Biological Journal of the Linnean Society, 2009, **97**, 346–361. With 3 figures

Varying rates of diversification in the genus *Melitaea* (Lepidoptera: Nymphalidae) during the past 20 million years

JULIEN LENEVEU[1]*, ANTON CHICHVARKHIN[2] and NIKLAS WAHLBERG[1,3]

[1]*Laboratory of Genetics, Department of Biology, University of Turku, 20014 Turku, Finland*
[2]*Institute of Biology and Soil Science, 100-Letiya Vladivostoka 159, Vladivostok 690022, Russia*
[3]*Department of Zoology, Stockholm University, 106 91 Stockholm, Sweden*

Received 19 November 2008; accepted for publication 22 November 2008

The influence of Quarternary glacial cycles on the extant diversity of Holarctic species has been intensively studied. It has been hypothesized that palaeoclimatic changes are responsible for divergence events in lineages. A constant improvement in DNA sequencing and modeling methods, as well as palaeoclimatic reconstruction, permit a deeper exploration of general causes of speciation in geological time. In the present study, we sampled, as exhaustively as possible, the butterflies belonging to the genus *Melitaea* (Lepidoptera: Nymphalidae), which are widely spread in the Palaearctic region. We conducted analyses to assess the phylogeny of the genus and estimated the timing of divergence and the most likely distribution of ancestral populations. The results obtained indicate that the systematics of the genus is in need of revision and that the diversity of the genus has been profoundly shaped by palaeoenvironmental changes during its evolutionary history. The present study also emphasizes that, when employed with caveats, major palaeoenvironmental events could represent very powerful tools for the calibration of the dating of divergences using molecular data.

ADDITIONAL KEYWORDS: colonization – divergence time – palaeoclimate – speciation rate – systematic.

INTRODUCTION

Discovering how rates of speciation vary over time and over space is one of the main goals for investigating the speciation process (i.e. the origins of biodiversity). It is fundamental to discover whether there exist some general causes of lineage divergence, causes which might be related to the questions of 'when' and 'where' speciation has occurred (Barraclough & Nee, 2001). The answer to the question of 'when' has long been the contribution of palaeontology to evolutionary biology: a replacement of species in changing paleoenvironments during the geological epochs on earth was noted in early studies (Agassiz & Gould, 1860; Cuvier, 1812). Subsequently, the development of radiometric dating during the first half of the 20th Century allowed more precise absolute dating of rocks and the fossils found within them (Laming, 1965). However, the fossil record is well-known to be incomplete and, in the last 20 years, with the rise of DNA sequencing, the molecular clock has become a useful tool for inferring accurate divergence times within extant taxa when calibrated with fossils (Glazko, Koonin & Rogozin, 2005; Donoghue & Benton, 2007).

The question of 'where' is far more delicate: it appears difficult to rigorously test alternative hypotheses concerning the geography of speciation when only information about the phylogeny and current range of extant species is available (Losos & Glor, 2003; Kodandaramaiah & Wahlberg, 2007). However, repeated patterns of phylogeography at the species level for various groups of organisms suggest that some taxonomic divergence events can be attributed to the climatic oscillations during the Quaternary (Avise, Walker & Johns, 1998; Santucci, Emerson &

*Corresponding author. E-mail: julien.leneveu@utu.fi

Hewitt, 1998; Albre, Gers & Legal, 2008). Populations have diverged or experienced drastic demographic changes, and subsequent local adaptation and/or genetic drift are considered to lead to genetic, morphological, and behavioural differentiation among lineages (Hewitt, 2004; Pérez Tris *et al.*, 2004). The geographical localization of the speciation events can be inferred according to the extensive knowledge about the timing and ecological effect of the climatic oscillations or, more generally, by taking into account information on extreme paleoenvironmental changes (Hewitt, 2000). Thus, a reconstructed phylogenetic hypothesis that includes most of the smaller evolutionary entities (species or populations) of a higher level group, when taken together with data on the chronology and geography, may allow inferences on the rate and the causes of biological diversification within the studied taxon.

The Quaternary glacial ages and its impact on the biogeography and evolution of European species have been intensively studied, but this is not the only relatively recent extreme environmental change in the area. The Late Miocene to Early Pliocene interval was characterized in Europe and West Asia by large-scale marine transgressions of the Mediterranean Sea, with the opening of aquatic corridor in alternation with episodes of regressions and the appearance of terrestrial passages (Agusti, Oms & Meulenkamp, 2006). The so-called Messinian age ends with a complete isolation of the Mediterranean Sea from the Atlantic Ocean approximately 300 000 years ago and the subsequent desiccation of the Mediterranean basin (Hsu, Ryan & Cita, 1973; Krijgsman *et al.*, 1999). These dramatic changes in sea level were accompanied by major climatic changes (Kovac *et al.*, 2006) and modifications in the vegetal cover (Francois *et al.*, 2006), and some biogeographical implications for mammals (Agusti, Garces & Krijgsman, 2006) as well as insects (Weingartner, Wahlberg & Nylin, 2006; Kodandaramaiah & Wahlberg, 2009) have been pointed out.

In the present study, we aim to infer the evolutionary history of the butterflies that belong to the genus *Melitaea* (Lepidoptera: Nymphalidae) and show the influence of paleoenvironmental changes on their diversity. The genus *Melitaea* comprises approximately 80 species, all restricted to the Palaearctic region. The taxonomy of the group has been the subject of numerous studies and species in the genus have been placed in up to four genera. The first phylogenetic study of the group showed that two of the genera (*Melitaea* and *Didymaeformia*) were para- or polyphyletic, and that the most stable solution was to include all the species in *Melitaea* (Wahlberg & Zimmermann, 2000). Subsequent studies have shown that the genus is very strongly supported and stable (Wahlberg, Brower & Nylin, 2005; Wahlberg & Freitas, 2007), but the relationships of species and species-groups in the genus have not been studied in detail.

MATERIAL AND METHODS

Taxonomic sampling

We sampled a total of 74 individuals of 65 described species belonging to the genus *Melitaea*. For some species, two or more individuals have been included in the final analysis (for details, see molecular methods and discussion below). In addition, we used 28 species belonging to the tribe Melitaeini as our outgroup, and rooted our tree using the genus *Euphydryas*. The DNA sequences for these outgroups were taken from previous studies (Wahlberg & Freitas, 2007). The sampled specimens are listed in Table 1. Most of the < 20 unsampled species are morphologically very close to sampled species, with the exception of the very rare *Melitaea yuenty* (found in south-east China), which is morphologically quite divergent from the rest of the species.

Molecular methods

We generally extracted DNA from two legs of dried adult butterflies using the DNEasy extraction kit (Qiagen); the final extract being eluted in 20–50 μL of eluting buffer. For a small number of specimens, DNA was extracted from the thorax of the adult or half of the head of larvae using the same extraction kit. Three genes have been sequenced for this study, including a mitochondrial gene cytochrome oxidase subunit I (*COI*, 1487 bp) and two nuclear genes, elongation factor-1α (*EF-1α*, 1240 bp) and wingless (*Wgl*, 403 bp).

