Genetic structure of *Parnassius mnemosyne* (Lepidoptera: Papilionidae) populations in the Carpathian Basin

Katalin Pecsenye[1], János P. Tóth[2], Judit Bereczki[1,2], Noémi Szolnoki[1], and Zoltán Varga[1]

[1] Department of Evolutionary Zoology and Human Biology, University of Debrecen, Debrecen, Egyetem tér 1., Hungary

[2] MTA-DE "Lendület" Behavioural Ecology Research Group, Department of Evolutionary Zoology and Human Biology, University of Debrecen, Debrecen, Egyetem ter 1. H-4032 Hungary

Corresponding author: K. Pecsenye

e-mail: pecskati@science.unideb.hu

Abstract

The pattern of genetic variation in a butterfly species depends on the past history of the given species and also on recent evolutionary processes affecting its populations. The aim of the present study was (i) to analyse the enzyme polymorphism in the Clouded Apollo populations of the Carpathian Basin to reveal the contemporary pattern of their genetic differentiation and (ii) to compare it with an expanded mtDNA haplotype network of the SE European populations. Allozyme polymorphism was analysed in 22 populations of four geographic regions: Transdanubian (TM) and North Hungarian Mountains (NM), Körös (KÖR) and Bereg – Apuseni – East Carpathian regions (BEAC). The results of the Bayesian clustering analyses based on allozymes supported the presence of three main genetic lineages in the Carpathian Basin: one was typical for TM, another was characteristic for NM and the third cluster was predominant in KÖR. The populations of BEAC harboured a mixture of two clusters. The mtDNA haplotype network suggested a fairly similar distribution: the peri-Alpine clade together with the West Balkan clade was detected in TM, while the East Balkan clade occurred in NM, partly in TR and in the two eastern regions of the Basin (KÖR and BAEC). The incongruities between the results of the mtDNA and allozyme studies can be explained by the different time scales represented by the two markers. The mtDNA haplotype network provided strong evidence concerning the existence of two Balkan lineages, which probably formed a "zone of admixture" in the Transdanubian and North Hungarian Mountains. The possibility of Last Glacial survival of *P. mnemosyne* in the Carpathian Basin and the conservation implications of these results are discussed.

Key words: allozyme polymorphism, mtDNA haplotypes, genetic clusters, phylogenetic lineages, refugia, colonisation

Introduction

The present geographical distribution of butterfly species and especially the pattern of genetic variation in their populations basically depend on two sets of processes: (i) the past history of the given species, and (ii) the current distribution of its suitable habitats. The first point implies that the pattern of genetic differentiation among contemporary populations of a species reflects its geographical dynamics, i.e. periodical retreats into refugia and expansions during glacial-interglacial cycle(s) (Gratton et al. 2008; Hewitt 1996, 2000; Schmitt et al. 2007). These dynamic processes were mostly revealed in the 'paradigmatic' species retreating during the glacial phases into southern refugia situated in the Mediterranean peninsulas, and thus, regularly subdivided into genetically differentiated lineages (Habel et al. 2005; Hewitt 1999, 2004; Schmitt and Hewitt 2004; Taberlet et al. 1998). Hence, the recent geographical pattern of genetic variation of these species is the consequence of their northwards expansion and amalgamation of haplotypes, formerly isolated in southern refugia (Hewitt 2001, 2004; Schmitt and Seitz 2001; Schmitt et al. 2002; Wahlberg and Saccheri 2007). The geographical pattern of genetic variation in "extra-Mediterranean" species (Malicky et al. 1983) was even much less understood, although several authors already hypothesised (Varga 1977) or demonstrated (e.g. Babik et al. 2004; Jaarola and Searle 2002; Kotlík et al. 2006; Pinceel et al. 2005; Stewart and Lister 2001; Todisco et al. 2010; Ursenbacher et al. 2006) the existence of "cryptic" refugia North of the Mediterranean peninsulas reviewed by Schmitt and Varga (2012). Our target species also belongs to this group of species, according to the survey of Gratton et al. (2008).

On the other hand, the present range of habitats required by a butterfly species is also vital in its contemporary distribution. Nevertheless, human activities have fundamentally rearranged the distribution of natural habitats over thousands of years and particularly dangerously during the last hundred years. Changes in land-use patterns, urbanization or increasing human recreational activities have constituted serious threats to many habitats, resulting in a severe

decline in biodiversity in Europe (New 1993; van Swaay and Warren 1999; van Swaay et al. 2006). Butterflies are highly sensitive to environmental change, as a result of their specialized ecology and coarse-grained perception of habitats (Maes and Dyck 2005; Maes and Van Dyck 2001; Parmesan et al. 1999; van Swaay and Warren 1999; van Swaay et al. 2006; Filz et al. 2013). In consequence of deteriorating habitat quality and the increasing fragmentation of natural environments, many European butterfly species have become endangered (Heath 1981; van Swaay 1990; Thomas 2005; Välimäki and Itämies 2003). In order to work out efficient strategies to preserve an endangered butterfly species, one must understand the structure of its genetic variation; that is to synthesise the knowledge of past history and present genetic composition of the species.

Information on these two aspects of genetic diversity may help to estimate the number and distribution of Evolutionarily Significant Units (ESU) in the species at risk. The concept of ESU was introduced by Ryder (1986), though its definition has changed over the past three decades (Crandall et al. 2000). First, it has been associated with genetically and adaptively differentiated populations (Ryder 1986). Then, ESU has been considered as a reproductively isolated set of populations (Waples 1991). Later, Moritz (1994) regarded it as a reciprocally monophyletic unit. A more operative concept of ESU has been introduced by Crandall et al. (2000), stressing the importance of the lack of ecological and genetic exchangeability between different ESUs. In practical nature conservation, however, an even more important unit is the functional conservation unit (FCU), which is a group of demographically or ecologically independent populations (Allendorf and Luikart 2009; Maes et al. 2004). FCUs are often defined within ESUs and represent populations that are important for the long term persistence of an ESU (Hughes et al. 1997; Vila et al. 2006). Though there are certain difficulties in the application of these terms, it is of outmost importance in conservation practice to assign populations to particular ESUs or FCUs.

