Ann. Zool. Fennici 43: 285–294 ISSN 0003-455X
Helsinki 28 June 2006

Allopatric origins of sympatric forms: the skippers *Carterocephalus palaemon palaemon*, *C. p. tolli* and *C. silvicolus*

Mirosław Ratkiewicz[1] & Bogdan Jaroszewicz[2]

[1)] *Institute of Biology, University of Białystok, Świerkowa 20B, PL-15-950 Białystok, Poland (e-mail: ermi@uwb.edu.pl)*
[2)] *Białowieża Geobotanical Station, Warsaw University, Sportowa 19, PL-17-230 Białowieża, Poland (e-mail: bogdanjaroszewicz@wp.pl)*

Received 1 Apr. 2006, revised version received 19 Apr. 2006, accepted 19 Apr. 2006

Ratkiewicz, M. & Jaroszewicz, B. 2006: Allopatric origins of sympatric forms: the skippers *Carterocephalus palaemon palaemon*, *C. p. tolli* and *C. silvicolus*. — *Ann. Zool. Fennici* 43: 285–294.

The taxonomic status and evolutionary origin of subspecies is often controversial. The *Carterocephalus palaemon* (Pallas, 1771) subspecies complex (Chequered skipper; Lepidoptera, Hesperiidae) may be a model for taxonomic clarification by means of genetic studies. We analysed two subspecies: *Carterocephalus palaemon palaemon* and *C. palaemon tolli* (the latter one endemic to the Białowieża Primeval Forest, Poland) at 34 allozyme loci and compared them with the northern chequered skipper *Carterocephalus silvicolus* (Meigen, 1829), a well-defined species. The overall F_{ST} value between *C. palaemon* complex and *C. silvicolus* was significant (0.700, $P < 0.001$). Significant genetic differentiation ($F_{ST} = 0.191$, $P < 0.001$) was found between sympatrically occurring forms: *C. p. palaemon* and *C. p. tolli*, confirming their phenotypic distinction. PCA and admixture analyses revealed the existence of three distinct gene pools (*C. silvicolus*, *C. p. palaemon* and *C. p. tolli*). The genetic differentiation between *C. palaemon* complex and *C. silvicolus* was similar to that found among the other well-defined species, while the F_{ST} value between the two genetic lineages of *C. palaemon* is typical for a butterfly species with subspecific structure. Observed differentiation between *C. p. palaemon* and *C. p. tolli* most likely results from long-lasting isolation between lineages rather than resulting from an initial divergence driven by strong disruptive selection.

Introduction

The mechanisms of speciation and the evolution of reproductive isolation are of great interest in evolutionary biology (Rice & Hostert 1993). With allopatric speciation, new species may be formed after the acquisition of a generalised incompatibility arising between physically separated populations. The uniform divergence across the genome that evolved in allopatry might be maintained following secondary contact due to the genetic differentiation at many loci, or may be eroded by introgression (Barton & Hewitt 1985). With sympatric speciation, new species are formed via allele frequency differentiation at only a few loci, without any physical barriers

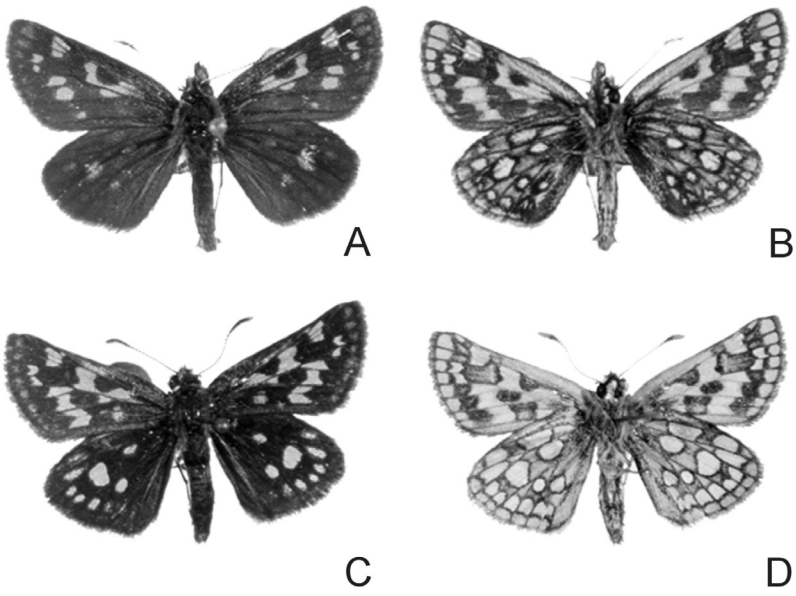

Fig. 1. Morphological differences between *C. palaemon tolli* (**A**: upper-side, **B**: underside) and *Carterocephalus p. palaemon* (**C**: upper-side, **D**: underside) in the Białowieża Primeval Forest, NE Poland.

to gene flow, while most of the genomes are not differentiated (Rice & Hostert 1993). As in the case of allopatric speciation, at the subspecies level a relatively small proportion of loci may be involved in the initial phase of differentiation. For example, the allele distribution at only one locus (*Idh-1*) out of 20 studied allowed the division of *Polyommatus coridon* into western and eastern genetic lineages that evolved in two different refugia (Adriato- and Ponto-Mediterranean regions; Schmitt & Seitz 2001b).

