Biological Journal of the Linnean Society, 2017, **XX**, 1–15. With 5 figures.

Mito-nuclear discordance helps to reveal the phylogeographic patterns of *Melitaea ornata* (Lepidoptera: Nymphalidae)

JÁNOS P. TÓTH[1,*], ZOLTÁN VARGA[2], RUDI VEROVNIK[3], NIKLAS WAHLBERG[4], ALEX VÁRADI[2] and JUDIT BERECZKI[2]

[1]*MTA-DE 'Lendület' Behavioural Ecology Research Group, 4032 Debrecen Egyetem tér 1, Hungary*
[2]*Department of Evolutionary Zoology and Human Biology, Institute of Biology and Ecology, University of Debrecen, 4032 Debrecen Egyetem tér 1, Hungary*
[3]*Department of Biology, Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, 1000 Ljubljana, Slovenia*
[4]*Department of Biology, Lund University, 223 62 Lund, Sweden*

Received 6 September 2016; revised 7 November 2016; accepted for publication 1 December 2016

Periodical changes of glacial and interglacial conditions have influenced the distribution of most living organisms and shape the separation of different genetic lineages significantly. We investigated the phylogeography of a nymphalid butterfly *Melitaea ornata*. Our main aim was to explore the existence, the origin, and the variability of different genetic lineages based on a multilevel approach. *M. ornata* and its close relatives (with a focus on *M. phoebe*) from the Palearctic were analysed based on five gene regions (COI, EF-1a, MDH, RPS5, and wingless) using Bayesian methods to infer the phylogeographic history. The DNA-based analyses have been complemented with species distribution modelling (SDM) and *Wolbachia* screening. The Bayesian inference analysis showed mito-nuclear discordance in *M. ornata*, which is split into an eastern and a western clade. Based on mitochondrial DNA , the western clade of *M. ornata* clusters together with *M. phoebe*, while the eastern clade is well-separated. In contrast to this, the combined nuDNA-based analysis revealed that *M. ornata* forms a monophyletic group which is clearly separated from *M. phoebe*. The timing of divergence analyses suggest that the split between *M. ornata* and *M. phoebe* is about 6 million years old based on both the COI and the concatenated nuclear genes. SDM predicted considerably larger area shifts for *M. phoebe* than for *M. ornata*. LGM refugia were predicted for both species to the Mediterranean peninsulas in Europe and several Middle-East and Asian localities. The prevalence of *Wolbachia* infection was 88.9% in *M. phoebe* and only 7% in *M. ornata*. Our results clearly indicate a lack of ongoing hybridization between *M. phoebe* and *M. ornata*, but argue for an ancient hybridization event in the Apennine Peninsula which strongly influenced the observed split between the two clades of *M. ornata*.

ADDITIONAL KEYWORDS: hybridization – Last Glacial Maximum – species distribution modelling – *Wolbachia*.

INTRODUCTION

The Earth's climate has shown a general cooling trend that began around 45 Mya in the Eocene with strong oscillations from the Miocene (Zachos, Dickens & Zeebe, 2008). Various records from the Quaternary show increasingly severe ice age cycles with shorter and shorter warm interglacials driven by the change of Earth's climate sensitivity to CO_2-based radiative forcing which was half as strong during the warm Pliocene as during the cold late Pleistocene epoch (0.8–0.01 Mya) (Lisiecki & Raymo, 2007; Martínez-Botí *et al.*, 2015). These periodical changes of glacial and interglacial conditions have influenced the distribution of most living organisms significantly. According to the classic refugia paradigm for Europe, the bulk of temperate species could survive in southern Europe during the glaciations, mostly in climatically buffered

*Corresponding author. E-mail: acutiformis@yahoo.com

parts of the Mediterranean peninsulas (de Lattin, 1967; Hewitt, 1996; Schmitt, 2007), although this paradigm has been greatly refined by the growing amount of new information. Past distributions of species are essentially inferred from four types of evidence (Tzedakis, Emerson & Hewitt, 2013): (1) plant and animal macrofossils; (2) pollen records; (3) genetic data; and (4) species distribution models.

The most comprehensive coverage comes from plant macrofossils and pollen spectra. These indicate that many species respond individually to climatic oscillations (Bhagwat & Willis, 2008; Stewart *et al.*, 2010) depending on differential responses to various ecological factors. These species-specific reactions produce a variety of distributional patterns by speeding up or slowing down range expansion and contraction, thus they can result in a significant difference in species composition of communities that could be essentially different from those observed today (Schmitt & Varga, 2012). Obviously, the change of vegetation has also an effect on the animal assemblages, although the fossil record is poor or not available for the majority of animal species, especially for insects. For these fossil-deficient groups, molecular methods and distribution modelling techniques could help us to infer their biogeographic history. The disjunct glacial distribution patterns have regularly resulted in the separation of different genetic lineages in the three major Mediterranean peninsulas, Anatolia (Comes & Kadereit, 1998; Taberlet *et al.*, 1998; Hewitt, 1999, 2000; Schmitt, 2007), North Africa, large Mediterranean islands, and also some microrefugial areas (Provan & Bennett, 2008; Dapporto, 2010; Husemann *et al.*, 2014).

Mitochondrial DNA (mtDNA) proved to be suitable for barcoding animal life and highly informative in phylogenetic and phylogeographic studies (Hebert, Ratnasingham & de Waard, 2003). Nevertheless, the use of nuclear (nuDNA) markers in addition to mtDNA improves the power of phylogenetic and phylogeographic hypothesis tests significantly and highlights the limitations of studies using only mtDNA markers. Results based solely on mtDNA could be very different from those which are inferred from nuDNA (e.g. Wahlberg *et al.*, 2009); a phenomenon known as mito-nuclear discordance. According to Toews & Brelsford (2012) the most frequent explanations of mito-nuclear discordance are incomplete lineage sorting, adaptive introgression of mtDNA, demographic disparities and sex-biased asymmetries, hybrid zone movement, human introductions, and *Wolbachia* infection.

Recently, it has become clear that patterns of mtDNA inheritance could be highly manipulated by the maternally inherited, intracellular bacteria genus *Wolbachia*. This microorganism could induce 'two barcodes – one species' (Kodandaramaiah *et al.*, 2013) or its opposite, 'one barcode – two species' phenomenon (Jiggins, 2003). In these cases, the results of mtDNA-based analyses could be misleading. Thus, a phylogeographic or a phylogenetic study using only mtDNA is unreliable without *Wolbachia* screening.

Multiple gene analyses are strongly recommended not only for phylogeography but also for phylogenetic studies in general (e.g. Wiemers & Fiedler, 2007; Wahlberg *et al.*, 2009). Moreover, the combined use of different data sets and methods – for example multiple gene analyses complemented with *Wolbachia* screening, morphometric data, and species distribution modelling (SDM) predictions – enables us to better understand how the various factors together shape the evolutionary relationships among taxa.

SDM could be useful to estimate the potential distribution at different time scales, thus it could provide insight into area dynamics even for fossil-deficient species. Therefore, SDM has become an important tool for historical biogeography usually combined with molecular methods or morphometry (e.g. Habel *et al.*, 2011; Schorr *et al.*, 2012; Rajaei *et al.*, 2013; Zinetti *et al.*, 2013; Tóth, Varga & Bereczki, 2016).

The demand for such multilevel approach has arisen in the case of the *Melitaea phoebe* species group, which has been studied intensively in the last few years. It became clear that a cryptic species, namely *M. ornata* Christoph, 1893, was concealed within the morphologically similar species *M. phoebe* from which it was first distinguished based on larval morphology (Russell *et al.*, 2005; Varga, Szabó & Kozma, 2005).

M. phoebe ([Denis & Schiffermüller, 1775]) is distributed widely in the Palearctic, although it seems less abundant under Mediterranean climate conditions. *M. ornata* also has a relatively large distribution but the most northern localities are found in the Carpathian Basin and in the Volgograd Region in Russia (Kuznetsov & Stradomsky, 2010). Owing to the application of geometric morphometrics and phylogenetic methods, two more *Melitaea* species have been identified in this group, namely *M. telona* Fruhstorfer, 1908 and *M. abbas* Gross & Ebert, 1975, and the relationships between these taxa have become much clearer through a series of recent studies (Leneveu, Chichvarkhin & Wahlberg, 2009; Tóth & Varga, 2011; Tóth *et al.*, 2014, 2016). One of the most surprising discoveries was that *M. telona* from the Levant region appears to be a distinct species also based on molecular data (Tóth *et al.*, 2014). Additionally, a more focused study on Iranian *Melitaea* specimens (Tóth, *et al.*, 2016) from several localities showed that *M. ornata* is a rare species in this region and found exclusively in the north-eastern part of the country, while the widely distributed sister species in Iran is *M. abbas*.