The phylogeography of *Parnassius mnemosyne* (Linnaeus, 1758) (Lepidoptera: Papilionidae) in Europe has been surveyed by Gratton et al. (2008). Their study using mtDNA (COI) sequences has revealed multiple glacial refugia of the species in the perialpine region (origin of western lineage), in the Hellenic Peninsula and in the Balkan/Carpathian region (origin of the eastern lineage). The haplotypes observed in the Pannonian/Carpathian region mostly belong to the eastern ancestral lineages and can be split into two sublineages corresponding to two colonisation routes starting from the Balkan and/or peri-Pannonian mountains: EN sublineage (northern route in eastern lineage) and EE sublineage (eastern route in eastern lineage). At the same time, one haplotype of the western lineage also occurs in the western part of Hungary. It thus appears that *P. mnemosyne* populations have at least three mtDNA clades in the Carpathian Basin. Our aim was to analyse the genetic structure of the *P. mnemosyne* populations in the Carpathian Basin using allozymes and mitochondrial COI sequence data to contrast the influence of past history and the more recent structure of genetic variation connected with the geographic distribution of suitable habitats in the surveyed region.

In the present study, we hence compared the sequence of the mitochondrial cytochrome c oxidase subunit I (COI) with enzyme polymorphism in several populations of *P. mnemosyne* in the Carpathian Basin to reveal the pattern of genetic differentiation.

Materials and Methods

Study species

Our target species, the Clouded Apollo (*P. mnemosyne*) belongs to the protected species in most countries of Europe (Habitats Directive Annex IV; IUCN Red List) due to the decreasing trends in most of its populations and the fragmentation of its geographical range (van Swaay et al. 2006, 2010). It is predominantly a western Palaearctic temperate species, distributed in Europe from the North of the Iberian Peninsula across most parts of Central Europe up to 65°

N in Scandinavia, and in the central and northern part of European Russia. In the western and central part of its distribution, it seems to be a mountain species; consequently, it is completely missing in the UK, Denmark, the Benelux countries and in major parts of the Iberian Peninsula. The continental part of its distribution is more continuous, where populations are also known from lowland areas. Nevertheless, it is extinct or seriously declining in most lowlands of Central Europe.

The Clouded Apollo is a univoltine species. The imago flies from May till mid-June, in high mountains even to the end of July. It requires structured habitats; larvae live in habitats where the food plants grow (in Central European mesic forests *Corydalis cava,* or in rocky ravine forests *C. solida*, or in the Balkan high rupicolous mountains *Corydalis* spp.). At the same time, the imago prefers sunny, flower-rich glades, forest-fringes or meadows for mating and feeding. The dispersal ability of *P. mnemosyne* is fairly limited with individuals hardly ever moving more than 2-3 km and with average home ranges of about 200 m (Konvička and Kuras 1999; Välimäki and Itämies 2003). Requiring continuous supply of open-canopy structures historically maintained by coppicing or forest pasture, the Clouded Apollo is very sensitive to habitat changes (Freese et al. 2006; Konvička and Kuras 1999; Warren and Key 1991). In the Carpathian Basin, the once large, continuous and mostly traditionally used forested areas (Magyari 2002) have been greatly reduced or replaced by monocultural plantations due to intensification. In addition to the extensive deforestation in the mountainous areas, an extreme fragmentation of the gallery forests has occurred in the lowland due to the control of the Tisza River and its tributaries (Meglécz et al. 1999). As a consequence, Clouded Apollo populations experienced a decline rate in the Carpathian Basin similar to that in Western Europe. Nevertheless, besides several small isolates, some large, dense populations still exist, mostly in the Transdanubian and North Hungarian Mountains (Meglécz et al. 1997, 1998, 1999).

Samples

Parnassius mnemosyne samples were collected from 12 geographic areas of the Carpathian Basin grouped into the following four regions: Transdanubian Mountains (TM): Danube Bend (3 populations), Velencei hills (1 population), Bakony Mts. (3 populations), Keszthelyi hills (1 population) and South Tolna (1 population); North Hungarian Mountains (NM): Zempléni Mts. (1 population), Aggteleki Karst (1 population), Bükk Mts. (3 populations); Körös region (KÖR: 4 populations); Bereg – Apuseni – East Carpathian region (BAEC): Bereg lowland (1 population), Apuseni Mts. (1 population), East Carpathian Mts. (2 populations) (Table 1, Fig. 1). Altogether, 525 individuals were collected from 22 populations between 1998 and 2014. Imagos were collected in May and June after the main egg laying period and stored at –80°C. Sample sizes varied between 11 and 47, according to the size of populations. Enzyme polymorphisms were investigated in all 525 individuals, while 20 of them were selected for DNA sequencing (Table S1).

Enzyme studies

Allozyme polymorphisms were studied at 14 different loci by vertical polyacrylamide gel electrophoresis. Thoraxes were homogenized in 300µl extraction buffer and these samples were used to study *Got*, *Gpdh*, *G6pgdh*, *Hk*, *Idh*, *Mdh*, *Me*, *Pgi*, *Pgm*, *SodA* and *SodB*. Abdomens were homogenized in 150µl extraction buffer and these extracts were used to analyse *Acon*, *Est* and *Lap*. The extraction buffer, the electrophoresis buffer systems and running conditions, together with the staining solutions were used as in Bereczki et al. (2005). Genotypes of the different individuals were scored.