We carried out polymerase chain reactions in a 20-μL reaction volume that included 1 μL of DNA extract. Primers and cycling profiles are given in Table 2. The primers contain a standard tail and we subsequently used universal forward and reverse primers for the sequencing reactions (Wahlberg & Wheat, 2008). Sequences were checked and aligned by eye using BIOEDIT software (Hall, 1999). Alignment and phylogeny inferences are fundamentally interdependent and the use of independent analysis could lead to biased and overconfident estimations (Lunter *et al.*, 2005), but the very low taxonomic level investigated in the present study (genus level phylogeny) implies a high level of similarity between sequences and a subsequent low risk of misalignment.

We coded heterozygous positions for the nuclear genes *EF-1α* and *Wgl* as ambiguities using the Nomenclature Committee of the International Union of Biochemistry's codes (IUPAC). The GenBank accession numbers are given in Table 1. Initial

Table 1. Specimen list

Species	Voucher code	Collection locality	Range	Accession numbers (GenBank)		
				COI	*EF-1α*	*Wgl*
Euphydryas phaeton	NW13-3	USA: Maryland	NA	AF187797	AY788747	AF153975
Euphydryas desfontainii	NW70-4	Spain: Cataloña	NA	AY090226	AY090193	AY090159
Anthanassa texana	NW12-6	USA: Texas	NA	AF187806	AY788716	X
Chlosyne acastus	NW35-15	USA: Colorado	NA	AF187735	AY788725	AY788486
Chlosyne cyneas	NW38-17	Ecuador: Sucumbios	NA	AY187757	AY788726	AY788487
Chlosyne gaudealis	NW37-2	France: La Selva	NA	AF187770	AY788727	AY788488
Eresia eunice	NW92-5	Brazil: Bertioga	NA	AY788624	AY788738	AY788499
Eresia quintilla	NW76-3	Ecuador: Esmeraldas	NA	AY788627	AY788741	AY788502
Higginsius fasciata	NW87-1	Peru: Cuzco	NA	AY788630	AY788749	AY788510
Janatella leucodesma	NW85-16	Panama	NA	AY788641	AY788761	AY788521
Mazia amazonica	NW76-6	Ecuador	NA	AY788654	AY788773	AY788533
Microtia elada	NW7-1	USA: Texas	NA	AY788659	AY788786	AY788546
Phyciodes graphica	NW67-9	Mexico: Jilotepec	NA	AY156684	AY788790	AY788550
Phyciodes picta	NW34-7	USA: Colorado	NA	AF187800	AY788796	AY788556
Poladryas arachne	NW27-4	USA: California	NA	AF187783	X	X
Melitaea acraeina	NW139-5	Uzbekistan: Komsomolabad	B	FJ462229	FJ462289	FJ462164
Melitaea aetherie	NW103-12	Morocco: Moyen Atlas	C	FJ462230	FJ462290	FJ462165
Melitaea ala	AC4-2	China: Tian-Shan	B	FJ462231	FJ462291	FJ462166
Melitaea ambigua	NW10-1	Mongolia: Tov Aimak	D	AF187736	FJ462292	FJ462167
Melitaea ambrisia	NW139-3	Uzbekistan: Kuramin Mt	B	FJ462232	FJ462293	FJ462168
Melitaea amoenula	NW23-15	India: Taglong Ladak	B	AF187737	FJ462294	FJ462169
Melitaea arcesia	NW10-9	Mongolia: Tov Aimak	B,D	AF187741	FJ462295	FJ462170
Melitaea arduina	NW23-5	Greece: Pisoderi	A,B	AF187742	AY788774	AY788534
Melitaea asteria	NW142-19	Italy: Trentino	A	FJ462233	FJ462296	X
Melitaea athalia	NW76-14	Sweden: Vallentuna	A	FJ462234	FJ462297	FJ462171
Melitaea athene	NW15-4	Kazakhstan: Zaisan	B	AF187799	FJ462298	FJ462172
Melitaea aurelia	NW23-2	France: Dijon	A	AF187745	FJ462299	FJ462173
Melitaea avinovi	NW122-11	China: Pamir	B	FJ462235	FJ462300	X
Melitaea bellona	NW144-10	China: Wudu	B	FJ462236	FJ462301	FJ462174
Melitaea britomartis	NW15-13	Russia: Saratov	A	AF187748	FJ462302	FJ462175
Melitaea cassandra	AC3-9	Russia: Suusamyk Range	B	FJ462237	FJ462303	FJ462176
Melitaea casta	NW85-3	Iran: Küh-e-Sorkh	B	FJ462238	FJ462304	FJ462177
Melitaea caucasogenita	NW24-12	Turkey: Posof	A	FJ462239	FJ462305	FJ462178
Melitaea celadussa	AC6-14	France: Aude	A	FJ462240	FJ462306	FJ462179
Melitaea centralasiae	NW19-5	Russia: Djirga	B	FJ462241	FJ462307	FJ462180
Melitaea chitralensis	AC4-11	China: Pamir	B	FJ462242	FJ462308	FJ462181
Melitaea chuana	NW142-18	China: Tibetan Plateau	B	FJ462243	FJ462309	FJ462182
Melitaea cinxia	JL3-2	Morocco: Atlas	C	EF680410	FJ462310	X
Melitaea cinxia	NW73-14	Sweden: Stockholm	A,B	AY788656	AY788776	AY788536
Melitaea collina	JL2-7	Iran: Lorestan Prov.	A	FJ462244	FJ462311	FJ462183
Melitaea consulis	NW85-5	Iran: Küh-e-Garbos	A	FJ462245	FJ462312	FJ462184
Melitaea deione	NW150-13	Morocco: Atlas	C	FJ462247	FJ462314	FJ462185
Melitaea deione	JL126	France: Alpes	A	FJ462246	FJ462313	FJ462186
Melitaea deserticola	NW34-12	Lebanon: Bouârej	A	AF187759	FJ462315	FJ462187
Melitaea deserticola	JL3-10	Morocco: Atlas	C	FJ462248	FJ462316	FJ462188
Melitaea diamina	NW10-24	Mongolia: Ulaanbaatar	A,B	AF187761	FJ462317	FJ462189
Melitaea didyma	NW99-12	Russia: Kilmesh	B	FJ462249	FJ462318	FJ462190
Melitaea didyma	AC6-7	Spain: Cataloña	A	FJ462251	FJ462320	FJ462192
Melitaea didyma	NW107-5	Morocco: Moyen Atlas	C	FJ462253	FJ462322	FJ462194
Melitaea didyma	AC7-8	France: Aude	A	FJ462252	FJ462321	FJ462193
Melitaea didyma	AC3-3	Russia: Suusamyk Range	B	FJ462250	FJ462319	FJ462191
Melitaea didymoides	NW28-14	China: Hebei	D	FJ462254	FJ462323	FJ462195