The distribution of *M. ornata* at different time scales has already been studied using the Maximum Entropy (MaxEnt) SDM method (Tóth *et al.*, 2013). Combining the results of genitalia morphometrics and SDM, refugial

areas were hypothesized to the Apennine Peninsula, southern Balkans, Asia Minor, Levant, and northern Iran. The authors suggested that there are four separated morphotypes in *M. ornata*: the western type survived mainly in the Apennine Peninsula and colonized the Carpathian Basin after the Last Glacial Maximum (LGM), the eastern type survived in the southern Balkans and colonized Southeast Europe, (including southern Russia) and Kazakhstan. Additionally, two narrowly distributed morphotypes were described from northern Iran and from the Levant region.

Here we investigate the phylogeography of *M. ornata* based on one mitochondrial (COI) and four nuclear genes (EF-1a, MDH, wingless, and RPS5). Our main goals were (1) to reveal the potential refugia of the studied species and reconstruct the possible post-glacial re-colonization routes based on molecular evidence, (2) to explore the existence, the origin, and the variability of different genetic lineages, and (3) to compare these results to those which were obtained from genitalia morphometry and SDM.

MATERIAL AND METHODS

DNA STUDIES

We sampled 77 specimens of *M. phoebe* and *M. ornata* from the Palearctic, mostly from the Ponto-Mediterranean region. The specimens were identified based on wing pattern characteristics following the identification key in Tóth & Varga (2011) and partly also based on genitalia slides. The sequences of their close relatives, that is *M. punica*, *M. telona*, *M. abbas*, *M. tangigharuensis*, *M. sarvistana*, *M. aetherie*, *M. arduinna*, *M. collina*, and *M. consulis* as well as distantly related *M. trivia*, *M. romanovi*, and *M. cinxia*, were downloaded from NCBI database (Geer *et al.*, 2010) and from the NSG DNA sequence database VoSeq (Peña & Malm, 2012) or re-used from our previous studies (Table S1). Our sampling covered the whole known range of *M. ornata*, although the European part has been sampled at a much better resolution than the Asian.

DNA was extracted from the head or the proximal end of the abdomen following the protocol in Bereczki *et al.* (2014). The mitochondrial cytochrome c oxidase subunit I (COI), the nuclear elongation factor 1α (EF-1α), malate dehydrogenase (MDH), ribosomal protein S5 (RpS5), and wingless (wg) were amplified by specific primers modified at their 5′-end to include the universal sequencing primer T7 promoter (Wahlberg & Wheat, 2008). Amplification from 1 μL of DNA extracts was carried out in 25 μL final reaction volumes containing 5× PCR buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 mg/mL BSA, 0.02 units/μL of Taq DNA polymerase (Phusion Hot Start II High-Fidelity, Thermo Scientific), and 0.3 μM of each primer. Amplification was carried out in an ABI Veriti thermal cycler programmed for initial denaturation for 2 min at 98 °C; 40 cycles of 10 s at 98 °C, 30 s at the locus specific annealing temperature, 1 min 30 s at 72 °C; and final elongation of 10 min at 72 °C. The success of PCR amplification was checked by running 2 μL of product on 1% agarose gels stained with GelRed Nucleic Acid Stain (Biotium Inc., Fremont, CA, USA). PCR products were sequenced by the commercial service provider Macrogen Inc. (Seoul, South Korea).

The presence of *Wolbachia* was checked by polymerase chain reaction (PCR) by the amplification of the highly conservative 16S ribosomal RNA gene with *Wolbachia* specific primers W-Spec of Werren & Windsor (2000), following the protocols described by these authors. We used positive (definitely infected samples) and negative controls (master mix without any DNA sample) in each reaction. The success of PCR amplification was checked by running 2 μL of product on 1% agarose gels. *Wolbachia* strain identification was carried out by the amplification of the highly variable *Wolbachia* surface protein (WSP) following the instructions of *Wolbachia* MLST database (http://pubmlst.org/wolbachia/). After sequencing we defined the strains using the reference sequences of *Wolbachia* MLST database.

DNA sequences were edited and revised manually by Chromas Lite v. 2.01, then aligned using 'decipher' package (Wright, 2016) using R. Two datasets with differing taxon sampling were analysed, one using MrBayes 3.2.6 (Ronquist *et al.*, 2012) and the other using BEAST 2.3.2 (Bouckaert *et al.*, 2014).

The MrBayes analyses included all sequences from *M. ornata*, *M. phoebe*, and *M. telona*, with *M. tangiguharensis* as the outgroup. In this analysis we were mainly interested in the intraspecific structure of the data focusing on *M. phoebe* and *M. ornata*. The MrBayes analyses were conducted on single-gene, nuclear genes only, and all-gene datasets. The multiple gene datasets were partitioned by genes. The different models of molecular evolution were sampled for each gene (both single and combined data) by using the model-jumping feature through the command 'nst=mixed' and site rate heterogeneity was taken into account with the gamma + invariant sites parameters. Two independent MCMC runs, each with four simultaneous chains (one cold and three heated) for each analysis, were run for 10 million generations and the sampling of trees and parameters was set to every 1000 generations. Convergence of the two runs was determined by the stationary distribution plot of the log-likelihood values against number of generations and confirmed by the average standard deviation of split frequencies which were lower than 0.05 in all cases. We discarded the first 2 500 000 generations as burn-in and trees were summarized under the 50% majority rule method.

Table 1. The parameters of DNA variability. *N*, the number of sequences; *V*, the number of variable sites; PI, the number of parsimony informative sites; *h*, the number of haplotypes; HD, haplotype (gene) diversity; π, nucleotide diversity (per site); θ_w, theta (per site) from the number of variable sites; neutrality tests, Tajima' D, Fu & Li's D, significance: **P* < 0.05; ***P* < 0.02.

	Mitochondrial	Nuclear (2263 bp)			
	COI (1328 bp)	**EF1α** (901 bp)	**MDH** (596 bp)	**RPS5** (427 bp)	**Wingless** (339 bp)
Melitaea ornata – East					
N	26	52	48	54	46
V/PI	28/12	32/27	5/5	15/11	8/7
h/HD	16/0.954	22/0.938	5/0.738	16/0.763	11/0.881
Π	0.00342	0.00688	0.00201	0.00641	0.00599
θ_w	0.00553	0.00811	0.00189	0.00771	0.00604
Tajima' D	−1.40703	−0.50276	0.15223	−0.51012	−0.02479
Fu & Li's D	−1.85205	0.36308	1.10043	−0.30340	0.06040
Melitaea ornata – West					
N	19	36	32	34	34
V/PI	7/6	11/9	5/5	12/7	4/4
h/HD	6/0.778	12/0.827	6/0.649	11/0.843	3/0.522
Π	0.00137	0.00404	0.00241	0.00488	0.00578
θ_w	0.00151	0.00321	0.00208	0.00687	0.00289
Tajima' D	−0.31656	0.81189	0.41489	−0.92197	2.49110*
Fu & Li's D	0.68695	−0.01164	1.13769	−0.96221	1.04416
Melitaea ornata –Total					
N	45	88	80	88	80
V/PI	59/44	35/31	7/7	22/16	9/9
h/HD	22/0.946	31/0.943	8/0.812	25/0.805	12/0.862
Π	0.01366	0.00663	0.00250	0.00616	0.00760
θ_w	0.01033	0.00835	0.00237	0.01067	0.00596
Tajima' D	1.13946	−0.65277	0.13440	−1.26540	0.72993
Fu & Li's D	−0.13674	0.16352	1.20971	−0.93464	0.70648
Melitaea phoebe					
N	32	62	52	60	52
V/PI	18/4	22/15	7/4	33/28	12/11
h/HD	12/0.794	24/0.817	9/0.801	30/0.920	15/0.843
Π	0.00142	0.00263	0.00233	0.01039	0.00668
θ_w	0.00356	0.00544	0.00297	0.01758	0.00783
Tajima' D	−2.04207*	−1.62299	−0.58443	−1.34420	−0.43072
Fu & Li's D	−3.26510**	−0.71717	−0.85941	0.66209	0.93179