The distinction between these two evolutionary mechanisms (sympatric *versus* allopatric divergence) is important for understanding the speciation process. However, the issue of uniform *vs*. restricted differentiation among natural populations is rarely addressed. For example, the AFLP analysis performed on the intertidal snail *Littorina saxatilis* populations was consistent with expectations of sympatric speciation (Wilding *et al*. 2001). However, the pattern observed could also be the result of differential introgression after secondary contact. Sympatric evolution was also tested in the case of the host-adapted forms of *Tetramesa hyalipennis* reared from the grasses *Elymus repens* and *E. farctus* (Al-Barrak *et al*. 2004). The two forms of *T. hyalipennis* appeared to be closely related and, since the populations examined were all sympatric without obvious physical barriers to reproduction, the authors concluded that some degree of sympatric evolution has occurred. For the host races of the larch budmoth *Zeiraphera diniana*, Emelianov *et al*. (1995) showed that genetic differentiation between larch and pine forms was much greater than those between localities within larch and within pine forms. Moreover, only three out of 24 allozyme loci were involved in the differences between the two forms, probably in sympatry. If so, host-adapted forms of *Tetramesa hyalipennis* and the host races of the larch budmoth *Zeiraphera diniana* could be rare examples of sympatric speciation in insects.

The butterfly species chequered skipper *Carterocephalus palaemon* (Pallas, 1771) is widely distributed and within its holarctic range it contains several subspecies and morphological forms (Higgins & Riley 1970, Howe 1975: 507–508, Devyatkin 1997). In Europe, a nominative form *Carterocephalus palaemon palaemon* is present (Buszko & Masłowski 1993). The Białowieża Primeval Forest is the only place known where two different subspecies of chequered skipper: *C. p. palaemon* and *C. p. tolli* occur sympatrically. The latter was described as a subspecies endemic for that forest (Krzywicki 1967). The main differences between the studied *C. p. tolli* and nominative *C. p. palaemon* are (Fig. 1): larger wingspan (*C. p. palaemon* to *C. p. tolli* is like 1:1.125), drastic reduction of yellow pattern

on the upperside of both wings, white spots on the underside of the rear wings (*C. p. palaemon* has yellow ones), darker underwings, and violet shade of the brown background of the upperside of the wings. The front wings of *C. p. tolli* are longer and narrower than *C. p. palaemon*; they are not as pointed and the outer edges are more round than in the nominative form (Krzywicki 1967).

Besides morphological differences, differences in behaviour and ecology exist between the subspecies. *C. p. tolli* prefers more humid habitats than the nominative form. *C. p. tolli* butterflies usually rear a few days later than those of *C. p. palaemon*. *C. p. tolli* has lower light demands and thus it is able to settle on smaller clearings, narrower roads and penetrate the interior of thinned woodstands. Males of *C. p. tolli* are searching for females "patrolling" light spots inside of woodstands. In contrast, males of *C. p. palaemon* are rather "perchers" (*sensu* Scott 1974); they never patrol the area and do not enter the interior of woodstands (Jaroszewicz 2000). Table 1 presents the morphological, ecological, and behavioural differences between the studied forms.

Due to its sympatric occurrence with *C. p. palaemon*, the taxonomic status of *C. p. tolli* was controversial from the moment it was described by Krzywicki (1967). Natural selection is presumed to be the driving force behind the occurrence of phenotypically and genetically divergent populations in sympatry. From conventional definitions under the phylogenetic species concept (Templeton 1989), two sympatric forms should, if conspecific, be a part of the same evolutionary unit and thus have the same gene frequencies at each locus. Even if these forms were subject to some sort of assortative mating or selection pressure, the gene frequencies should, due to recombination, remain similar (Wright 1978). Thus, significant differentiation at any locus between two sympatric forms demonstrates a barrier to gene flow and at least partial prezygotic or postzygotic isolation. In this context, it is of great interest to establish whether genetic differentiation between *C. p. palaemon* and *C. p. tolli* in the Białowieża Forest is the sole product of divergent natural selection in sympatry as proposed by Krzywicki (1967). If so, the divergence of a restricted set of gene loci is expected. Alternatively, genetic divergence may be partly caused by unique histories and the two forms that have evolved in allopatry established a secondary contact in the Białowieża Forest.

Material and methods

Individuals of *Carterocephalus p. palaemon*, *C. p. tolli* and *C. silvicolus* (Meigen, 1829) were sampled at the end of May 1999 in the southern part of the Białowieża Primeval Forest (NE Poland, at the Polish/Belorussian border; 23°50´E, 52°42´N). Owing to its geographical location, nature preservation, and diversified landscape, the Białowieża Forest plays an important role in shaping the species ranges at the edge

Table 1. Morphological and behavioural differences between *Carterocephalus palaemon palaemon* and *C. p. tolli* in Białowieża Primeval Forest, NE Poland (after Krzywicki 1967, Jaroszewicz 2000).

Feature	*C. p. palaemon*	*C. p. tolli*
Wingspan (ratio 1:1.125)	lesser	larger
Shape of front wings	wider and shorter	narrower and longer
Top of front wings	pointed	not pointed
Edge of front wings	curved	straighter
Wings upper side background	brown	brown with violet
Wings underside background	yellow/light brown	darker
Yellow pattern of upper side	well-developed	reduced
Spots on underside of hind wings	yellow	white/silver
Habitat	fresh	dump
Light demands	higher	lower
Mating behaviour of males	perching	perching and patrolling
Rearing	earlier	later

of western and eastern Europe (Faliński 1994, Gutowski & Jaroszewicz 2004). This area is also important for the distribution of butterflies' ranges (Buszko *et al*. 1996).