Table 1. *Continued*

Species	Voucher code	Collection locality	Range	Accession numbers (GenBank)		
				COI	*EF-1α*	*Wgl*
Melitaea elizabethae	AC4-7	China: Pamir	B	FJ462255	FJ462324	FJ462196
Melitaea enarea	NW113-15	Tadjikistan: W. Pamir	B	FJ462256	FJ462325	FJ462197
Melitaea fergana	AC3-12	Kyrgyzstan: Fergamsky Mts.	B	FJ462257	FJ462326	FJ462198
Melitaea gina	NW85-4	Iran: Dascht-e-Arjan	B	FJ462258	X	X
Melitaea infernalis	NW36-1	Kazakhstan	B	FJ462259	FJ462327	FJ462199
Melitaea interrupta	NW17-3	Russia: Arkhyz	A	FJ462260	FJ462328	FJ462200
Melitaea latonigena	NW25-3	Russia: Utitzcina	D	FJ462261	FJ462329	FJ462201
Melitaea leechi	NW67-7	China	B	FJ462262	FJ462330	FJ462202
Melitaea ludmilla	AC3-11	Russia: Suusamyk Range	B	FJ462263	FJ462331	FJ462203
Melitaea lutko	NW15-3	West China	B	FJ462264	FJ462332	FJ462204
Melitaea lunalata	AC5-3	China: Tian Shan	B	FJ462265	FJ462333	FJ462205
Melitaea maracandica	AC5-1	China: Pamir	B	FJ462266	FJ462334	FJ462206
Melitaea menetriesi	NW113-12	Russia: Kamchatka	D	FJ462267	FJ462335	FJ462207
Melitaea minerva	NW113-3	Uzbekistan: Kuramin Mts.	B	FJ462268	FJ462336	FJ462208
Melitaea ninae	NW113-10	Kirgizstan	B	FJ462269	FJ462337	FJ462209
Melitaea pallas	AC4-9	China: Tian Shan	B	FJ462270	FJ462338	FJ462210
Melitaea parthenoides	JL1-2	France: Pyrénées	A	FJ462271	FJ462339	FJ462211
Melitaea permuta	NW139-4	Uzbekistan: Gissar Mts	B	FJ462272	FJ462340	FJ462212
Melitaea persea	NW120-11	Iran: Ardabil	A,B	FJ462273	FJ462341	FJ462213
Melitaea phoebe	NW15-14	Russia: Saratov	B	FJ462274	FJ462342	FJ462214
Melitaea phoebe	AC6-6	Spain: Cataloña	A	FJ462275	FJ462343	FJ462215
Melitaea plotina	NW113-7	Russia: Urulga	D	FJ462277	FJ462345	FJ462217
Melitaea protomedia	NW40-6	China: Peking	D	FJ462278	FJ462346	FJ462218
Melitaea punica	JL3-7	Morocco	C	FJ462276	FJ462344	FJ462216
Melitaea romanovi	NW99-9	Russia: S Buryatia	D	FJ462280	FJ462348	FJ462220
Melitaea saxatilis	NW120-8	Iran: Tehran	A,B	FJ462281	FJ462349	FJ462221
Melitaea scotosia	NW27-11	China: Hebei Prov.	D	AF187804	AY788780	AY788740
Melitaea shandura	AC5-16	China: Pamir	B	FJ462282	FJ462350	FJ462222
Melitaea sibina	NW140-10	Kirgizstan	B	FJ462283	FJ462351	FJ462223
Melitaea solona	NW113-1	Kirgizstan	B	FJ462284	FJ462352	FJ462224
Melitaea sultanensis	NW113-13	Kirgizstan: Trans Alai	B	FJ462285	FJ462353	FJ462225
Melitaea sutschana	NW19-9	Russia: Chita Region	D	AF187805	FJ462354	FJ462226
Melitaea telona	AC5-11	Lebanon	A,B	FJ462279	FJ462347	FJ462219
Melitaea trivia	AC7-3	Spain: Cataloña	A	FJ462280	FJ462355	FJ462227
Melitaea trivia	NW23-6	Greece: Pisoderi	A	AF187810	AY788782	AY788542
Melitaea varia	NW24-13	France: Alpes	A	AF187812	AY788783	AY788543
Melitaea wiltshirei	NW140-12	Iran: Hamadan	B	FJ462288	FJ462356	FJ462228

Letters in the 'range' column correspond to geographic areas used in the analysis of historical biogeography. A, Western Palaearctic, excluding North African zone. B, Central Palaearctic, including Tibetan Plateau. C, Northern Africa. D, Eastern Palaearctic. Note that the specimens used as outgroups have not been included in the analysis (NA).

screening of specimens to include in the study was performed by comparing *COI* sequences (DNA 'barcodes') of a large number of individuals (http://nymphalidae.utu.fi/Vouchers.htm), with any highly divergent individuals identified morphologically as belonging to the same species being chosen for further sequencing. Thus, some species have multiple individuals sampled in this study.

Phylogenetic analysis

Inferences about the relationships between the individuals were estimated in two different ways. First, we searched for the most parsimonious tree from the three genes equally weighted combined dataset using the program TNT, version 1.1 (Goloboff, Farris & Nixon, 2008). We performed the analysis using the

Table 2. Primers used for amplifying and sequencing DNA

Gene			Primers	
Cytochrome oxydase I (*COI*)	Part I	F	LCO	5′-GGT CAA CAA ATC ATA AAG ATA TTG G-3′
		R	HCO	5′-TAA ACT TCA GGG TGA CCA AAA AAT CA-3′
	Part II	F	Jerry	5′-CAA CAY TTA TTT TGA TTT TTT GG-3′
		R	Pat	5′-ATC CAT TAC ATA TAA TCT GCC ATA-3′
Elongation factor-1α (*EF-1α*)	Part I	F	Starsky	5′-C ACA TYA ACA TTG TCG TSA TYG G-3′
		R	Luke	5′-C ATR TTG TCK CCG TGC CAK CC-3′
	Part II	F	CHO	5′-GTC ACC ATC ATY GAC GC-3′
		R	Verdi	5′-GAT ACC AGT CTC AAC TCT TCC-3′
	Part III	F	EF51.9	5′-CAR GAC GTA TAC AAA ATC GG-3′
		R	EFrcM4	5′-ACA GCV ACK GTY TGY CTC ATR TC-3′
Wingless (*Wgl*)		F	LepWG1	5′-GAR TGY AAR TGY CAY GGY ATG TCT GG-3′
		R	LepWG2	5′-ACT ICG CAR CAC CAR TGG AAT GTR CA-3′

new technology search (running successively the Sectorial, Ratchet, Drift, and Tree Fusing algorithm for 100 random addition rounds) (Goloboff, 1999). Support was estimated for the resulting clades using the bootstrap resampling method based on 1001 replications. Second, we carried out a Bayesian analysis on the dataset using MrBayes, version 3.1.2 (Ronquist & Huelsenbeck, 2003). The search was performed on the combined dataset with parameter values estimated separately for each gene region, using the General Time Reversible (GTR) model of sequence evolution and a variation in the rate following a gamma distribution. The analysis was run twice simultaneously for 10 000 000 generations with every 1000 trees sampled. We discarded the first 500 000 generations (500 samples) as burn-in (based on visual inspection of the convergence and stability of the log likelihood values of the two independent runs). This analysis provides, at the same time, an estimation of the relative genetic distance between clades (branch length) and a branch support (a posteriori probability) taking into account the a priori parameters.

Dating of lineages divergence and rate of speciation

To estimate the age of divergence between taxa, we carried out a relaxed clock method Bayesian analysis using BEAST, version 1.4.6 (Lunter *et al.*, 2005). This software allows a phylogenetic analysis of the dataset using the maximum sum of clade credibility topology tree inferred from the Bayesian analysis. The xml input file was primarily created by BEAUti, version 1.4.6 (included in the package) and then edited by hand to include two partitions (one for the mitochondrial gene and one for the two nuclear genes combined). The output file contains the estimated dating including the 95% credibility intervals.