Measures of genetic variation were calculated for each population sample using GENALEX v. 6.5 (Peakall and Smouse 2006) and FSTAT v.1.2 (Goudet 1995): average number of alleles per locus (n_A), average effective number of alleles (n_E), allelic richness (*A*r) calculated for N=11, average observed heterozygosity (H_o) and proportion of polymorphic loci using the 95%

criterion (P_{95}). The genetic variation of the populations was also characterised by the distribution of alleles between two frequency classes: fixed (p=1) and variable (0<p<1) alleles. The allele pool of the regions was sorted to common (all: detected in all regions), partially common (other: detected in more than one region) and specific (spec: detected in the given region only) alleles. All characteristics of polymorphism were compared among the four regions by performing ANOVA running GLIM 4 (Francis et al. 1994).

For the analysis of the structure of the populations, a principal component analysis (PCA) was first computed on the basis of their allele frequency distribution using PAST v. 1.56 (Hammer et al. 2001). Next, we applied a Bayesian clustering method (Pritchard et al. 2000). Here, we estimated the most probable number of genetically differentiated groups (K) in our populations and assigned the individuals to these groups. STRUCTURE V. 2.3.2 was run without population priors with initial burn in 100,000 and running length 500,000 to carry out these analyses. K-values were assumed to be between 1 and 22. In the evaluation of the results, ΔK (Evanno et al. 2005) was computed using STRUCTURE HARVESTER V. 0.6.91 (Earl and von Holdt 2011). To describe the genetic composition of the Transdanubian and North Hungarian Mountains, the correlation between genetic and geographic distances was tested by a Mantel test (Mantel 1967). GENALEX V. 6 (Peakall and Smouse 2006) was used to carry out these tests with 999 permutations. Finally, we analysed the populations of the Eastern regions (NM, KÖR, BAEC) and searched for the presence of outlier loci. In this analysis, the locus-specific effects of differentiation (e.g. selection) are separated from genome-wide ones (e.g. random drift, gene flow) (Stinchcombe and Hoekstra 2007). The F_{ST} values observed among the investigated populations are compared to the random distribution of fixation indices obtained as a function of between population heterozygosity (Beaumont and Nichols 1996). LOSITAN (Antao et al. 2008) was computed to find outlier enzyme loci using the data sets where the populations were pooled according to regions. The confidence limit was set either 0.95 or 0.99, while the false

discovery rate was 0.1 or 0.05. The number of simulations was 300,000 in each run and the computation was repeated five times for all four combinations of the parameters.

DNA studies

DNA was extracted by homogenizing the heads of butterflies following the protocol in Bereczki et al. (2014). COI was amplified by specific primers modified at their 5′-end to include the universal sequencing primer T7promoter (Wahlberg and Wheat 2008). We followed the above amplification protocol and the guidelines of the Nymphalidae Systematics Group (http://nymphalidae.utu.fi/). DNA sequences were edited and revised manually by Chromas Lite v. 2.01. The COI sequence of the 20 individuals selected from the samples of this study were aligned to particular sequences accessible in Gratton et al (2008) using MEGA v. 6 (Tamura et al. 2013). Based on the aligned COI sequences, a haplotype network was constructed in the R computing environment (R Core Team 2014) using PEGAS package v. 0.6 (Paradis 2010). Haplotype distribution in the geographical space was visualised using QGIS (QGIS Development Team 2014).

Results

Level of enzyme polymorphism

Clouded Apollo populations of the Carpathian Basin were highly polymorphic at the investigated enzyme loci with 2.45 alleles per locus on average and 0.188 as mean frequency of heterozygotes (Table 2: n_A and H_o). Though most average measures of polymorphism tended to be lower in the Northern Range, none of these differences was significant (Table 2).

The allele pool of *P. mnemosyne* in the Carpathian Basin amounted to 87 at the 14 loci analysed. The total number of alleles differed largely among the four regions (Figure 2). The highest number was observed in the Transdanubian Mountains (TM) and in the Bereg –

Apuseni – East Carpathian (BAEC) region (~70% of the allele pool in the Carpathian Basin), while it was relatively low in the North Hungarian Mountains (NM) and in the Körös region (KÖR) (~50% of the allele pool in the Carpathian Basin) (Fig. 2). The distribution of alleles among the common, partially common and specific classes was also different in the four regions. The highest portion of the region specific alleles was observed in the Transdanubian Mountains (TM: ~20%), whereas the lowest occurred in the North Hungarian Mountains (NM: ~2%) (Fig. 2).

The genetic structure of P. mnemosyne populations in the Carpathian Basin

The first two axes of the principal component analysis accounted for 54.1% of the total variation in allele frequencies. The first axis was mostly explained by the differences at the *Pgi* and *Pgm* loci, while the *Est* and *Hk* loci contributed most to the second axis. Each population was assigned to one out of four distinct clouds corresponding to the four geographic regions (Fig. 3).

Based on the ΔK values, Bayesian clustering analysis estimated K=2 the most likely number of genetic clusters for the Clouded Apollo populations of the Carpathian Basin (Table S2). The genetic differences between the two clusters were characterised by their unique alleles (Table S3: K=2). The bar plot of the membership coefficients of the investigated *P. mnemosyne* individuals indicated three distinct regions with different average membership coefficients (Fig. 4: K=2). Cluster 2 was predominantly present in the individuals of the North Hungarian and Transdanubian Mountains (NM and TM: $F_{2,20}$=32.56; P<0.001), while cluster 1 was characteristic for the individuals of the Körös and BAEC regions. Nevertheless, the average frequency of cluster 1 differed greatly between these two regions; it amounted to 94.5% in the Körös region, whereas it was only ~28% in the Bereg – Apuseni – East Carpathian (BAEC) region (Fig. 4: K=2 – KÖR vs. BAEC).

