Original article

Differentiation in the marbled white butterfly species complex driven by multiple evolution forces

Jan Christian Habel1*, Roger Vila2, Raluca Vodă2,3, Martin Husemann4, Thomas Schmitt5,6, Leonardo Dapporto2,7

1Terrestrial Ecology Research Group, Department of Ecology and Ecosystem Management, School of Life Sciences Weihenstephan, Technische Universität München, D-85350 Freising, Germany

2Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37, ESP-08003 Barcelona, Spain

3Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, I-10123 Turin, Italy

4General Zoology, Institute for Biology, Martin-Luther-University Halle-Wittenberg, D-06120 Halle (Saale), Germany

5Senckenberg German Entomological Institute, D-15374 Müncheberg, Germany

6Zoology, Institute for Biology, Faculty of Natural Sciences I, Martin Luther University Halle-Wittenberg, D-06099 Halle, Germany

7Department of Biology, University of Florence, I-50019 Florence, Italy

*Corresponding author:

Jan Christian Habel, Terrestrial Ecology Research Group, Department of Ecology and Ecosystem Management, School of Life Sciences Weihenstephan, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, D-85350 Freising, Germany; E-Mail: Janchristianhabel@gmx.de

Running title: Diversification in a butterfly species complex

ABSTRACT

Aim Genetic and phenotypic data may show convergent or contrasting spatial patterns. Discrepancies between markers may develop in response to different evolutionary forces. Here, we analyse taxonomic relationships among closely related butterfly taxa of the marbled

white species group and test for potential evolution forces, including genetic and phenotypic characters, driving inter- and intraspecific differentiation.

Location Western Palaearctic (including north-western Africa).

Methods We compared the distribution of the mitochondrial cytochrome c oxidase subunit I gene (COI) sequences, of some allozyme loci, and of the shape of wings and genitalia obtained by applying landmark-based techniques for three butterfly species, Melanargia galathea (distributed across central and eastern Europe), M. lachesis (Iberia) and M. lucasi (North Africa).

Results All the markers showed a strong spatial structure but also discordance among patterns. COI sequences, wing shape and genitalia indicated a main split between M. galathea and M. lucasi. A lower differentiation between M. galathea and M. lachesis was found in wing shape and reflected in two mutations of the COI gene, while allozymes indicated a strong divergence. Within M. galathea, allozyme data and COI, but not morphology, revealed the existence of a slightly differentiated lineage in the Italian Peninsula, France and Switzerland. Based on COI, Melanargia lucasi revealed two sub-groups, a western and an eastern Maghreb lineage.

Main conclusions Long-term isolation of Melanargia populations between North Africa and Europe led to divergence between M. galathea and M. lucasi. This was followed by a recent differentiation among populations presumably isolated during Pleistocene cold periods, such as M. lachesis in Iberia. These lineages are characterized by a low tendency to overlap in secondary sympatry. The different patterns of the four markers may arise from divergent evolutionary processes and pressures: wings may be mainly affected by environmental selection, genital structures by sexual selection, whereas long-term isolation and drift may drive divergence of mitochondrial DNA and allozymes.

Keywords: Biogeography, geographic isolation, geometric morphometrics, parapatric differentiation, refugia, reproductive isolation, sexual selection

INTRODUCTION

Differentiation of genetic lineages and their diversification into distinct species is a process predominantly occurring in allopatry and determined by the presence of barriers to gene flow typically established by dispersal constraints, climatic oscillations and/or tectonic events (Hewitt, 2004; Mittelbach & Schemske, 2015). The persistence of such barriers is often only temporal and populations can disperse from their original areas, resulting in secondary contacts between different genetic lineages or taxa. The establishment of secondary sympatry

is a fundamental process accruing local diversity (Mittelbach & Schemske, 2015). With growing availability of large genetic datasets, it is becoming evident that the maintenance of allopatry among species is not only determined by physical and ecological barriers and that most terrestrial taxa show a delay to secondary sympatry slowing down the accumulation of species in communities (Pigot & Tobias, 2015). Different mechanisms have been advocated to explain the lack of overlapping distributions among closely related species and lineages; these include density dependent processes, competition for similar resources, reproductive interference and environmental preferences (Waters, 2011; Vodă et al., 2015). On the other hand, delay to secondary sympatry may facilitate the last stages of the speciation process by reducing gene flow before reproductive barriers have been fully established.

Speciation is also related to particular life history traits of species, such as philopatry and low dispersal ability (Claramunt et al., 2012), hybridization and introgression (Mallet, 2007), fast responses to particular ecological pressures (Wade, 2001). These factors may be differently reflected in phenotypic and genotypic traits within and between species. Accordingly, the identification of a diversified population as a different species can be a difficult exercise. Hence, modern integrative taxonomy and biogeography should be based on the comparison of results obtained from multiple marker systems to fully understand the evolutionary patterns and the degree of diversification in species groups.

In this study we perform molecular and morphological analyses on three closely related Nymphalid butterfly taxa forming the Melanargia galathea (Linnaeus, 1758) species group in the western Palaearctic. For these analyses we collected individuals from populations spread over major parts of its western Palaearctic distribution range, encompassing two different continents (North Africa and Europe) and a large island (Sicily). We applied two molecular markers (mitochondrial DNA (mtDNA) and allozyme polymorphisms), and two phenotypic markers (wing shape and genitalia). This set of markers were selected to disentangle divergent selection and evolution regimes shaping inter- and intraspecific differentiation independently form each other: mitochondrial DNA, a rather conservative marker is assumed to be mainly driven from geographic isolation, while allozyme polymorphisms may also be affected from environmental selection (Vodă et al., 2015); the two phenotypic traits also depend on diverging evolution drivers, with wing shapes being strongly affected from environmental conditions while genitalia structures are assumed to be mainly influenced from sexual selection. These four markers were used to test products from diverging evolution drivers (e.g. geographic, ecological and sexual isolation and selection), which might have produced multiple genetic and phenotypic variation across the M. galathea group. We further review the current taxonomic status of the M. galathea species, according our results.

MATERIAL AND METHODS

Study species

The Marbled White butterfly M. galathea is widely distributed in southern and central Europe. Its northern distribution margin runs through central England, Belgium and northern Germany; it reaches the Baltic Sea in Poland. In Iberia, M. galathea only occurs in a restricted area south of the Pyrenees, while in the rest of the Iberian Peninsula it is replaced by its close relative M. lachesis (Kudrna et al., 2015). Melanargia lachesis occurs all across Iberia; North of the Pyrenees, the taxon is restricted to the Mediterranean parts of the French Roussillon (Bozano, 2002), where it locally overlaps with M. galathea. In this potential area of syntopy, the distribution of the two taxa remains mostly micro-allopatric apart from some few areas where the two species hybridize (Dennis et al., 1991; Fernández-Rubio, 1991b). Melanargia lucasi is endemic to the hill and mountain areas of the Maghreb region (Morocco, Algeria and Tunisia) (Bozano, 2002).

Melanargia galathea and M. lucasi had recently been suggested to be independent species based on the divergence in mitochondrial markers (COI and 16S) and male genitalia (Nazari et al., 2010). Previously, these taxa were considered to be distinct populations of M. galathea, supported by a high similarity in allozyme patterns (Habel et al., 2011) mirrored by the high similarity of the wingless nuclear gene (Nazari et al., 2010).