Because there are no fossils of butterflies belonging to the genus *Melitaea* available, the age constraint has been given in accordance with the recent study on the subfamily Nymphalinae (Wahlberg, 2006), which estimated the origin of the genus *Melitaea* at 21.7 Mya (Bayesian relaxed clock; a priori parameters (0.002, 0.02); 95% confidence interval = 15.5–29.4 Mya). Bayesian analysis was performed using the same model (GTR + Γ) that was used previously in the MrBayes analysis, allowing different rate of substitution for the mitochondrial and nuclear genes. The rate of mutation was assumed to be variable among the tree (relaxed clock model), with no relation a priori between a lineage's rate and that of its ancestor (uncorrelated) and with an underlying lognormal distribution. This a priori is particularly advantageous here: BEAST adjusts the parameters among the processed generations and gives an estimation of how clock-like the dataset is via the standard deviation parameter and coefficient of variation parameter (Drummond *et al.*, 2007).

The rate of speciation was represented by plotting the logarithm of the number of lineages against the relative time of each node since the root node. Under the constant speciation rate model, the probability of divergence event per time is equal over time and among species, and a straight line should be expected (Barraclough & Nee, 2001). Nevertheless, we believe that a comparison of the curve with a unique simulated 'straight' line as seen, for example, in McKenna & Farrell (2006) is not appropriate because the aim in this type of study is to observe multiple change in the rhythm during the focal period of time.

Biogeographic analysis

The method used to estimate the most likely ancestral distribution states over the phylogeny is based on

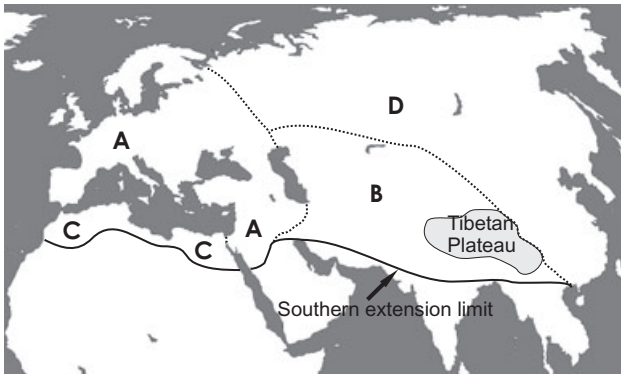

Figure 1. Map showing the maximal extension of the sampled species belonging to the genus *Melitaea* in Eurasia and Northern Africa. The map shows also the subdivisions used in our study. A, Western Palaearctic, excluding North African zone. B, Central Palaearctic, including Tibetan Plateau. C, Northern Africa. D, Eastern Palaearctic.

the null hypothesis of vicariance as an explanation for diversification events. We used the program DIVA, version 1.1 (dispersal-vicariance analysis; Ronquist, 1997), which was previously used in a number of recent studies on butterflies (Braby & Pierce, 2007; Wahlberg & Freitas, 2007; Kodandaramaiah & Wahlberg, 2009). The program assigns a cost of 0 for vicariance and sympatric lineage divergence and a cost of 1 for dispersal and extinction and the least cost ancestral state reconstruction is assumed to be the most probable. This means that the inferred historical dispersal events are based on a conservative hypothesis.

The total present geographical distribution of the genus *Melitaea* was divided in to four zones according to the knowledge of the distribution of extant endemic species (Fig. 1). The estimated ancestral range was restricted during the DIVA analysis (command 'maxareas') to the observed present maximal taxon area distribution (i.e. two zones).

RESULTS

The systematics of *Melitaea*

The parsimony analyses conducted with TNT on the combined dataset of the three genes resulted in 16 equally parsimonious phylogenetic trees of 5535 steps long (CI = 0.29, RI = 0.61). The strict consensus tree is compared with the Bayesian phylogenetic inference in Figure 2. The two topologies are broadly congruent, although the Bayesian inference appears to produce more resolution for some nodes where the polytomies are resolved with strong posterior probabilities. The same nomenclature has been used to refer to different taxonomic levels (Higgins, 1981), often in a confusing manner. To facilitate discussion, we refer to informal species groups by the names defined in Figure 2B. We have chosen these groups based on their stability to method of analysis and the robustness of their monophyly.

The genus *Melitaea* as circumscribed by Wahlberg & Zimmermann (2000) is, by all analysis made for the present study, strongly supported as a monophyletic group with respect to the outgroups used. It is primarily subdivided into two sister clades that are robust, which we call the *Melitaea* and *Didymaeformia* clades (Fig. 2). The *Melitaea* clade brings together 27 sampled species that were previously placed in the 'subgenera' *Melitaea* and *Mellicta*. *Melitaea* (*sensu* Higgins) is inferred to be polyphyletic after our analyses in both regard to the position in the branch (Fig. 2) and to the fact that *Melitaea romanovi*, *Melitaea avinovi*, *Melitaea arduinna*, and *Melitaea lutko*, traditionally part of this subgroup are spread in the *Didymaeformia* clade. It has been here divided into five clades: the widespread *Melitaea cinxia* (*cinxia* group), the *diamina* group (including *Melitaea protomedia*), the *arcesia* group (that includes *Melitaea chuana*, *Melitaea bellona*, *Melitaea amoenula*, as well as *Melitaea leechi*), the *minerva* group (*Melitaea elizabethae*, *Melitaea solona*, *Melitaea ludmilla*, *Melitaea sultanensis*, and *Melitaea pallas*) and, finally, the *athalia* species group [comprising all species in the *Mellicta* of Higgins (1955)]. The *athalia* group is sister to the *minerva* group; it is monophyletic and is clearly a subgroup of the *Melitaea* clade; a result that is globally congruent with Wahlberg & Zimmermann (2000). The group has previously been described by Higgins (1955), who provided an overview of this homogenous clade. A noteworthy result in the *athalia* group is that *Melitaea celadussa*, which is usually considered a subspecies of *Melitaea athalia* (Higgins, 1941, 1955; Lafranchis, 2000), is not directly related to *M. athalia*. Our analyses show that *M. athalia* is more closely related to *Melitaea caucasogenita* and *Melitaea ambigua*, this species group being sister to the *Melitaea deione*/*Melitaea britomartis* branch, and finally *M. celadussa* being sister to that clade (Fig. 2).