These results, however, partly contradicted those of PCA where the Transdanubian and North Hungarian Mountains were clearly separated from each other. So we also tested the possibility of having 3 clusters (K=3) in the populations of the Carpathian Basin (Table S2). We again characterised the genetic differences among the three clusters by cluster specific alleles. The results were suggestive (Table S3: K=2 vs. K=3). Cluster 1 assuming K=3 (K3Cl1) had the same 4 unique alleles as detected in cluster 1 at K=2 (K2Cl1). Moreover, clusters 2 and 3 assuming K=3 shared one of the unique alleles of cluster 2 at K=2 (K2Cl2). We, therefore, considered them as clusters 2A and 2B (K3Cl2A and K3Cl2B). Nonetheless, these two 'subculsters' also had unique alleles (Table S3: K=3). The bar plot representing the membership probabilities of the individuals in these three clusters suggested four regions in accordance with the results of PCA (Fig. 4: K=3). Cluster 1 (K3Cl1) was the most frequent in the Körös region ($F_{2,20}$=4.36; P<0.05), cluster 2A (K3Cl2A) was predominant in the Transdanubian Mountains ($F_{2,20}$=18.5; P<0.001); and cluster 2B (K3Cl2B) was characteristic for the North Hungarian Mountains ($F_{2,20}$=3.62; P<0.05). At the same time, the individuals of the Bereg – Apuseni – East Carpathian (BAEC) region were mostly composed of clusters 1 (K3Cl1: 34.5%) and 2B (K3Cl2B: 53.9%); that is the two clusters typical for the North Hungarian Mountains (NR) and the Körös region (KÖR).

In the subsequent part of the analyses we searched for the correlation between the genetic and geographic distances in the populations of the Transdanubian and North Hungarian Mountains and carried out a Mantel test using the matrix of Cavalli-Sforza and Edwards chord distances. The results indicated a highly significant isolation by distance structure in this complex region (Fig. S1).

Finally we were interested how the genetic composition of the two Eastern regions (KÖR and BAEC) and the North Hungarian Mountains evolved. Clouded Apollo lives in different habitats in these three regions. Considering the climatic and other ecological differences among the different habitat types, we were searching for the genetic signature of diversifying

selection. That is, we were looking for loci with unexpectedly high F_{ST} values between all pairs of the three regions. The null distribution of fixation indices was obtained as a function of the between region heterozygosity using neutral coalescent simulations as implemented in LOSITAN. The four combinations of the two parameters (confidence limit and false discovery rate) with five replicates for every combination yielded 20 simulations for each region pair. In the comparison of the Körös and BAEC regions, we found altogether four loci as outliers, but only one of them (*Pgm*) proved to be consistently significant in all 20 simulations (Fig. 6A, Table S4). When the samples of the North Hungarian Mountains and the Körös region were included in the analyses, we again detected one locus (*Mdh*) with significantly higher than expected F_{ST} value in all 20 runs, regardless of the combination of parameters (Fig. 6B, Table S4). These two putative outlier loci exhibited private alleles in cluster 1 (K2Cl1 and K3Cl1) which was predominant in the Körös region. In the comparison of the two regions, however, where *P. mnemosyne* populations occupy similar mountainous forest habitats (NM and BAEC), we could not obtain any obvious result. Altogether, four loci exhibited significantly higher than expected F_{ST} value in the 20 runs, but none of these loci was significant in all simulations or at least consistent over the five repeats of one combination of the two parameters (Table S4). Thus, the results indicate that we were able to detect candidate enzyme loci (*Pgm* and *Mdh*) exposed to divergent selection in the analyses of those populations only which live in different habitats.

COI

In the combined data set (our own sequences and ones selected from those of Gratton et al. 2008), we could identify 26 unique haplotypes based on 927 bp of COI sequences which contained 22 parsimony informative sites. The haplotypes were plotted as a network and geo-referenced pie charts (Supplementary 1). Although the different haplotypes are separated by only a few mutations, strong geographical structure was recognised. In the Carpathian Basin,

three haplogroups were present. The western lineage (W), the eastern lineage which could be split into two sublineages: one of them distributed in the western part (EW), while the other in the eastern part of the Carpathian Basin (EE1). Unique haplotypes were also recognised in the North Hungarian Mountains.

Discussion

Phylogeography

The Clouded Apollo populations of the Carpathian Basin have a relatively high level of genetic diversity concerning both COI sequence variation and allozyme polymorphisms. We revisited the results of Gratton et al. (2008) regarding the mtDNA haplotypes of these and surrounding populations. Their geographical repartition clearly indicates two main core areas in Central and Southern Europe: the south-eastern peri-Alpine region and the mountains of the eastern Balkans. Both areas are characterised by numerous unique haplotypes separated from the most frequent ones (W, EW and EE) by a single or a few substitutions only, building typical star-like structures. Since the accumulation of these mutations probably required relatively large population sizes and also considerable time, we assume that Clouded Apollo survived more than just one glacial-interglacial cycle in these regions, subsequently populating several surrounding areas using different expansion routes (Supplementary 1).

According to these data, the western lineage (W) colonised a major part of the Alps and western Central Europe but only a restricted portion of the western part of the Carpathian Basin. From the East Balkan core area, tracks of two major haplotype groups exist to the Carpatho-Pannonian region. The West Balkan clade (EW) crossed several mountains of the western Balkans reaching the Transdanubian Mountains of Hungary, but also the eastern pre-Alpine zone of Austria. The East Balkan clade (EE1) crossed the South Carpathians (Iron Gate) and reached different parts of the East Carpathians, Transylvania and eastern Hungary

(North Hungarian Mountains and eastern part of the Great Hungarian Plain). Unique haplotypes were, however, also recognised in the North Hungarian Mountains. Furthermore, some derived haplotypes of the East Balkan clade (EE1) seem to have crossed the Porta Hungarica (i.e. the north-western gate of the Carpathian Basin) and probably correspond to the EN lineage described by Gratton et al. (2008).