The status of M. lachesis is less clear. Based on DNA barcoding data, M. lachesis and M. galathea are only slightly diverged at the COI locus, and thus, species status may not be justified for M. lachesis (Dincă et al., 2015). Nevertheless, these two taxa are phenotypically clearly distinct by wing patterns and show a divergence of about 2.5% in the nuclear wingless gene (Nazari et al., 2010). Furthermore, molecular analyses based on allozyme polymorphisms showed that different genetic groups of M. galathea exist in central Europe and the Balkans (Schmitt et al., 2006; Habel et al., 2011), (similarly, M. lucasi has been found to be subdivided into a western and an eastern lineage in the Maghreb (Habel et al., 2011).

Data sets

All allozyme data were taken from previous studies (Schmitt et al., 2006; Habel et al., 2011). Specimens for genitalia and wing analyses were taken from the Siegbert Wagener collection, kindly provided by the Alexander Koenig Museum (Bonn, Germany), from the Museum of Zoology and Natural History of the University of Florence (Italy) and from Roger Vila´s tissue collection. Newly generated COI sequences were also based on Roger Vila´s tissue collection.

Cytochrome c Oxidase subunit I (COI)

In total 223 COI sequences were analysed, of which 132 represent newly generated data. Details on samples and sampling localities are provided as Table S1 in Appendix S1. Total genomic DNA was extracted for nine of these specimens using Chelex 100 resin (100–200

mesh, sodium form; Bio-Rad Laboratories GmbH, California, USA) under the following protocol: one leg was removed and introduced into 100 µL of 10% Chelex solution; 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 3000 rpm. A 658 bp fragment at the 5' end of the mitochondrial COI gene was amplified by polymerase chain reaction (PCR) using the primers LepF1 (5´-ATTCAACCAATCATAAAGATATTGG-3´) and LepR1 (5´-TAAACTTCTGGATGTCCAAAAAATCA-3´) (Hebert et al., 2004). PCR was performed in 25 µL volume reactions 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, 5 U/µL) and 2 µL of extracted DNA. The typical thermal cycling profile was: denaturation at 92ºC for 60 s, followed by five cycles of 92ºC for 15 s, 49ºC for 45 s and 62ºC for 150 s, and by 35 cycles of 92ºC for 15 s, 52ºC for 45 s and 62ºC for 150 s and a final extension at 62ºC for 420 s. PCR products were purified and sequenced by Macrogen Inc. (Amsterdam, The Netherlands). The rest of 123 sequences generated for this study were obtained from Biodiversity Institute of Ontario, Canada. In this case, a glass fibre protocol (Ivanova et al., 2006) was employed to extract DNA; PCR and DNA sequencing were carried out following standard DNA barcoding procedures for Lepidoptera (deWaard et al., 2008). To examine patterns of genetic variation, we constructed a haplotype network with PopART (http://popart.otago.ac.nz) using the TCS algorithm (Clement et al., 2000). Several loops indicating ambiguous connections were broken taking into account frequency and genetic distance criteria (Excoffier & Langaney, 1989).

Phylogenetic inference and dating

Bayesian inference (BI) on the basis of COI haplotypes was employed to infer phylogenetic relationships and estimate divergence times, by using the software BEAST 1.8.0 (Drummond et al., 2012). As outgroup we used published COI sequences of four other species in the genus: M. larissa, M. russiae, M. occitanica and M. ines. jMODELTEST 2.1.4 (Darriba et al., 2012) to select the best-fitting DNA substitution models according to the Akaike information criterion (AIC). As a result, the GTR+I+G model was used. The gamma distribution was estimated automatically from the data using six rate categories. A fixed substitution rate prior of 0.0115 per site per My was used, based on that estimated for the entire mitochondrial genome of several arthropods (Brower, 1994). An uncorrelated relaxed clock (Drummond et al., 2006) and a constant population size under a calibrated yule model were established as priors. Two independent chains were run for 50 million generations each, sampling values every 1000 steps. A conservative burn-in of 500,000 generations was applied for each run after checking Markov chain Monte Carlo (MCMC) convergence through graphically monitoring likelihood values in TRACER 1.6 (available at: http://beast. bio.ed.ac.uk/Tracer). Independent runs were combined in LOGCOMBINER 1.8.0, as implemented in BEAST, and all parameters were analysed using TRACER 1.6 to determine whether they had also reached stationarity. Tree topologies were assessed using TREEANNOTATOR 1.8.0 in the BEAST package to generate a tree with median node heights. FIGTREE 1.4.2 (available at: http://beast.bio.ed.ac.uk/FigTree) was used to visualize the tree along with node posterior probabilities and age deviations.

Allozymes

A total of 15 allozyme loci were analysed for 1,158 individuals from 32 populations of the three taxa (23 populations of M. galathea, 1 population of M. lachesis and 8 populations of M. lucasi). Additional information about sampling sites and sample sizes are given in Table S2 in Appendix S2). Further information on the analytical procedure (allozyme loci analysed and respective running conditions) is given in Habel et al. (2011). Based on the raw data, we calculated pairwise FST values with ARLEQUIN v. 3.0 (Excoffier et al., 2005).

Wing shape

We analysed the wing patterns of 347 male individuals (M. galathea, N=248; M. lachesis, N=41; M. lucasi, N=58) from 66 sites covering major parts of the western Palaearctic range (Table S3 in Appendix S3). Standardized digital images were taken of all butterflies' fore and hind wings using a Canon M1 digital camera (50 mm lens). Shapes of the left fore and hind wings were quantified by 13 homologous landmarks on each wing (26 in total). To ensure repeatability and to minimize measurement errors, landmarks were exclusively placed on wing vein intersections or locations where a wing vein meets the edge of the wing. Landmarks were digitized using the program TPS-DIG 2.12 (Rohlf, 2008). A schematic overview of all landmarks on the wings is given as fig. S1 in Appendix S4.

Generalized Procrustes Analysis (GPA) was applied to the landmark data in order to remove variation in scale and orientation and to superimpose the objects in a common coordinate system (Adams et al., 2004). Partial warps were calculated using the shape residuals from GPA. By applying principal components analyses (PCA) to partial warps, relative warps (PCs) were obtained and used as variables in a partial least squares discriminant analysis (PLSDA) (Mitteroecker & Bookstein, 2011).

Genitalia

A total of 123 males (M. galathea, N=75; M. lachesis, N=25; M. lucasi, N=23) were examined. An overview of all specimens and sample sites is given as Table S4 in Appendix S5. Genitalia were dissected using standard procedures and the tegumen and valves were photographed using a Nikon Coolpix 4500 camera, mounted on a binocular microscope. The number of distal processes on the valves (cornuti) was counted. Further, a combination of landmarks and sliding semi-landmarks (Bookstein, 1997) was applied to the tegumen (13) and valva (12) outline: 4 points on the outlines of both tegumen and valves that could be precisely identified in all individuals were considered as landmarks (type II and type III landmarks, Bookstein, 1997), whereas the other points were allowed to slide along the outline trajectory. GPA and PLSDA were applied in similar ways as described for the wing shape.