The *Didymaeformia* clade comprises Higgins' (1941) *didyma*, *fergana*, *collina* and *phoebe* species-groups, which he subsequently split into the genera *Didymaeformia* and *Cinclidia* (Higgins, 1981) (38 species included here); these subdivisions found no support in the present study, being respectively paraphyletic and contained within a larger clade as previously found by Wahlberg & Zimmermann (2000). *Melitaea trivia*, previously included in the *Didymaeformia* genus and part of the *collina* group (paraphyletic) (together with *M. collina* and *Melitaea consulis*),

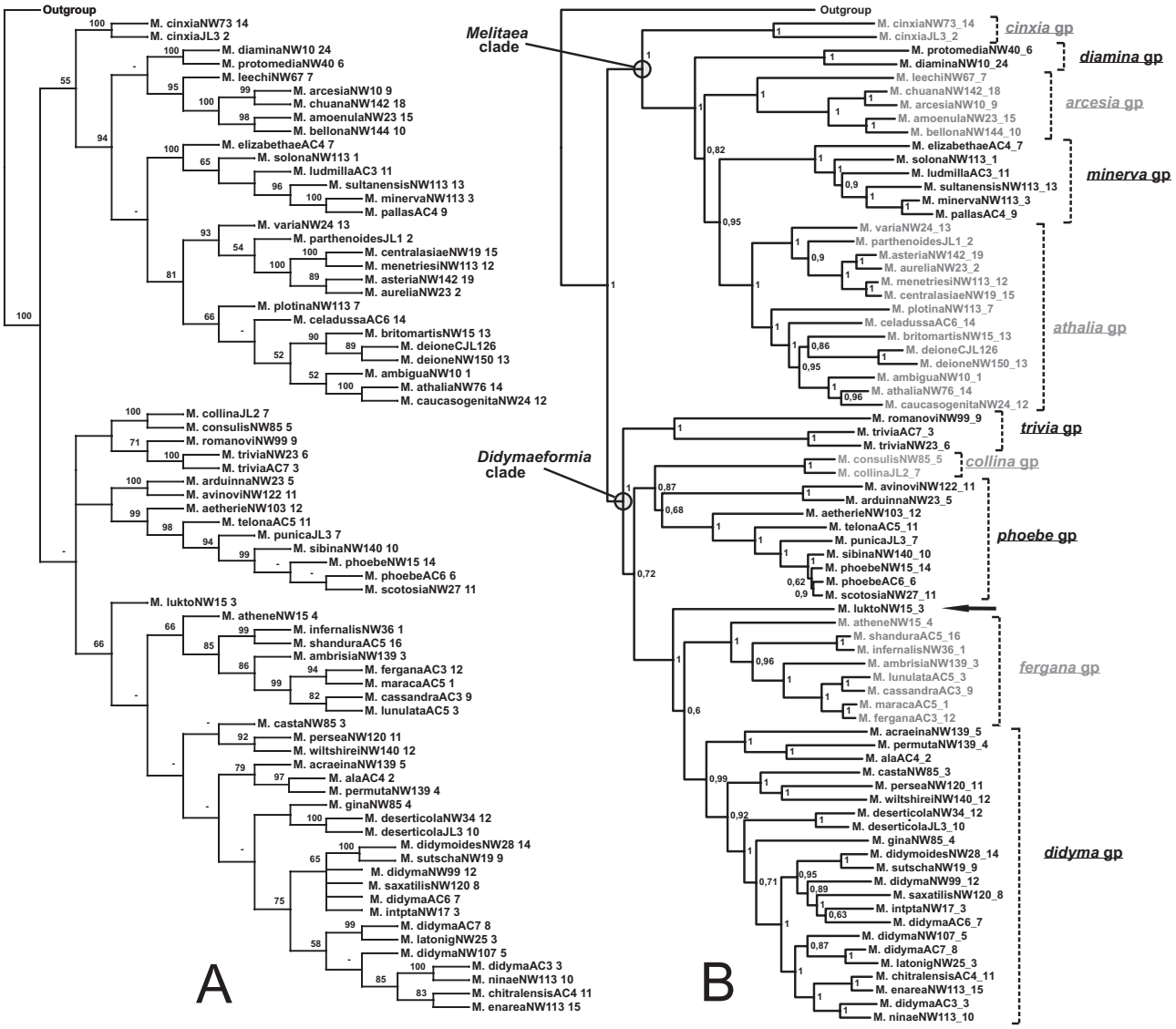

Figure 2. Comparison between parsimony and Bayesian phylogenetic inferences for the genus *Melitaea*. A, strict consensus of 16 most parsimonious trees, where the number to the left of each node is the bootstrap value estimated via resampling method based on 1001 replications (values below 50 are not shown). B, Bayesian topology. Values to the right of the nodes are posterior probabilities. The topology of the Bayesian tree will be used as a reference for the taxonomic inferences (see Discussion); proposed subdivisions are shown on the tree. The genus *Melitaea* is subdivided into two main clades: The *Melitaea* clade (five species groups) and the *Didymaeformia* clade (four species groups plus one isolated individual).

appeared to be allied in the basal position of that group with *M. romanovi*, which was expected to be part of the previously circumscribed genus *Melitaea* (*sensu* Higgins). Higgins' *phoebe* group (*Melitaea phoebe*, *Melitaea scotosia*, *Melitaea aetherie*, *Melitaea sibina*, *Melitaea telona*, and *Melitaea punica*) appear to be sister to *M. arduinna* + *M. avinovi*, which have previously been placed in Higgins' genus *Melitaea*. Our result is not strongly supported, but appears to be stable to method of analysis and we thus tentatively include the two species in the *phoebe* group. *Melitaea punica* and *M. telona* have often been considered synonymous and subspecies of *M. phoebe*. Our results suggest that both are independent lineages genetically quite distinct from *M. phoebe* + *M. sibina* + *M. scotosia*. The latter three are in fact almost identical genetically. In his first revision of the genus, Higgins had also included *M. collina* and *M. consulis* in the *phoebe* group (Higgins, 1941) and our Bayesian results (Fig. 2B) show that this might be

appropriate but, because the position of the two species is not stable, we have decided to refer to it as a separate species group.

The next clade up consists of three lineages corresponding to the *fergana* and *didyma* groups, as well as the independent lineage leading to *M. lutko*. The latter anomalous butterfly has been found only in a very restricted area in the Central Palaearctic and it has not been placed in any group with confidence on the basis of morphological characters (Higgins, 1941). Our analysis shows that the species is sister to the clade formed by the *fergana* + *didyma* groups, although this position is not strongly supported in either the parsimony or the Bayesian analysis (Fig. 2).

The eight following species: *Melitaea shandura*, *Melitaea infernalis*, *Melitaea ambrisia*, *Melitaea lunalata*, *Melitaea fergana*, *Melitaea maracandica*, *Melitaea cassandra*, and *Melitaea athene* form a monophyletic group strongly supported by the different analysis. The first six species had been grouped together by Higgins (1941) (as the '*fergana* group', an informal name that has been retained in the present study) on the basis of the very restricted area where they are found, great mountain ranges of central Asia and morphological characters. The recently described *M. cassandra* (by Kolesnishenko and Churkin in 2001) has been identified in the same region and our analyses suggest that it is closely related to *M. lunalata.* The general shape of the genitalia of the butterflies belonging to the *fergana* group suggests a close ancestral relationship with the *didyma* group and indeed the phylogeny shows that they are sister groups.

Our intraspecific *Melitaea didyma* sampling revealed a complex pattern of relationships; the individuals do not group together; the five included individuals are spread with good support values in a subgroup of the didyma group that contains also *Melitaea latonigena*, *Melitaea saxatilis*, *Melitaea interrupta*, *Melitaea chitralensis*, *Melitaea enarea*, *Melitaea ninae*, *Melitaea didymoides*, and *Melitaea sutschana*. This clade is contained in a larger clade '*didyma* group' together with *Melitaea acraeina*, *Melitaea permuta*, *Melitaea ala*, *Melitaea casta*, *Melitaea persea*, *Melitaea wiltshirei*, *Melitaea deserticola*, and *Melitaea gina* as previously proposed by Higgins (1981).