This geographical pattern highlights two remarkable points. (i) The North Hungarian Mountains (Bükk Mts., Zemplén Mts. and Aggtelek Karst) exhibits an unexpectedly high haplotype diversity (i.e. EE1 and EN, and two minor derived haplotype attached to EE1). It suggests that a secondary survival centre of the last glacial period (LGM) could exist in these mountains. This pattern is similar to that obtained for the beech (*Fagus sylvatica*) (Magri 2008; Magri et al. 2006), for the Woodland Ringlet (Besold and Schmitt 2015) or, even more precisely, for the Bank vole (*Clethrionomys glareolus*) (Filipi et al. 2015; Kotlík et al. 2006). (ii) Furthermore, we uncovered a "zone of admixture" (*sensu* Hampe and Petit 2005) of the two East Balkan mtDNA clades (EE1 and EW) co-occurring with the pre-Alpine (W lineage) in the Transdanubian Mountains. This seem to be parallel with several European nemoral species possessing a relatively high genetic diversity in southern Central Europe, but showing essential genetic differentiation in the fringing southern, rear edge of their range (Bilton et al. 1998; Petit et al. 2003).

Considering the allozyme patterns of Clouded Apollo populations in the Carpathian Basin, STRUCTURE analysis for K=3 mostly agreed with that of mtDNA. Based on allozyme data, we detected three main clusters in the Carpathian Basin: cluster 1 was characteristic for the eastern regions of the Basin (KÖR and BAEC), cluster 2A was predominant in the Transdanubian Mountains, while cluster 2B occurred in the North Hungarian Mountains. The three mtDNA haplogroups showed a fairly similar distribution: the peri-Alpine haplogroup (W) together with the West Balkan clade (haplogroup EW) was detected in the Transdanubian Mountains, while the East Balkan clade (haplogroup EE1) primarily occurred in the two eastern regions of the

Carpathian Basin (KÖR and BAEC), in the North Hungarian Mountains, but to a lesser extent also in the Transdanubian Mountains.

There was, however, also some incongruence between mtDNA and allozyme patterns. The most obvious was observed in the genetic composition of the Transdanubian and North Hungarian Mountains. The haplotype network suggested that the North Hungarian Mountains were mostly colonised by the East Balkan clade (EE1) but it also shows the presence of a derived unique haplotype EN. Whereas three different haplogroups were present in the Transdanubian range: West Balkan (EW), East Balkan (EE1) and the peri-Alpine (W) clades. Bayesian clustering analysis for K=2, however, showed similar genetic composition for these two mountain ranges in allozyme patterns. The most plausible explanation of this apparent inconsistency might be that the two markers represent different time scales.

The differentiation at the mtDNA level might reflect the genetic fingerprints of the last glacial cycles, i.e. (i) survival in the core areas (peri-Alpine refugia and East Balkan Mountains) possibly for more than a single glacial cycle, (ii) survival of the LGM in the secondary centres of genetic differentiation and (iii) routes of subsequent colonisation waves. The co-occurrence of some derived unique satellite haplotypes of the EE and EN clades in the North Hungarian Mountains suggests the Last Glacial persistence of Clouded Apollo in this region. On the other hand, the co-occurrence of different haplotypes in the populations of the Transdanubian Mountains is probably the result of post-glacial overlap of haplotypes derived from three different sources (W, EW and EE1). In contrast, the allozyme pattern reflects much more recent evolutionary processes. We suppose intense gene flow in the forested landscape of the Transdanubian and North Hungarian Mountains (Cassel-Lundhagen et al. 2009) among widely distributed and abundant populations of Clouded Apollo, resulting in a similar genetic composition of populations within this extended region. The significant isolation by distance structure detected in the Transdanubian and North Hungarian Mountains is clear evidence supporting this hypothesis.

The other incongruity between mtDNA pattern and allozyme STRUCTURE results concerned the genetic structure of the Bereg – Apuseni – East Carpathian (BAEC) and Körös regions. Both regions were colonised by the East Balkan clade (EE1), but the average membership coefficients differed among enzyme loci. In the Körös region, cluster 1 was predominant, while the BAEC region exhibited a mixture of clusters 1 and 2B. Moreover, the BAEC region had a considerable portion of the species allele pool in the Carpathian Basin and a quite sizable number of specific alleles of the latter. The background of the incongruity between the two markers can again be associated with the different temporal scales addressed by them. Although the origin of the mtDNA clades and the colonisation routes had probably been similar in these two regions, the present evolutionary processes indicated by the enzyme loci were possibly different. In the Körös region, the isolated Clouded Apollo populations inhabit fragmented gallery forests, while their habitats are mountain forests in the BAEC region, and also in the North Hungarian Mountains. That is, the isolated populations of these lowland forests can be considered as peripheric both in geographical and ecological sense (e.g. Cassel-Lundhagen et al. 2009) since the range of the Clouded Apollo does not extend to the inner part of the Pannonian lowland. Moreover, these riverine hardwood forests exist as 'micro-climatic islands' within the summer-dry forest-steppe climate of the lowland. Thus, the presence of significant outlier loci in the KÖR region, compared to the populations of other regions (BAEC and NM), could possibly be explained by climatic adaptation. Nevertheless we cannot rule out the possibility of bottle neck in the Körös region either. At present this region is quite isolated from the large BEAEC region and only consists of a few populations.

