This manuscript is contextually identical with the following published paper:

Szivák, I., Mikes, T., Szalontai, B. et al. (2017) Ecological divergence of Chaetopteryx rugulosa species complex (Insecta, Trichoptera) linked to climatic niche diversificationHydrobiologia. 794: 31-47. doi:10.1007/s10750-016-3068-0

The original published PDF available in this website:

http://link.springer.com/article/10.1007%2Fs10750-016-3068-0

Ecological divergence of *Chaetopteryx rugulosa* species complex (Insecta, Trichoptera) linked to climatic niche diversification

Ildikó Szivák[1], Tamás Mikes[2,3], Bálint Szalontai[4], Mladen Kučinić[5], Ivan Vučković[6], Edit Vadkerti[7], Péter Kisfali[8], Steffen U. Pauls[2*], Miklós Bálint[2*]

1 Balaton Limnological Institute, MTA Centre for Ecological Research, H-8237 Tihany, Klebelsberg Kuno u. 3, Hungary

2 Senckenberg Biodiversity and Climate Research Centre, D-60325 Frankfurt am Main, Senckenberganlage 25, Germany

3 Verven 38, N-4014 Stavanger, Norway

4 Life and Brain Center, University of Bonn Medical Center, D - 53127 Bonn, Sigmund Freud Str. 25., Germany

5 Department of Biology, University of Zagreb, HR-10000 Zagreb, Rooseveltov trg 6, Croatia

6 Elektroprojekt, Civil and Arhitectural Dept., Environmental Protection and Waste Management, Alexandra von Humbolta 4, HR-10000 Zagreb, Croatia

7 College of Eötvös József, Water Supply and Environmental Engineering Institute, H-6500 Baja, Bajcsy-Zsilinszky Endre u. 14., Hungary

8 Department of Medical Genetics, University of Pécs, H-7624 Pécs, Szigeti út 12, Hungary

* equal contributions

28 Corresponding authors:

29 Ildikó Szivák: ildiko.szivak@okologia.mta.hu, +36 70 5472966

30 Miklós Bálint: mbalint@senckenberg.de, +49 (0)69 7542 1856

Abstract

Climate is often considered to be an important, but indirect driver of speciation. Indeed, environmental factors may contribute to the formation of biodiversity, but to date this crucial relationship remains largely unexplored. Here we investigate the possible role of climate, geological factors and biogeographical processes in the formation of a freshwater insect species group, the *Chaetopteryx rugulosa* species complex (Trichoptera) in the Western Balkans. We used multi-locus DNA sequence data to establish a dated phylogenetic hypothesis for the group. The comparison of the dated phylogeny with the geological history of the Western Balkans shows that lineage formation coincided with major past Earth surface and climatic events in the region. By reconstructing present-day habitat conditions (climate, bedrock geology) we show that the lineages of *C. rugulosa* species complex have distinct climatic but not bedrock geological niches. Without exception all splits associated with Pliocene/Pleistocene transition led to independent, parallel split into 'warm' and 'cold' sister lineages. This indicates a non-random diversification on the *C. rugulosa* species complex associated with late Pliocene climate in the region. We interpreted the results as the diversification of the species complex was mainly driven by ecological diversification linked to past climate change, along with geographical isolation.

Keywords

phylogeny; climate; Trichoptera; topography formation; *Chaetopteryx*; molecular clock

Introduction

Throughout the Pleistocene, climatic variations exerted a strong influence on the biogeography of Europe. In particular, the southern areas played a key role in both persistence (Hewitt, 2000, 2011) and formation of genetic lineages and even species (Levsen et al., 2012; April et al., 2013). Southern European areas often represented the refugia of many taxa during the cold and dry climates of ice ages. It is assumed that much of the endemic biota of southern Europe was formed during the repeated range expansion and contraction of species distribution (Malicky, 1983; Gómez & Lunt, 2007). These range expansions and contractions were mediated by glacial – interglacial cycles (Hewitt, 2000). Climate is generally perceived as an indirect, but important driver of species formation. On the other hand, climate may directly drive evolution (Oppold et al., 2016) and control the formation of new lineages and species via adaptation (Graham et al., 2004; Rissler & Apodaca, 2007; Kalkvik et al., 2012). Although climate and Earth surface processes in southern Europe have had a highly complex history, the role of environmental factors shaping biodiversity has received much less attention than the role of geographical isolation (Dijkstra et al., 2014; Previšić et al., 2014b).

The Balkan Peninsula is one of the three major southern European refugial areas, and it is recognized as a European biodiversity hotspot (Griffiths et al., 2004; Hewitt, 2011). The Balkan Peninsula is particularly renowned for high levels of endemism (Griffiths et al., 2004; Sotiropoulos et al., 2007), also in freshwaters (Bănărescu, 2004; Bilandžija et al., 2013; Oláh & Kovács, 2013; Previšić et al., 2009, 2014a, 2014b, 2016; Vitecek et al., 2015). Many of the endemic Mediterranean species and lineages are small-range endemics. The diversification of these lineages is often explained within allopatric speciation induced by range shifts caused by past climate change (Gómez & Lunt, 2007; Hewitt, 2011). However, strong differences are also common among the abiotic habitat conditions of allopatric lineages and species (e.g. Previšić et al., 2014b).

The link between ecological trait divergence and speciation is increasingly recognized for many taxa (Funk et al., 2006; Pauls et al., 2008; Statzner & Dolédec, 2011; Múrria et al., 2012; Bilandžija et al., 2013; Zhang et al., 2014) and ecological divergences may be important drivers of speciation (ecological speciation, Rundle & Nosil, 2005; Schluter, 2009). The local adaptation to different ecological conditions is accompanied by increasing reproductive isolation, since divergent selection acts on populations (Rundle & Nosil, 2005; Schluter, 2009). Strong gradients in selective pressures can lead to lineage divergence and new biological species (Schluter, 2009), and climatic conditions have been shown to speed up evolution (Oppold et al., 2016). Lineage divergence is promoted by adaptive life-history differences, such as the timing of emergence, or shift in diet (e.g. Feder, 1998; McPeek & Wellborn, 1998; Elmer et al., 2010a) or sexual selection,

79 particularly in mate choice. Consequently, ecological speciation could occur among closely related lineages
80 due to environmental niche divergence (e.g. Evans et al., 2009; Dormann et al., 2010; Statzner & Dolédec,
81 2011; Zhang et al., 2014) and under any geographic arrangement of populations (e.g. allopatric, parapatric,
82 sympatric; Graham et al., 2004; Rundle & Nosil, 2005; Nosil, 2012). In summary, ecological speciation could
83 be distinguished from other models of speciation based on the existence of ecologically mediated
84 divergences among populations that cause reproductive isolation (Rundle & Nosil, 2005). In the other type of
85 speciation models chance events play a central role such as speciation by incidental divergence in isolation,
86 hybridization, genetic drift and population bottlenecks (Graham et al., 2004; Rundle & Nosil, 2005).

87 Research that integrates phylogenetic hypotheses with geographic and ecological data already
88 provided new insights into speciation and diversification (Graham et al., 2004). Several authors suggest to
89 combine the spatial analysis of present-day environmental data (e.g. climate, bedrock geology) and
90 phylogeny to quantify environmental niche differences between genetically defined lineages across
91 geographic scales (e.g. Graham et al., 2004; Rissler & Apodaca, 2007; Huang et al., 2013; Zhang et al.,
92 2014). By examining the spatial pattern of environmental parameters it is possible to assess whether
93 ecologically mediated divergent selection is consistently associated with speciation (Graham et al., 2004).

94 Several major lines of argumentation are used to co-interpret phylogenetic and environmental data.
95 First, if allopatric sister lineages strongly segregate in environmental niche space, ecological divergences
96 may be interpreted as important drivers of speciation (Graham et al., 2004; Zhang et al., 2014). Conversely,
97 if allopatric sister linages have identical or nearly identical environmental niche space, this is considered to
98 support geographical isolation**,** but not ecological divergence as the main important factor in speciation
99 (Graham et al., 2004). However, differences in present-day ecological conditions of species may arise by
100 several mechanisms (e.g. ecological diversification along an environmental gradient or speciation by non-
101 ecological processes in isolation with subsequent ecological differentiation), and this means that the
102 segregation of sister lineages in environmental space is not ultimate evidence for ecological speciation.
103 Second, the coincidence of lineage split with changes in environmental conditions (e.g. climate change,
104 bedrock geology) is often interpreted as support for the role of environmental heterogeneity in shaping
105 biodiversity (Espeland et al., 2008; Espeland & Johanson, 2010; Previšić et al., 2014b), but such
106 coincidences are also insufficient to completely rule out speciation by non-ecological processes in isolation
107 with subsequent ecological differentiation. Third, parallel evolution of traits (e.g. the repeated, coinciding,
108 paraphyletic adaptation of lineages to distinct environmental factors) is considered a strong evidence for

ecological speciation (Schluter, 2001, 2009; Rundle & Nosil, 2005; Nosil, 2012) and supported by numerous examples (Rundle et al., 2000; McKinnon et al., 2004; Elmer et al., 2010b; Kautt et al., 2012). Since changes in climatic conditions or bedrock geology are expected to occur at more or less the same time across an entire region, the simultaneous, repeated, independent radiation of lineages to distinct environmental niches at this time point supports ecological speciation regarding a major, common environmental change. In contrast, in case of speciation by non-ecological processes in isolation we expect that the present environmental conditions of the lineages should reflect ancient adaptations that are monophyletically preserved along the phylogeny. Although some post-speciation local adaptations may happen sometimes after the radiation, their timing is likely to be more randomly distributed across the phylogeny. Although none of these lines of argumentation are ultimate proofs of ecological versus other forms of speciation (e.g. incidental divergence in isolation), each of them is regularly used (alone, or together) to interpret the coincidence of phylogenetic patterns and environmental conditions in the both speciation frames (e.g. Graham et al., 2004; Rundle & Nosil, 2005; Espeland & Johanson, 2010). Given the climatic and geological heterogeneity of the Balkan Peninsula and the ubiquity of cryptic diversity and endemism in this region, we hypothesize that ecological divergences played an important role in the formation of the Balkan freshwater biodiversity, and expect to find phylogenetic patterns that point toward this.

We focus on a freshwater insect species complex from the Western Balkans to understand whether abiotic environmental factors may have contributed to lineage diversification. First we used multi-locus DNA markers to establish a dated phylogeny of the target radiation (i.e. *Chaetopteryx rugulosa* species complex) and compared the timing of lineage formation with that of the past Earth surface and climatic events. The coincidence of lineage splits and diversification events of environmental conditions indicate the potential role of environmental change and ecological differentiation in the formation of the *C. rugulosa* species complex. In the second step, we characterized the present-day habitat conditions (climatic and bedrock geological features) of the taxa and contrasted these with the reconstructed phylogeny to quantify environmental niche differences between genetically defined lineages. We expect that the lineages of *C. rugulosa* species complex have distinct environmental niches indicating the importance of ecological divergences in speciation. However, both patterns of lineage divergence may arise from the geographic isolation of populations in areas that are also different environmentally. Finally, we tested whether the lineages associated with distinct environmental niches evolved monophyletically (consistent with speciation by non-ecological processes) or paraphyletically, i.e. independently at multiple times (consistent with ecological

speciation).