A total of 34 *Carterocephalus* specimens: 11 of *C. silvicolus*, 12 *C. p. palaemon* and 11 *C. p. tolli*, identified by morphology, were used for the study. Butterflies were stored at –80 °C until electrophoresed. Twenty-five enzyme systems representing 34 loci were analysed using starch gel (Murphy *et al*. 1996) and cellulose acetate electrophoreses (Richardson *et al*. 1986). All enzyme systems were scored from thorax extractions, and extractions from the head were used for the phosphoglucomutase (PGM) enzyme. The full list of loci studied is available from the authors. A locus was considered polymorphic if more than one allele was detected. Diversity parameters were calculated using FSTAT version 2.9.3 (available at http://www2.unil.ch/popgen/softwares/fstat.htm). The following measures of diversity were calculated: percent of polymorphic loci (*P*), the allelic richness (*A*), and the average observed (H_O) and expected (H_E) heterozygosities. Deviations from Hardy-Weinberg expectations (HWE) were tested in every sample using GENEPOP version 3.4 (Raymond & Rousset 1995). A Markov chain method was used to estimate the exact *P* value. Tests on linkage disequilibria were performed with the help of GENEPOP. Wherever significance testing occurred, the sequential Bonferroni technique was employed (Rice 1989).

Genetic distances among the samples were measured using Nei's (1978) genetic distance. *F*-statistics were estimated for each polymorphic locus and over all loci. Means and standard deviations of F_{ST} estimates were obtained by jacknifing over loci (FSTAT). This may allow one to detect outlier loci (Luikart *et al*. 2003), i.e., loci under evolutionary forces. We used simulations to predict the expected distribution of genetic differentiation (F_{ST}) across loci for a given average divergence. Simulation was necessary because differentiation across loci is influenced by genetic drift and experimental sampling. We used 10 000 simulations in order to identify markers that may indicate deviation from neutral expectations in pairwise comparisons of diverging populations using DETSEL software (Vitalis *et al*. 2002).

Evidence of group distinction was obtained using two different approaches. Principle component analysis (PCA) was performed on individual genotypes to investigate spatial patterns of genetic variation. To accomplish this we used PCAGEN version 1.2 (available at http://www2.unil.ch/popgen/softwares/pcagen.htm). Population structure was assessed using a model-based Bayesian procedure, implemented in STRUCTURE version 2.0 (Pritchard *et al*. 2000). This model was designed to identify the *K* (unknown) populations (genetic clusters) of the origin of individuals, and simultaneously to assign the individuals to the populations. No prior population information concerning the origin of individuals was used. The results presented in this study are based on 1 000 000 iterations, following a burn-in period of 100 000 iterations.

Results

No significant genotypic disequilibrium was detected for any locus pair of the three studied taxa. No significant deviations from Hardy-Weinberg equilibrium were observed, except for the glucosephosphate isomerase (*Gpi-1*) locus in *C. p. tolli*. Thirty-four studied loci possessed 71 alleles and 19 of these loci were polymor-

Table 2. *Carterocephalus* samples examined: *C. silvicolus*, *C. palaemon palaemon*, *C. palaemon tolli*. Number of individuals (*N*), allelic richness (*A*), percentage of polymorphic loci (*P*), observed (H_O) and expected heterozygosity (H_E) of examined *Carterocephalus* samples from Białowieża Forest, NE Poland.

No.	Population	*N*	*A*	*P* (%)	H_O	H_E
1	*C. silvicolus*	11	1.12	11.8	0.02	0.03
2	*C. palaemon palaemon*	12	1.65	41.2	0.14	0.16
3	*C. palaemon tolli*	11	1.62	41.2	0.11	0.15

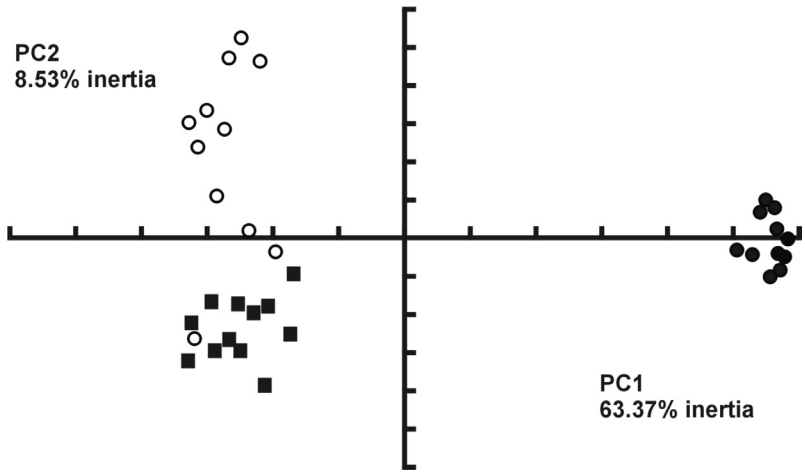

Fig. 2. Principal component analysis (PCA) plot showing the genetic differences among three different clusters. The two principal components used in this plot (PC1 and PC2) account for 71.9% of the total variance. Filled circles: *C. silvicolus*, squares: *C. p. palaemon*, open circles: *C. p. tolli*.

phic. The parameters of genetic diversity (i.e. allelic richness, percentage of polymorphic loci, expected and observed heterozygosity) were much higher in the two subspecies of *C. palaemon* than in *C. silvicolus* (Table 2).