BEAST 2.3.2 (Bouckaert *et al.*, 2014) was used to estimate the age of divergence between taxa. The BEAST analyses included representatives of major lineages of *M. ornata* and *M. phoebe*, as well as representatives of all related species (based on Tóth *et al.*, 2014): *M. punica*, *M. telona*, *M. abbas*, *M. tangigharuensis*, *M. sarvistana*, *M. aetherie*, *M. arduinna*, *M. collina*, and *M. consulis*, including outgroups *M. trivia*, *M. romanovi*, and *M. cinxia*. The five gene dataset was assigned as two partitions (one for the mitochondrial gene and one for the four nuclear genes combined). The two partitions were unlinked for substitution models, clock models, and topology. 'bModelTest' package was used for Bayesian site model selection. We applied the lognormal relaxed clock model, and the Yule prior. Two independent analyses were run for 10 million generations and the sampling of trees and parameters were sampled every 1000 generations. As in Leneveu *et al.* (2009), a secondary calibration point was selected from Wahlberg (2006) to calibrate the age of the divergence between *M. cinxia* and the rest of the taxa included. Thus, the calibration point was set to 21 Mya using a normal distribution with ±3 My SD). The independent runs were combined with LogCombiner using a 25% burn-in. The summarized trees from MrBayes and from BEAST analysis

were plotted using FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

Additionally, haplotype networks were constructed using the R computing environment (R Core Team, 2013) with the 'Pegas' package v.0.8 (Paradis, 2010). One of the advantages of Pegas is that it is able to extract haplotypes taking into account base ambiguities. We performed the haplotype analysis gene by gene but also for concatenated nuclear genes. The distribution of haplotypes in geographical space was visualized using QGIS 2.14 (QGIS Development Team, 2015).

We also estimated the level of genetic diversity by the following parameters: the number of variable and informative sites (*V* and PI), haplotype number and diversity (*h* and HD) (Librado & Rozas, 2009), nucleotide diversity (π) (Nei & Li, 1979; Nei, 1987; Tajima, 1983), and the number of segregating sites (θ_w) (Watterson, 1975; Nei, 1987). In cases of nuclear genes, PHASE haplotype reconstruction option was used because of the presence of ambiguous sites (Stephens *et al.*, 2001; Stephens & Donnelly, 2003). Neutrality tests were performed on all genes to determine the departure from the neutral model of molecular evolution. Under neutrality, both Tajima's D and Fu & Li's D are expected to be zero. Positive D values indicate an excess of intermediate-frequency variants and can be due to the operation of natural selection. In contrast, a value significantly less than zero indicates a higher-than-expected number of low-frequency variants and might be the consequence of a recent selective sweep or processes such as background selection (Tajima, 1989; Aris-Brosou & Excoffier, 1996). The parameters of gene diversity and neutrality tests were calculated using the program DnaSP v. 5.10.01 (Librado & Rozas, 2009).

Species distribution modelling

We used the MaxEnt modelling method to predict the potential distribution of *M. ornata* and *M. phoebe* using BIOCLIM variables (Busby, 1991). MaxEnt is a widely used method for predicting species distributions using presence-only data (Phillips, Dudík & Schapire, 2004; Warren & Seifert, 2010). MaxEnt's predictive performance is consistently competitive with the highest performing methods (Elith *et al.*, 2011).

Non-overlapping 181 presence points were used for modelling the distribution of *M. phoebe* and 204 for modelling *M. ornata*. In the case of *M. ornata*, we used the same presence point dataset as in Tóth et al. (2013) with the necessary corrections, for example taking into consideration the recent results on separate species status of *M. telona* and *M. abbas* (see Introduction). Presence data for *M. phoebe* were collated mainly from the Hungarian Natural History Museum and the Bavarian State Collection of Zoology, Munich. Additionally, some data were also used from the literature (Tshikolovets, Bidzilya & Golovushkin, 2002; Tshikolovets, Yakovlev & Kosterin, 2009).

The climate variables were downloaded from WorldClim database (www.worldclim.com). Although MaxEnt is more robust in controlling for correlations between variables than stepwise regression (Elith *et al.*, 2011), strongly correlated variables ($r > 0.75$) are recommended to be excluded from the analysis (see Elith, Kearney & Phillips, 2010; Stohlgren *et al.*, 2010). ENMtools 1.4 was used to calculate the level of correlations (Warren, Glor & Turelli, 2010). To assess which predictors provide the most useful information by itself, we applied the jackknife test using MaxEnt. Jackknife and the correlation tests were considered during variable selection.

For *M. ornata* we used the same variable set as presented in Tóth et al. (2013): altitude, Bio8 (mean temperature of the wettest quarter), Bio10 (mean temperature of warmest quarter), Bio15 (precipitation seasonality), Bio16 (precipitation of wettest quarter), and Bio18 (precipitation of the warmest quarter). For *M. phoebe* five variables were selected: altitude, Bio7 (temperature annual range), Bio10 (mean temperature of warmest quarter), Bio19 (precipitation of coldest quarter), Bio17 (precipitation of driest quarter), and Bio15 (precipitation seasonality). The obvious sampling bias in presence points was counterbalanced using bias grids. These grids were produced by deriving a Gaussian kernel density map of the occurrence locations (for more detail see Elith *et al.*, 2010).

The discrimination ability of the model was evaluated by area under the curve (AUC) metric. The value of AUC varies between 0.0 and 1.0, where 1.0 is considered a perfect prediction and 0.5 or less is considered to be a prediction that is not better than random (Fielding & Bell, 1997; Franklin & Miller, 2009). The distribution models were projected back to the LGM, that is ~21 000 years before present. For the projections we used the predictions of two different global circulation models (MIROC-ESM and CCSM).

The results were visualized on a binary presence (1) absence (0) raster using the 10 percentile training presence threshold rule. To evaluate the area dynamics of the studied species, we used these binary rasters for current climate and the LGM scenarios. The presence values for the LGM have been changed from 1 to 2 followed by grid overlaying which results in four possible values for each cell: (1) where the species potentially occurred during the LGM but currently does not; (2) areas outside of the realized niche: areas that are suitable neither under current conditions nor under LGM conditions; (3) areas where the species could potentially occur in both present and LGM climates; and (4) areas where the species potentially occurs

currently, but which were not suitable during the LGM (for more details of the methodology see Scheldeman & Zonneveld, 2010).

RESULTS

In this study, 77 specimens of *M. phoebe* and *M. ornata* were analysed on a Eurasian scale (Table S1). Based on the results of DNA studies, all the specimens were identified correctly. The final dataset contained 3654 bp of concatenated DNA sequence from one mitochondrial (COI) and four nuclear genes (EF-1a, MDH, RPS5, and wingless) with 242 variable (6.7%) and 121 parsimony informative sites (3.4%).

The Bayesian inference analysis using MrBayes showed mito-nuclear discordance. Based on COI, *M. ornata* splits into two groups. These two groups show a strong geographical pattern, thus we name them the 'western clade' and 'eastern clade'. If we use only mtDNA, the western clade of *M. ornata* clusters together with *M. phoebe* specimens, while the eastern clade is well-separated from *M. phoebe* (Fig. 1). In contrast, the combined nuclear DNA-based analysis shows that all *M. ornata* specimens form a monophyletic group which is clearly separate from *M. phoebe* (Fig. 1).

The maximum clade credibility trees obtained from the BEAST analyses suggest that the split between *M. phoebe* and *M. ornata* lineages is about 6 million years old (95% credibility interval: 3.3–9.6 Mya) based on both the COI and the concatenated nuclear genes (EF-1a, MDH, RpS5, and wingless) (Fig. 2).