Materials and methods

Taxon and population sampling

The *Chaetopteryx rugulosa* species complex (Limnephilidae, Trichoptera) has caused taxonomic difficulties in the past (e.g. Malicky, 2004; Oláh et al., 2012; Malicky, 2014). The strong variation in male genital characters (traditionally used in caddisfly taxonomy) led to the inflated use of varying taxonomic ranks (subspecies, species groups, species clusters, Malicky et al., 1986; Malicky, 2004; Oláh et al., 2012). Based on fine scale morphological studies on male and female genitalia, Oláh et al. (2012) described seven new species in the *C. rugulosa* species complex, raised all taxa to species rank and distinguished three species subgroups. Malicky (2014) was concerned about the validity of species distinguished on the basis of paramere variation because this is an inherently variable structure in *Chaetopteryx* populations (e.g. Kučinić et al., 2013) and synonymized many of these species. Kučinić et al. (2013) did not find any molecular evidence (based on mtCOI sequences) supporting the validity of several of the new species either. We follow the most current taxonomy of the *C. rugulosa* species complex provided by Malicky (2014) which recognizes 13 species of which we included 9 in this study.

We sequenced specimens from 9 taxa covering all subgroups (2-4 species per subgroup: *C. irenae*, *C. schmidi*, *C. rugulosa*) sensu Oláh et al. (2012). The sequenced populations cover the distribution area of the species complex: the south-eastern Alps (*C. clara* McLachlan, *C. noricum* Malicky, *C. rugulosa* Kolenati), Slovenian Karst (*C. goricensis* Malicky & Krusnik, *C. irenae* Krusnik & Malicky), North Dinaric Alps (*C. marinkovicae* Malicky & Krusnik), SW Hungary (*C. mecsekensis* Nógrádi), central part of Croatia (*C. rugulosa* Kolenati, *C. bucari* Kučinić, Szivák & Delić) and Southern Carpathians (*C. schmidi* Botosaneanu) (Fig. 1).

We obtained one to seven specimens per species for the phylogenetic analysis (with an average of three specimens per species: 29 specimens from 20 populations of the nine species Table S1). We sequenced or downloaded several outgroup taxa of varying putative phylogenetic distance: *Chaetopteryx bosniaca* Marinković*, C. fusca* Brauer*, C. major* McLachlan*, C. Moretti* Lodovici & Valle*, C. villosa* Fabricius; C*haetopterygopsis maclachlani* Stein*, Chaetopteryx aproka* Oláh (from the same tribe); *Limnephilus centralis* Curtis (same subfamily); *Metanoea rhaetica* Schmid, *Drusus alpinus* Meyer-Dür*, D. discolor* Rambur and *D.*

168 *rectus* McLachlan (same family; see Table S1 for the GenBank accession numbers and references).

169 We obtained habitat information for the populations of the *C. rugulosa* species complex in two steps:
170 first, we selected localities where the presence of the *C. rugulosa* species complex's taxa was genetically
171 confirmed (see Kučinić et al., 2013). As the genetic data fully confirmed the morphological identifications
172 from the literature regarding the taxonomy of Malicky (2014), in the second step we also included published
173 localities where the taxa of the *C. rugulosa* species complex were morphologically identified. Overall, the
174 habitat conditions of 79 populations were considered in this analysis (Fig. 1, Table S2).

175 *DNA isolation, PCR amplifications and sequencing*

176 We extracted genomic DNA from either legs or abdominal segments of adults. We incubated the legs and
177 the abdomens in extraction buffer (0.5% SDS; 0.1M NaCl; 10mM Tris-HCl, pH 7.4; 1mM EDTA and
178 proteinase K (Sigma Aldrich, USA; 200µg/ml)) at 37°C over night rotating the samples at 100 rpm. We
179 collected the homogenates and extracted twice with equal volumes of chloroform. DNA was precipitated
180 using 0.5 volume of 5M NaCl and 0.6 volume of isopropanol. We incubated samples at -20°C for 20 min and
181 centrifuged, then washed the pellets with 70% ethanol, dried at 37°C and dissolved in triple distilled water.
182 DNA concentrations were checked spectrophotometrically in a NanoDrop ND-1000 spectrophotometer
183 (Thermo Scientific, USA).

184 We amplified nuclear genes using specific primers for the elongation factor (ChElongF [5'-
185 GAAAGTTCGAGAAGGAGGCC]; ChElongR [3'-CCTTGAACCAGGGCATCTTG]) and wingless (ChWgF [5'-
186 ACTTGCTGGATGCGCCTGCC]; ChWgR [3'- ACCCTCTTCCGCAGCACATGAG]) genes with initial
187 denaturation at 94 °C for 5 min followed by 35 cycles of incubations at 94°C for 45 s, 57°C (in case of
188 elongation factor) or 45 °C (in case of wingless) for 45 s, and 72°C for 90 s. Each 23 µl PCR reaction
189 contained 1.0 mM of each primers, 2.3 µl 10 x Taq Buffer, 0.5 mM dNTP mix, 4 µl of gDNA, 1.25 U
190 DreamTaq^{TM} DNA Polymerase (Fermentas, USA) and 12.5 µl water. PCR products were recovered from
191 1.5 % agarose gel buffered with TBE using GeneJET Gel Extraction Kit (Fermentas, USA) following
192 manufacturer instructions. Samples were sequenced with a BigDye Terminator Cycle Sequencing Kit
193 (Applied Biosystems, USA). The mitochondrial COI barcodes were produced by the Canadian Centre for
194 DNA Barcoding, University of Guelph, Canada. Standard barcoding protocols were followed for DNA
195 extraction (Ivanova et al., 2006), PCR amplification and COI sequencing (Hajibabaei et al., 2005; deWaard et
196 al., 2008). Full-length COI DNA barcodes were amplified using the two primer sets LepF1/LepR1 (Hebert et

al., 2004) and LCO1490/HCO2198 (Folmer et al., 1994). COI barcodes and detailed specimen information are deposited in the Barcode of Life Data Systems (BOLD – http://www.boldsystems.org/; Ratnasingham & Hebert, 2007) "*Chaetopteryx* of Europe" project. Nuclear DNA sequences are deposited in the Genbank. All access codes are available in Table S1.

Phylogenetic analyses

We obtained mitochondrial and nuclear sequences for nine taxa of the *C. rugulosa* species complex, and *C. aproka, C. major* and *C. bosniaca* (overall 37 specimens from 22 populations). Sequences were edited manually and aligned using the program Geneious 5.4 (Drummond et al., 2011). Our datasets consisted of a 617 bp-long sequence matrix of the *mtCOI* gene, a 431 bp-long matrix of the nuclear *EF-1α* gene and a 321 bp-long matrix of the nuclear *wingless* gene (Table S1). The *mtCOI* sequences of the group included 131 variable sites (21.2%). The *EF-1α* sequences of the group included 5 variable sites (1.2%). The *wingless* sequences of the group included 23 variable sites (7.2%).

Bayesian phylogenetic analyses were performed in MrBayes 3.2 (Ronquist et al., 2012). We selected the best-fitting models of DNA substitution for the three gene fragments and their codons with the Akaike information criterion (AIC) as implemented in ModelTest 3.7 (Posada & Crandall, 1998). The following models were selected: GTR+γ for codon positions 1, 2, and 3 of mtCOI; F81 for codon position 1, 3 and HKY for codon position 2 in wingless; HKY for codon position 1 and F81 for codon positions 2 and 3 in EF1-α. We conducted Bayesian tree construction with six chains, two independent runs and five million generations for each data set. The few ambiguous sites from the nuclear genes wingless and EF-1α were coded as ambiguity symbols. Trees were sampled every 1,000th generation. The first 50% of generations were discarded as burn-in. We plotted the log-likelihood scores of sample points against generation time using Tracer 1.5 (Rambaut & Drummond, 2009) to ensure that stationarity was achieved. We used the remaining trees to create a 50% majority-rule consensus tree with the 'sumt' option in MrBayes. Posterior probabilities were obtained for each clade. We examined the heterogeneity of the phylogenetic signal among data partitions (Buckley et al., 2002) by comparing the combined analysis topology with the 0.95 posterior intervals of the single gene analyses. As no conflicts were evident, we assumed that the three data sets were congruent and can be combined.

We dated the phylogeny in BEAST 1.7.2 (Heled & Drummond, 2010) using a species tree approach

for the mitochondrial and wingless gene fragments to assess coincidence of lineage splitting and environmental change event in the formation of the *C. rugulosa* species complex. We opted to exclude EF-1α from the dating as no sequences were available for most outgroups. We also excluded the outgroup species *Limnephilus centralis* as we had no wingless sequence for this species. *BEAST uses multilocus data to infer a species tree, without concatenating the sequence sets. The few ambiguous sites from the wingless gene were coded as ambiguity symbols. We used substitution models inferred previously with ModelTest for the two gene fragments (GTR+I+y for mtCOI; GTR+I for wingless). We dated our phylogeny with a relaxed uncorrelated lognormal molecular clock (Drummond et al., 2006). The dating was calibrated with the estimated timing of surface uplift of the Mecsek Mountains in SW Hungary, a small (ca. 17×30 km), fault-bounded, geomorphologically isolated basement unit (Fig.1). The Mecsek was uplifted about 3 My ago (Horváth & Cloetingh, 1996; Bada et al., 2001; Csontos et al., 2002b; Sebe et al., 2008, with references therein). In this period a major changeover of the tectonic stress field affected the central part of the intra-Carpathian area (Bada et al., 2001). This resulted in the uplift of small "inselbergs" (including the Mecsek Mt.), attaining topographic characters very comparable to modern conditions (Csontos et al., 2002a, 2002b). As all species of the *C. rugulosa* species complex inhabit springs at elevated topography, we assumed that *C. mecsekensis* (endemic to the Mecsek Mt.; Fig. 1) was separated from other taxa of the *C. rugulosa* species complex following the uplift of the Mecsek about 3 My ago (*S.D.*=0.5, normal distribution). As explicit priors were not available for the rate of the mtCOI and wingless clocks, we used CTMC rate reference distributions for these priors (Ferreira & Suchard, 2008). We generated 100 million trees in four independent *BEAST runs, and sampled these for every 10,000 generations to obtain 10,000 tree samples. The convergence of the runs was checked in Tracer 1.5 (Rambaut & Drummond, 2009). The trees from the four runs were combined after discarding 10% of the trees as burn-in. The last 10,000 trees were visualized with DensiTree (Bouckaert, 2010) to show uncertainties in tree topologies and branch lengths.

Phylogenetic association of present-day habitat conditions

We characterized the present-day habitat conditions (climatic and bedrock geological features) of every known population of each species. We calculated averaged habitat conditions for each species, since these were represented by multiple known populations in the study. We contrasted the averaged habitat conditions with the reconstructed phylogeny to quantify the genetic component of species-specific environmental niche differences. We assessed whether lineage divergences may be associated with Earth surface and

biogeographical history processes by contrasting the dated phylogenetic tree with, respectively, tectonic events and past climate change. Finally, we tested whether lineages with distinct habitat conditions formed paraphyletically or monophyletically, and whether this radiation might be associated with current ecological conditions.