Overall representation of distribution patterns

In order to establish whether the species boundaries are correct, we tested for correspondence among the COI structure, allozyme patterns, taxonomic classification and phenotypic patterns. For each marker, we obtained bi-dimensional ordinations of the variation among specimens. For the dissimilarity matrices based on p-distances for COI and on FST for allozymes, we applied principal coordinate analysis (PCoA) reducing the dimensionality of the original matrices to two dimensions. To the relative warps obtained by the analyses of wings and genitalia, we applied partial least squares discriminant analysis (PLSDA) using species as grouping variable that returned a bi-dimensional representation of the variation among specimens. PLSDA also identified which shape variables are responsible for the differentiation among taxa; these differences were visualized by thin plate splines. For wings and genitalia, we evaluated the degrees of diversification among species as a percentage of specimens that can be blindly attributed to their taxon by applying a Jackknife algorithm classifying each specimen individually. PLSDA and Jackknife analyses were carried out with the 'plsda' and 'predict' functions of the mixOmics R package (Lê Cao et al., 2011). The bi-dimensional representations obtained for different markers are not directly comparable since they have arbitrary orientation and scale. Moreover, the cases (specimens) are not the same for different markers. To allow a direct comparison among the patterns of different markers, we applied the method described by Dapporto et al. (2014). We aggregated specimens into populations on the basis of their membership to the same species and to the same square of 2 x 2 degrees of latitude and longitude and computed the population configurations by calculating, for each marker, the barycenter of the specimens belonging to the same species/square. Subsequently, we minimized the differences between population configurations for different markers by applying a series of Procrustes analyses. We used the COI data, for which most populations were sampled as a reference. The configurations for genitalia, wing shape and allozymes where rotated and scaled based on the average configurations of the shared populations between COI and the other markers. Single specimens for each marker were finally rotated by using the same parameters resulting in a model based on shared populations. Procrustes analyses were carried out by using the 'recluster.procrustes' function of the R package recluster (Dapporto et al., 2014). Note that the selection of COI as a reference configuration for all Procrustes analyses does not affect the overall results; moreover this analysis does not change the diversity patterns inside each configuration, but only rotates and translates them to minimize arbitrary differences in location and orientation. The correlations between the population configurations of all pairs of markers were tested with the 'protest' function of the vegan package for R. Finally, the bi-dimensional configurations for specimens were projected in the RGB colour space by using the recluster package (Dapporto et al. 2013). This procedures attributes red, yellow, green and blue colours to each corner of a bi-dimensional representation and the colour of each point in the graph is interpolated among these extreme values. As a consequence, each point in a RGB colour space receives a unique colour according to its position with more similar cases receiving more similar colours (Kreft & Jetz, 2010). Specimens belonging to the same species and the same area of 2 x 2 degrees for latitude and longitude were grouped and their individual RGB colours were plotted on a map as pie charts.

For COI, in order to highlight possible patterns of mutual exclusion to a smaller spatial scale, we also aggregated the specimens to 0.5 x 0.5 degree rectangles.

RESULTS

Genotypes

We detected two main groups based on the COI sequences: a European cluster with M. galathea and M. lachesis and a North African cluster with M. lucasi (Fig. 1a-e), separated by 25 mutations (3.8 % divergence) (Fig. 1e). The TCS haplotype network showed a clear geographic structure within each of the two main groups (Fig. 1e). However, when the PCoA configuration was projected in the red-green-blue (RGB) space, the genetic structure within clades was less evident (Fig. 1a). Therefore, we analysed the two groups separately (Fig. 1b-d). Melanargia lucasi was subdivided into two lineages separated by a minimum of eight mutations (1.2 %): the first lineage was represented by specimens from Morocco, while the second lineage contained specimens sampled in Algeria and Tunisia (Fig. 1b, e). The M. galathea-lachesis group showed a more complex pattern (Fig. 1b-f): three main groups were identified, separated by just one or two mutations; M. lachesis was confined to one of these lineages (Fig. 1e). Despite the low genetic differentiation, the three groups had clear spatial and taxonomic patterns. All specimens identified as M. lachesis based on wing patterns shared similar haplotypes and clustered together. No individuals identified as M. galathea clustered within this group (Fig. 1e). The two additional lineages consisted solely of individuals identified as M. galathea displaying a clear geographic structure: the first group occurs in Italy and southern France (red), the second in central and eastern Europe as well as western Asia (green) and northern Iberia (dark green) (Fig. 1b,e). When examined in a higher spatial resolution (0.5 x 0.5 degrees, i.e. about 40 x 55 km), the distribution pattern of the four lineages became more evident and the three lineages appeared allopatric in most of the collection sites (Fig. 1f).

The BI recovered a phylogenetic tree with a similar topology to the one previously published for the genus Melanargia by Nazari et al. (2010) (Fig. 2). Melanargia lucasi formed a well-differentiated clade, as previously shown (Nazari et al., 2010), with strong support (posterior probability, 1.0). Melanargia galathea and M. lachesis have been recovered as monophyletic, with a fairly good support (posterior probability, 0.85). The estimated divergence time between the three taxa and their most common ancestor was around 2 Ma (95% highest posterior density interval, HPD, 1.38–2.93 Ma).

Allozyme analyses revealed a clear split between M. galathea and M. lachesis, but the populations of M. lucasi clustered within M. galathea. However, we detected two distinct genetic clusters in Europe: a western European cluster that includes Italy, France and the North African populations and a south-eastern-central European cluster (Fig. 3a-b). The protest analysis revealed that the geographic pattern of variation of COI, wings and genitalia are significantly correlated to each other (Fig. 3c), while allozyme variation showed no significant spatial correlation with the three other markers.

Phenotypes

Wing-shape analyses produced a total of 44 relative warps (22 for each wing). PLSDA using species membership as a grouping variable produced three distinct clusters (Fig. 4a-c). Three relative warps (PCs) explained most of the variance (hind wing PC1: 31.5 %; fore wing PC2: 21.9 %; hind wing PC3: 10.5 %). The inspection of thin plate splines for these variables revealed that M. lucasi is characterized by more elongated fore and hind wings compared to M. galathea and M. lachesis, while M. lachesis has a more elongated discoidal cell (Fig. 4b). Jack-knifing showed that most of the specimens could be blindly assigned to their species on the basis of wing shape (Fig. 4b). The distribution of wing shape in geographic space was largely congruent with the differentiation pattern of COI (Fig. 4a).

The analyses of genitalia produced a total of 42 relative warps (22 for the tegumen and 20 for the valva). PLSDA using the 42 relative warps and the number of cornuti as variables and species membership as a grouping factor produced two distinct clusters (M. galathea + lachesis vs. M. lucasi) for the first component, but showed a tendency to separate M. galathea from M. lachesis for the second component (Fig. 5b). The main pattern of diversification was produced by the number of cornuti, lower in M. lucasi compared to M. galathea and M. lachesis. Further, the valvae were less elongated in M. lucasi compared to M. galathea and also less elongated compared to M. lachesis (valva PC1: 48.7% of variance); the tegumen showed a longer uncus in M. galathea (tegumen PC2: 17.5% of variance). Jack-knifing resulted in several wrong assignments of M. galathea individuals as M. lachesis based on genital shape (table in Fig. 5b)

DISCUSSION

In this study, we performed a comparative biogeographic analysis of the M. galathea species complex based on two genetic markers (COI and allozymes) and two phenotypic characters (wing and genitalia shape) for populations sampled across Europe and North Africa. All the markers revealed a strong spatial differentiation but in some cases their patterns were discordant. COI sequences and the morphological markers displayed mostly congruent differentiation patterns that were partly in contrast to data provided by allozymes and the nuclear wingless gene previously assessed by Nazari et al. (2010).