Conservation implications

In the short term protection of endangered species, the delineation of functional conservation units (FCU) is of pivotal importance. A functional conservation unit is defined as a population or a group of populations that have restricted gene flow from other populations of the species

(e.g. de Guia and Saitoch 2011; Fraser and Bernatchez 2001). According to allozyme pattern, we have recognized three genetic clusters in the Carpathian Basin. Thus, we have to assume three FCUs of *P. mnemosyne* implying that all three units have to be preserved. We detected a high genetic variability both in mtDNA COI sequence and allozymes in the Transdanubian and North Hungarian Mountains. These populations are relatively well-covered by protected areas in accordance with the Habitats Directive (National Parks and NATURA 2000 sites). The peri-Pannonian populations in the lowland gallery forests, however, are vulnerable or even endangered. In these regions, the only chance for the long-term preservation of the species is to maintain or restore structurally rich woodlands. This type of forest management is also vital for some other woodland butterflies with serious declining rates (*Leptidia morsei, Lopinga achine*) or regional threats (e.g. *Euphydryas maturna*) in the Carpathian Basin. The conservation of *P. mnemosyne* is especially critical in the Körös region, which only consists of four populations in a fairly restricted area. At the same time, this region is characterised by a specific allozyme cluster with four cluster specific alleles, not occurring in other regions of the Carpathian Basin. In order to preserve these populations the high game density has to be seriously reduced and the extensive management (especially the clear-cutting practice) of the hardwood gallery forests must be prevented. Considering the limited long-distance dispersal ability of the butterfly, it is necessary to form or maintain ecological corridors between the forest patches to enhance migration among them. This might facilitate the development of a network of populations and decrease their isolation, which is essential to maintain their variation.

Acknowledgements

The study was financed by the NKFP-3 B/023/2004 and OTKA K84071 projects. JB was supported by János Bolyai Scholarship of the Hungarian Academy of Sciences. The technical assistance of V. Mester in the electrophoretic work is very much respected. A. Crisan, L.

Peregovits, L. Rákosy, S. Szabó and P. Kozma contributed to the sampling at several sites. The support of the Nature Conservation Authorities of Hungary is highly appreciated.

References

Allendorf, F. W., & Luikart, G. (2009). *Conservation and the Genetics of Populations*. John Wiley & Sons.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a F(st)-outlier method. *BMC Bioinformatics*, *9*, 323. doi:10.1186/1471-2105-9-323

Babik, W., Branicki, W., Sandera, M., Litvinchuk, S., Borkin, L. J., Irwin, J. T., & Rafiński, J. (2004). Mitochondrial phylogeography of the moor frog, Rana arvalis. *Molecular Ecology*, *13*(6), 1469–1480. doi:10.1111/j.1365-294X.2004.02157.x

Beaumont, M. A., & Nichols, R. A. (1996). Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proceedings of the Royal Society of London B: Biological Sciences*, *263*(1377), 1619–1626. doi:10.1098/rspb.1996.0237

Bereczki, J., Pecsenye, K., Peregovits, L., & Varga, Z. (2005). Pattern of genetic differentiation in the Maculinea alcon species group (Lepidoptera, Lycaenidae) in Central Europe. *Journal of Zoological Systematics and Evolutionary Research*, *43*(2), 157–165. doi:10.1111/j.1439-0469.2005.00305.x

Bereczki, J., Tóth, J. P., 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*(1), 32–43. doi:10.1111/jzs.12034

Besold, J. & Schmitt, T. (2015). More northern than ever thought: Refugia of the Woodland Ringlet butterfly Erebia medusa (Nymphalidae: Satyrinae) in Northern Central Europe. *Journal of Zoological Systematics and Evolutionary Research, 53,* 67-75. doi:10.1111/jzs.12076

Bilton, D. T., Mirol, P. M., Mascheretti, S., Fredga, K., Zima, J., & Searle, J. B. (1998). Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London B: Biological Sciences*, *265*(1402), 1219–1226. doi:10.1098/rspb.1998.0423

Cassel-Lundhagen, A., Tammaru, T., Windig, J. J., Ryrholm, N., & Nylin, S. (2009). Are peripheral populations special? Congruent patterns in two butterfly species. *Ecography*, *32*(4), 591–600. doi:10.1111/j.1600-0587.2008.05685.x

Crandall, K. A., Bininda-Emonds, O. R. P., Mace, G. M., & Wayne, R. K. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, *15*(7), 290–295. doi:10.1016/S0169-5347(00)01876-0

de Guia, A. P. O., & Saitoch, T. (2011). Evolutionarily Significant Units of Gray-Sided Vole (Myodes rufocanus) in Hokkaido, Japan. *Philippine Journal of Science*, *140*(1), 41–50.

Earl, D. A., & vonHoldt, B. M. (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. doi:10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

Filipi, K., Marková, S., Searle, J. B., & Kotlík, P. (2015). Mitogenomic phylogenetics of the bank vole Clethrionomys glareolus, a model system for studying end-glacial

colonization of Europe. *Molecular Phylogenetics and Evolution, 82, Part A*, 245–257. doi:10.1016/j.ympev.2014.10.016

Filz, K.J., Engler, J.O., Stoffels, J., Weitzel, M., Schmitt, T. (2013): Missing the target? A critical view on butterfly conservation efforts on calcareous grasslands in south-western Germany. *Biodiversity and Conservation, 22,* 2223-2241. doi:10.1007/s10531-012-0413-0

Francis, B., Green, M., & Payne, C. (1994). *GLIM 4. The statistical system for generalised linear interactive modelling.* New York: Oxford University Press.

Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*, *10*, 2741–2752.

Freese, A., Benes, J., Bolz, R., Cizek, O., Dolek, M., Geyer, A., et al. (2006). Habitat use of the endangered butterfly Euphydryas maturna and forestry in Central Europe. *Animal Conservation*, *9*, 388–397.

Godfrey M. Hewitt. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*(1-2), 87–112. doi:10.1111/j.1095-8312.1999.tb01160.x

Goudet, J. (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of Heredity*, *86*(6), 485–486.