Haplotype networks

The different genes show very different levels of variability as is expected based on the length and the different evolutionary rates of each region. The most variable is COI for which 12 *M. phoebe* and 21 *M. ornata* haplotypes have been recognized. The most frequent *M. phoebe* haplotype is found both in *M. phoebe* and in *M. ornata* specimens, but only this haplotype was shared between the two species, all the remaining haplotypes are species-specific found either only in *M. phoebe* or only in *M. ornata* specimens. The distribution of the haplotypes shows a strong geographical pattern. In the case of *M. ornata*, the putatively hybrid haplotype occurs in southern Italy, Croatia, Bosnia–Herzegovina, and Hungary but it seems to be more widely distributed in *M. phoebe* as it is found in all sample sites east of the Alps. Although the haplotypes of the western clade of *M. ornata* show only few mutational differences compared to the *M. phoebe* haplotypes, we can recognize four distinct groups which are localized in four geographic regions: Sicily, southern Italy, Slovenia, and Hungary. All of them are characterized by unique haplotypes in our dataset. The eastern clade of *M. ornata* shows a high level of haplotype diversity based on COI which is highly correlated with geographic distribution, namely both the northern and the southern part of the Balkans, western Anatolia, Volgograd region, and eastern Kazakhstan have unique haplotypes (Fig. S1).

In the case of *M. phoebe*, the haplotypes also show East–West separation. The western clade includes the samples from northern Iberia, Apennine Peninsula, southern France, and Slovenia, while the eastern clade comprises the rest of the samples. The Iranian and Kyrgyzian samples represent unique haplotypes.

In contrast to the COI, *M. phoebe* and *M. ornata* do not share haplotypes across the nuclear gene regions. The Pegas analysis extracted 18 *M. ornata* and 14 *M. phoebe* haplotypes whose distribution shows strong geographical pattern which also clearly supports the existence of the eastern and the western clades (Fig. S2).

Wolbachia infection

Wolbachia screening was carried out for 70 specimens. The infection pattern was strikingly different in the two *Melitaea* species. The prevalence of *Wolbachia* infection was 88.9% in *M. phoebe* (24 infected specimens out of 27) and 7% in *M. ornata* (only 3 infected individuals out of 43). The infected *M. ornata* specimens were found in southern Italy, Hungary, and Macedonia (Fig. 3).

The most frequent *Wolbachia* surface protein (WSP) allele was the No. 694 while the allele No. 703 was represented only in a single *M. phoebe* individual (Table S1). The WSP sequences originating from two *M. ornata* specimens proved to be identical with those from *M. phoebe* (allele No. 702). Unfortunately, we failed to sequence WSP from the infected Hungarian *M. ornata* specimen.

An obvious geographical pattern is apparent in the WSP allele distribution (Fig. 4). The allele 694 is distributed in the Apennine Peninsula and also found in two individuals from the north-western Balkans, while the allele 702 dominates the eastern part of the distribution area of *M. phoebe*. The allele 702 is also found in two specimens of *M. ornata* from southern Italy (western clade) and from Macedonia (eastern clade). The allele 703 is found in a single *M. phoebe* specimen from Macedonia.

Genetic diversity of *M. ornata* and *M. phoebe*

The final COI alignment contained 77 sequences with a total length of 1328 bases out of which 66 (5.3%) sites

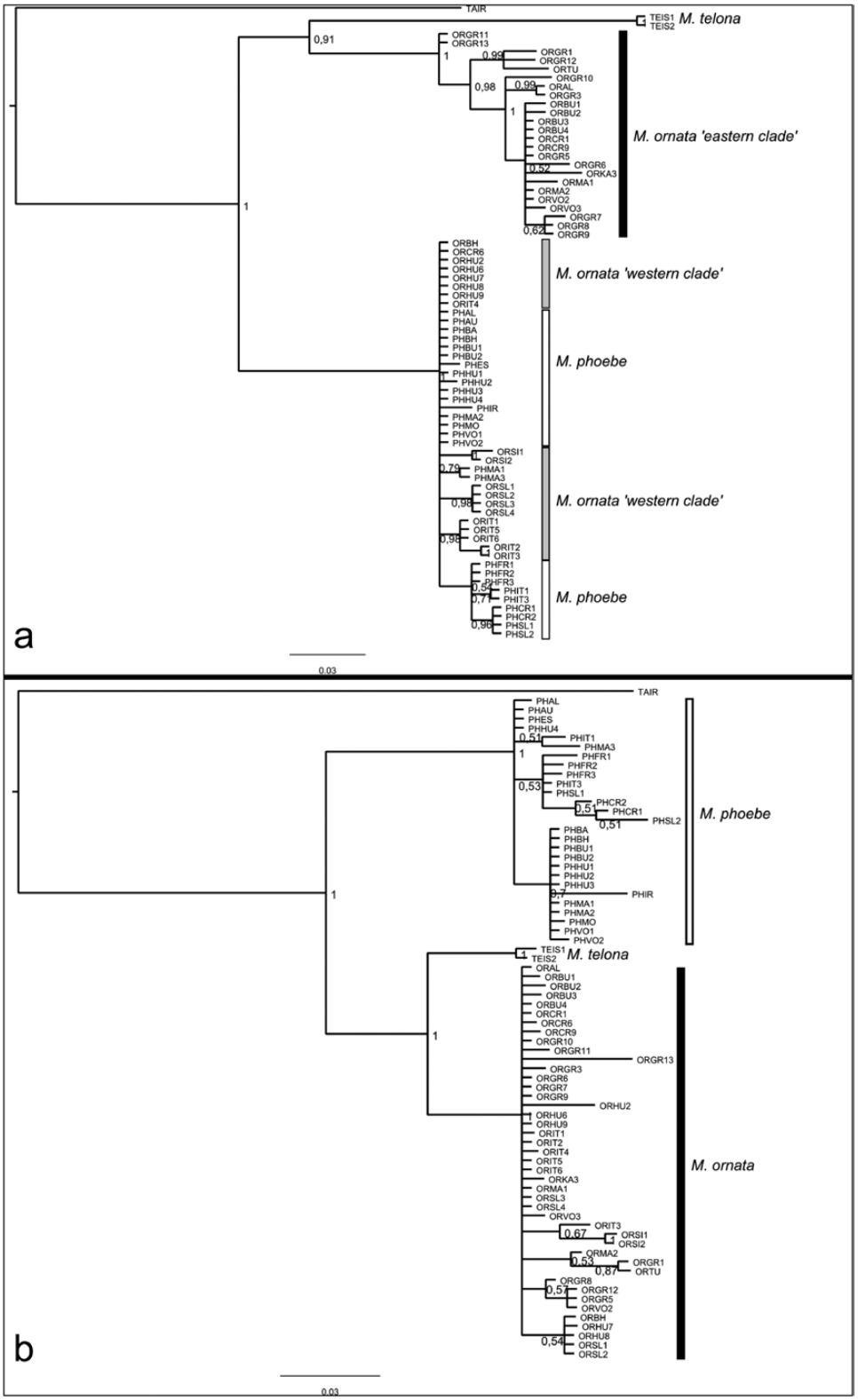

Figure 1. Bayesian inference trees based on (A) COI and (B) combined nuclear genes. *Melitaea tangigharuensis* (TAIR) used as outgroup.