Morphological and mitochondrial data gave strong evidence for an inter-specific split between the European M. galathea and the North African M. lucasi, which was not clearly reflected by allozymes and the wingless gene. Conversely, the differentiation between the two European taxa, M. galathea and M. lachesis, was particularly strong according to allozymes and wingless data; divergence in mitochondrial and morphological markers was less pronounced, but nevertheless evident, hence justifying the acceptance of M. lachesis as a valid species. COI sequences and allozymes suggested an additional split within the Maghreb region into a western and an eastern lineage. However, for the allozyme data, the eastern Maghreb populations clustered together with populations from Sicil and Italan Peninsul,, while it was an independent group within M. lucasi for COI. Furthermore, at least two additional allozyme

clusters, an Italian and a Balkan one, were detected, showing similarities with the COI haplo-groups that were not detected by wing and genital morphology.

The incongruence we detected for some of the markers may have several non-exclusive reasons, which we discuss below. Specifically, we focus on the biogeographic history of the populations and species and highlight potential evolutionary drivers, which may lead to such diverging signatures. Finally, we revise the current taxonomic status of the three taxa we analysed in this study.

Concordant and divergent patterns

Our analysis of COI revealed that M. galathea and M. lachesis are represented by one major haplo-group with three geographic sub-lineages separated by a comparatively low number of mutations. In European butterflies, a divergence of 0.3 % in COI, similar to what we found between the Iberian lineage of M. galathea and M. lachesis, is usually associated with divergence into interfertile and not always allopatric lineages often showing little or no phenotypic differentiation (Dincă et al., 2015). Introgression has been shown to occur even among well-recognized species of European butterflies (Mallet, 2005), leading to discordances between species-specific wing patterns and genetic differentiation (e.g. the cases of Lysandra (Talavera et al., 2013) and Iphiclides (Dincă et al., 2015)). From this perspective, the congruence of COI sequences, wing shape and allozyme with the wing colour pattern on the basis of which we attributed all the specimens to M. galathea or M. lachesis, is rather surprising; especially considering that we analysed several populations from the contact zone of both taxa (Fig. 1f). This suggests that, despite rare reports of specimens with intermediate wing patterns between M. galathea and M. lachesis, gene flow appears to be mostly absent.

Strict spatial segregation and chequerred distribution patterns among lineages were not only recorded between M. galathea and M. lachesis, but appear to be a general feature of the M. galathea species group. The three main COI lineages, comprising M. galathea and M. lachesis, detected in Europe and in the Middle East show clear geographic boundaries and have been found to coexist in the same 0.2 x 0.2 degree rectangle only in two areas (for M. galathea and M. lachesis and between two M. galathea lineages, Fig. 1f). The geographic distributions of the COI and allozyme lineages strongly support repeated contractions of populations to the three main southern European peninsulas (i.e. Iberia, Italy, Balkans) during cold stages of the Quaternary followed by interglacial expansions to central Europe. The occurrence of contact zones along the main mountain chains of Europe (i.e. Alps and Pyrenees) and a strong contribution of south-eastern European populations to central Europe, facilitated by the absence of mountains isolating the Balkan Peninsula, is one of the most recurrent paradigms in the phylogeography of European species (Hewitt, 1996, 2004; Schmitt, 2007).

The comparatively low levels of divergence displayed by COI and limited diversification in genital shape support a relatively recent split between M. lachesis and M. galathea. The high variation in allozymes and the strong differentiation of the wingless gene was unexpected in this context (but see below), as it is well known that mtDNA tends to diverge faster than nuclear DNA (Avise, 2009). In fact, the high divergence found for the wingless gene is unusual in closely related taxa and could suggest that evolution at this locus may not reflect neutral patterns. Regardless of this, the differentiation at the nuclear level was supported by the wing shape pattern, according to which 97.6% of all M. lachesis specimens could be blindly assigned to the correct species. It must be noted that, we only analysed a single population of M. lachesis for allozymes and only few wingless sequences were available (Nazari et al., 2010); more detailed analyses of these markers could reveal different patterns of variation.

A completely different diversification exists between the taxa living in allopatry in North Africa and Europe. Melanargia lucasi displayed a high COI divergence (minimum p-distance: 3.8 %) with respect to M. galathea and M. lachesis suggesting that M. lucasi represents a distinct species, in concordance with Nazari et al. (2010). Additionally, the genitalia and wing shape of M. lucasi differ from the two European taxa. The wingless gene, allozymes and wing colour patterns are relatively similar to M. galathea, although patterns of high divergence between European and North African species and lineages are the rule and not the exception (Husemann et al., 2014). This differentiation has likely evolved during long geographic separation between the two continents, which were only connected during the Messinian Salinity Crisis 5.9-5.3 Million years before present (Manzi et al., 2013). Genetic imprints of that period, or even older, were detected in many animal groups (Habel et al., 2012 with references therein). The minimum current distance between Africa and Europe is about 15 km at the Strait of Gibraltar, a relatively short sea barrier considering the dispersal abilities of many organisms, in particular flying insects. In fact, recent colonizations (i.e. less than 100,000 years) between North Africa and Iberia have been detected for different taxa (e.g. Cosson et al., 2005) and many Atlanto-Mediterranean butterfly species are found on both sides of this strait (Tolman & Lewington, 2008). However, this clearly is not the case of the group studied here since the opposed sides of the Gibraltar strait host either M. lucasi (North Africa) or M. lachesis (Iberia), which were distinct in all studied markers.

The genetic and morphological differences in the M. galathea species group are less pronounced across the Strait of Sicily, separating Tunisia and Sicily by a current minimum distance of 140 km (Habel et al., 2011). However, due to the lowering of the sea level during cold stages of the Quaternary, North Africa and Sicily were geographically considerably closer than at present, and reconstructions of coastlines estimated that this channel was only 50 km (Manzi et al., 2013). Thus, one explanation for the similarity in allozyme and wing pattern between M. galathea and M. lucasi might be a relatively recent exchange of individuals and hence of genetic information across the Strait of Sicily with subsequent introgression that could explain the discordance with mitochondrial markers.

Potential drivers leading to discordant differentiation

The different patterns shown by the four markers used in this study suggest that several mechanisms have been involved in the process of differentiation for these species. We further discuss potential drivers and scenarios.

MtDNA is mostly advocated as a neutral marker evolving in a clock-wise manner (Hewitt, 1996; Avise, 2009). We found the strongest differentiation for COI between M. galathea and M. lucasi. This split occurred more than 2 Ma, and is most probably the first one in the M. galathea species group, reflecting the fact that Europe and Africa represent two distinct geographic entities constantly being separated by sea since the Pliocene. In contrast, M. galathea and M. lachesis seem to represent a different case: although separated only by two mutations in COI (p-distance: 0.3 %), M. lachesis is consistently distinguished from its congeners by COI sequences as well as wing colouration and allozyme patterns.

These differences may be explained by several non-exclusive factors, including climatic and environmental preferences, restricted dispersal behaviour and density dependent phenomena (Pigot & Tobias, 2015; Vodă et al., 2015b). However, one might speculate that the diverging ecological preferences (M. galathea and M. lucasi in southern Europe and North Africa avoid extremely warm and dry environments, whereas M. lachesis is present under such conditions; personal observations of the authors) might be one major driver of this inconsistency among markers. Although allozymes are known as a suitable neutral, biogeographic marker, in many cases, several examples have suggested that these markers may be under thermal selection (e.g. Karl et al., 2009). Hence, the differential habitat use might have evoked a strong selection pressure on allozymes, enhancing the divergence between these two groups.