Six loci were species-diagnostic for *C. silvicolus* and were fixed for the following alleles: *Acp-2*^{B}, *Es-D*^{A}, *Hbdh*^{B}, *6Pgd*^{D}, *Pgm-1*^{B} and *Pgm-2*^{C}. Fifteen private alleles, representing 11 loci were found to be present exclusively either in *C. p. palaemon* or *C. p. tolli*. The eight alleles were *C. p. palaemon*-specific and seven others were present exclusively in *C. p. tolli*. All allele frequencies are presented in the Appendix. The Nei's (1978) genetic distances between *C. silvicolus* and *C. p. palaemon* and between *C. silvicolus* and *C. p. tolli* were 0.413 and 0.425, respectively. The Nei's distance between *C. p. palaemon* and *C. p. tolli* was 0.045. Global genetic differentiation among *C. silvicolus* and *C. palaemon* was significant (F_{ST} = 0.700 ± 0.067, 95% CI = 0.553–0.813, P < 0.001). Substantial genetic differentiation between *C. p. palaemon* and *C. p. tolli* was found (F_{ST} = 0.191 ± 0.046, 95% CI = 0.096–0.266, P < 0.001). Four (*Aat-2*, *Ak-2*, *Hk-1* and *Me-1*) out of 16 variable loci in *C. p. palaemon* and *C. p. tolli* showed significant genetic differentiation. The F_{ST} values for these four loci were larger than the mean (Table 3) and were identified as differentiated using simulations.

Principle component analysis (PCA) generated two principle components which cumulatively explained 71.9% of the total genetic diversity. The PCA showed a clear separation between *C. silvicolus* and both *C. p. palaemon* and *C. p. tolli* on PC-1, while *C. p. palaemon* and *C. p. tolli* were clearly distinct on PC-2 (Fig. 2). Two possible hybrids between them were identified. A single individual that had *C. p. tolli* morphology was identified as *C. p. palaemon* using the allelic pattern in PCA analysis (Fig. 2).

The Bayesian genetic analyses showed K = 3 populations (Ln PrX/K = –487), suggesting that the pooled *Carterocephalus* "population" was

Table 3. Average F_{ST}, F_{IS} and F_{IT} values for each locus and over all 16 variable loci between *C. palaemon palaemon* and *C. palaemon tolli* occurring in Białowieża Primeval Forest, NE Poland.

Locus	F_{ST}	F_{IS}	F_{IT}
6Pgd	–0.011	0.691	0.687
Mpi	0.085	0.247	0.312
Me-1	0.368#***	0.242	0.521
Pgm-1	0.066	–0.086	–0.014
Gpi	0.087	0.173	0.245
Idh-1	0.104	0.636	0.674
Idh-2	–0.072	0.528	0.494
αGpd	0.075	–0.066	0.014
Ak-2	0.385#**	0.162	0.485
Aat2	0.317#****	0.444	0.620
Aat1	0.029	0.080	0.110
Mdh-2	–0.002	–0.001	–0.022
Pep-1	0.155	–0.034	0.126
Hk-1	0.343#**	–0.274	0.163
Acon	0.081	0.072	0.147
Pgm-2	0.175	0.110	0.265
All loci	0.191 ± 0.046***	0.199 ± 0.059***	0.352 ± 0.061

F_{ST} values significantly larger than a mean, * P < 0.05, ** P < 0.01, *** P < 0.001 (Bonferroni corrected).

heterogeneous and may contain at least three genetically distinct groups. We therefore estimated the proportion of membership (q) of each predefined group into three clusters representing the ''cryptic'' genetic populations. Cluster 1 grouped the *C. p. palaemon* sample, cluster 2 grouped the *C. p. tolli* sample, and cluster 3 grouped the *C. silvicolus* sample (with q ranging from 0.890 to 0.996, Table 4). These results are in accordance with the nongenetic classification, and are independent of any prior population information. The admixture approach identified two admixed (hybrid) individuals, and a single individual that was incorrectly classified as a *C. p. tolli* by its phenotype (data not shown).

Discussion

The objectives of this study were to compare the genetic diversity among three taxa belonging to *Carterocephalus* genus (*C. p. palaemon*, *C. p. tolli* and *C. silvicolus*) and to estimate the extent of genetic differentiation among them. We detected reduced genetic variability in *C. silvicolus* and the presence of three genetically differentiated clusters.