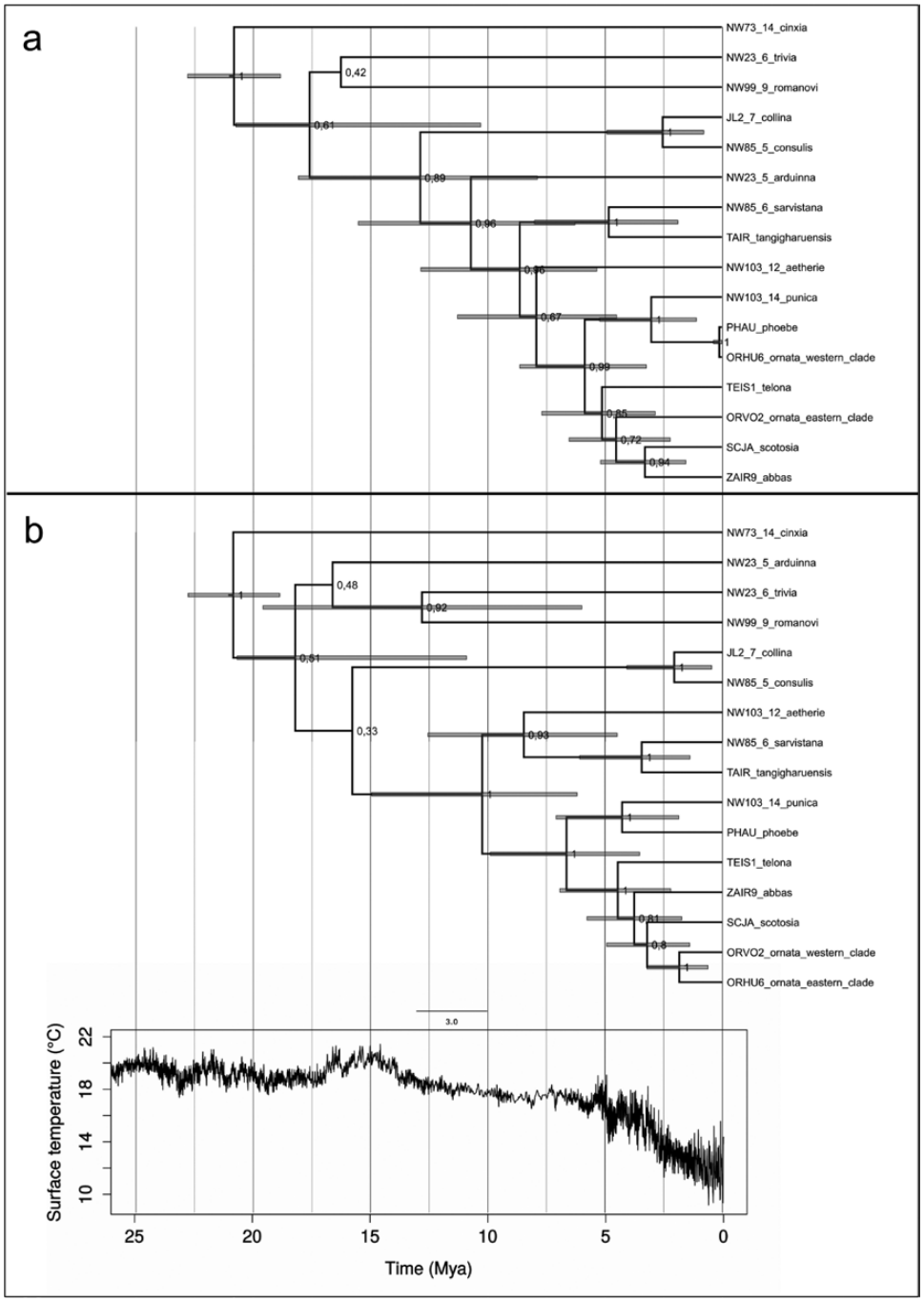

Figure 2. Maximum clade credibility trees obtained in BEAST (A) COI and (B) nuclear. Node bars representing the age 95% credible intervals and posterior probabilities of each clade labelled on the branches. Only one specimen per species (or clade) was used in this analysis. Surface temperature from Zachos *et al.* (2008).

were variable and 44 (3.5%) were parsimony informative. The number of these sites was proportionally higher in *M. ornata* than in *M. phoebe*. Similarly, nucleotide diversity and the number of segregating sites were greater by one order of magnitude in *M. ornata* than in *M. phoebe*. In total, DnaSP revealed 32 unique haplotypes out of which 21 were found in *M. ornata*, 10 were found in *M. phoebe*, and one mitotype was shared between the two species. Significant deviation from the neutral model of molecular evolution was established only in *M. phoebe* where both Tajima's D and Fu & Li's D were negative (Table 1).

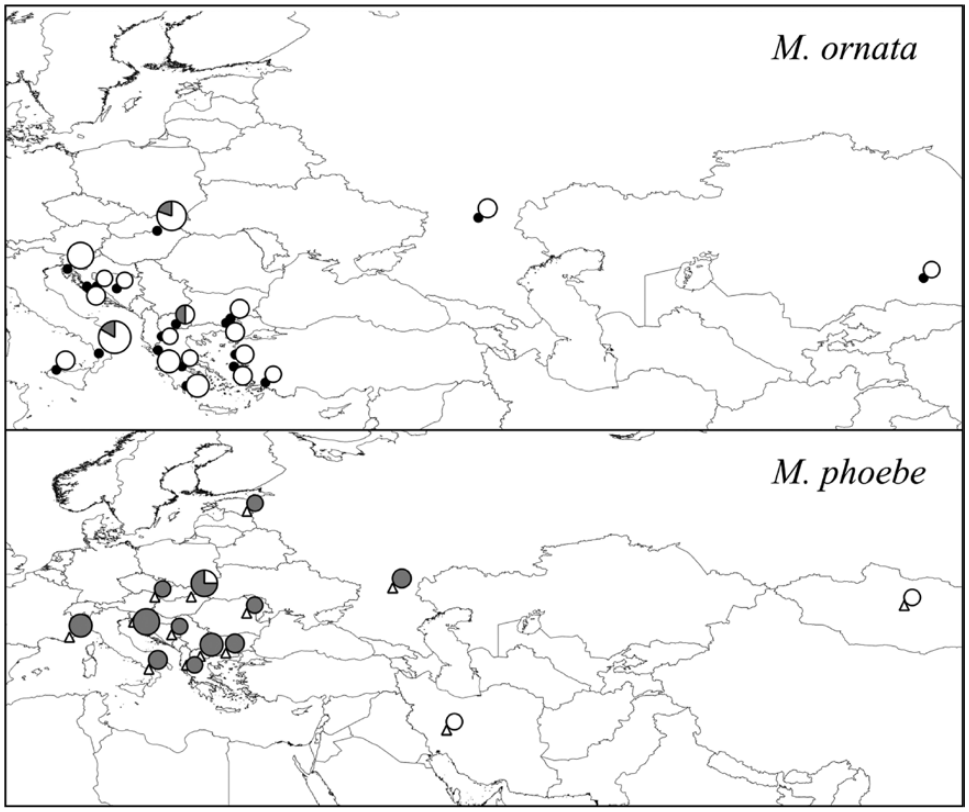

Figure 3. The results of *Wolbachia* presence tests are presented in georeferenced pie charts for *M. phoebe* (triangle) and *M. ornata* (dot). Grey colour indicates *Wolbachia* presence. The level of *Wolbachia* infection is much higher in *M. phoebe* than in *M. ornata*.

Altogether 2263 bases were analysed from four different nuclear gene regions (EF-1α, MDH, RPS5, and wingless) in 75 individuals. The diversity of these DNA fragments varied, and in general, EF-1α and RPS5 showed higher variability than the other two nuclear genes. Neutrality tests indicated significant departure from the neutral model only in the western clade of *M. ornata* where Tajima's D was positive (Table 1).

Species distribution modelling

The MaxEnt models yielded a good fit for the known distribution of *M. phoebe* (AUC = 0.871, SD = 0.049) and *M. ornata* (AUC = 0.919, SD = 0.039). The predictions of the species distribution models showed very different area dynamics for the two studied species (Fig. 5). Considerably larger area shifts were predicted for *M. phoebe* than for *M. ornata*. During the LGM, significant area retraction was predicted for both species but in the case of *M. ornata* the range predicted for the present proved to be clearly smaller than in *M. phoebe*, which could colonize more northern localities. LGM refugia were predicted for both species in the Mediterranean peninsulas (Iberia, the Apennines, and the Balkans), in Europe, and furthermore in North Africa, Anatolia, the Levant region, Elburs Mts., Zagros Mts, and Central Asia. Additionally, refugial areas were predicted for *M. phoebe* in Eastern Asia.

DISCUSSION

In this study, we have inferred the phylogeographic history of *M. ornata* as well as shed some light on the phylogeography of *M. phoebe*. Both mtDNA and nuDNA confirmed a split between the western and the eastern clade of *M. ornata*. It was however surprising that the western clade of *M. ornata* showed a high level of genetic similarity with *M. phoebe* based on mitochondrial DNA (COI). In contrast, when we examine the nuclear DNA (EF-1a, MDH, RPS5, and wingless) all *M. ornata* specimens were clearly separated from *M. phoebe* specimens forming reciprocally monophyletic groups, while still supporting the western/eastern split within both species.