We assembled present-day climatic (related to temperature and precipitation) and geological (bedrock type) datasets for 79 populations of the *C. rugulosa* species complex (Table S2). We downloaded 31 climate-related data layers (resolution 30 arc-seconds, Table S3) from the WoldClim database (Hijmans et al., 2005). We queried the climatic layers with the locality dataset in R ver. 2.14.0 (R Core Team, 2013) with the 'rgdal' (Keitt et al., 2010) and 'dismo' packages (Hijmans et al., 2012). We assembled a climatic data matrix for each site (31 continuous climatic variables × 79 sites). We overlaid the localities with geological maps of the Balkan Peninsula and read the bedrock types at the drainage area above the sites to obtain a data matrix of geological feature (10 dummy bedrock geological variables × 79 sites; Table S2). Before the analyses, the climatic variables were transformed depending on their scale of measurement to reach their normality and reduce heteroscedasticity (see Table S3).

We measured the phylogenetic signal of each species' environmental preference by selecting environmental parameters (climate and bedrock geological traits) associated with lineages on the phylogenetic tree using Blomberg's K-statistic (Blomberg et al., 2003). We calculated the mean values of each climatic variable and the mode values of bedrock geological parameters for habitats of each species. In this way in the case of each species we got one mean value for each of the 31 climatic variables. In our study none of the species was co-occurred with one other at the same sampling site. The Blomberg's K statistic compares the observed signal in a trait to the signal expected under a Brownian motion model of trait evolution on a phylogeny (Blomberg et al., 2003), a high K statistic value (>1) being indicative of a strong phylogenetic signal in the observed trait. The statistical significance of phylogenetic signal can be evaluated by comparing observed variance of independent variable contrasts of the trait to a null model of mixing species labels across the tips of the phylogeny (Kembel et al., 2010). The inferred phylogenetic signal is a quantitative measure of whether a phylogeny predicts more ecological similarity among related species than expected by chance alone (Blomberg & Garland, 2002).

We used Linear Discriminant Analysis (LDA; Fisher, 1936) to test whether the lineages of the *C. rugulosa* species complex inhabit areas with distinct abiotic conditions and have distinct environmental niches. The LDA determines to what extent an independent set of explanatory variables (here: habitat

conditions) explains a pre-defined grouping of objects (here: the populations grouped by species). The LDA computes discriminant functions (canonical axes) from standardized explanatory variables. The functions quantify the relative contributions of the explanatory variables to the discrimination of the objects (Legendre & Legendre, 2012). A cross-validation procedure within the LDA allows for a probabilistic assignment of objects to each group. The LDA was run only on the eight climatic variables that contained statistically significant phylogenetic signal (Table 1), since all bedrock geology and the remaining 23 climatic variables were no statistically significant. We did not eliminate highly correlated climatic variables which were statistically significantly associated with the phylogenetic signal, since we intended to keep all climatic variables that had biological/evolutionary signal. Since we are aware of the potentially harmful influence of collinearity on multivariate ordinations methods (such as LDA) (Dormann et al., 2013), we executed a reduced LDA model without highly correlated variables. In order to associate the discriminant functions with climatic conditions, we calculated Pearson correlations between the discriminant axes and the climatic variables. The discriminant functions of the LDA are environmental niches for each species which include the effects of several individual environmental variables and reflect the strongest environmental gradient existing between the habitats of populations belonging to nine species. We obtained the species mean scores on the first and second discriminant functions and tested phylogenetic signal of these values on the dated phylogenetic tree using Blomberg's K-statistic. In order to verify the segregation of species' climatic niches in multidimensional space (based on the LDA model) we conducted multivariate analysis of variance (MANOVA) using species scores on all canonical axes as explanatory variables and the genetically defined species as grouping variable. Additionally, we performed a Kruskal-Wallis variance analysis to examine the climatic niche segregation along the first discriminant function. Results from the LDA analysis enabled us to group the lineages of *C. rugulosa* species complex into two climatically distinct niches along the first canonical axis. In order to verify the segregation of these climatic niches we performed Mann-Whitney nonparametric U test using the species scores on the first canonical axis as explanatory variables and two climatic niches as factor. All statistical analyses were performed in software R ver. 2.14.0 (R Core Team, 2013) using the packages 'picante' (Kembel et al., 2010) for Blomberg's K-statistic, 'MASS' (Venables & Ripley, 2002) for LDA, 'vegan' (Oksanen et al., 2013) for MANOVA, Kruskal-Wallis variance analysis and Mann-Whitney U test.

Finally, to assess the possibility of parallel evolution we tested whether the lineages associated with two distinct climatic niches evolved monophyletically or paraphyletically. We set up topological constraints on

316 the group's phylogeny to reflect two models of lineage evolution. First, we enforced the monophyly of
317 lineages associated with relatively warm conditions. Second, we allowed for the paraphyletic evolution of
318 these lineages. We compared the two phylogenetic models in MrBayes v.3.2 (Ronquist et al., 2012) using
319 the stepping-stone algorithm (Xie et al., 2011). We used 50 steps with 100,000 generations each, for a total
320 of 5,000,000 generations. We monitored convergence by recording diagnostics in every 50,000 generations.

321

322 **Results**

323 The Bayesian phylogenetic inference shows that the *Chaetopteryx rugulosa* species complex is
324 monophyletic (Fig. 2). The species complex can be divided into three major clades (clade A, B, C). The
325 Istrian *Chaetopteryx marinkovicae* (clade A) is basal to the other lineages which form two clades (clade B, C,
326 Fig. 2). These are located to the North and East from the habitats of *C. marinkovicae* (Fig. 1). The 'clade B'
327 consists of the most southerly species *C. schmidi* and *C. bucari* (Figs. 1, 2), and it is basal to the 'clade C'.
328 The 'clade C' consists of *C. mecsekensis*, *C. noricum*, *C. rugulosa, C. clara*, *C. irenae*, *C. goricensis* (Fig. 2).
329 All six species of the 'clade C' occur north of the Kupa-Sava River systems (Fig. 1). The 'clade C' can be
330 further divided into two groups: one consists of *C. mecsekensis*, *C. rugulosa*, *C. noricum* and other consists
331 of the *C. clara, C. goricensis, C. irenae*.

332 We estimated a mean substitution rate of 1.14% / My (S.E. 0.015) for *mtCOI*, and a mean substitution
333 rate of 0.31% (S.E. 0.007) for *wingless* by calibrating the group's phylogeny with the uplift of the Mecsek Mt.
334 The time span to the most recent common ancestor of the *C. rugulosa* species complex was 12 Mya (95%
335 highest probability densities 4.8 – 20.5 My, Fig. 3). The split of the 'clade B' and 'clade C' dates to 5.9 Mya
336 (95% HPD 2.6 – 10). The split between the two groups from the 'clade C' dates to 4.5 Mya (95% HPD 2.1 –
337 7.5). In the 'clade B' *C. schmidi* split from *C. bucari* 2.3 Mya, (95% HPD 0.5 – 4.7). In the 'clade C' *C. irenae*
338 split from *C. clara* and *C. goricensis* 2.9 Mya (95% HPD 0.8 – 5.2). The date of divergence of *C.*
339 *mecsekensis* from *C. rugulosa* and *C. noricum* (input as a prior at 3 Mya in the analysis) was shifted to 2.8
340 Mya (95% HPD 1.8 – 3.8). The splits between *C. goricensis* and *C. clara* (0.4 Mya, 95% HPD 0.01 – 0.9),
341 and *C. rugulosa* and *C. noricum* (0.2 Mya, 95% HPD 0 – 0.6) are relatively recent.

342 Blomberg's K-statistic indicates that none of the bedrock geological variables are significantly
343 associated with the phylogenetic signal. Eight out of 31 climatic variables were statistically significantly
344 ($p \leq 0.05$) associated with the phylogenetic tree (Table 1). The strongest phylogenetic signals were associated

345 with the precipitation seasonality (Blomberg's K=1.15) and the precipitation of the driest month (Blomberg's
346 K=0.78).

347 The first two discriminant functions (canonical axes) of the LDA explained 84.3% of the total variability
348 in species occurrence data. The first discriminant function was negatively associated with the temperature
349 variables (e.g. min. temperature of coldest month), and positively with the temperature annual range (Table
350 1; Fig. 4). The second discriminant function was negatively related to the variability of temperature and
351 precipitation (e.g. temperature and precipitation seasonality), and positively to the amount of precipitation
352 (e.g. precipitation of driest month; Table 1; Fig. 4). Globally, 91.1% of the populations was correctly grouped
353 into pre-defined species by the cross-validation procedure. This indicates that each member of the species
354 complex has distinct environmental niches (Fig. 5). Similarly, the MANOVA analysis also showed that the
355 climatic niches of the species are statistically significantly segregated in the multidimensional space of LDA
356 (Wilks' α=0.00026, $p<0.001$).

357 The relative contribution of the first canonical axis to the discrimination of the populations into pre-
358 defined species was high (60.2%) and it was also statistically significantly linked to the phylogenetic signal
359 (Blomberg' s K=0.88, p=0.025). The relative contribution of second canonical axis was lower (24.1%) and the
360 Bloomberg's K analysis did not indicate a link toward the phylogenetic signal (K=0.27, p=0.19). The climatic
361 niche of each species are statistically significantly segregated along the first canonical axis based on the
362 Kruskal-Wallis variance analysis (χ^2=67.7, $p<0.001$). Finally, the Mann-Whitney nonparametric U test
363 confirmed that the 79 populations can be assigned into two climatically distinct niches along the first
364 canonical axis (W=1378, $p<0.001$; Fig. 4). Five of the nine lineages (*C. goricensis, C. irenae, C.*
365 *marinkovicae, C. mecsekensis, C. schmidi*) were associated with the mean and high temperatures in
366 December and in the coldest quarter of the year along the first axis (referred to as the 'warm' group here).
367 The other four lineages (*C. bucari, C. clara, C. noricum, C. rugulosa*) formed another group which was
368 positively associated with the first axis and characterized by high variability in annual temperature and lower
369 temperature values in December and in the coldest quarter of the year (referred to as the 'cold' group here).

370 The elimination of the collinear climatic variable was not influenced on the main results of the LDA
371 model (not shown here). In the case of the reduced LDA model (with 5 climatic variables) the first two
372 canonical axes of the LDA explained 90.47% of the total variability in species occurrence data. Globally,
373 89.87% of the populations were correctly grouped into pre-defined species by the cross-validation
374 procedure. The climatic niches of the species are statistically significantly segregated based on the results of

MANOVA (Wilks' α=0.00177, $p<0.001$) and Kruskal-Wallis variance analysis ($\chi 2$=67.7, $p<0.001$). The Mann-Whitney nonparametric U test also confirmed the segregation of the *C. rugulosa* species complex's lineages into two climatically distinct niches along the first canonical axis.

The stepping-stone comparison of marginal likelihoods of topologically constrained and unconstrained phylogenies shows that lineages split multiple times in association with 'warm' vs. 'cold' climatic conditions (Figs. 2, 4). The harmonic mean of the marginal log-likelihood was -6382 for phylogenies if the monophyly of 'warm' associated sister lineages was enforced. The harmonic mean of the marginal log-likelihood was -6304 for phylogenies with polyphyletic 'warm' associated sister lineages giving a Bayes Factor of 78 ($2\ln(B_{10})$). The second model is 78 ln units better than the first one and the better model is very strongly supported by a Bayes Factor test. Thus, these results indicate strong evidence for 'warm' associated sister lineages do not form a monophyletic group. Implying the idea that climatically defined sister lineages split paraphyletically, namely independently multiple times during the radiation of the *C. rugulosa* species complex. The dated phylogeny indicates that this parallel split occurred around the Pliocene/Pleistocene transition.