As the ratio of dark and light coloration on butterfly wings should strongly influence their thermoregulation, wing colour pattern differences between M. lachesis on the one hand and the two other taxa on the other hand, might also be triggered by the differential habitat use of these two groups. This hypothesis is supported by previous study on the evolution of wing shapes in butterflies, revealing that phenotypic differentiation may occur fast and at small geographic scales (Habel et al., 2013). Such shifts have been shown along altitudinal and latitudinal changes, but also along other environmental gradients to specific environmental conditions (e.g. for Pararge aegeria: Vandewoestijne & Van Dyck, 2011). The genitalic structures are not under thermal, but likely under sexual selection, potentially explaining difference in the patterns. However, more detailed and experimental studies are necessary to disentangle these aspects.

The taxonomy of the M. galathea species group revisited

The phenotypic and genetic data revealed significant splits among the three Melanargia taxa and the commonly used taxonomy is incongruent with our results: in the past, M. lucasi was mostly treated as a subspecies of M. galathea (e.g. Tolman & Lewington, 2008) and only recently taxonomists accepted this taxon as a distinct species (Nazari et al., 2010, Tshikolovets, 2011). All markers, except allozymes, support M. lucasi as a valid species endemic to the Maghreb region. The two taxa were likely separated since the Pliocene, but with possible events of introgression later on.

In contrast, M. lachesis was broadly accepted as a distinct species (Habel et al., 2011), although recent results challenged this hypothesis (Dincă et al., 2015). The comparatively weak split based on genitalia, mtDNA, wing shape, but strong differentiation in allozyme, nuclear markers and the strong tendency for chequered distribution patterns indicate locally restricted gene flow between M. lachesis and M. galathea. Therefore, all three entities should be accepted as distinct species, but with different evolutionary histories and ages.

Our study provides further support that multiple marker studies yield a more comprehensive understanding of biogeographic dynamics and the resulting taxonomy. However, at the same time, our data underline that delineating a specific driver leading to specific structures remains difficult.

Acknowledgements. Funding for this research was provided by the project “Barcoding Italian Butterflies”, by the European Union’s Seventh Framework programme for research and innovation under the Marie Skłodowska-Curie grant agreement No 609402-2020 researchers: Train to Move (T2M) postdoctoral fellowship to R. Vodă, by European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant (project no. 658844 Eco-PhyloGeo) to L. Dapporto and by the Spanish Ministerio de Economía y Competitividad. We are grateful to two anonymous referees for critical comments on a draft version of this article.

Biosketch:

Jan Christian Habel is assistant professor at the Department of Terrestrial Ecology at the Technical University Munich, Germany. He is working in the field of conservation biology, molecular ecology and biogeography, on various taxa.

Authors’ contributions: J.C.H., T.S. and L.D. created the study setup, J.C.H., R.V., R.V. and L.D. collected and analysed the data; all authors interpreted the data and wrote the manuscript.

Editor: Aristeidis Parmakelis

References

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

Bookstein, F.L. (1997) Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Medical Image Analysis 1, 225-243.

Bozano, G.C. (2002) Guide to the Butterflies of the Palearctic Region Satyridae, Part 3: Subtribes Melanargiina and Coenonympha, Genera Melanargia, Coenonympha, Triphysa and Sinonympha. Milano.

Brower, A.V. (1994) Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences USA 91, 6491–6495.

Claramunt, S., Derryberry, E.P., Remsen, J.V. & Brumfield, R.T. (2012) High dispersal ability inhibits speciation in a continental radiation of passerine birds. Proceedings of the Royal Society London B Biological Sciences 279, 1567-1574.

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

Cosson, J.F., Hutterer, R., Libois, R., Sarà, M., Taberlet, P. & Vogel, P. (2005) Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the western Mediterranean: a case study with the greater white-toothed shrew, Crocidura russula (Mammalia: Soricidae). Molecular Ecology 14, 1151–1162.

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., 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.

Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9, 772

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). Biological Journal of the Linnean Society 101, 1-49.

Descimon, H. & Mallet, J. (2009) Bad species. In: Ecology of butterflies in Europe, 219-249.

deWaard, J.R., Ivanova, N.V., Hajibabaei, M. & Hebert, P.D.N. (2008) Assembling DNA barcodes: analytical protocols. In: Martin C, (ed.) Methods in molecular biology: environmental genetics. Totowa, NJ: Humana Press. pp. 275–293.

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.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29, 1969– 1973.

Excoffier, L. & Langaney, A. (1989) Origin and differentiation of human mitochondrial DNA. American Journal of Human Genetics 44, 73-85.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1, 47–50.

Fernández-Rubio, F. (1991) Guia de mariposas diurnas de la Peninsular Ibérica, Baleares, Canarias, Azores y Madeira (Papilionidae, Pieridae, Danaidae, Satyridae y Hesperiidae). Ediciones Pirámide, Madrid.

Gómez, A. & Lunt, D.H. (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Phylogeography of southern European refugia (pp. 155-188). Springer Netherlands.

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

Habel, J.C., Husemann, M., Schmitt, T., Zachos, F.E., Honnen, A.-C., Petersen, B., Parmakelis, A. & Stathi, I. (2012) Microallopatry caused diversification in Buthus scorpions (Scorpiones: Buthidae) of the Atlas Mountains (NW Africa). PLoS ONE 7, e29403.

Habel, J.C., Rödder, D., Scalercio, S., Meyer, M. & Schmitt, T. (2010) Strong genetic cohesiveness between Italy and the Maghreb in four butterfly species. Biological Journal of the Linnean Society 99, 818-830.

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

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

Hewitt, G.M. (2004) Genetic consequences of climatic oscillation in the Quaternary. Philosophical Transactions of the Royal Society London B Biological Sciences 359, 183–195.

Husemann, M., Llucia-Pomares, D. & Hochkirch, A. (2013) A review of the Iberian Sphingonotini with description of two novel species (Orthoptera: Acrididae: Oedipodinae). Zoological Journal of the Linnean Society 168, 29.60.

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. Journal of Biogeography 41, 81-94.

Ivanova, N.V., Dewaard, J.R. & Hebert, P.D.N. (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes 6, 998–1002.

Karl, I., Schmitt, T. & Fischer, K. (2009) Genetic differentiation between alpine and lowland populations of a butterfly is related to PGI enzyme genotype. Ecography 32, 488-496.

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

Kudrna, O., Pennerstorfer, J. & Lux, K. (2015) Distribution Atlas of European Butterflies and Skippers. Peks, Schwanfeld.

Lê Cao, K.-A., Boitard, S. & Besse, P. (2011) Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. BMC Bioinformatics 12, 253.

Mallet, J. (2005) Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20, 229-237.

Mallet, J. (2007) Hybrid speciation. Nature 446, 279-283.

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.

Mittelbach, G.G. & Schemske, D.W. (2015) Ecological and evolutionary perspectives on community assembly. Trends in Ecology and Evolution 30, 241-247.

Mitteroecker, P. & Bookstein, F. (2011) Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. Journal of Evolutionary Biology 38, 100-114.

