K. 2007. Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology,* **4**: 8.

Wiemers M. 2003. Chromosome differentiation and the radiation of the butterfly subgenus *Agrodiaetus* (Lepidoptera: Lycaenidae: Polyommatus) – a molecular phylogenetic approach. Doctoral thesis, Rheinischen Friedrich-Wilhelms-Universität, Bonn.

Wiemers M., Stradomsky B.V. & Vodolazhsky D.I. 2010. A molecular phylogeny of Polyommatus s. str. and Plebicula based on mitochondrial COI and nuclear ITS2 sequences (Lepidoptera: Lycaenidae) *Eur. J. Entomol.* **107**: 325–336.

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.

Wilson E.O. 1992. The diversity of life. Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Wilson J-J., Sing K.-W. & Sofian-Azirun M. 2013. Building a DNA Barcode Reference Library for the True Butterflies (Lepidoptera) of Peninsula Malaysia: What about the Subspecies? *PLoS ONE,* **25**: doi:

10.1371/journal.pone.0079969.

Zapata F.A., Gaston K.J. & Chown S.L. 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. *J. of Anim Ecol.*, **72**: 677-690.

Zinetti F., Dapporto L., Vovlas A., Chelazzi G., Bonelli S., Balletto E. & Ciofi C. 2013. When the Rule Becomes the Exception. No Evidence of Gene Flow between Two *Zerynthia* Cryptic Butterflies Suggests the Emergence of a New Model Group. *PLoS ONE,* **8**: e65746. doi:10.1371/journal.pone.0065746.

Zuckerkandl E. & Pauling L. 1962. Molecular disease, evolution, and genetic heterogeneity. In: Kasha M. & Pullman M.B. (eds.). Horizons in Biochemistry. Academic Press, New York.