Discussion

Genetic markers confirm all morphologically defined taxa of the *C. rugulosa* species complex to be monophyletic and distinct (Fig. 2). Our 3 gene phylogeny is coincident with the mtCOI phylogeny shown by Kučinić et al. (2013). Our phylogeny also confirms the species-level validity of *C. schmidi* (see also Oláh et al., 2012) and the outgroup position of *C. aproka* established by Oláh et al. (2012) in their morphology-based revision of *C. rugulosa* species complex. The monophyly of the species complex is also supported. The sampled species clearly group into three distinct clades (clade A, B, C), with *C. marinkovicae* (clade A) being basal to the rest of the group. The 'clade B' and 'clade C' contains 2 and 6 well-supported taxa, respectively and are separated along the Kupa-Sava-Drava River systems. The molecular differences are weak, but statistically significant between the most recently (0.4 Mya and 0.2 Mya) diverged, closely related species pairs (*C. clara* - *C. goricensis*, *C. rugulosa* - *C. noricum*) (Fig. 2; see also Kučinić et al., 2013: 12, table 3). The molecular phylogeny does not confirm the morphology-based taxonomic inferences of the within-group relationships (see Oláh et al., 2012). A possible reason for this is that morphology-based taxonomic relationships among caddisflies are inferred mostly on the basis of fine-scale genital morphology, namely paramere variation. Fine scale genital variation may be suited to identify species or population structure (e.g.

404 Malicky, 2016), but can be very misleading in the inference of evolutionary relationships (Pauls et al., 2008;
405 Malicky, 2014).

406 We found higher taxonomic richness in the western area of distribution. Correspondingly, Oláh et al.
407 (2012) found 4 taxa (*C. bucari, C. papukensis* Oláh & Szivák*, C. mecsekensis, C. schmidi*) at the eastern
408 part of the Western Balkan and 12 taxa at the western part after extensive field work, which covered the
409 almost all potential habitats from Romania to Slovenia. This emphasizes the importance of the western area
410 in the group's formation.

411 We dated the phylogeny of the species complex and compared the timing of lineage formation with the
412 timing of past Earth surface and climatic events to evaluate the potential role of environmental change and
413 ecological differentiation in the formation of the species complex. We estimated a mutation rate of 1.14% per
414 million years for the mitochondrial COI gene, and 0.31% for the nuclear *wingless* gene by calibrating the *C.*
415 *rugulosa* species complex phylogeny by the uplift of the Mecsek Mt. The *mtCOI* mutation rate estimated in
416 this study is very similar to the 2.3% pairwise sequence divergence reported by Brower (1994), but lower
417 than the revised rate recently suggested by Papadopoulou et al. (2010) for insects. The difference
418 emphasizes the presence of taxon-specific variation in mitochondrial mutation rates which needs to be
419 considered in phylogenetic dating.

420 Four time periods seem to be important in the formation of the *C. rugulosa* species complex (Fig. 3).
421 These time periods coincide with important past Earth surface and climatic events that shaped the Balkan
422 Peninsula. The disjunction of *C. marinkovicae* from other taxa in the *C. rugulosa* species complex was dated
423 at ~12 Mya. This species is present only in a few isolated areas of the North Dinaric Alps. In our
424 interpretation *C. marinkovicae* split from the remaining *C. rugulosa* group ancestors due to the plate
425 tectonics-driven mountain formation of this period. The Dinaric Alps were separated from the Alps by a major
426 marine gateway, the Transtethyan Trench Corridor (or Slovenian Strait) until late Langhian time (~14 My ago,
427 Fig. 6, Kováč et al., 2007). The closure of the Transtethyan Trench Corridor in the late Langhian overlaps
428 with the protracted tectonic shortening at the Dinaride thrust front (Mikes et al., 2008; Ustaszewski et al.,
429 2008). Development of high-elevation topography is typically inherent to such shortening events (Burbank &
430 Anderson, 2012), thus the formation of the North Dinaride topography likely involved a major pulse in
431 topography formation in this period. The disjunction of *C. marinkovicae* from other lineages of the *C.*
432 *rugulosa* species complex coincides with the tectonically-driven formation of topography in the North Dinaric
433 Alps subsequent to the closure of the Transtethyan Trench Corridor. The second disjunction was dated at

~5.9 Mya, marking the separation of the ‘clade B’ and ‘clade C’. The timing of the split coincides with the shoreline retreat of the Pannon Lake (Fig. 6). The retreat of the lake lasted until 4.5 My ago, when the lake basin became entirely filled up by sediments (Magyar et al., 1999). We postulate that the eastward retreat of an east-west trending branch of the remnant lake (Magyar et al., 1999) was accompanied by both significant rearrangements of drainage catchments and the decrease of topographic relief at the basin margin. These pronounced landscape changes potentially rendered habitat connectivity impossible and led to the formation of new lineages in allopatry (Griffiths et al., 2004). The formation of other spring-inhabiting Trichoptera from Western Balkan region was also dated at this time period (*Drusus* species, Previšić et al., 2009). The third period of disjunctions was dated at ~2.3 – 2.9 Mya. This coincides with the Pliocene/Pleistocene transition and the onset of pronounced climatic oscillations (Fig. 3, Lisiecki & Raymo, 2005). Three groups of the *C. rugulosa* species complex were formed during this period (Fig. 3). Finally, the last disjunctions within the species complex are recent: they coincide with major glacial cycles of the last phase of the Pleistocene. In conclusion, the oldest disjunctions within the *C. rugulosa* species complex are likely connected to major geological processes shaping the Western Balkan topography, highlighting the role of tectonically driven geographic separation. However, in our interpretation the more recent radiations of the species complex are linked to climatic events of the Pleistocene, being likely driven by ecological diversification inducted by climatic conditions.

First, habitat condition analysis shows that the lineages of the *C. rugulosa* species complex inhabit areas with distinct climatic conditions (but not different bedrock geology). These climate niches are determined by temperature parameters during the emergence period of species, annual variability in the temperature and precipitation, and the amount of precipitation in the driest period of the year. The strong segregation of allopatric sister lineages in environmental niche space points toward the role of ecological divergences as important drivers of speciation (Graham et al., 2004; Zhang et al., 2014), but it is not ultimate evidence for ecological speciation. The differences in present-day ecological conditions of species may arise by several mechanisms, for instance ecological diversification along an environmental gradient or speciation by non-ecological processes in isolation with subsequent ecological differentiation. Second, the more recent radiations of the species complex coincide with major past climatic events (Pleistocene climatic oscillations), and such coincidences are often interpreted as reflecting ecological speciation (i.e. environmental heterogeneity shapes biodiversity, Espeland et al., 2008; Espeland & Johanson, 2010; Previšić et al., 2014b), but one has to keep in mind that such coincidences are also insufficient to rule out speciation by

non-ecological processes in isolation with subsequent ecological differentiation. Third, the radiation of lineages into 'cold' and 'warm' habitats occurred at several parallel occasions about the same time (around the Pliocene/Pleistocene transition) during the history of the *C. rugulosa* species complex (Fig. 3). In our opinion this is the strongest support of the ecological diversification in the group: the simultaneous, repeated, paraphyletic radiation of lineages having distinct traits associated with recent environmental factors (here: local climate) is generally indicative of ecological speciation (Rundle et al., 2000; Rundle & Nosil, 2005; Schluter, 2009; McKinnon et al., 2004; Elmer et al., 2010b; Nosil, 2012), because as Schluter states it, "such repetition is unlikely to result from chance; environmental selection pressures must therefore be the cause of speciation" (Schluter, 2009). If only speciation by non-ecological processes is responsible for the formation of the more recent lineages, we would expect that environmental adaptations are more or less monophyletically preserved along the phylogeny or that they are randomly distributed in terms of timing and phylogeny and the descendants of the lineage pairs live in 'warm'/'cold' habitat pairs only by chance. In summary, although the results are no final proof against speciation by non-ecological processes in allopatry, they all point toward the plausible importance of ecological speciation. Conclusive evidence should come from experiments that compare species fitness under diverse climatic conditions, completed with cross-breeding experiments that evaluate reproductive isolation and hybrid fitness (Malicky, 1996; Malicky & Pauls, 2012).

Climatic conditions often have major impact on the survival of species and intraspecific genetic lineages (Lehrian et al., 2010; Bálint et al., 2011; Pauls et al., 2013). Freshwater species may be particularly affected by climate, as climate strongly influences the yield of springs and the frequency of floods and dry periods. Climate also contributes to the definition of microclimatic habitat characteristics (air temperature of emergence sites, snow patches). The reproduction of Chaetopterigini caddisflies is timed to the late autumn – early winter. This period is prone to rapid fluctuations in temperature and precipitation, thus adaptation to these fluctuations can seriously influence the long-term survival of populations (Hoffmann & Sgro, 2011; Pauls et al., 2013). In our interpretation local climatic adaptation potentially influenced the diversification of the *C. rugulosa* species complex, although we recognize that such adaptations must be proven experimentally, and that our analysis is limited by lacking samples for all species in the complex.

General evidence of ecological diversification is rapidly accumulating, e.g. in studies showing differentiation during past climatic events (Espeland et al., 2008; Espeland & Johanson, 2010; Previšić et al., 2014b), ongoing adaptation to habitat conditions (Evans et al., 2009; Dormann et al., 2010; Zhang et al.,

494 2014) or multiple radiation and paraphyletic formation of lineages adapted to current habitat conditions
495 (McKinnon et al., 2004; Elmer et al., 2010b; Kautt et al., 2012). Our results further support that climate may
496 influence the formation of biodiversity and provide directions for future experiments and genome analyses.
497 Generally, more studies linking phylogenetic inferences with ecological trait evolution are likely to shed light
498 on possible examples of ecological diversification, particularly in taxonomically and ecological diverse groups
499 of aquatic insects (Dijkstra et al., 2014). Ideally such identified cases can then be studied with evolutionary
500 ecology experiments to foster our understanding of the involved processes, particularly as more genomic
501 data become available for non-model organisms (Pauls et al., 2014).

502

503 **Acknowledgements**

504 We are grateful to János Oláh, Hans Malicky, Ákos Uherkovich, Sára Nógrádi, Ana Previšić, Gorazd Urbanic,
505 Marco Valle, Omar Lodovici, Antun Delić, Iva Mihoci, Aleksandar Popijač and Matija Bučar who provided us
506 with valuable specimens, location data and general advice. We express our gratitude to Xin Zhou for
507 handling the mitochondrial barcoding of our specimens within the Barcoding of Life initiative. I.Sz. was
508 partially funded through an incoming research grant of the Biodiversity and Climate Research Centre in the
509 frame of the research funding program Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer
510 Exzellenz (LOEWE) of Hesse's Ministry of Higher Education, Research, and the Arts. M.B. and S.P. are
511 funded by the research funding program Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer
512 Exzellenz (LOEWE) of Hesse's Ministry of Higher Education, Research, and the Arts, as well as the Fonds
513 zur Förderung der wissenschaftlichen Forschung (FWF) grant P 23687-B17.