Biogeography of *Melitaea* through recent geologic ages

The analysis conducted on the Bayesian topology estimated the origin of species diversity in *Melitaea* somewhere in Central Palaearctic during the early Miocene (Fig. 3). The two main sister clades described above began diverging soon after during the early Burdigalian at approximately 20 Mya, given our calibration point of 21 Mya for the crown group *Melitaea*.

According to our current dataset, the *Melitaea* clade subdivided first to give birth to the currently widespread *M. cinxia* lineage. The phylogeography of *M. cinxia* has been recently studied by Wahlberg & Saccheri (2007), who showed that *M. cinxia* individuals found in Morocco were highly divergent to sampled Eurasian individuals. In the present study, the split between North African and Eurasian populations is inferred to have occurred during the late Miocene (Messinian age). The next divergence event is estimated to have occurred during the middle Miocene and gave birth to another widespread species group: the *diamina* group. *Melitaea protomedia* has been assumed to be a subspecies of *Melitaea diamina* (restricted to Central Asia) and our analysis shows that the most recent ancestor of the sampled individuals lived during the Messinian. The next important node is situated in the late Miocene (approximately 10 Mya): the morphologically homogenous *arcesia* group diversified (the origin of the group is older, Langhian Age) occupying exclusively the Tibetan Plateau in the central Palaearctic zone (Figs 1 & 3). Approximately at the same time, the *minerva* group spread in a wider area in the Central Palaearctic.

The *athalia* group is sister clade to the *minerva* group, and the analysis suggests a colonization event from the Central Palaearctic toward the Western Palaearctic during the middle Serravallian. The *athalia* group shows basically the same temporal pattern of divergence as the previously described groups but for species principally distributed in the Western part of Eurasia: the group subdivided first during the early Tortonian into two sister groups for which the main diversification events took place during the Messinian.

The *Didymaeformia* clade gives first birth to *M. trivia* and *M. romanovi* lineages (*trivia* group). The divergence between these two species occurred early during the Langhian, they are positioned at the end of a very long branch and they both exhibit particular current biogeographic patterns. *Melitaea trivia* is regarded as 'an archaic form' by Higgins (1941), its distribution is relatively large, and genetic divergence between local populations is important as shown by our sampling. Included individuals of this species from Eastern and Western Europe (grouped in the analysis under Western Palaearctic) are estimated to have evolved individually from approximately 3.9 Mya. *Melitaea romanovi* inhabits a quite limited area in the steppes near the Baikal Lake, where it is always present at low density.

The pattern of very long branches is also found at the basal position of the *collina* and *phoebe* groups,

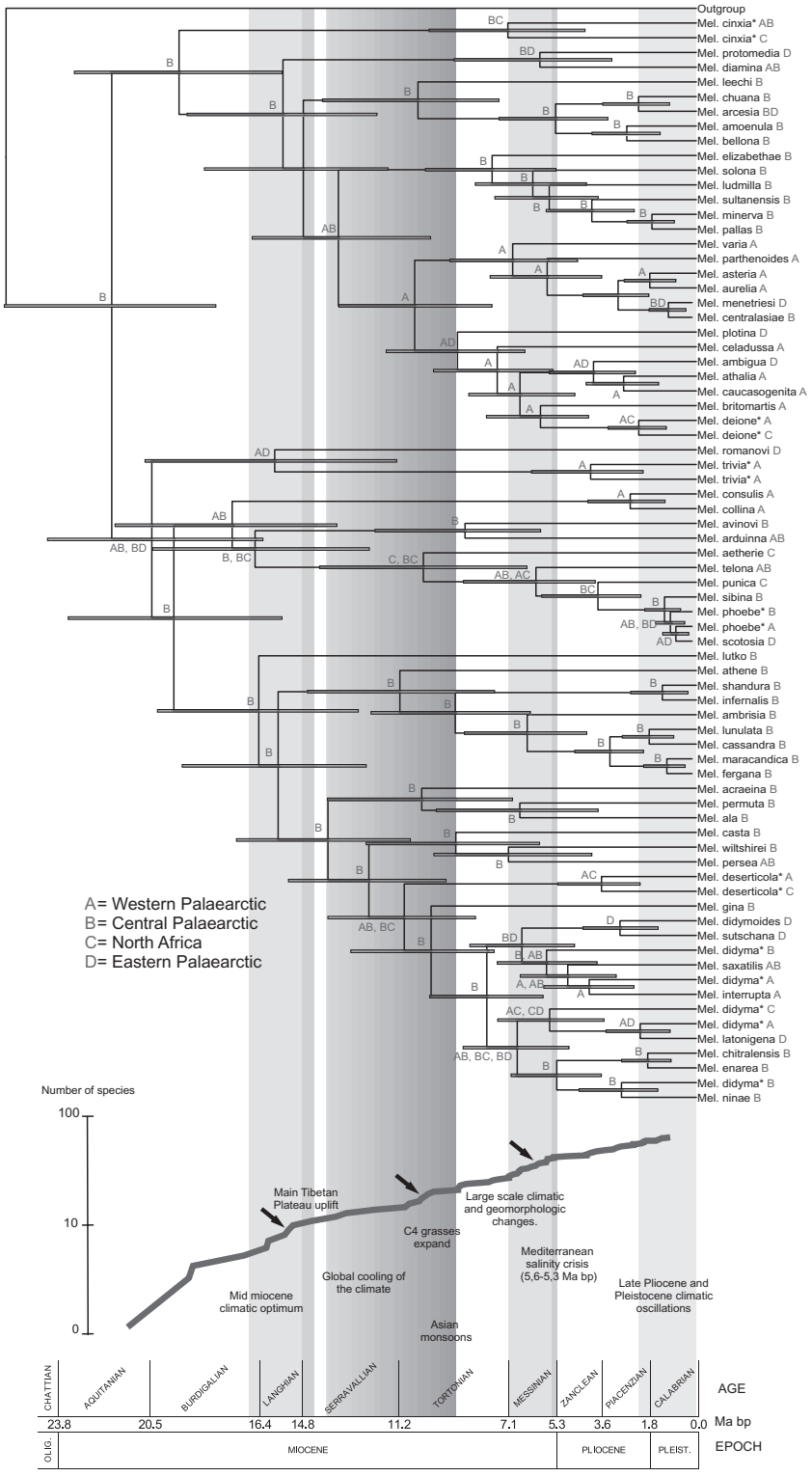

Figure 3. Estimated timing and location of the events of speciation during the evolutionary history of the genus *Melitaea*. The Bayesian ultrametric tree shows the 95% a posteriori confidence intervals for the dating of each node. Grey letters (cf. legend on the left of the tree) following the names of the taxa define their current distribution and ancestral estimated ranges are shown for every node. The curve below shows the cumulative number of species through time (logarithmic scale); turns in the slope reflect changes in the rhythm of speciation or extinction. Vertically shaded areas on the figure indicate the principal environmental changes that occurred in the area during the studied period.

with *M. consulis* and *M. collina* sharing a most recent common ancestor with the rest of the group that lived during the late Burdigalian (Fig. 3). This common ancestor is estimated to have been widespread in the Palaearctic, whereas the most recent ancestor of the *collina* group was exclusively living in the Western Palaearctic. Another long branch leads to *M. arduinna* and *M. avinovi*, whose lineage is estimated to have diverged from the rest of the *phoebe* group in the early Langhian mainly in the Central Palaearctic. The rest of the species belonging to the *phoebe* group show a variable pattern of distribution, with the ancestral distributions being inferred as relatively widespread. Interestingly, the divergences of *M. phoebe*, *M. sibina*, and *M. scotosia* are inferred to be more recent (Pleistocene) than the divergence of *M. punica* (Pliocene), which is often considered a subspecies of *M. phoebe*.