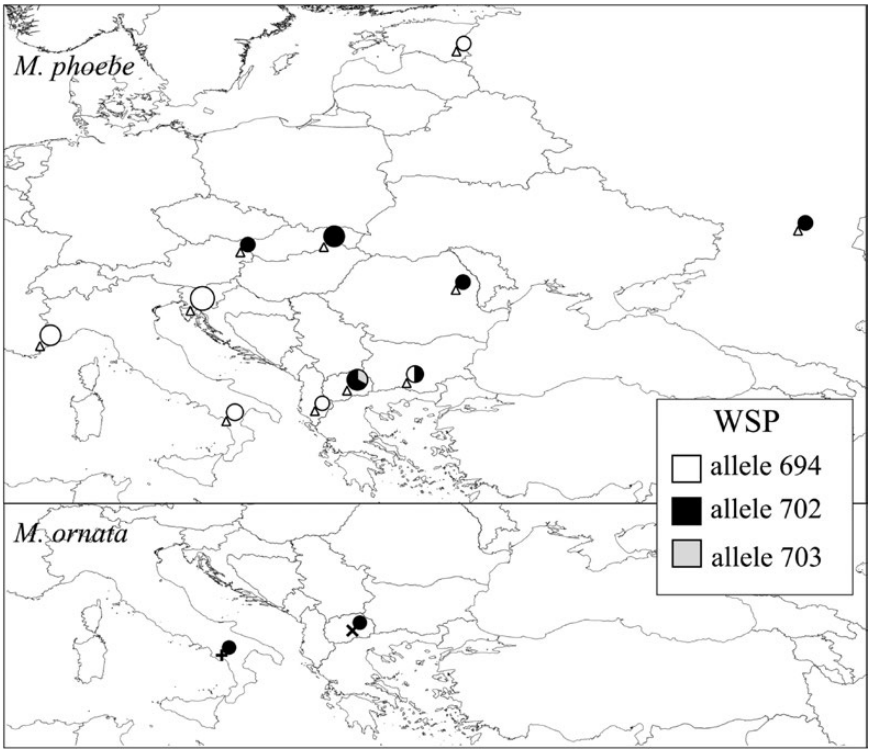

Figure 4. Distribution of *Wolbachia* surface protein (WSP) alleles in *M. phoebe* and *M. ornata*.

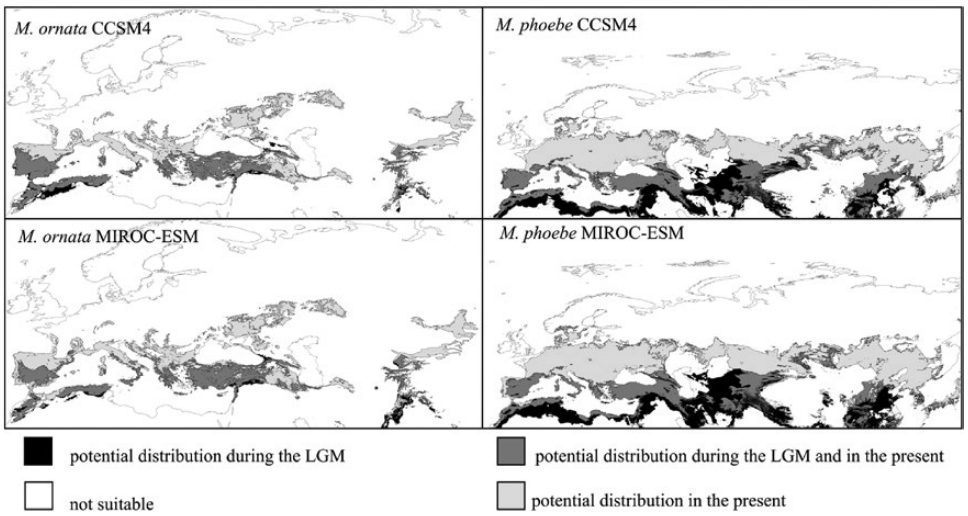

Figure 5. Predicted area dynamic for *M. phoebe* and *M. ornata* considering two time scales (present and the Last Glacial Maximum) using two climate models (MIROC and CCSM4).

Interspecific interaction and introgression

Based on the results, it seems that the mito-nuclear discordance is a result of ancient hybridization between *M. phoebe* and *M. ornata*. Several authors have proposed that hybridization occurred between *M. ornata* and *M. phoebe* (Varga, 1967; Bálint & Ilonczai, 2001; Russell, Pateman & Verovnik, 2014), but none of the authors have made any appropriate tests to validate their hypothesis. Most recently, Russell, Pateman & Verovnik (2014) obtained egg batches from a female sampled from populations of *M. ornata*-like butterflies in south-western Slovenia, which were reared in the

United Kingdom. Based on the heavy mortality of ova, larvae, pupae, and imagoes, they suggested that the female was *M. ornata* fertilized by a male *M. phoebe* from a nearby population, producing very few hybrid offspring. Russell, Pateman & Verovnik (2014) also examined the morphology of the few surviving imagoes but without any measurements or statistical analysis. High mortality alone does not prove that the examined offspring were hybrids. For example, high humidity that was recorded by the authors during the experiment could also cause high mortality especially for a dry-adapted Mediterranean species such as *M. ornata*. In this way, Russell, Pateman & Verovnik (2014) have not shown any clear evidence of ongoing hybridization between the two species.

Further evidence seems also to contradict this presumption. Firstly, our results showed that most of the *M. phoebe* individuals were infected with *Wolbachia*, while only a few infected specimens were found in *M. ornata*. This very characteristic pattern has two consequences: (1) an ongoing hybridization between *M. ornata* and *M. phoebe* is very unlikely, as much higher levels of *Wolbachia* infections would be expected in *M. ornata* as *Wolbachia* is maternally inherited to the next generation; (2) the low level of genetic difference in COI between *M. phoebe* and the western clade of *M. ornata* could not be the result of a recent *Wolbachia* infection, although we cannot exclude the effect of a historical infection, which has disappeared from *M. ornata*. Secondly, if hybridization between *M. phoebe* and the western lineage of *M. ornata* is in progress we would expect shared haplotypes also in nuclear regions and would not expect that vast majority of unique mitochondrial haplotypes are species-specific.

Further conclusions can be drawn from the examination of the mtDNA-based haplotype network. Only one mitotype was shared both in *M. ornata* and in *M. phoebe*, all the other haplotypes were species-specific. These results clearly indicate a lack of ongoing hybridization, since these differences between haplotypes could evolve only in genetic isolation.

On the other hand, the high level of similarity in COI between *M. phoebe* and western clade of *M. ornata* indicates past hybridization between the two species. The fact that there are several unique species-specific haplotypes both in *M. phoebe* and in the western lineage of *M. ornata* as well as only one mitotype is shared between these two species suggest that the hybridization event did not proceed during the LGM but during an earlier glaciation. These results are in accordance with the BEAST estimation on divergence time which infers the split of the eastern and the western clade of *M. ornata* at about 2 Mya (95% credibility interval: 0.5–3 Mya) based on nuclear genes. It is very probable that this hybridization happened during a cooling period when both species were forced to retreat into refugial areas. This process was most likely accompanied by the reduction of the overall population size of both species. It is highly possible that hybridization event took place in southern Italy and – as a heritage of this event – all individuals from the western lineage share *phoebe*-like COI haplotypes.

The hybridization event seems to have been unidirectional since *phoebe*-like mtDNA was observed in the western lineage of *M. ornata*, but we did not find any individual among *M. phoebe* specimens whose mtDNA would be *ornata*-like. Additionally, we did not find any sign of hybridization in nuDNA. This unidirectional hybridization is not rare in animals. Out of 80 studies that analysed the mtDNA of at least five hybrid individuals, 50 showed that all hybrids contained the mtDNA of only one of the two parental species (Wirtz, 1999). Wirtz's (1999) hypothesis for this pattern that females are generally the choosier sex; therefore, most hybrid mating should happen between females of the less common species and males of the more common species, which can lead to increased mitochondrial introgression relative to nuclear introgression.

Glacial refugia

SDM showed that the southern parts of the Mediterranean peninsulas could have served as important refugia during the glacial periods in Europe for both *M. ornata* and *M. phoebe*, but they are also predicted to be suitable areas at present. The exception is the Iberian Peninsula where only a well-separated subspecies of *M. phoebe* occurs (*M. phoebe occitanica*). Refugial areas were also predicted for both species in Anatolia, the Elburz Mts. (Iran), and the slopes of several central Asian mountains. Additionally, refugia were also predicted in the Far East in the case of *M. phoebe*. The predicted area stability in the Mediterranean peninsulas does not mean that the abundance of these species were stable over time. In fact, the present distribution pattern of the two species shows that *M. phoebe* is the less frequent species in the predicted refugial areas. It is very probable that during the glacial periods the ratio of *M. phoebe* to *M. ornata* was significantly different than currently. Our SDM analyses also suggests that *M. phoebe* has a much higher expansivity during the post-glacial warming and possibly also during the anthropogenous changes of vegetation, for example logging of forests or extension of the traditional land use. The fact that *M. phoebe* regularly produces two or three generations per year in warmer areas can result in a faster population growth and also a potentially higher speed of expansion, compared to single generation in *M. ornata*.