514 **References**

April, J., R. H. Hanner, A.-M. Dion-Côté & L. Bernatchez, 2013. Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. Molecular Ecology 22: 409–422.

Bada, G., F. Horváth, S. Cloetingh, D. D. Coblentz & T. Tóth, 2001. Role of topography-induced gravitational stresses in basin inversion: The case study of the Pannonian basin. Tectonics 20: 343–363.

Bálint, M., S. Domisch, C. H. M. Engelhardt, P. Haase, S. Lehrian, J. Sauer, K. Theissinger, S. U. Pauls & C. Nowak, 2011. Cryptic biodiversity loss linked to global climate change. Nature Climate Change 1: 313–318.

Bănărescu, P. M., 2004. Distribution pattern of the aquatic fauna of the Balkan Peninsula In Griffiths, H. I., B. Kryštufek & J. M. Reed (eds), Balkan Biodiversity. Springer, Netherlands: 203–217.

Bilandžija, H., B. Morton, M. Podnar & H. Ćetković, 2013. Evolutionary history of relict Congeria (Bivalvia: Dreissenidae): unearthing the subterranean biodiversity of the Dinaric Karst. Frontiers in Zoology 10: 5.

Blomberg, S.P. & T. Garland, 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. Journal of Evolutionary Biology 15: 899–910.

Blomberg, S. P., T. Garland & A. R. Ives, 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57: 717–745.

Bouckaert, R. R., 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26: 1372–1373.

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

Buckley, T. R., P. Arensburger, C. Simon & G. K. Chambers, 2002. Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. Systematic Biology 51: 4–18.

Burbank, D. W. & R. S. Anderson, 2012. Tectonic geomorphology. J. Wiley & Sons, Chichester, West Sussex; Hoboken, NJ.

Csontos, L., L. Benkovics, F. Bergerat, J.-L. Mansy & G. Wórum, 2002a. Tertiary deformation history from seismic section study and fault analysis in a former European Tethyan margin (the Mecsek-Villány area, SW Hungary). Tectonophysics 357: 81–102.

Csontos, L., E. Márton, G. Wórum & L. Benkovics, 2002b. Geodynamics of SW-Pannonian inselbergs (Mecsek and Villány Mts, SW Hungary): Inferences from a complex structural analysis. Stephan Mueller Special Publication Series 3: 227–245.

deWaard, J. R., N. V. Ivanova, M. Hajibabaei & P. D. Hebert, 2008. Assembling DNA barcodes. Analytical protocols. Methods in molecular biology 410: 275–293.

Dijkstra, K.-D. B., M. T. Monaghan & S. U. Pauls, 2014. Freshwater biodiversity and insect diversification. Annual Review of Entomology 59: 143–163.

Dormann, C. F., B. Gruber, M. Winter & D. Herrmann, 2010. Evolution of climate niches in European mammals? Biology Letters 6: 229–232.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R García Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell & S.Lautenbach, 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 27–46.

Drummond, A. J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer & A. Wilson, 2011. Geneious v5.4. Available from: http://www.geneious.com/.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips & A. Rambaut, 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4:e88.

Elmer, K. R., T. K. Lehtonen, A. F. Kautt, C. Harrod & A. Meyer, 2010a. Rapid sympatric ecological ifferentiation of crater lake cichlid fishes within historic times. BMC Biology 8: 60.

Elmer, K. R., H. Kusche, T. Lehtonen & A. Meyer, 2010b. Local variation and parallel evolution: Morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. Philosophical Transactions of the Royal Society B: Biological Sciences 365: 1763–1782.

Espeland, M. & K. A. Johanson, 2010. The effect of environmental diversification on species diversification in New Caledonian caddisflies (Insecta: Trichoptera: Hydropsychidae). Journal of Biogeography 37: 879–890.

Espeland, M., K. A. Johanson & R. Hovmöller, 2008. Early Xanthochorema (Trichoptera, Insecta) radiations in New Caledonia originated on ultrabasic rocks. Molecular Phylogenetics and Evolution 48: 904–917.

Evans, M. E. K., S. A. Smith, R. S. Flynn, J. Michael & M. J. Donoghue, 2009. Niche Evolution, and Diversification of the “Bird-Cage” Evening Primroses (Oenothera, Sections Anogra and Kleinia). The American Naturalist 173: 225–240.

Feder, J. L., 1998. The apple maggot fly, Rhagoletis pomonella: flies in the face of conventional wisdom about speciation? In Howard, D. & S. Berlocher (eds), Endless Forms: Species and Speciation. Oxford University Press, Oxford: 130–144.

Ferreira, M. A. R. & M. A. Suchard, 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. Canadian Journal of Statistics 36: 355–368.

Fisher, R. A., 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics 7: 179–188.

Folmer, O., M. Black, W. Hoeh, R. Lutz & R. Vrijenhoek, 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Funk, D. J., P. Nosil & W. J. Etges, 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. Proceedings of the National Academy of Sciences of the United States of America 103: 3209–3213.

Gómez, A. & D. Lunt, 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula In Weiss, F. (ed), Phylogeography of Southern European Refugia. Springer, Netherlands: 155–188.

Graham, C.H., S. R. Ron, J. C. Santos, C. J. Schneider & C. Moritz, 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution 58: 1781–1793.

Griffiths, H. I., B. Krystufek & J. M. Reed (eds), 2004. Balkan Biodiversity: Pattern and Process in the European Hotspot. Springer, Netherlands.

Hajibabaei, M., J. R. deWaard, N. V. Ivanova, S. Ratnasingham, R. T. Dooh, S. L. Kirk, P. M. Mackie & P. D. Hebert, 2005. Critical factors for assembling a high volume of DNA barcodes. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 360: 1959–1967.

Harzhauser, M., O. Mandic, T. A. Neubauer, E. Georgopoulou & A. Hassler, 2015. Disjunct distribution of the Miocene limpet-like freshwater gastropod genus Delminiella. Journal of Molluscan Studies 1: 8.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen & W. Hallwachs, 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America 101: 14812–14817.

Heled, J. & A. J. Drummond, 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27: 570–580.

Hewitt, G. M., 2000. The genetic legacy of the Quaternary ice ages. Nature 405: 907–913.

Hewitt, G. M., 2011. Mediterranean Peninsulas: The Evolution of Hotspots. In Zachos, F. E., & J. C. Habel

(eds), Biodiversity Hotspots. Springer, Berlin, Heidelberg: 123–147.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones & A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965–1978.

Hijmans, R. J., S. Phillips, J. Leathwick & J. Elith, 2012. dismo: Species distribution modeling. R package version 0.7-17. Available from: http://CRAN.R-project.org/package=dismo.

Hoffmann, A. A. & C. M. Sgro, 2011. Climate change and evolutionary adaptation. Nature 470: 479–485.

Horváth, F. & S. Cloetingh, 1996. Stress-induced late-stage subsidence anomalies in the Pannonian basin. Tectonophysics 266: 287–300.

Huang, Y., X. Guo, S. Y. W. Ho, H. Shi, J. Li, J. Li, B. Cai, & Y. Wang, 2013. Diversification and demography of the oriental garden lizard (Calotes versicolor) on Hainan Island and the adjacent mainland. PLoS One 8: e64754.

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

Kalkvik, H. M., I. J. Stout, T. J. Doonan & C. L. Parkinson, 2012. Investigating niche and lineage diversification in widely distributed taxa: phylogeography and ecological niche modeling of the *Peromyscus maniculatus* species group. Ecography 35: 54–64.

Kautt, A. F., K. R. Elmer & A. Meyer, 2012. Genomic signatures of divergent selection and speciation patterns in a ‘natural experiment’, the young parallel radiations of Nicaraguan crater lake cichlid fishes. Molecular Ecology 21: 4770–4786.

Keitt, T. H., R. Bivand, E. Pebesma & B. Rowlingson, 2010. rgdal: Bindings for the geospatial data abstraction library. Available from: http://CRAN.R-project.org/package=rgdal.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg & C. O. Webb, 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26: 1463–1464.

Kováč, M., A. Andreyeva-Grigorovich, Z. Bajraktarević, R. Brzobohatý, S. Filipescu, L. Fodor, M. Fodor, A. Nagymarosy, N. Oszczypko, D. Pavelić, F. Rögl, B. Saftić, L. Sliva & B. Studencka, 2007. Badenian evolution of the Central Paratethys Sea: paleogeography, climate and eustatic sea-level changes. Geologica Carpathica 58: 579–606.

Kučinić, M., I. Szivák, S. U. Pauls, M. Bálint, A. Delić & I. Vučković, 2013. *Chaetopteryx bucari* sp. n., a new species from the *Chaetopteryx rugulosa* group from Croatia (Insecta, Trichoptera, Limnephilidae) with molecular, taxonomic and ecological notes on the group. ZooKeys 320: 1–28.

Lehrian, S., M. Bálint, P. Haase & S. U. Pauls, 2010. Genetic population structure of an autumn emerging caddisfly with inherently low dispersal capacity and insights into its phylogeography. Journal of the North American Benthological Society 29: 1100–1118.

Legendre, P. & L. Legendre, 2012. Numerical Ecology, Third Edition. Elsevier, Amsterdam, Oxford.

Levsen, N. D., P. Tiffin & M. S. Olson, 2012. Pleistocene speciation in the genus *Populus* (Salicaceae). Systematic Biology 61: 401–412.

Lisiecki, L. E. & M. E. Raymo, 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. Paleoceanography 20, PA1003, doi:10.1029/2004PA001071.

Magyar, I., D. H. Geary & P. Müller, 1999. Paleogeographic evolution of the Late Miocene Lake Pannon in Central Europe. Palaeogeography, Palaeoclimatology, Palaeoecology 147: 151–167.

Malicky, H., 1983. Chorological patterns and biome types of European Trichoptera and other freshwater insects. Archiv für Hydrobiologie 96: 223–244.

Malicky, H., 1996. Das Problem der allopatrischen Arten bei europäischen Köcherfliegen (Insecta: Trichoptera). Natura Croatica 5: 11–23.

Malicky, H., 2004. Atlas of European Trichoptera. Springer, Dordrecht.

Malicky, H., 2014. Comments on two recently published papers on *Cheumatopsyche* (Hydropsychidae) and *Chaetopteryx* (Limnephilidae). Baueria 41: 51–53.

Malicky, H., 2016. Die mitteleuropäische Verbreitung zweirer Morphotypen von *Allogamus auricollis* (Trichoptera, Limnephilidae) mit phänologischen und bionomischen Notizen. Braueria 43: 29–38.

Malicky, H., C. Krusnik, G. Moretti & S. Nógrádi, 1986. Ein Beitrag zur Kenntnis der *Chaetopteryx rugulosa* Kolenati, 1848, - Gruppe (Trichoptera, Limnephilidae). Entomofauna - Zeitschrift für Entomologie 7: 1–27.

Malicky, H. & S. U. Pauls, 2012. Cross-breeding of *Chaetopteryx morettii* and related species, with molecular and eidonomical results (Trichoptera, Limnephilidae). Annales de Limnologie - International Journal of Limnology 48: 13–19.

Mandic, O., A. de Leeuw, J. Bulić, K. F. Kuiper, W. Krijgsman & Z. Jurišić-Polšak, 2012. Paleogeographic evolution of the Southern Pannonian Basin: $^{40}Ar/^{39}Ar$ age constraints on the Miocene continental series of Northern Croatia. International Journal of Earth Sciences 101: 1033–1046.