Nazari, V., Hagen, W.T. & Bozano, G.C. (2010) Molecular systematics and phylogeny of the Marbled Whites (Lepidoptera: Nymphidae, Satyrinae, Melanargia Meigen). Systematics Entomology 35, 132-147.

Pigot, A.L. & Tobias, J.A. (2015) Dispersal and the transition to sympatry in vertebrates. Proceedings of the Royal Society London B Biological Sciences 282, 20141929.

Rohlf, J. (2008) TPSDIG, Version 1.39. Department of Ecology and Evolution, State University of New York, Stony Brook, NY.

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

Schmitt, T., Habel, J.C., Zimmermann, M. & Müller, P. (2006) Genetic differentiation of the Marbled White butterfly, Melanargia galathea, accounts for glacial distribution patterns and postglacial range expansion in southeastern Europe. Molecular Ecology 15, 1889-1901.

Talavera, G., Lukhtanov, V.A., Rieppel, L., Pierce, N.E. & 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.

Toews, D.P., Mandic, M., Richards, J.G. & Irwin, D.E. (2014) Migration, mitochondria, and the Yellow-rumped warbler. Evolution 68, 241-255.

Tolman, T. & Lewington, R. (2008) Collins butterfly guide: the most complete field guide to the butterflies of Britain and Europe. Collins, London.

Tshikolovets, V. (2011) Butterflies of Europe and the Mediterranean area. Vadim Tshikolovets, Pardubice.

Vandewoestijne, S. & Van Dyck, H. (2011) Flight morphology along a latitudinal gradient in a butterfly: do geographic clines differ between agricultural and woodland landscapes? Ecography 34, 876-886.

Vodă, R., Dapporto, L., Dincă, V. & Vila, R. (2015a) Cryptic matters: overlooked species generate most butterfly beta-diversity. Ecography 38, 405-409.

Vodă, R., Dapporto, L., Dincă, V. & Vila, R. (2015b) Why do cryptic species tend not to co-occur? A case study on two cryptic pairs of butterflies. PLoS One 10, e0117802.

Wade, M.J. (2001) Evolution: infectious speciation. Nature 409, 675-677.

Warren, J.H. (1998) Wolbachia and speciation. In: (Howard, D. J., & Berlocher, S. H.) Endless forms: species and speciation, 245-260.

Waters, J.M. (2011) Competitive exclusion: phylogeography's 'elephant in the room'? Molecular Ecology 20, 4388–4394.

SUPPORTING INFORMATION

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

Appendix S1: Sampling localities for mtDNA analyses

Appendix S2: Sampling localities for allozyme analyses.

Appendix S3: Sampling localities for wing-vein analyses.

Appendix S4: Land-marks selected for wing-vein analyses.

Appendix S5: Sampling localities for genitalia analyses.

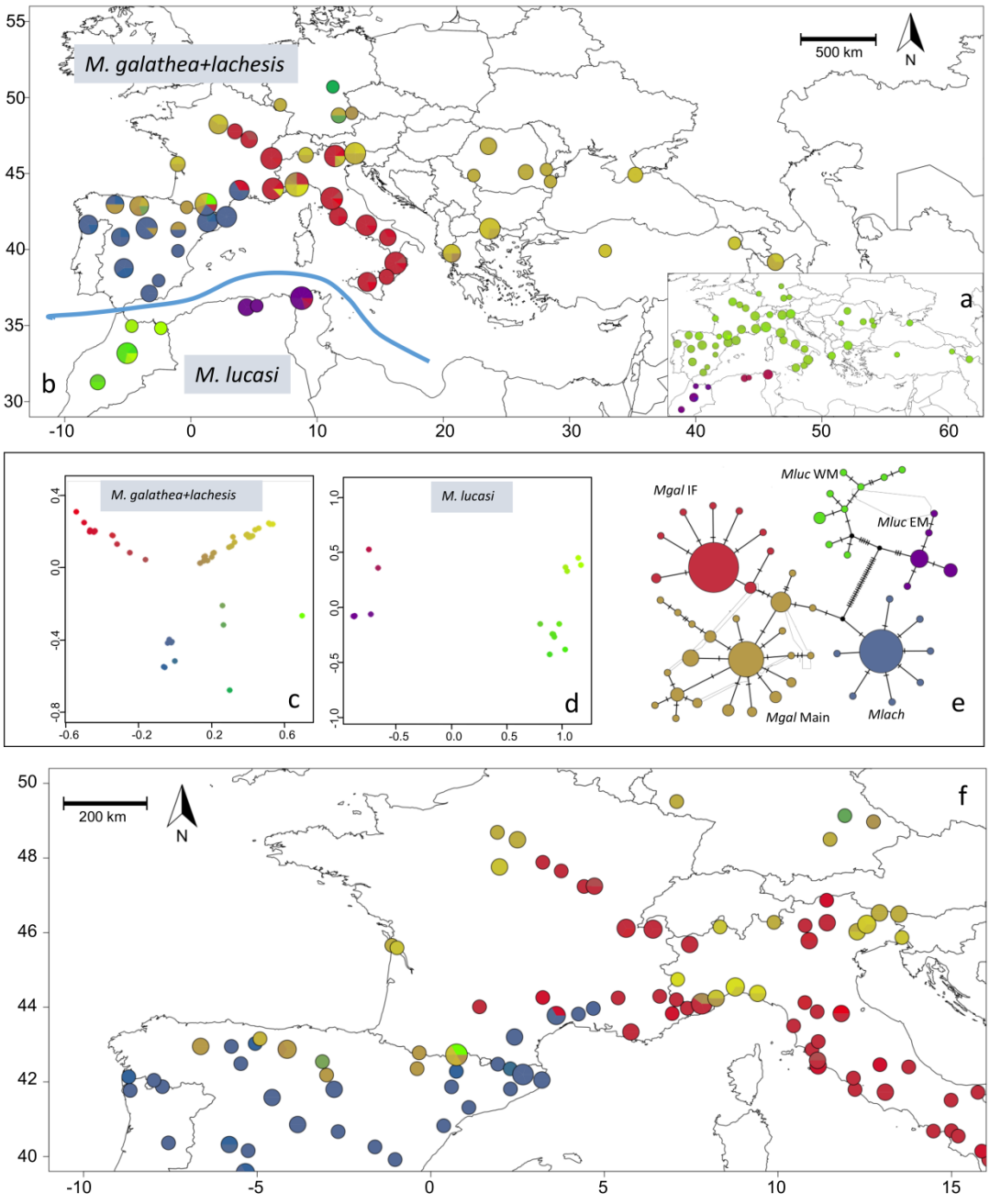

Figure 1: Spatial patterns of differentiation for the three species Melanargia galathea, M. lachesis and M. lucasi based on COI sequences over the western Palaearctic region, including North Africa. Colours indicate their similarity. a) PCoA configurations projected in a RGB space for all samples (and species) analysed; b-d) for populations assigned to specific taxonomic groups; e) haplotype network revealing remarkable genetic differentiation with clear geographic structure within each of the two main groups; f) with higher spatial resolution detecting three distinct distribution patterns.