The predicted refugial areas are also supported by the haplotype diversity distribution, although we

have only limited samples from the Asian part of the distribution for both species (Figs S1 and S2). Based on the genetic data, the western lineage of *M. ornata* colonized the Apennine Peninsula, the north western parts of the Balkan Peninsula, and the Carpathian Basin. Several unique haplotypes were found in the southern part of the Apennine Peninsula and in Sicily, which also supports the hypothesis that these areas served as refugia during the ice ages.

Post-glacial range expansions

The Carpathian Basin belongs to the northernmost regions where *M. ornata* occurs at present. Populations living here have to survive under marginally suitable climate conditions (Tóth *et al.*, 2013), which probably induced the striking food plant specialization of these populations (Tóth *et al.*, 2015). These traits along with the haplotype distribution pattern favour a hypothesis of long distance colonization from a distant core area, that is from southern Italy, probably through some intermediate areas in the north-western Balkans (evidenced by occurrence in Slovenia and Croatia). These findings are in accordance with the earlier results concerning the post-glacial areal dynamics of the Carpathian Basin. Recent biogeographical studies have shown that this region was post-glacially re-populated not only from 'paradigmatic' refugia by long-distance dispersals but also from near-lying 'peri-Pannonian' core areas from climatically buffered regions (reviewed in Varga, 2010; Schmitt & Varga, 2012). These results generally highlight the primary importance of the Balkan Peninsula as the most important glacial refugium and source for re-population (e.g. Schmitt & Seitz, 2002; Habel, Schmitt & Müller, 2005; Schmitt, 2007). In this respect, our surveys join the few examples where the source of the post-glacial re-population was a south-western refugium (Petit *et al.*, 2002; Bihari *et al.*, 2011; Sramkó *et al.*, 2014).

The eastern clade of *M. ornata* is distributed from the Balkans to the easternmost part of Kazakhstan. Unfortunately, we could analyse only a few individuals from Anatolia, but based on these specimens we expect high level of divergence from the populations of the Balkans. The Balkan refugia are also supported by the distribution pattern of some unique haplotypes.

In Croatia both lineages of *M. ornata* have been observed. Additionally, a single specimen from Bosnia and Herzegovina had an eastern COI haplotype, while its nuDNA clearly belongs to the western clade. These results indicate that the two lineages of *M. ornata* meet in north-western part of the Balkan Peninsula. The exploration of this potential contact zone of eastern and western populations would be a fascinating topic for further studies but, unfortunately, the data of the present study do not allow us to make any robust conclusions on this.

The separation of an eastern and a western clade is supported in both species but the Carpathian Basin was re-populated from different sources in the case of *M. ornata* (south-western origin) and *M. phoebe* (south-eastern origin). Interestingly, the distribution of the main *Wolbachia* strains in *M. phoebe* also support the two separate clades as the allele 694 is mostly distributed in the Apennine Peninsula, while the allele 702 dominates the area east of the Alps. The eastern strain is associated with highly reduced haplotype diversity, while this phenomenon is not experienced in the western strain. Interestingly, the allele 702 has also been found in two specimens of *M. ornata* (from southern Italy and Macedonia). Further research is required to explain this pattern and clarify the potential role of *Wolbachia* in the spread of the hybrid COI haplotype in the ancient refugial populations of *M. ornata*.

The outlined historical dynamics of *M. ornata* transmits a more general message. Our target butterfly superficially seems to represent a typical 'paradigmatic' species restricted into Mediterranean peninsular refugia during the glacial maxima and expanded post-glacially to the North (see also the results of distribution modelling). The picture unravelled in this study seems to be somewhat more complicated as *M. ornata* shows an extended but scattered range into the continental steppic belt, in contrast to the majority of Mediterranean (sensu lato) species (e.g. the *Maniola jurtina* species complex, (Grill, Gkiokia & Alvarez, 2006; Kreuzinger *et al.*, 2015) which exhibit a high level of diversity around the Mediterranean basin. The observed pattern can be readily compared to the genetic structure of *M. cinxia* (Wahlberg & Saccheri, 2007) where – on a large scale – south-western, south-eastern, and continental clades have been distinguished; however, the continental clade of the *M. cinxia* has a much more continuous distribution into Central Asia which was also predicted as a primary core area of diversification of *Melitaea* clades (Leneveu *et al.*, 2009).

ACKNOWLEDGEMENTS

Acknowledgements are due to Jana Šlancarová, Peter Russell, and Gennadiy Kuznetsov, who kindly provided important specimens for this study. We highly appreciate the useful comments, ideas, and suggestions by Thomas Schmitt and Leonardo Dapporto and an anonymous reviewer.

FUNDING

The study was financed by the Hungarian Scientific Research Fund (OTKA K84071 and K109223). JB was supported by János Bolyai Scholarship of the Hungarian Academy of Sciences.

REFERENCES

Aris-Brosou S, Excoffier L. 1996. The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution* 13: 494–504.

Bálint Z, Ilonczai Z. 2001. Újabb adatok a Magyarországon védett nappali lepkék ismeretéhez (Lepidoptera: Lycaenidae, Nymphalidae) [New data on the proteced Hungarian butterflies (Lepidoptera: Lycaenidae, Nymphalidae)]. *Természetvédelmi Közlemények* **9:** 209–218.

Bereczki J, Tóth JP, Sramkó G, Varga Z. 2014. Multilevel studies on the two phenological forms of Large Blue (Maculinea arion) (Lepidoptera: Lycaenidae). *Journal of Zoological Systematics and Evolutionary Research* 52: 32–43.

Bhagwat SA, Willis KJ. 2008. Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *Journal of Biogeography* **35:** 464–482.

Bihari P, Sipos B, Melika G, Fehér B, Somogyi K, Stone GN, Pénzes Z. 2011. Western Palaearctic phylogeography of an inquiline oak gall wasp, *Synergus umbraculus*. *Biological Journal of the Linnean Society* **102:** 750–764.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Busby JR. 1991. BIOCLIM: a bioclimate analysis and prediction system. *Plant Protection Quarterly* **6:** 8–9.

Comes HP, Kadereit JW. 1998. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* **3:** 432–438.

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

Elith J, Kearney M, Phillips S. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* **1:** 330–342.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17:** 43–57.

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

Franklin J, Miller JA. 2009. *Mapping species distributions.* New York: Cambridge University Press.

Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH. 2010. The NCBI BioSystems database. *Nucleic Acids Research* **38:** D492–D496.

Grill A, Gkiokia E, Alvarez N. 2006. Evolutionary history and patterns of differentiation among European *Maniola* butterflies (Lepidoptera: Satyrinae). *European Journal of Entomology* **103:** 613–618.

Habel JC, 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.

Habel JC, Schmitt T, Müller P. 2005. The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: the phylogeography of the Marbled White butterfly (Satyrinae, Lepidoptera). *Journal of Biogeography* **32:** 1489–1497.

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. Series B: Biological Sciences* **270:** S96–S99.

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

Hewitt GM. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68:** 87–112.

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

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.

Jiggins FM. 2003. Male-killing Wolbachia and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. *Genetics* **164:** 5–12.

Kodandaramaiah U, Simonsen TJ, Bromilow S, Wahlberg N, Sperling F. 2013. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecology and Evolution* **3:** 5167–5176.

Kreuzinger AJ, 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. *Ecology and Evolution* **5:** 46–58.

Kuznetsov GV, Stradomsky BV. 2010. About finding *Melitaea telona* Fruhstorfer, 1908 (Lepidoptera: Nymphalidae) in Volgograd region. *Caucasian Entomological Bulletin* **6:** 193–194.

de Lattin G. 1967. *Grunddriß der Zoogeographie.* Gustav Fischer Verlag: Jena.

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

Leneveu J, Chichvarkhin A, Wahlberg N. 2009. Varying rates of diversification in the genus *Melitaea* (Lepidoptera: Nymphalidae) during the past 20 million years. *Biological Journal of the Linnean Society* **97:** 346–361.

Lisiecki LE, Raymo ME. 2007. Plio–Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. *Quaternary Science Reviews* **26:** 56–69.