Genetic diversity within clusters

Our results indicate pronounced genetic differences between *C. silvicolus* and both taxa of *C. palaemon*. The population of the northern chequered skipper, *Carterocephalus silvicolus,* was characterized by low expected heterozygosity, allelic richness, and percentage of polymorphic loci (H_E = 3%, A = 1.12, P = 11.8%) in comparison with *C. palaemon* (H_E = 15%–16%, A = 1.62–1.65, P = 41.2%). The level of allozyme diversity in *C. silvicolus* was considerably lower than in other common and widespread European butterfly species; e.g. *Melanargia galathea/lachesis* (H_E = 13%–21%, A = 1.83–2.22, P = 33.3%–44.4%; Habel *et al.* 2005), *Zygaena exulans* (H_E = 6%–17%, A = 1.27–2.27, P = 26.7%–86.7%; Schmitt & Hewitt 2004a) and *Erebia medusa* (H_E = 14.3%, A = 1.77, P = 45.4%; Schmitt & Seitz 2001a). The variability in *C. silvicolus* is similar to the values reported for *Coenonympha hero* (H_E = 0.3%–9.6%, A = 1.1–1.5, P = 10%–50%; Cassel & Tammaru 2003) — a rare species with a low effective population size. The low genetic diversity in *C. silvicolus* might be due to its peripheral position close to the southern and western limits of the distribution range and might reflect genetic impoverishment. Cassel and Tammaru (2003) — in a study of *C. hero* — showed that low genetic variability may be present in the species' core area and that the post-glacial colonization history could affect intraspecific diversity. Thus, reasons for the low genetic variation in the studied population of northern chequered skipper remain unclear and more populations from different parts of the species' range should be studied. Regardless of the factors involved, the low genetic diversity in *C. silvicolus* may result in reduced adaptability due to genetic impoverishment (Schmitt & Hewitt 2004b).

Both taxa of *C. palaemon* show upper intermediate genetic diversity values (*see* Table 2) when compared with values found in common and widespread butterfly taxa in Europe (Schmitt & Seitz 2001a, Schmitt & Hewitt 2004a, Habel *et al.* 2005). However, very common butterfly species have even larger diversity values, e.g. *Aglais urticae* (H_E = 24.8%, A = 2.84, P = 76%; Vandervoestijne *et al.* 1999), *Polyommatus coridon/hispana* (H_E = 20.5–27.5%, A = 2.65–2.79, P = 65%–77%; Schmitt *et al.* 2005) and *P. icarus* (H_E = 17.7%, A = 2.98, P = 78%; Schmitt *et al.* 2003). The high genetic diversity in *C. p. palaemon* is not surprising, as it is a widespread form. The variability in *C. p. tolli* was also high and this fails to support Krzywicki's

Table 4. Bayesian clustering analyses for the pooled *Carterocephalus* samples (34 individuals, 34 loci) from Białowieża Primeval Forest (NE Poland), performed using STRUCTURE (Pritchard *et al.* 2000). The proportion of membership of each predefined group in each of three inferred clusters (q_1, q_2 and q_3).

Sample	Cluster		
	1 (q_1)	2 (q_2)	3 (q_3)
C. silvicolus	0.002	0.002	0.996
C. palaemon palaemon	0.970	0.022	0.008
C. palaemon tolli	0.107	0.890	0.003

(1967) hypothesis of sympatric speciation in the Białowieża Forest and endemism of this form. The population of *C. p. tolli* must have a large effective population size and/or may experience significant gene flow with other populations. At present, there are no data regarding the occurrence of *C. p. tolli* outside the Białowieża Forest in Poland, however, the existence of undetected populations cannot be ruled out.

Genetic differentiation among clusters

The observed high differentiation between *C. silvicolus* and *C. palaemon* (F_{ST} = 0.700 ± 0.067, Nei's genetic distance = 0.413–0.425) supports their status as individual species. Such remarkable genetic differences suggest a very long period of independent evolution. The divergence between *C. silvicolus* and *C. palaemon* is much larger than, for example, between *Polyomnatus coridon* and *P. hispana* which are considered to be "good species" (F_{ST} = 0.16, Nei's D = 0.082; Schmitt *et al.* 2005).

The samples of *C. palaemon* split into two genetic lineages (*C. p. palaemon* and *C. p. tolli*). Both forms are characterized by many private alleles (*see* Appendix). The mean Nei's (1978) genetic distance (D = 0.045) and genetic differentiation (F_{ST} = 0.191) between these two lineages is similar to other butterfly species with an accepted subspecific structure, e.g. *Erebia medusa* (F_{ST} = 0.149, D = 0.051–0.117; Schmitt & Seitz 2001a), *Polyommatus coridon* (F_{ST} = 0.06, D = 0.041; Schmitt & Seitz 2001b), *P. coridon/hispana* complex (Schmitt *et al.* 2005). The F_{ST} value between *C. p. palaemon* and *C. p. tolli* was higher than in the majority of other studies with a few exceptions, e.g. *Erebia epiphron* (F_{ST} = 0.291, D = 0.071; Schmitt *et al.* 2006) or *Aglaope infausta* (F_{ST} = 0.404, Nei's D = 0.137; Schmitt & Seitz 2004). Thus, we assume that the samples of *C. palaemon* belong to two different subspecies or may be "semispecies" of a superspecies complex. This pattern of genetic distinctiveness is also supported by morphometry as *C. p. palaemon* and *C. p. tolli* differ with respect to the several morphological characters (Table 1 and Fig. 1; Krzywicki 1967). The studied forms also differ with respect to ecological preferences, time of rearing and mating tactics (Table 1; Krzywicki 1967, Jaroszewicz 2000).