The *fergana* group is estimated to originate in the Central Palaearctic during the Langhian with a phylogeographical pattern very similar to the *arcesia* group in the *Melitaea* clade. It is also characterized by species that are restricted to mountain zone in the Central Palaearctic. It diverged from its sister group (the *didyma* group) approximately 15 Mya during the main uplift of the Himalaya.

As described previously, our analyses highlight the complexity in the pattern of relationship between the different populations of *M. didyma*; for example, one individual sampled in Southern France is closely related, with a time to most recent common ancestor (tmrca) close to 2 Mya, to the individual of *M. latonigena* (sometimes regarded as a subspecies of *M. didyma*) caught in the Eastern Siberian mountains. The populations they belong to are together sister to the North African population of *M. didyma* (estimated tmrca 5 Mya). The same pattern of genetic divergence has also been identified between the two individuals of *M. deserticola* that have been included. They have been sampled in Syria (zone incorporated in the Western Palaearctic) and North Africa and their tmrca is by the analysis estimated to be close to the latter (i.e. the drying out of the Mediterranean Sea). Our results also indicate that *M. ninae* (also known as *M. pseudoala*) was long ago genetically isolated from *M. ala*, apparently more closely related to *M. didyma*.

Variation in the rate of speciation through time

As shown above, it appears that the events of speciation are not randomly distributed in time. In the very first part of the curve, we notice that the rhythm of speciation starts relatively high and decreases slightly after two million years. This 'push of the past' is likely to be an artefact resulting from the fact that we are considering those clades that survived to the present day, and these are the ones that, on average, got off to a flying start (Nee *et al.*, 1994). A second change in the rhythm of speciation is estimated to have occurred during the Langhian, contemporary with the main Tibetan Plateau uplift (Coleman & Hodges, 1995). This is the time when the *arcesia* and *fergana* groups, which comprise exclusively butterflies found in the mountain ranges associated with the Tibetan Plateau uplift, originated. Subsequently the speciation rate remained slower and stable for approximately 4 Mya. The next turn is situated during the early Tortonian and, subsequently, another one takes place in the middle Tortonian at a time when it is believed that the climate globally cooled down, the ice sheet appeared in Antarctica as well as the monsoon in Asia (Zachos *et al.*, 2001). Several turns are visible on the curve during the Messinian: the rhythm of speciation increase by pulse and remains high until the end of the Age. The Messinian is well known for the dramatic climatic and geomorphological changes that occurred beginning with opening of terrestrial communication between Eurasia and North Africa. The Mediterranean Sea dried up completely during 5.6 and 5.3 Mya, after which the Straits of Gibraltar reopened and has remained open until present (see below). Our analysis shows that, out of the five events of divergence between Africa and Eurasia, four could have taken place during the relatively brief Messinian period. The curve remains more flat for the last 5 million years, although unsampled species (which are likely to be closely related to sampled species, see Material and methods) mean that the rates of speciation are underestimated at the tips of the tree.

DISCUSSION

About the methodology

Interpretations, subjectively made from the results, are based on the reliability of the inferences, which are objectively constructed from non-exhaustive data by algorithms. The reliability of the inferences made can be assessed by the indices of confidence given by the analyses. When estimating the phylogenetic relationships of organisms, these indices are represented by node support (e.g. bootstrap values for parsimony analysis and posterior probabilities for Bayesian inferences). In the present study, the topology of the parsimony and Bayesian based inferences are very similar, but the average support of the nodes appears to be stronger in the latter, although the comparability of the two different forms of support has been the subject of debate (Douady *et al.*, 2003). The bootstrap values can be used to evaluate the relative amount of

favourable and contradictory evidences in the data with the given assumptions used to interpret those (Goloboff *et al.*, 2003). In the present study, some nodes are poorly supported by the parsimony analysis, but they are positioned after short branches (Figs 2, 3) and this is likely to be a lack of phylogenetic signal in the data rather than a conflict in the sequences. Inferences about timing of divergence reinforce this point of view because the low bootstrap values (below 50) are placed in areas of rapid speciation on the Bayesian tree. For example, one of the major differences between parsimony and Bayesian inferences is the position of the *arcesia* group, which in the most parsimonious trees is placed as sister to the *diamina* group with no support, but as sister to the *minerva* + *athalia* groups with some support in the Bayesian analyses. The three groups are estimated to have diverged within only 0.8 Mya from the most recent common ancestor of the whole clade (Fig. 3). Similar patterns are found at other points of conflict between the parsimony and Bayesian inferences.

The parameters of the models used in the Bayesian inferences are explored and optimized during the Markov Chain Monte Carlo (e.g. mean rate of evolution and standard deviation of this rate among branches, relative rate of nucleotides substitution). It is interesting to note that inferences from Bayesian analysis do not just measure the support of the most probable topology given priors and adjusted parameters: once a relatively stable phylogenetic topology has been reached, the posterior probability calculated for every node also gives an estimation of stability of that node under limited variation of previously optimized parameters (Drummond & Rambaut, 2004). Our discussion about the evolutionary history of the genus *Melitaea* will be centered essentially on the inferences made out of the Bayesian analysis (Fig. 3).

Systematics and classification of *Melitaea*

We have included the majority of species of *Melitaea* in our study and have discovered the major lineages that lead to well-supported and stable clades (Fig. 2). Almost all missing species can easily be placed into one of the informally named clades based on morphology alone, although their phylogenetic position within the clade will remain unknown until those species are sampled. The exception is the rare species *M. yuenty*, which is morphologically very distinctive and has not been placed with any species group as it recalls various extant taxa (Higgins, 1941). Our phylogenetic hypothesis will form the basis of a rigourous classification of the genus. Although we do find two stable clades that we refer to as the *Melitaea* and *Didymaeformia* clades, we do not recommend that the two clades be elevated to genus level. It is clear that the basal branches of both clades are weakly supported and composed of short branches, which may lead to different relationships when more data are analysed. By contrast, the entire clade comprising *Melitaea*, as we circumscribe it, is a very stable and well-supported entity in the tribe Melitaeini and should be retained as a single genus.

Several patterns of phylogenetic relationships were rather surprising. First, the position of *M. celadussa* as sister to the *M. athalia* and *M. deione* clades clearly indicates that it is a separate species and not a subspecies of *M. athalia*, as has always been assumed. *Melitaea celadussa* and *M. athalia* are known to have a narrow hybrid zone where their ranges meet (Higgins, 1955); thus, the two species have not attained complete reproductive isolation, despite diverging from each other possibly 7 Mya. A more detailed study of the species pair would be necessary to discover whether there is gene flow between the species, but unpublished *COI* sequences of 14 *M. celadussa* and 12 *M. athalia* specimens from throughout their ranges suggest that mitochondrial DNA does not introgress (N. Wahlberg, unpubl. data). Our analyses also suggest that *M. arduinna* + *M. avinovi* are sister to the *phoebe* group, in contrast to previous studies (Wahlberg & Zimmermann, 2000; Wahlberg *et al.*, 2005). This position makes sense when looking at host plant use because both the *phoebe* group and *M. arduinna* are known to utilize *Centaurea* (Asteraceae) species as host plants (i.e. the host plant of *M. avinovi* is unknown), which is unique in all Melitaeini, as well as almost all Nymphalidae (Wahlberg, 2001; Janz, Nylin & Wahlberg, 2006; Nylin & Wahlberg, 2008).