McKinnon, J. S., S. Mori, B. K. Blackman, L. David, D. M. Kingsley, L. Jamieson, J. Chou & D. Schluter, 2004: Evidence for ecology's role in speciation. Nature 429: 294–298.

McPeek, M. A. & G. A. Wellborn, 1998. Genetic variation and reproductive isolation among phenotypically

divergent amphipod populations. Limnology and Oceanography 43: 1162–1169.

Mikes, T., M. Báldi-Beke, M. Kázmér, I. Dunkl & H. von Eynatten, 2008. Calcareous nannofossil age constraints on Miocene flysch sedimentation in the Outer Dinarides (Slovenia, Croatia, Bosnia-Herzegovina and Montenegro). Geological Society, London, Special Publications 298: 335–363.

Múrria, C., N. Bonada, M. A. Arnedo, C. Zamora-Muñoz, N. Prat & A. P. Vogler, 2012. Phylogenetic and ecological structure of Mediterranean caddisfly communities at various spatio-temporal scales. Journal of Biogeography 39: 1621–1632.

Nosil, P., 2012. Ecological Speciation. Oxford University Press, Oxford.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens & H. Wagner, 2013. vegan: Community Ecology Package, Available from: http://CRAN.R-project.org/package=vegan.

Oláh, J., T. Kovács, I. Sivec, I. Szivák & G. Urbanic, 2012. Seven new species in the *Chaetopteryx rugulosa* species group: applying the phylogenetic species concept and the sexual selection theory (Trichoptera: Limnephilidae). Folia Historico Naturalia Musei Matraensis 36: 51–79.

Oláh, J. & T, Kovács, 2013: New species and new records of Balkan Trichoptera II. Folia Historico Naturalia Musei Matraensis 37: 109–121.

Oppold, A-M., J. A. M. Pedrosa, M. Balint, J. B. Diogo, J. Ilkova, J. L. T. Pestana & M. Pfenninger, 2016. Support for the evolutionary speed hypothesis from intraspecific population genetic data in the non-biting midge *Chironomus riparius.* Proceedings of the Royal Society B 283: 20152413.

Papadopoulou, A., I. Anastasiou & A. P. Vogler, 2010. Revisiting the insect mitochondrial molecular clock: The Mid-Aegean Trench calibration. Molecular Biology and Evolution 27: 1659–1672.

Pauls, S. U., W. Graf, P. Haase, H. T. Lumbsch & J. Waringer, 2008. Grazers, shredders and filtering carnivores–the evolution of feeding ecology in Drusinae (Trichoptera: Limnephilidae): insights from a molecular phylogeny. Molecular Phylogenetics and Evolution 46: 776–791.

Pauls, S. U., C. Nowak, M. Bálint & M. Pfenninger, 2013. The impact of global climate change on genetic diversity within populations and species. Molecular Ecology 22: 925–946.

Pauls, S. U., M. Apl, M. Bálint, P. Bernabò, F. Čiampor Jr, Z. Čiamporová-Zaťovičová, D.S. Finn, J. Kohout, F. Leese, V. Lencioni, I. Paz-Vinas & M. T. Monagham, 2014: Integrating molecular tools into freshwater ecology: developments and opportunities. Freshwater Biology 59: 1559–1576.

Posada, D. & K. A. Crandall, 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817–

818.

Previšić, A., C. Walton, M. Kucinić, P. T. Mitrikeski & M. Kerovec, 2009. Pleistocene divergence of Dinaric Drusus endemics (Trichoptera, Limnephilidae) in multiple microrefugia within the Balkan Peninsula. Molecular Ecology 18: 634–647.

Previšić, A., W. Graf, S. Vitecek, M. Kučinić, M. Bálint, L. Keresztes, S.U. Pauls & J. Waringer, 2014a. Cryptic diversity of caddisflies in the Balkans: the curious case of Ecclisopteryx species (Trichoptera: Limnephilidae). Arthropod Systematics & Phylogeny 72: 309–329.

Previšić, A., J. Schnitzler, M. Kučinić, W. Graf, H. Ibrahimi, M. Kerovec & S. U. Pauls, 2014b. Micro-scale vicariance and diversification of Western Balkan caddisflies linked to karstification. Freshwater Science 33: 250–262.

Previšić, A., A. Gelemanović, G. Urbanič & I. Ternjej, 2016. Cryptic diversity in the Western Balkan endemic copepod: Four species in one? Molecular Phylogenetics and Evolution 100: 124–134.

Rambaut, A. & A. J. Drummond, 2009. Tracer v1.5, Available from: http://beast.bio.ed.ac.uk/Tracer.

Ratnasingham, S. & P. D. N. Hebert, 2007. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7: 355–364.

R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.Vienna, Austria, http://www.R-project.org/.

Rissler, L. J. & J. J. Apodaca, 2007. Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). Systematic Biology 56: 924–942.

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

Rundle H. D., L. Nagel, J. Wenrick Boughman & D. Schluter, 2000. Natural selection and parallel speciation in sympatric sticklebacks. Science 287: 306–308.

Rundle, H. D. & P. Nosil, 2005. Ecological speciation. Ecology Letters 8: 336–352.

Schluter, D., 2001. Ecology and the origin of species. Trends in Ecology & Evolution 16: 372–380.

Schluter, D., 2009. Evidence for ecological speciation and its alternative. Science 323: 737–741.

Sebe, K., G. Csillag & G. Konrád, 2008. The role of neotectonics in fluvial landscape development in the Western Mecsek Mountains and related foreland basins (SE Transdanubia, Hungary). Geomorphology 102:

55–67.

Sotiropoulos, K., K. Eleftherakos, G. Dzukic, M. Kalezic, A. Legakis & R. Polymeni, 2007. Phylogeny and biogeography of the alpine newt *Mesotriton alpestris* (Salamandridae, Caudata), inferred from mtDNA sequences. Molecular Phylogenetics and Evolution 45: 211–226.

Statzner, B. & S. Dolédec, 2011. Phylogenetic, spatial, and species-traits patterns across environmental gradients: the case of Hydropsyche (Trichoptera) along the Loire River. International Review of Hydrobiology 96: 121–140.

Ustaszewski, K., S. M. Schmid, B. Fügenschuh, M. Tischler, E. Kissling & W. Spakman, 2008. A map-view restoration of the Alpine-Carpathian-Dinaridic system for the Early Miocene. Swiss Journal of Geosciences 101: 273–294.

Venables, W. N. & B. D. Ripley, 2002. Modern Applied Statistics with S. Springer, New York, USA.

Vitecek, S., M. Kučinić, J. Oláh, A. Previšić, M. Bálint, L. Keresztes, J. Waringer, S.U: Pauls & W. Graf, 2015. Description of two new filtering carnivore Drusus species (Limnephilidae, Drusinae) from the Western Balkans. Zookeys 513: 79–104.

Xie, W., P. O. Lewis, Y. Fan, L. Kuo & M. H. Chen, 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Systematic Biology 60: 150–160.

Zhang, Y., C. Chen, L. Li, C. Zhao, W. Chen & Y. Huang, 2014: Insights from ecological niche modelling on the taxonomic distinction and niche differentiation between the black-spotted and red-spotted tokay geckoes (*Gekko gecko*). Ecology and Evolution 4: 3383–3394.

Tables

Table 1.

Association of environmental preference on the phylogenetic tree measured by the phylogenetic signal of climatic variables (using Blomberg's K-statistic). The high K value (>1) indicate strong phylogenetic signal of the observed trait and *p-value* indicate the statistical significance based on variance of phylogenetically independent contrasts relative to tip shuffling randomization. LD1 and LD2 represent the contribution of climatic variables to the first two linear discriminant functions. *P*-values indicate statistical significance based on Pearson correlation: *** $p<0.001$; ** $p<0.01$; * $p<0.05$.

Climatic parameters	Abbr.	Blomberg's K	*p-value*	LD1	LD2
Temperature seasonality (standard deviation *100)	bio_4	0.586	0.049	0.25*	-0.75***
Min. temperature of coldest month	bio_6	0.429	0.05	-0.48***	0.16
Temperature annual range	bio_7	0.697	0.011	0.53***	-0.52***
Mean temperature of coldest quarter	bio_11	0.446	0.038	-0.47***	0.24*
Precipitation of driest month	bio_14	0.78	0.017	-0.11	0.75***
Precipitation seasonality (coefficient of variation)	bio_15	1.15	0.002	0.12	-0.59***
Max. temperature of December	tmax_12	0.459	0.041	-0.45***	0.31**
Mean temperature of December	tmean_12	0.438	0.050	-0.48***	0.25*

524 **Figure captions**

525

526 **Fig. 1** Distribution of the *Chaetopteryx rugulosa* species complex with the 79 populations used for habitat
527 condition analyses. Different symbols indicate the locations of different species (for locality data see Table
528 S2). Red symbols mark relatively 'warm' habitats; blue symbols indicate relatively 'cold' habitats.

529

530 **Fig. 2** Phylogeny of the *C. rugulosa* species complex. The phylogeny is based on three genes *(mtCOI,*
531 *nuWG, nuEF-1α)*. Lineages with black circles on the nodes are supported by Bayesian posterior probabilities
532 *pp* ≥ 0.95. The posterior probabilities are shown on the nodes. 'Warm' habitat conditions are marked with
533 light gray, 'cold' habitat conditions are marked with dark gray.

534

535 **Fig. 3** Dated phylogeny of the *C. rugulosa* species complex, and the evolution of global temperature
536 reflected by benthic oxygen isotope records (Lisiecki & Raymo, 2005). Estimated average dates of lineage
537 formation are indicated in million years. Less likely node heights appear blurred on the graph. 'Warm' habitat
538 conditions are marked with light gray, 'cold' habitat conditions are marked with dark gray.

539

540 **Fig. 4** Climatic habitat condition analysis of the *C. rugulosa* complex's species using LDA. The axes mark
541 linear functions that discriminate among the populations of the *C. rugulosa* species on the basis of climatic
542 habitat conditions. For climatic variable codes see Table 1. Different symbols indicate the different species.
543 Red symbols indicate relatively 'warm' habitats; blue symbols indicate relatively 'cold' habitats.

544

545 **Fig. 5** Plot of cross-validation table for climatic niche segregation of *C. rugulosa* complex's species based on
546 LDA analysis. Correctly classified populations are placed on the diagonal. The square size equals to the
547 percentage proportion of populations of posterior group assignment based on posterior probabilities. Rows
548 correspond to *a priori* defined species, while columns correspond to inferred species. Squares with broken
549 lines show different climate niches ('warm' and 'cold').

550

Fig. 6 Paleogeography map of the Central Paratethys and the Transtethyan Trench Corridor. In the Middle Miocene (16-14 My ago) extensions of the Central Paratethys and the Transtethyan Trench Corridor are marked with light green (Mandic et al., 2012; Harzhauser et al., 2015). The retreating extension of Lake Pannon from 8.0 Mya to 4.5 Mya are marked with different colours and line styles: dash-dot-dot blue line 8 My ago; dash green line 6.5 My ago, and solid brown line 4.5 My ago (on Magyar et al., 1999).