The most likely explanation for the genetic differentiation between *C. p. palaemon* and *C. p. tolli* is assortative mating that may have originated in allopatry. This is supported by the relatively high (25%) number of outlier loci that behave unlike most other loci (e.g. those with an extremely high F_{ST}). Thus, in the case of *C. p. palaemon* and *C. p. tolli*, a fairly large portion of their variable genomes has acquired generalised differences. Other studies have shown a lower proportion of highly differentiated loci, e.g. host races of larch budmoth (12.5% out of 24 allozyme loci and 17.6% out of the 1291 AFLP loci; Emelianov *et al.* 1995, 2003) and parapatric morphs of intertidal snail (5% out of 306 AFLP loci; Wilding *et al.* 2001).

The alternative hypothesis of an initial divergence of a small part of a genome driven by strong disruptive selection seems to be less probable and the differentiation observed does not seem to be the sole product of divergent natural selection in sympatry. Similarly, the sympatric origin of *Polyommatus slovacus* with *P. coridon* was not supported using an allozyme study (Schmitt *et al.* 2005). The length of time that habitats have been suitable for *C. palaemon* in the Białowieża Primeval Forest (ca. 8500 years BP, Faliński 1994) seems too short to generate the observed genetic divergence between *C. p. palaemon* and *C. p. tolli*. If so, the two forms must have originated in different refugia, as hypothesized for other butterfly species (Schmitt & Hewitt 2004b). This is in direct contrast to the hypothesis of Krzywicki (1967). Schmitt and Seitz (2001a) supposed that differentiation of *Erebia medusa* (F_{ST} = 0.149, D = 0.051–0.117) into four genetic lineages took place in different refugia during the Würm glacial period. The somewhat larger F_{ST} value (0.191) between two lineages of *C. palaemon*. may indicate that differentiation between them predates last glacial period. Similarly to *E. medusa*, *C. p. palaemon* could have survived the last glacial period in a refugium situated in the eastern Balkans, while *C. p. tolli* had a more eastern refugium in Asia (southern Urals, e.g. *C. hero*; Cassel & Tammaru 2003) or Caucasus. The latter hypothesis can be further investigated with more samples, especially from the eastern parts of *C. palaemon* range. If the

hypothesis is correct, *C. p. tolli* would be a relic of limited postglacial expansion that was able to survive (until now) in the colder and more humid areas of the Białowieża Primeval Forest.

PCA and admixture analyses

The traditional F_{ST}-based approach for estimating divergence was reinforced by PCA and admixture analyses by detecting dispersers genetically. Both PCA and genetic admixture analyses were in accordance with the morphological classification, however they performed better than phenotypic analysis. When samples were assigned to three inferred clusters without using any prior population information, the *C. p. palaemon*, *C. p. tolli* and *C. silvicolus* individuals were assigned to clusters 1, 2 and 3, respectively, with the average proportion of individual memberships $q \geq 0.89$ and without significant ancestry in the other clusters, except for *C. p. tolli*. Three *C. p. tolli* specimens had significant ancestry in cluster number 1 (*C. p. palaemon*). When the ancestry of misassigned individuals was further investigated, results suggested that two butterflies recognised as *C. p. tolli* by their wing pattern might be *C. p. palaemon–C. p. tolli* hybrids. This genetic distinction was thus important because many hybrids may be unrecognized in morphologically uniform subgroups such as skipper butterflies (Mallet 2005). Another *C. p. tolli* individual had 98% probability of belonging to the population of *C. p. palaemon*. Therefore, this individual may represent a misidentified sample, as *C. p. palaemon* is known to produce dark individuals, similar to *C. p. tolli*, in specific environmental conditions (Krzywicki 1967, Jaroszewicz 2000).

The identification of about 8.7% hybrid individuals (2 out of 23) suggests that despite a long period of sympatry and syntopy, hybridization is rather negligible and is most probably limited to particular areas at the ecological edges of the *C. p. palaemon* and *C. p. tolli* distribution. Despite some hybridization, the gene pools of *C. p. palaemon* and *C. p. tolli* seem to be distinct. Selection in the sympatric populations of *C. palaemon* in Białowieża Forest may be driving and maintaining a subspecific divergence and may lead to assortative mating and ultimately to speciation.

We noted some directionality in the introgression of *C. p. palaemon* into the *C. p. tolli* gene pool ($q = 0.107$), while introgression in the other direction (from *C. p. tolli* to *C. p. palaemon*) was almost absent ($q = 0.022$). Observed directional introgression (*C. p. palaemon* into *C. p. tolli* gene pool) suggests that *C. p. palaemon* represents a more closed system of co-adapted gene complexes than *C. p. tolli*. Such asymmetric patterns of introgression are commonly observed in hybrid zones between species (Harrison 1993). Asymmetric introgression may be explained by differences in ecological demands of both subspecies. We may assume that crossbreeding between them may take place during dry years, when habitats of *C. p. palaemon* become too dry, causing butterflies to disperse in search of humid places that are normally occupied by *C. p. tolli*.

Genetic analyses indicate a distinction of ecological forms (subspecies) that occur in sympatry, and suggest the resolution of evolutionary questions that concern this very special case. We suggest that, although limited, some introgression is still continuing after separation into ecologically, genetically, and morphologically differentiated subspecies. This is in line with the conclusion of Emelianov *et al*. (2003) that states: "[...] a long and relatively stable sympatric phase of genetic divergence in the presence of gene flow is a probable feature of speciation".