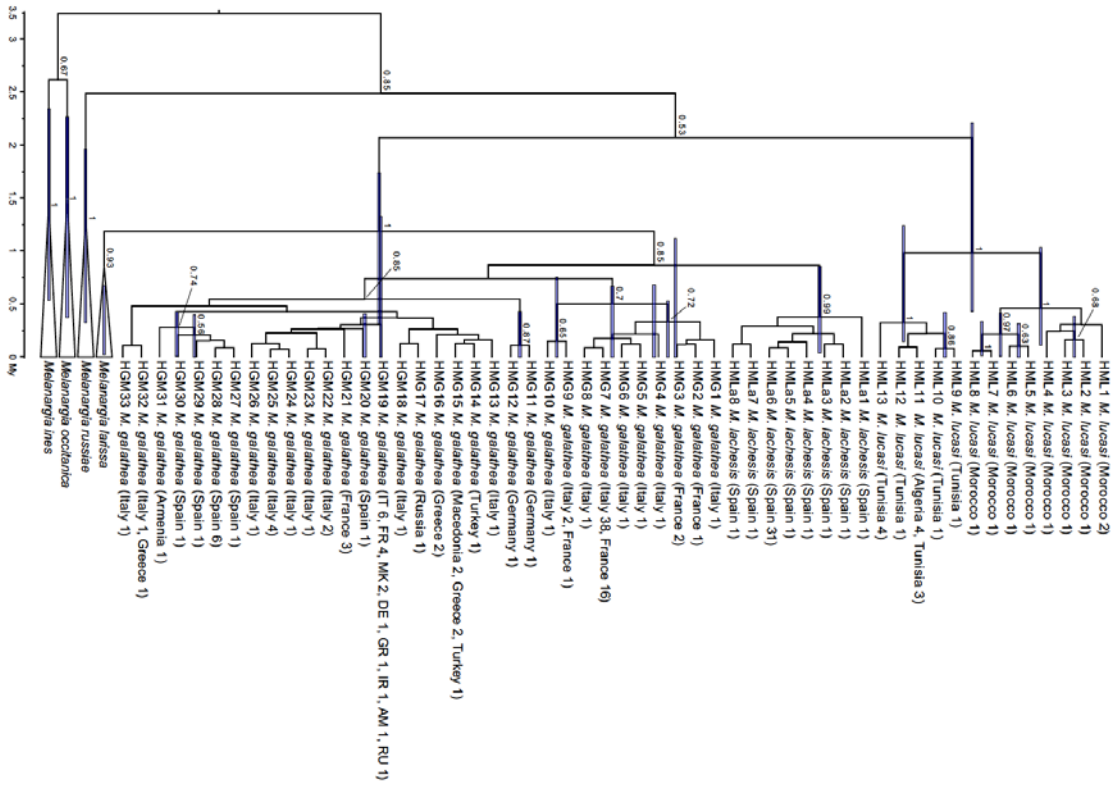

Figure 2: Phylogenetic tree with a similar topology to the one previously published for the genus Melanargia by Nazari et al. (2010).

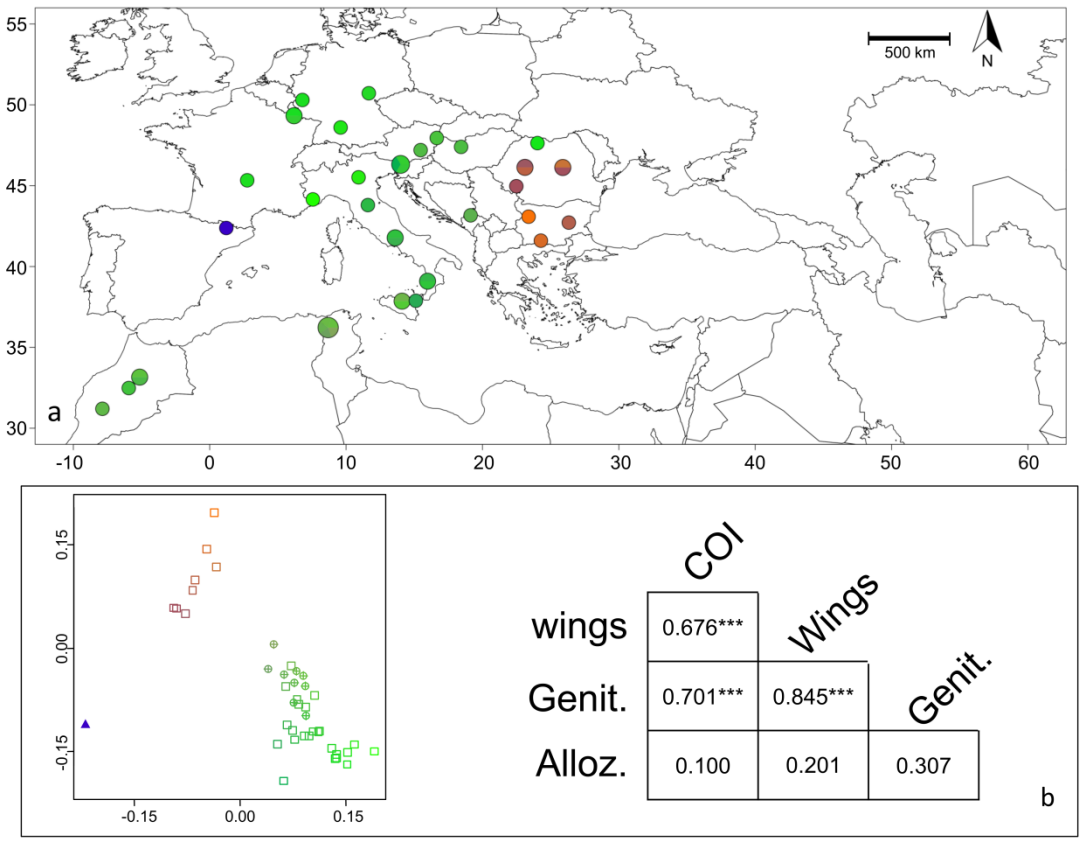

Figure 3: Spatial patterns of differentiation for the three species M. galathea, M. lachesis and M. lucasi based on allozyme polymorphisms; colours indicate their similarity. a) PCoA configurations projected in a RGB space; b) results from PCoA, c) inlet table with the results of a protest analysis incorporating all four characters analysed. ***: p ≤ 0.001.