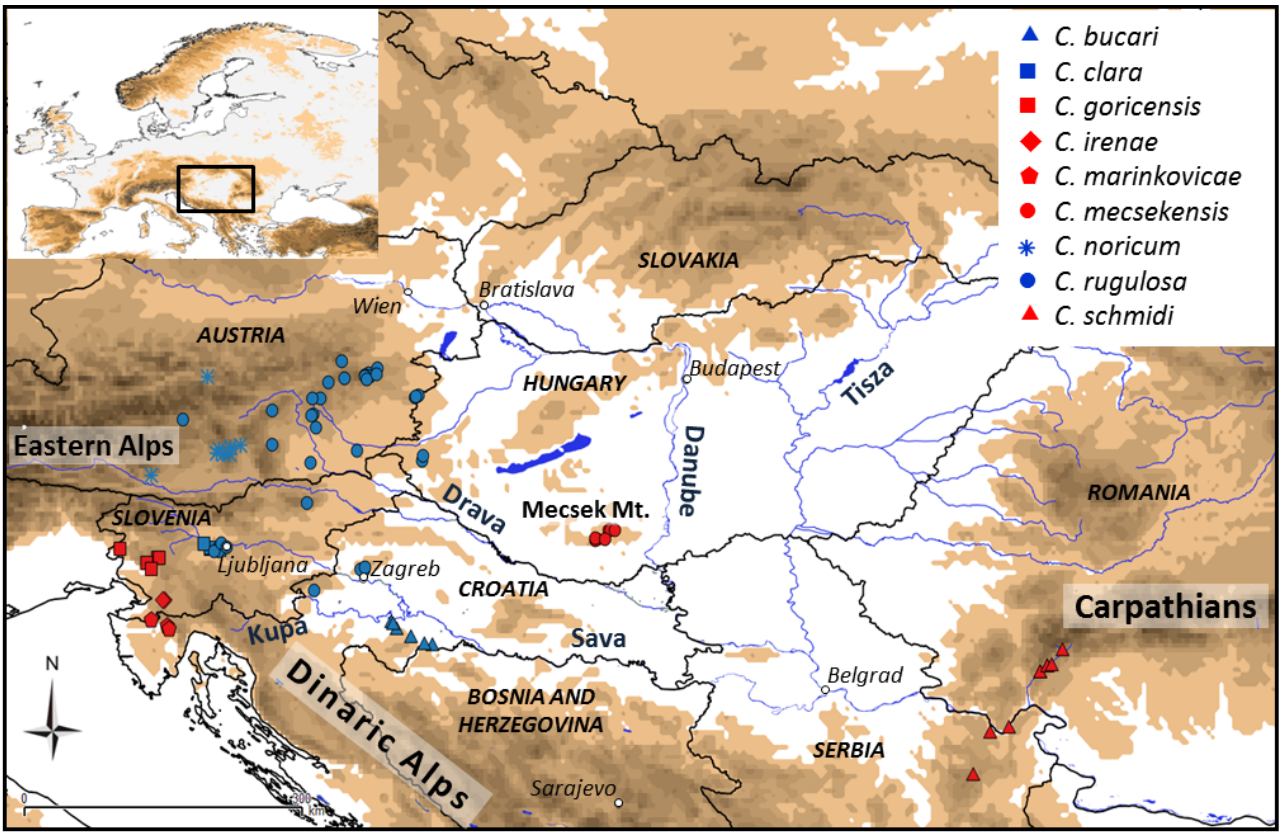
Supplementary Materials

Table S1. Collection data, Genbank and Barcoding of Life sequence accession codes and collector information of *Chaetopteryx rugulosa* species complex and outgroup specimens.

Table S2. Locality data of all populations within *C. rugulosa* species complex. These localities were used to compile the climatic dataset. The table includes the bedrock geological characteristics of the habitats. Abbreviations: fgsed - Fine-grained sedimentary rocks, sand - Sandstone, lime - Limestone, dolomite, coal - Coal, volc – Volcanic and volcaniclastic rocks, felig - Felsic igneous rocks, intig - Intermediate igneous rocks, mafig - Mafic igneous rocks, mbas - Metapelitic rocks, mpel - Metabasic rocks.

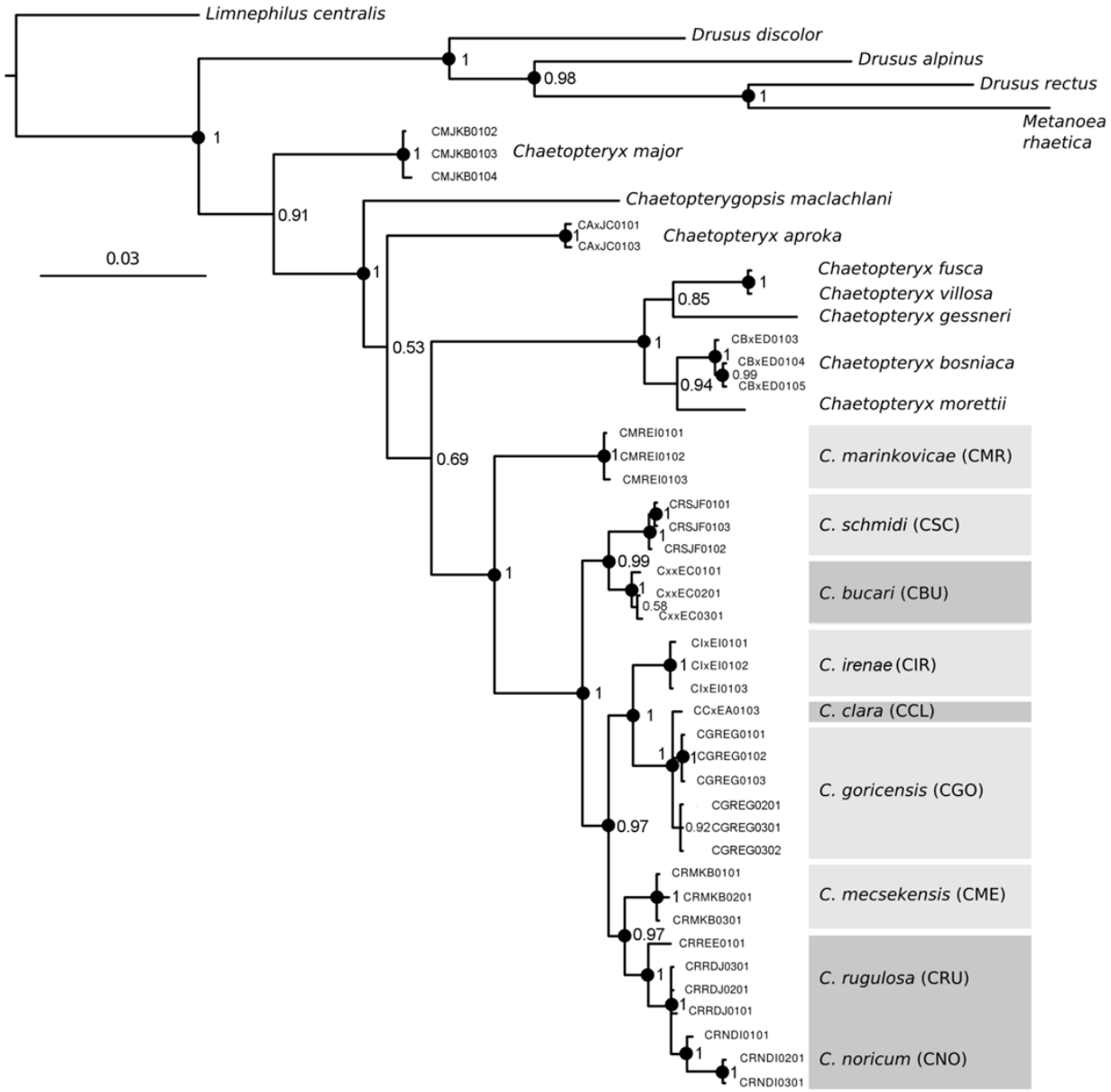
Table S3. Climatic data layers used to infer the climatic conditions in the habitats of populations within *C. rugulosa* species complex.