Gratton, P., Konopiński, M. K., & Sbordoni, V. (2008). Pleistocene evolutionary history of the Clouded Apollo (Parnassius mnemosyne): genetic signatures of climate cycles and a "time-dependent" mitochondrial substitution rate. *Molecular Ecology*, *17*(19), 4248–4262. doi:10.1111/j.1365-294X.2008.03901.x

Habel, J. C., Schmitt, T., & Müller, P. (2005). The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: the phylogeography of the Marbled White butterfly (Satyrinae, Lepidoptera). *Journal of Biogeography*, *32*(8), 1489–1497. doi:10.1111/j.1365-2699.2005.01273.x

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*(1), 9.

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, *8*(5), 461–467. doi:10.1111/j.1461-0248.2005.00739.x

Heath, J. (1981). Threatened Rhopalocera (butterflies) in Europe. *Veroffentlichungen fur Naturschutz und Landschaftspflege in Baden-Wurttemberg. Beiheft*. http://agris.fao.org/agris-search/search.do?recordID=US201302579729. Accessed 14 December 2015

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907–913. doi:10.1038/35016000

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, *58*(3), 247–276. doi:10.1006/bijl.1996.0035

Hewitt, G. M. (2001). Speciation, hybrid zones and phylogeography — or seeing genes in space and time. *Molecular Ecology*, *10*(3), 537–549. doi:10.1046/j.1365-294x.2001.01202.x

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *359*(1442), 183–195. doi:10.1098/rstb.2003.1388

Hughes, J. B., Daily, G. C., & Ehrlich, P. R. (1997). Population Diversity: Its Extent and Extinction. *Science*, *278*(5338), 689–692. doi:10.1126/science.278.5338.689

Jaarola, M., & Searle, J. B. (2002). Phylogeography of field voles (Microtus agrestis) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, *11*(12), 2613–2621. doi:10.1046/j.1365-294X.2002.01639.x

Konvička, M., & Kuras, T. (1999). Population Structure, Behaviour and Selection of Oviposition Sites of an Endangered Butterfly, Parnassius Mnemosyne, in Litovelské

Pomoravíl. Czech Republic. *Journal of Insect Conservation*, *3*(3), 211–223. doi:10.1023/A:1009641618795

Kotlík, P., Deffontaine, V., Mascheretti, S., Zima, J., Michaux, J. R., & Searle, J. B. (2006). A northern glacial refugium for bank voles (Clethrionomys glareolus). *Proceedings of the National Academy of Sciences*, *103*(40), 14860–14864. doi:10.1073/pnas.0603237103

Maes, D., & Dyck, H. V. (2005). Habitat quality and biodiversity indicator performances of a threatened butterfly versus a multispecies group for wet heathlands in Belgium. *Biological Conservation*, *123*(2), 177–187. doi:10.1016/j.biocon.2004.11.005

Maes, D., & Van Dyck, H. (2001). Butterfly diversity loss in Flanders (north Belgium): Europe's worst case scenario? *Biological Conservation*, *99*(3), 263–276. doi:10.1016/S0006-3207(00)00182-8

Maes, D., Vanreusel, W., Talloen, W., & Dyck, H. V. (2004). Functional conservation units for the endangered Alcon Blue butterfly Maculinea alcon in Belgium (Lepidoptera: Lycaenidae). *Biological Conservation*, *120*(2), 229–241. doi:10.1016/j.biocon.2004.02.018

Magri, D. (2008). Patterns of post-glacial spread and the extent of glacial refugia of European beech (Fagus sylvatica). *Journal of Biogeography*, *35*(3), 450–463. doi:10.1111/j.1365-2699.2007.01803.x

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., et al. (2006). A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, *171*(1), 199–221. doi:10.1111/j.1469-8137.2006.01740.x

Magyari, E. (2002). Holocene biogeography of Fagus sylvatica L. and Carpinus betulus L. in the Carpathian-Alpine Region. *Folia Historico – Naturalia Musei Matrensis*, *26*, 15–35.

Malicky, H., Ant, H., Aspöck, H., De Jong, R., Thaler, K., & Varga, Z. (1983). Argumente zur Existenz und Chorologie mitteleuropäischer (extramediterran-europäischer) Faunen-Elemente. *Entomologia generalis*, *9*(1-2), 101–119.

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, *27*(2 Part 1), 209–220.

Meglécz, E., Nève, G., Pecsenye, K., & Varga, Z. (1999). Genetic variations in space and time in Parnassius mnemosyne (L.) (Lepidoptera) populations in north-east Hungary: implications for conservation. *Biological Conservation*, *89*(3), 251–259. doi:10.1016/S0006-3207(99)00006-3

Meglécz, E., Pecsenye, K., Peregovits, L., & Varga, Z. (1997). Allozyme variation in Parnassius mnemosyne (L.) (Lepidoptera) populations in North-East Hungary: variation within a subspecies group. *Genetica*, *101*(1), 59–66. doi:10.1023/A:1018368622549

Meglécz, E., Pecsenye, K., Varga, Z., & Solignac, M. (1998). Comparison of Differentiation Pattern at Allozyme and Microsatellite Loci in Parnassius Mnemosyne (Lepidoptera) Populations. *Hereditas*, *128*(2), 95–103. doi:10.1111/j.1601-5223.1998.00095.x

Moritz, C. (1994). Defining "Evolutionarily Significant Units" for conservation. *Trends in Ecology & Evolution*, *9*(10), 373–375. doi:10.1016/0169-5347(94)90057-4

New, T. R. (1993). Introduction to the biology and conservation of the Lycaenidae. In *Conservation biology of Lycaenidae (Butterflies)*. (Vol. 8, pp. 1–21.). Gland, Switzerland: IUCN.