Acknowledgements

We thank Thomas Schmitt for comprehensive discussion and many valuable comments on the previous version of the manuscript. This project was supported by State Committee for Scientific Research Grant KBN No. 6P04C 074 18 to B. Jaroszewicz.

References

Al-Barrak, M., Loxdale, H. D., Brookes, C. P., Dawah, H. A., Biron, D. G. & Alsagair, O. 2004: Molecular evidence using enzyme and RAPD markers for sympatric evolution in British species of *Tetramesa* (Hymenoptera: Eurytomidae). — *Biol. J. Linn. Soc.* 83: 509–525.

Barton, N. H. & Hewitt, G. M. 1985: Analysis of hybrid zones. — *Ann. Rev. Ecol. Syst.* 16: 113–148.

Buszko, J. & Masłowski, J. 1993: *Atlas motyli Polski. Część I. Motyle dzienne* (*Rhopalocera*): 256–257. Grupa Image, Warszawa.

Buszko, J., Kokot, A., Palik, E. & Śliwiński, Z. 1996: Motyle większe (*Macrolepidoptera*) Puszczy Białowieskiej. — *Parki nar. Rez. przyr*. 15(4): 3–46.

Cassel, A. & Tammaru, T. 2003: Allozyme variation in central, peripheral and isolated populations of the scarce heath (Coenonympha hero: Lepidoptera, Nymphalidae): Implications for conservation. — *Cons. Genet*. 4: 83–93.

Devyatkin, A. L. 1997: Family *Hesperiidae* Latreille, 1809. — In: Tuzov, V. K. (ed.), *Guide to the butterflies of Russia and adjacent territories (Lepidoptera, Rhopalocera)*, vol. 1: 105–133, 270–271, 425. Pensoft, Sofia–Moskwa.

Emelianov, I., Mallet, J. & Baltensweiler, W. 1995: Genetic differentiation in *Zeiraphera diniana* (Lepidoptera, Tortricidae, the larch budmoth) — polymorphism, host races or sibling species? — *Heredity* 75: 416–424.

Emelianov, I., Marec, F. & Mallet, J. 2003: Genomic evidence for divergence with gene flow in host races of the larch budmoth. — *Proc. R. Soc. Lond. B* 271: 97–105.

Faliński, J. B. 1994: Concise geobotanical atlas of Białowieża Forest. — *Phytocoenosis* 6: 3–34.

Gutowski, J. M. & Jaroszewicz, B. 2004: Białowieża Primeval Forest as a refuge of the European entomofauna. — *Wiad. Entomol*. 23 (Suppl. 2): 67–87.

Habel, J. C., Schmitt, T. & Müller, P. 2005. The fourth paradigm pattern of post-glacial range expansion of European terrestrial species: the phylogeography of the marbled white butterfly (Satyrinae, Lepidoptera). — *J. Biogeog*. 32: 1–9.

Harrison, R. G. 1993: *Hybrid zones and the evolutionary process*. — Oxford University Press, Oxford.

Higgins, L. G. & Riley, N. D. 1970: *A field guide to the butterflies of Britain and Europe*. — Collins, London.

Howe, W. H. 1975: *The butterflies of North America*. — Doubleday and Corp, Garden City, New York.

Jaroszewicz, B. 2000: *Zagadnienie odrębności taksonomicznej Carterocephalus palaemon tolli Krzywicki, 1967 w Puszczy Białowieskiej*. — Ph.D. thesis, University of Toruń.

Krzywicki, M. 1967: Fauna *Papilionoidea* i *Hesperioidea* (*Lepidoptera*) Puszczy Białowieskiej. — *Annales zool*. 25: 1–213.

Luikart, G., England, P. R., Tallmon, D., Jordan, S. & Taberlet, P. 2003: The power and the promise of population genomics: from genotyping to genome typing. — *Nature Review in Genetics* 4: 981–994.

Mallet, J. 2005: Hybridization as an invasion of the genome. — *Trends Ecol. Evol*. 20: 229–237.

Murphy, R. W., Sites, J. W., Buth, D. G. & Haufler, C. H. 1996: Proteins: isozyme electrophoresis. — In: Hillis, D. M., Moritz, C. & Mable, B. K. (eds.), *Molecular systematics*: 51–120. Sinauer, Sunderland.

Nei, M. 1978: Estimation of average heterozygosity and genetic distance from a small number of individuals. — *Genetics* 89: 583–590.

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

Raymond, M. L. & Rousset, F. 1995: GENEPOP (Version 1.2), a population genetics software for exact tests and ecumenism. — *J. Heredity* 86: 248–249.

Rice, W. R. 1989: Analysing tables of statistical tests. — *Evolution* 43: 223–225.

Rice, W. R. & Hostert, E. E. 1993: Laboratory experiments on speciation: what have we learned in 40 years? — *Evolution* 47: 1637–1653.

Richardson, B. J., Baverstock, P. R. & Adams, M. 1986: *Isoenzyme electrophoresis. A handbook for animal systematics and population studies*. — Academic Press. Inc., San Diego, California.

Schmitt, T. & Hewitt, G. M. 2004a: Molecular biogeography of the arctic-alpine disjunct burnet moth species *Zygaena exulans* (Zygaeindae, Lepidoptera) in the Pyrenees and Alps. — *J. Biogeogr*. 31: 885–893.