570 **Figure 1**
571



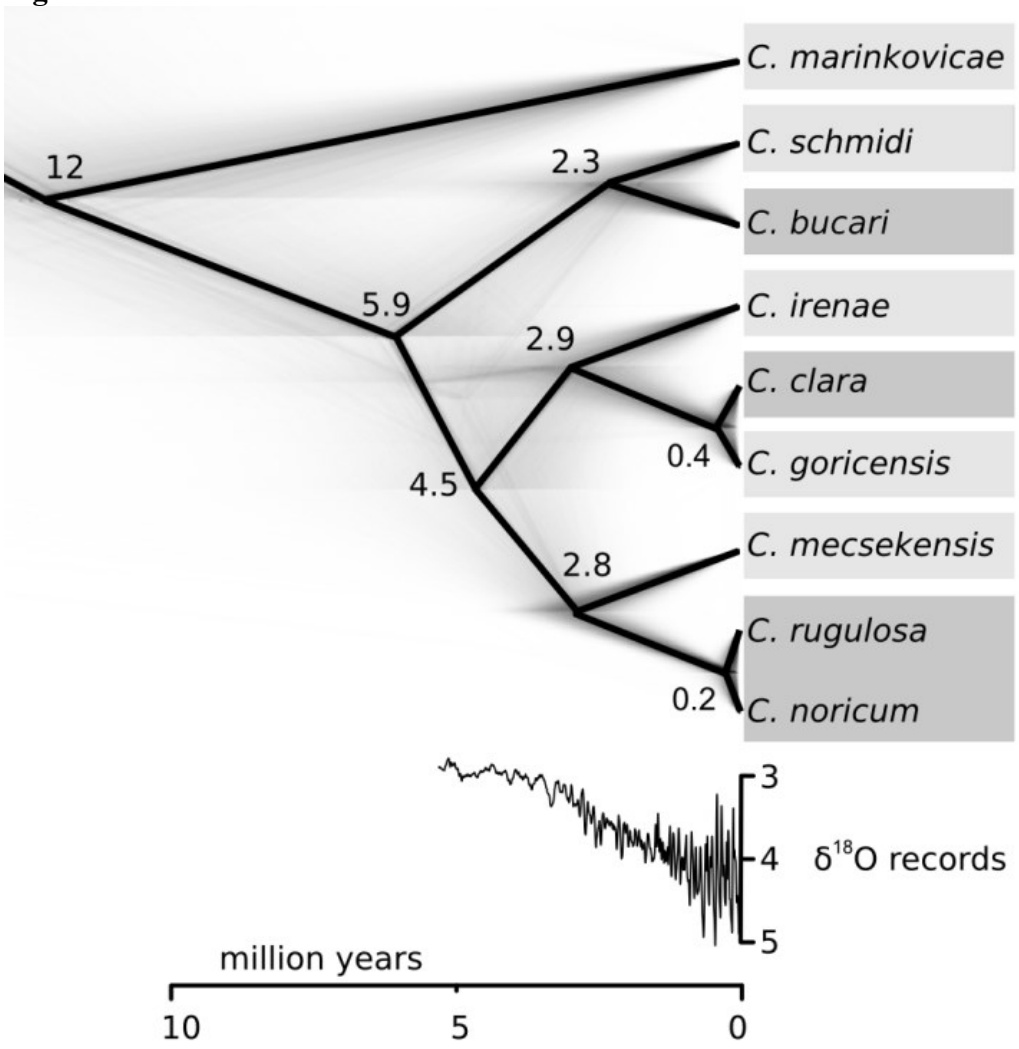
572
573

574 **Figure 2**
575



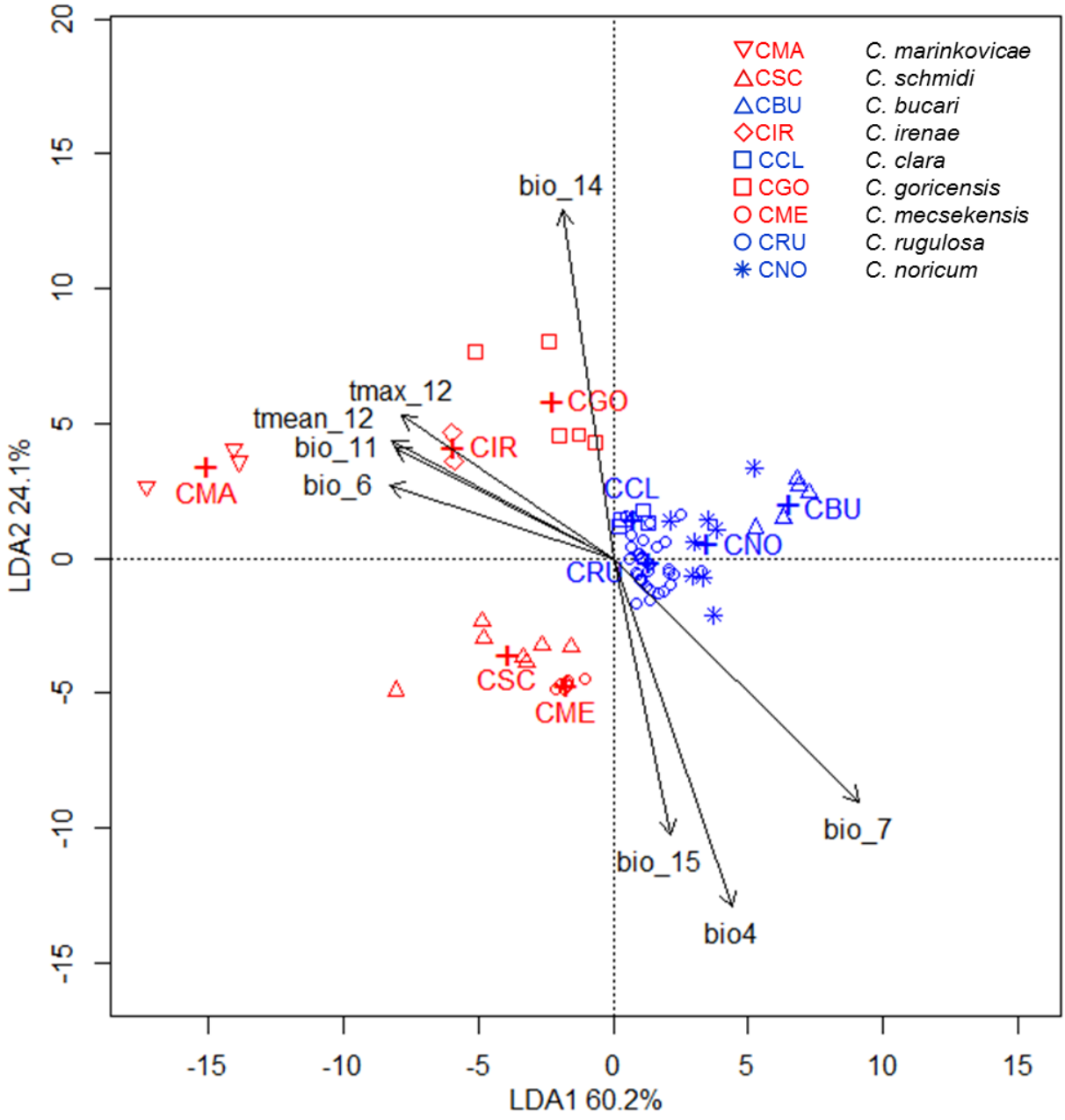
576
577

578 **Figure 3**

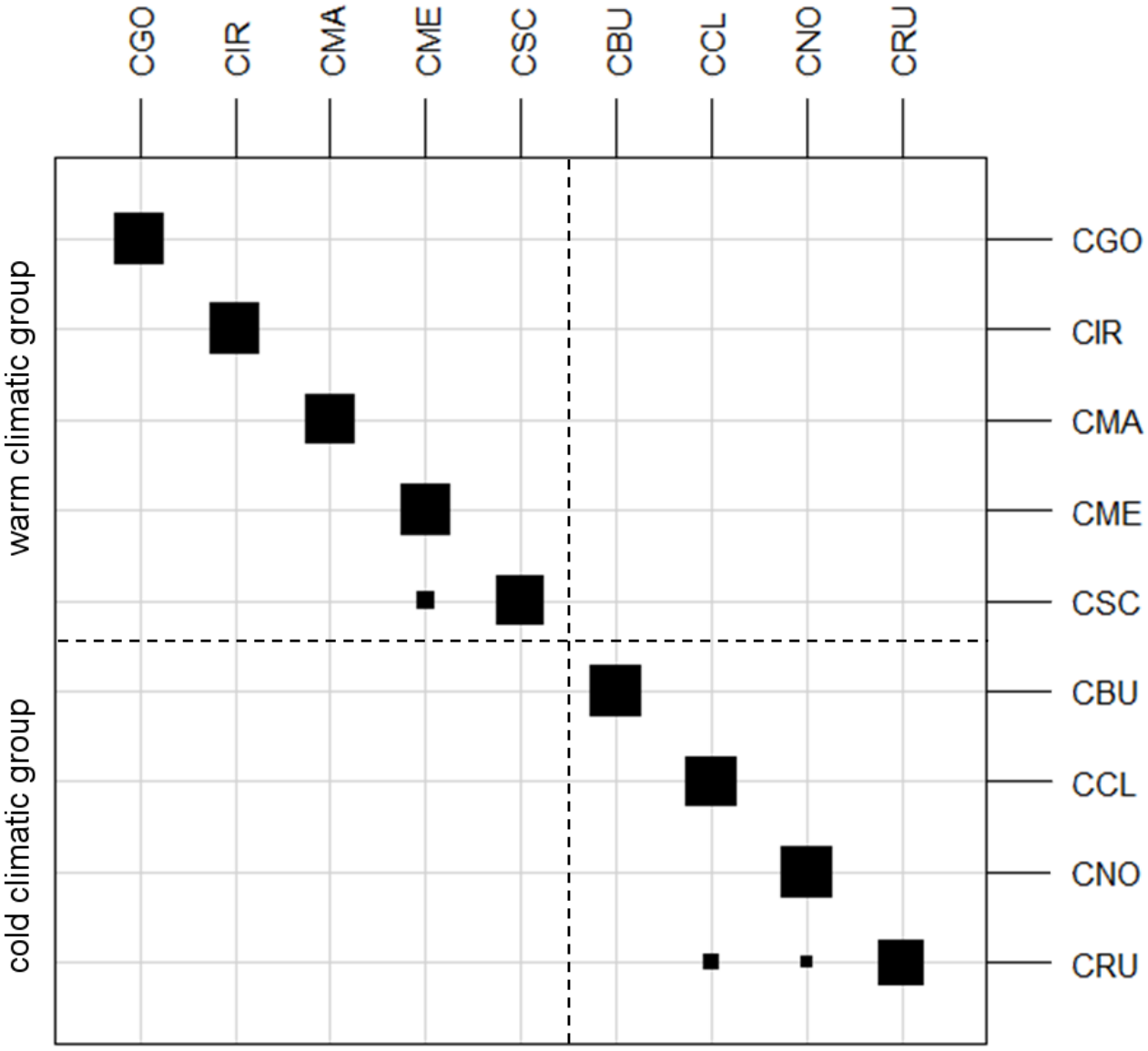


579
580

Figure 4


CMA C. marinkovicae
CSC C. schmidi
CBU C. bucari
CIR C. irenae
CCL C. clara
CGO C. goricensis
CME C. mecsekensis
CRU C. rugulosa
CNO C. noricum
bio_14
tmax_12
tmean_12
bio_11
bio_6
bio_7
bio_15
bio4
LDA1 60.2%
LDA2 24.1%

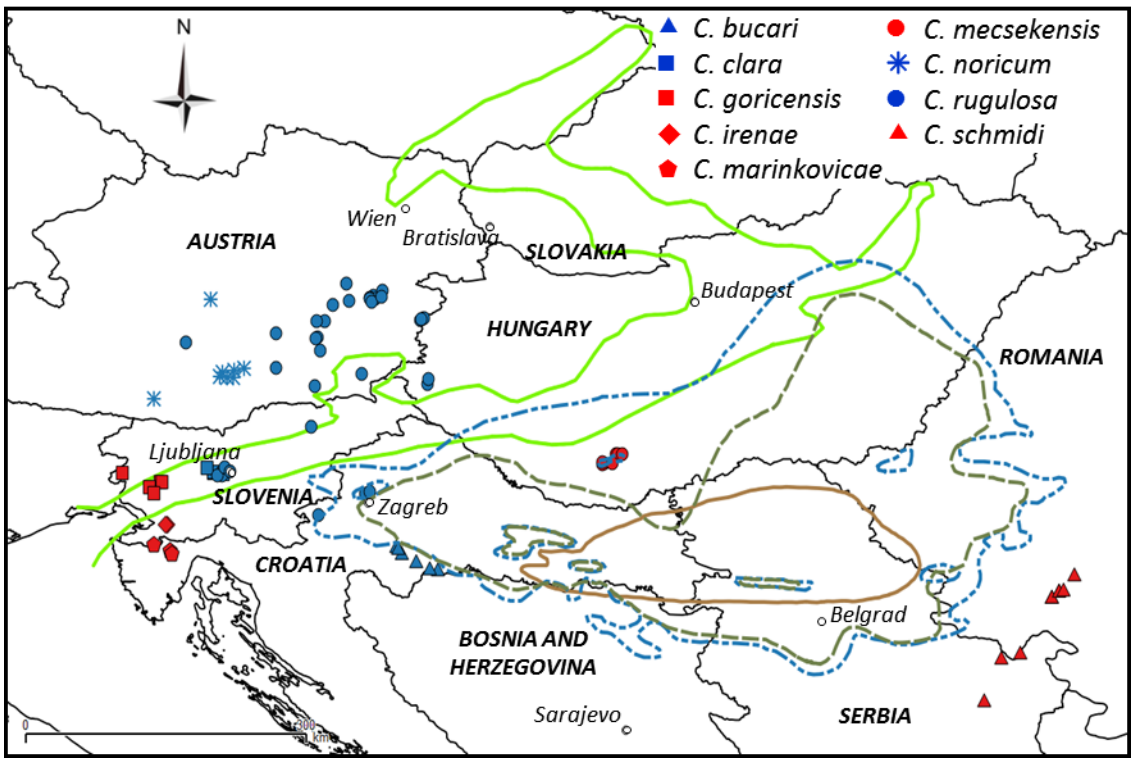
585 **Figure 5**
586



587
588

589 **Figure 6**

590



591