Paradis, E. (2010). pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*(3), 419–420. doi:10.1093/bioinformatics/btp696

Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J. K., Thomas, C. D., Descimon, H., et al. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, *399*, 579–583.

Peakall, R., & Smouse, P. E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288–295. doi:10.1111/j.1471-8286.2005.01155.x

Petit, R. J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., et al. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science (New York, N.Y.)*, *300*(5625), 1563–1565. doi:10.1126/science.1083264

Pinceel, J., Jordaens, K., Pfenninger, M., & Backeljau, T. (2005). Rangewide phylogeography of a terrestrial slug in Europe: evidence for Alpine refugia and rapid colonization after the Pleistocene glaciations. *Molecular Ecology*, *14*(4), 1133–1150. doi:10.1111/j.1365-294X.2005.02479.x

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

QGIS Development Team. (2014). *Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org*.

R Core Team. (2014). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.*

Ryder, O. A. (1986). Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, *1*(1), 9–10. doi:10.1016/0169-5347(86)90059-5

Schmitt, T., Giessl, A., & Seitz, A. (2002). Postglacial colonisation of western Central Europe by Polyommatus coridon (Poda 1761) (Lepidoptera: Lycaenidae): evidence from population genetics. *Heredity*, *88*(1), 26–34. doi:10.1038/sj.hdy.6800003

Schmitt, T., & Hewitt, G. M. (2004). The genetic pattern of population threat and loss: a case study of butterflies. *Molecular Ecology*, *13*(1), 21–31.

Schmitt, T., Rákosy, L., Abadjiev, S., & Müller, P. (2007). Multiple differentiation centres of a non-Mediterranean butterfly species in south-eastern Europe. *Journal of Biogeography*, *34*(6), 939–950. doi:10.1111/j.1365-2699.2006.01684.x

Schmitt, T., & Seitz, A. (2008). Intraspecific allozymatic differentiation reveals the glacial refugia and the postglacial expansions of European Erebia medusa (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, *74*(4), 429 – 458. doi:10.1111/j.1095-8312.2001.tb01404.x

Schmitt, T., & Varga, Z. n. (2012). Extra-Mediterranean refugia: The rule and not the exception? *Frontiers in Zoology*, *9*(1), 22–33. doi:10.1186/1742-9994-9-22

Stewart, J. R., & Lister, A. M. (2001). Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, *16*(11), 608–613. doi:10.1016/S0169-5347(01)02338-2

Stinchcombe, J. R., & Hoekstra, H. E. (2007). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, *100*(2), 158–170. doi:10.1038/sj.hdy.6800937

Taberlet, Fumagalli, Wust-Saucy, Cosson, & Taberlet. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*(4), 453–464.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729. doi:10.1093/molbev/mst197

Thomas, J. A. (2005). Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1454), 339–357. doi:10.1098/rstb.2004.1585

Todisco, V., Gratton, P., Cesaroni, D., & Sbordoni, V. (2010). Phylogeography of Parnassius apollo: hints on taxonomy and conservation of a vulnerable glacial butterfly invader. *Biological Journal of the Linnean Society*, *101*(1), 169–183. doi:10.1111/j.1095-8312.2010.01476.x

Ursenbacher, S., Carlsson, M., Helfer, V., Tegelström, H., & Fumagalli, L. (2006). Phylogeography and Pleistocene refugia of the adder (Vipera berus) as inferred from

mitochondrial DNA sequence data. *Molecular Ecology*, *15*(11), 3425–3437. doi:10.1111/j.1365-294X.2006.03031.x

Välimäki, P., & Itämies, J. (2003). Migration of the clouded Apollo butterfly Parnassius mnemosyne in a network of suitable habitats – effects of patch characteristics. *Ecography*, *26*(5), 679–691. doi:10.1034/j.1600-0587.2003.03551.x

van Swaay, C. A. M. (1990). An assessment of the changes in butterfly abundance in The Netherlands during the 20th Century. *Biological Conservation*, *52*(4), 287–302. doi:10.1016/0006-3207(90)90073-X

van Swaay, C. M., & Warren, M. (1999). *Red Data Book of European Butterflies (Rhopalocera)*. Council of Europe.

van Swaay, C. M., Cuttelod, A., Collins, S., Maes, D., Lopez Munguira, M., Šašić, M., et al. (2010). European red list of Butterflies. *http://library.wur.nl/WebQuery/clc/1939351*. http://library.wur.nl/WebQuery/clc/1939351. Accessed 16 December 2015

van Swaay, C. M., Warren, S., & Löis, G. (2006). Biotope use and trends of European butterflies. *Journal of Insect Conservation*, *10*, 189–209.

Varga, Z. (1977). Das Prinzip der areal-analytischen Methode in der Zoogeographie und die Faunenelemente-Einteilung der europäischen Tagschmetterlinge (Lep.: Diurna). *Acta Zoologica Debrecina*, *14*, 223–285.

Vila, M., Lundhagen, A. C., Thuman, K. A., Stone, J. R., & Bjorklund, M. (2006). A new conservation unit in the butterfly Erebia triaria (Nymphalidae) as revealed by nuclear and mitochondrial markers. *Annales Zoologici Fennici*, *43*, 72–79.

Wahberg, 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., & Wheat, C. W. (2008). Genomic Outposts Serve the Phylogenomic Pioneers: Designing Novel Nuclear Markers for Genomic DNA Extractions of Lepidoptera. *Systematic Biology*, *57*(2), 231–242. doi:10.1080/10635150802033006

Waples, R. S. (1991). Pacific salmon, oncorhynchus spp., and the definition of... *Marine Fisheries Review*, *53*(3), 11.

Warren, M. S., & Key, R. S. (1991). Woodlands: past, present and potential for insects. In *The conservation of insects and their habitats.* (pp. 155–212.). London: Academic Press.