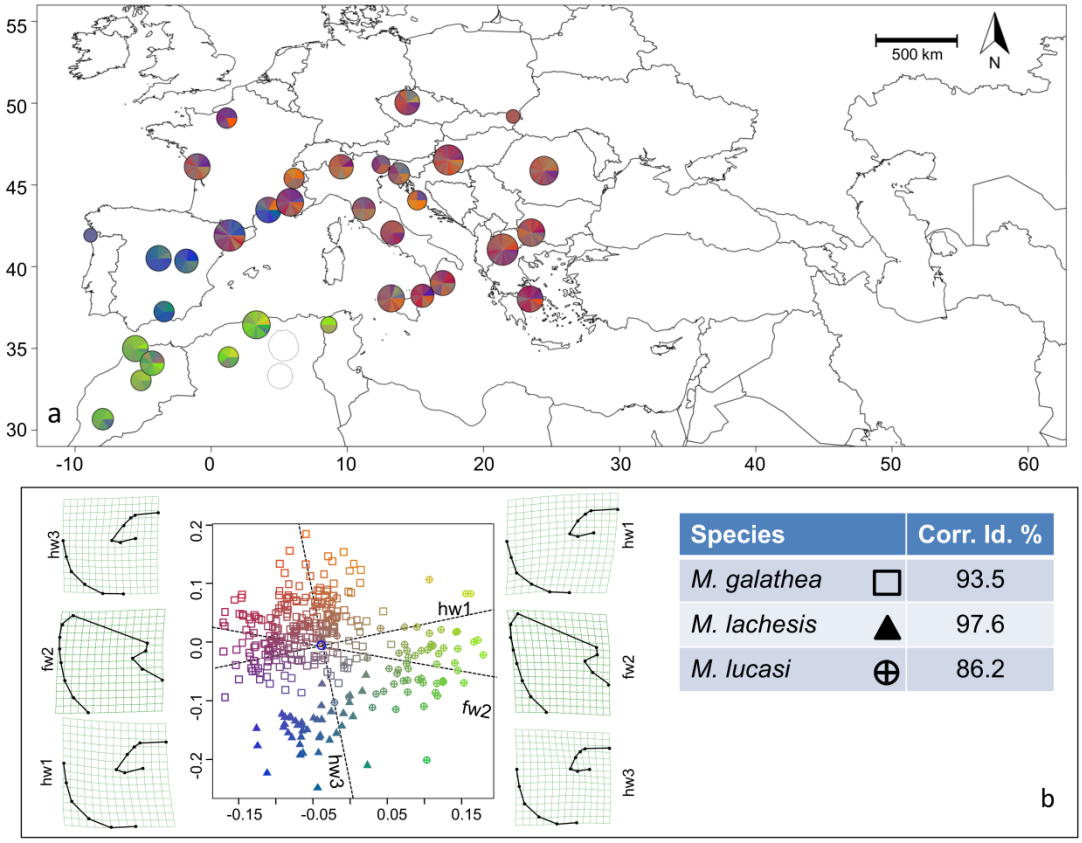

Figure 4: Spatial patterns of differentiation for the three species M. galathea, M. lachesis and M. lucasi based on wing-shape structures; colours indicate their similarity. a) Results from partial least square discriminant analysis (size of circles reflect sampling sizes) and b) PCs based on three relative warps; the inlet table presents results from Jack-knifing.

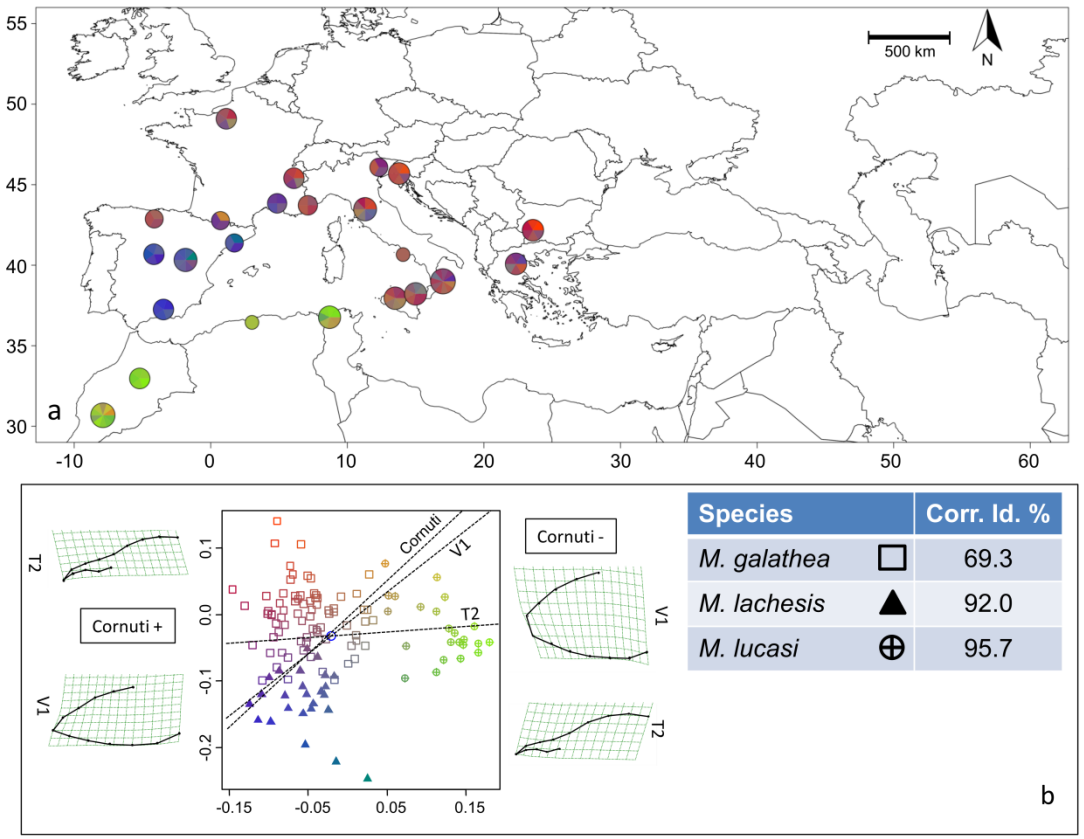

Figure 5: Spatial patterns of differentiation for the three species M. galathea, M. lachesis and M. lucasi based on genitalia structures; colours indicate their similarity. a) Results from partial least square discriminant analysis (size of circles reflect sampling sizes) and b) PCs based on three relative warps; the inlet table presents results from Jack-knifing.

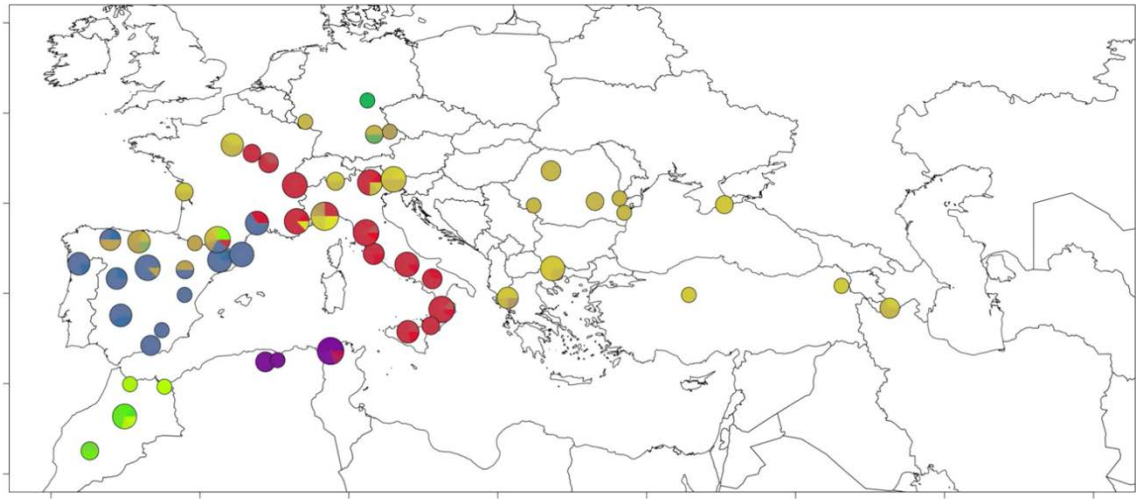

Figure 6: Spatial patterns of differentiation for the three species M. galathea, M. lachesis and M. lucasi based on COI sequences; colours indicate their similarity, using PCoA configurations projected in a RGB space for all samples (and species) analysed. Potential glacial refugial regions (R) are indicated with cycles, postglacial range expansions are shown by arrows.