Martínez-Botí MA, Foster GL, Chalk TB, Rohling EJ, Sexton PF, Lunt DJ, Pancost RD, Badger MP, Schmidt DN. 2015. Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. *Nature* **518:** 49–54.

Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76: 5269–5273.

Nei M. 1987. *Molecular evolutionary genetics*. Columbia University Press: New York.

Paradis E. 2010. pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics* **26:** 419–420.

Peña C, Malm T. 2012. VoSeq: a voucher and DNA sequence web application. *Plos One* **7:** e39071.

Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl UM, van Dam B, Deans JD, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JS, König AO, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Popescu F, Slade D, Tabbener H, de Vries SGM, Ziegenhagen B, de Beaulieu J-L, Kremer A. 2002. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management* **156:** 49–74.

Phillips SJ, Dudík M, Schapire RE. 2004. A maximum entropy approach to species distribution modeling. Proceedings of the twenty-first international conference on Machine learning. Banff, Alberta, Canada, ACM. 83.

Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* **23:** 564–571.

QGIS Development Team. 2015. *QGIS Geographic Information System: Open Source Geospatial Foundation Project*. http://hub.qgis.org/projects/quantum-gis/wiki/QGIS_Citation_Repository

Rajaei H, Rödder D, Weigand AM, Dambach J, Raupach MJ, Wägele JW. 2013. Quaternary refugia in southwestern Iran: insights from two sympatric moth species (Insecta, Lepidoptera). *Organism Diversity & Evolution* **13:** 409–423.

R Core Team. 2013. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Russell P, Pateman J, Gascoigne-Pees M, Tennent WJ. 2005. Melitaea emipunica (Verity, 1919) stat. nov: a hitherto unrecognised butterfly species from Europe (Lepidoptera: Nymphalidae). *Entomologist's Gazette* **56:** 67–70.

Russell P, Pateman J, Verovnik R. 2014. First record of Melitaea ornata Christoph, 1893, from Slovenia, with notes on its confirmed distribution and hybridisation with *M. phoebe* ([Denis & Schiffermüller], 1775). *Entomoloist's Gazette* **65:** 135–153.

Scheldeman X, Zonneveld M van. 2010. *Training manual on spatial analysis of plant diversity and distribution*. Rome: Biodiversity International.

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

Schmitt T, Seitz A. 2002. Postglacial distribution area expansion of *Polyommatus coridon* (Lepidoptera: Lycaenidae) from its Ponto-Mediterranean glacial refugium. *Heredity* **89:** 20–26.

Schmitt T, Varga Z. 2012. Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology* **9:** 1–12.

Schorr G, Holstein N, Pearman PB, Guisan A, Kadereit JW. 2012. Integrating species distribution models (SDMs) and phylogeography for two species of Alpine Primula. *Ecology and Evolution* **2:** 1260–1277.

Sramkó G, Molnár A, Hawkins JA, Bateman RM. 2014. Molecular phylogeny and evolutionary history of the Eurasiatic orchid genus *Himantoglossum* s.l. (Orchidaceae). *Annals of Botany* **114:** 1609–1626.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics* 68: 978–989.

Stephens M, Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *The American Journal of Human Genetics* 73: 1162–1169.

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

Stohlgren TJ, Ma P, Kumar S, Rocca M, Morisette JT, Jarnevich CS, Benson N. 2010. Ensemble habitat mapping of invasive plant species. *Risk Analysis* **30:** 224–235.

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

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105: 437–460.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585–595.

Toews DP, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21:** 3907–3930.

Tóth JP, 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. *Systematic Entomology* **39:** 749–757.

Tóth JP, Bereczki J, Végvári Z, Juhász E, Varga Z. 2015. Different host plant utilization ability of two closely related *Melitaea* species (Lepidoptera: Nymphalidae). *European Journal of Entomology* **12:** 120–125.

Tóth JP, Varga K, Végvári Z, Varga Z. 2013. Distribution of the Eastern knapweed fritillary (*Melitaea ornata* Christoph, 1893) (Lepidoptera: Nymphalidae): past, present and future. *Journal of Insect Conservation* **17:** 245–255.

Tóth JP, Varga Z. 2011. Inter- and intraspecific variation in the genitalia of the '*Melitaea phoebe* group' (Lepidoptera, Nymphalidae). *Zoologischer Anzeiger – A Journal of Comparative Zoology* **250:** 258–268.

Tóth JP, Varga Z, Bereczki J. 2016. Long-term survival and diversification of an endemic *Melitaea* species in mountains of Iran and adjacent areas. *Journal of Zoological Systematics and Evolutionary Research* 54: 106–115.

Tshikolovets V, Bidzilya A, Golovushkin M. 2002. *Butterflies of Transbaikal Siberia*. Tshikolovets: Kiev-Brno.

Tshikolovets VV, Yakovlev R, Kosterin O. 2009. *Butterflies of Altai, Sayans and Tuva (South Siberia)*. Tshikolovets: Kiev-Brno.

Tzedakis PC, Emerson BC, Hewitt GM. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution* **28:** 696–704.

Varga Z. 1967. A *Melitaea phoebe* délkelet-európai populációinak taxonómiai elemzése, két új alfaj leírásával [Contributions to the taxonomy of *Melitaea phoebe* Schiff. in Southerastern Europe with description of two new subspecies]. *Acta biologica Debrecina* **5:** 119–137.

Varga Z. 2010. Extra-Mediterranean refugia, post-glacial vegetation history and area dinamics in eastern central Europe In: Habel JC, Assmann T, eds. *Relict species: phylogeography and conservation biology.* London: Springer.

Varga Z, Szabó S, Kozma P. 2005. *Melitaea ogygia kovacsi* Varga, 1967 (Lepidoptera, Nymphalidae) in the Pannonian region: taxonomy, bionomy, conservation biology. In: Kühn E, Feldmann R, Thomas J, Settele J, ed. *Studies on the ecology and conservation of butterflies in Europe*. Sofia-Moscow: Pensoft Publishers, 65–68.

Wahlberg N. 2006. That awkward age for butterflies: insights from the age of the butterfly subfamily Nymphalinae (Lepidoptera: Nymphalidae). *Systematic Biology* **55:** 703–714.

Wahlberg N, Saccheri I. 2007. The effects of Pleistocene glaciations on the phylogeography of *Melitaea cinxia* (Lepidoptera: Nymphalidae). *European Journal of Entomology* **104:** 675–684.

Wahlberg N, Weingartner E, Warren AD, Nylin S. 2009. Timing major conflict between mitochondrial and nuclear genes in species relationships of *Polygonia* butterflies (Nymphalidae: Nymphalini). *BMC Evolutionary Biology* **9:** 1–16.

Wahlberg N, Wheat CW. 2008. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of lepidoptera. *Systematic Biology* **57:** 231–242.

Warren DL, Glor RE, Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33:** 607–611.

Warren DL, Seifert SN. 2010. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications* **21:** 335–342.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7: 256–276.

Werren JH, Windsor DM. 2000. Wolbachia infection frequencies in insects: evidence of a global equilibrium? *Proceedings of the Royal Society B: Biological Sciences* **267:** 1277–1285.

Wiemers M, Fiedler K. 2007. Does the DNA barcoding gap exist? A case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology* **4:** 8.

Wirtz P. 1999. Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour* **58:** 1–12.

Wright ES. 2016. Using DECIPHER v2.0 to analyze big biological sequence data in R. *The R Journal* 8: 352–359.

Zachos JC, Dickens GR, Zeebe RE. 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451:** 279–283.

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.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Specimens used in this study.
Figure S1. Haplotype networks used in the phylogeographic study of *M. ornata* and *M. phoebe* based on COI. Species and different linages are marked with different symbols.
Figure S2. Haplotype networks used in the phylogeographic study of *M. ornata* and *M. phoebe* based on concatenated nuclear genes.
Figure S3. Haplotype networks used in the phylogeographic study of *M. ornata* and *M. phoebe* based on EF-1a.
Figure S4. Haplotype networks used in the phylogeographic study of *M. ornata* and *M. phoebe* based on MDH.
Figure S5. Haplotype networks used in the phylogeographic study of *M. ornata* and *M. phoebe* based on RPS5.
Figure S6. Haplotype networks used in the phylogeographic study of *M. ornata* and *M. phoebe* based on wingless.