Schmitt, T. & Hewitt, G. M. 2004b: The genetic pattern of population threat and loss: a case study of butterflies. — *Mol. Ecol*. 13: 21–31.

Schmitt, T. & Seitz, A. 2001a: Intraspecific allozymatic differentiation reveals the glacial refugia and the postglacial expansion of European *Erebia medusa* (Lepidoptera: Nymphalidae). — *Biol. J. Linn. Soc*. 74: 429–458.

Schmitt, T. & Seitz, A. 2001b: Allozyme variation in *Polyommatus coridon* (Lepidoptera: Lycaenidae): identification of ice-age refugia and reconstruction of post-glacial expansion. — *J. Biogeogr.* 28: 1129–1136.

Schmitt, T. & Seitz, A. 2004: Low diversity but high differentiation: the population genetics of *Aglaope infausta* (Zygaenidae: Lepidpteda). — *J. Biogeogr*. 31: 137–144.

Schmitt, T., Giessl, A. & Seitz, A. 2003: Did *Polyommatus icarus* (Lepidoptera, Lyceanidae) have distinct glacial refugia in southern Europe? Evidence from population genetics. — *Biol. J. Linn. Soc*. 80: 529–538.

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

Schmitt, T., Varga, Z. & Seitz, A. 2005: Are *Polyommatus hispana* and *Polyommatus slovacus* bivoltine *Polyommatus coridon* (Lepidoptera: Lycaenidae)? The discriminatory value of genetics and taxonomy. — *Org. Div. & Evol*. 5: 297–307.

Scott, J. A. 1974: Mate-locating behavior of butterflies. — *Amer. Midl. Nat*. 91: 103–117.

Templeton, A. R. 1989: The meaning of species and speciation: a genetic perspective. — In: Otte, D. & Endler, J. A. (eds.), *Speciation and its consequences*: 3–27. Simauer Associates, Sunderland, Massachusetts.

Vanderwoestijne, S., Nève, G. & Baguette, M. 1999: Spatial and temporal population genetic structure of the butterfly *Aglais urticae* L. (Lepidoptera, Nymphalidae). — *Mol. Ecol*. 8: 1539–1543.

Vitalis, R., Dawson, K., Boursot, P. & Belkhir, K. 2002: DETSEL 1.0: a computer program to detect markers responding to selection. — *J. Heredity* 94: 429–431.

Wilding, C. S., Butlin, R. K. & Grahame, J. 2001: Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. — *J. Evol. Biol*. 14: 611–619.

Wright, S. 1978: *Evolution and the genetics of populations, vol. 4, Variability within and among natural populations*. — University of Chicago Press, Chicago.

Appendix. Allele frequencies of 19 variable loci in three samples of *Carterocephalus*: *C. silvicolus* (CS), *C. palaemon palaemon* (CP) and *C. palaemon tolli* (CT) occurring in Białowieża Primeval Forest, NE Poland.

Locus	Allele	CS	CP	CT	Locus	Allele	CS	CP	CT
Aat-1	A	0.000	0.115	0.200	*Idh-1*	A	0.000	0.000	0.150
	B	1.000	0.885	0.700		B	1.000	1.000	0.850
	C	0.000	0.000	0.100	*Idh-2*	A	0.000	0.846	0.850
Aat-2	A	0.000	0.654	0.150		B	1.000	0.154	0.150
	B	0.000	0.154	0.650	*Mdh-2*	A	0.000	0.962	0.950
	C	1.000	0.192	0.000		B	1.000	0.038	0.000
	D	0.000	0.000	0.200		C	0.000	0.000	0.050
Acon	A	0.000	0.000	0.300	*Me-1*	A	0.000	0.346	0.150
	B	1.000	0.500	0.400		B	0.318	0.500	0.050
	C	0.000	0.500	0.250		C	0.682	0.154	0.800
	D	0.000	0.000	0.050	*Mpi*	A	0.000	0.115	0.050
Acp-2	A	0.000	1.000	1.000		B	0.000	0.115	0.100
	B	1.000	0.000	0.000		C	0.045	0.462	0.800
Ak-2	A	0.000	0.846	0.350		D	0.935	0.231	0.000
	B	0.955	0.154	0.650		E	0.000	0.077	0.050
	C	0.045	0.000	0.000	*Pep-1*	A	1.000	0.308	0.050
Est-D	A	1.000	0.000	0.000		B	0.000	0.692	0.950
	B	0.000	1.000	1.000	*6Pgd*	A	0.000	0.000	0.000
α*Gpd*	A	0.000	0.000	0.100		B	0.000	0.808	0.900
	B	1.000	1.000	0.900		C	0.000	0.038	0.100
Gpi	A	0.000	0.077	0.000		D	1.000	0.000	0.000
	B	0.773	0.462	0.300	*Pgm-1*	A	0.000	0.885	1.000
	C	0.227	0.346	0.700		B	1.000	0.000	0.000
	D	0.000	0.115	0.000		C	0.000	0.115	0.000
Hbdh	A	0.000	1.000	1.000	*Pgm-2*	A	0.000	0.231	0.000
	B	1.000	0.000	0.000		B	0.000	0.731	1.000
Hk-1	A	0.000	0.885	0.450		C	1.000	0.000	0.000
	B	1.000	0.115	0.550		D	0.000	0.038	0.000