Within the *phoebe* group, noteworthy patterns include the separation of *M. telona* early on from the other species related to *M. phoebe*. The specimen used here (voucher code NW34-11) has earlier been used in previous studies under the name *M. punica* (Wahlberg & Zimmermann, 2000; Wahlberg *et al.*, 2005; Wahlberg & Freitas, 2007). However, in the present study, it is clear that *M. punica* from North Africa (which is where the type locality for the name is) is a separate entity to *M. telona*, with the former being more closely related to *M. phoebe*. It may well be that the name *M. punica* should be restricted to populations found in North Africa, although a larger sampling of specimens from there and from the Middle East are necessary to ascertain this. The close relationship between *M. phoebe*, *M. sibina*, and *M. scotosia* requires further investigation because it appears that they may represent populations or environmental forms of one species, rather than independent species.

The systematics of the *didyma* group is in dire need of revision. Our sampling does not allow for any

conclusions, other than what is considered to be *M. didyma* may well be a series of cryptic species. The clade of most interest is the one that began diverging during the late Tortonian, after the ancestor of *M. gina* had branched off the stem lineage (Fig. 3). This clade is very difficult morphologically, with many forms being described (Higgins, 1941), and our molecular results do not shed much light on species delimitations either. A larger sampling of populations of all species would be necessary to find any consistent patterns.

Evolutionary history of the genus *Melitaea*

Our results on the times of divergences within the genus *Melitaea* are of course contingent on the calibrations that we used for estimating these times. In this case, we have used secondary calibration points taken from a study that was based on fossil evidence (Wahlberg, 2006). This approach has been used to investigate the evolutionary history of other groups of species in the subfamily Nymphalinae, to which *Melitaea* belongs, such as *Junonia* (Kodandaramaiah & Wahlberg, 2007) and the subtribe Phyciodina (Wahlberg & Freitas, 2007).

Our analysis has highlighted a repeated pattern in the evolutionary history of the genus *Melitaea*, which is that narrowly distributed species are found in clades that are often restricted to particular geographic areas. If the events of range extension are rare, it is likely that some intermittent mechanisms are acting and understanding them appears to be essential. Thus, our phylogenetic reconstruction shows, for example, that the species of two groups are found only in mountainous zones: the *fergana* group is constituted only of populations of butterflies from Central Asiatic mountain ranges and the *arcesia* group comprises almost exclusively species which are endemic to the Tibetan Plateau region. Furthermore, the dating estimates that these two independent groups are contemporary with the main Himalayan and Tibetan Plateau uplift. We interpret this, on one hand, as the mountain ranges having been colonized early in their geological history and, on the other hand, the 'door of the colonization process' has been closed after the ancestral populations became established there.

The uplift of the Himalayan range has been contemporary with a dramatic cooling in the climate in Eurasia and evidence correlates the two events together (Sharma *et al.*, 1999; Lavé & Avouac, 2001; Fang *et al.*, 2002). However, prior to the cooling, the climate on Earth had reached a temperature optimum for the Miocene which lasted for approximately 3 million years with temperatures on average 5 °C higher than presently (late Burdigalian/Langhian ages) (Zachos *et al.*, 2001; Bohme, 2003). During this period, the number of species increased significantly in the genus and, by assuming a constant rate of extinction in the lineages, we can interpret this correlation as a diversification of lineages in conjunction with a rise in the temperature. A possible explanation for the rise in species numbers could be that the change in the temperatures opened new niches for the populations of butterflies (e.g. changes in host plant distributions and opening of high elevation mountain passes to the butterflies) and, consequently, selective forces or drift could have acted on newly geographically/ecologically isolated populations. This explanation is reinforced by the previous observation which concerned the mountainous populations (*arcesia* and *fergana* groups). In both cases, it appears that the settlement of one population prevented other closely-related species to settle afterwards; just as if the quantitatively limited suitable niches were all rapidly exploited. It has been argued that the first population that fills a newly available environment disperses and reproduces exponentially, whereas those arriving behind can only do it logistically (Hewitt, 2000).

As noted previously, between the late Langhian and the end of the Serravallian the temperatures dropped globally on earth, by 4 °C in 6 million years (Zachos *et al.*, 2001). Concerning the evolutionary history of the *Melitaea* butterflies, the number of species remained almost stable with no noticeable turn in the slope during the major part of the period. The next upturn takes place at the beginning of the Tortonian (Fig. 3) when the continuous decrease in the temperature coincides with the aridification of the climate in Eurasia and a subsequent expansion of grassland (Pagani, Freeman & Arthur, 1999; Maki *et al.*, 2003). The displacement of largely C_3 vegetation, probably semi-deciduous forest, by C4 grasslands is probably continentwide (Quade *et al.*, 1995). It is likely that the opening of the vegetal cover has had a positive impact on butterfly populations in general (Peña & Wahlberg, 2008), and *Melitaea* species in particular, because they are generally found in open habitats such as meadows and disturbed habitats (Wahlberg, 2001).

This mechanism of isolation of populations based on major paleoenvironmental change can also be identified as an ongoing process in the more proximate part of the ultrametric tree. Repeated patterns of divergence between populations from Eurasia and North Africa can be observed, all taking place in a relatively recent geologic time (after 8 Mya). The use of a unique calibration point at the base of the phylogeny, the obvious impossibility to include extinct taxa, the unexhaustive sampling of extant lineages, the limited amount of molecular data, and the

approximation of the algorithm used for the analysis, comprise some of the reasons that prevented us from obtaining a precise date for the nodes. Nevertheless, it is probable that the Messinian age (7.1 to 5.3 Mya) represents a key period for the colonization of North Africa by species of the genus *Melitaea*. This period has been intensively studied and it is known the Mediterranean Sea completely dried up for 300 000 years between 5.6 and 5.3 Mya (the Mediterranean salinity crisis) as a consequence of the closure of the Straits of Gibraltar (Krijgsman *et al.*, 1999); terrestrial communications would have been possible at that moment.

To test the hypothesis that the divergence events can be related to the Mediterranean salinity crisis, we performed another dating analysis using the end of the Messinian age (that corresponds to the opening of the Straits of Gibraltar) as a reference for the nodes which are suspected to reflect speciation event directly linked to the crisis. Thus, the tmrca has been set to 5.3 Mya for the two included individuals of *M. cinxia*, the two included *M. deserticola*, and *M. didyma* (voucher code NW107-5) in relation with *M. didyma* (voucher code AC7-8); all the other parameters were kept as for the previous analysis. The results are not presented here but they are not significantly different to those shown in Figure 3, with the new estimates lying within the credibility intervals of the previous estimates. It should be noted that the split for the populations of *M. deione* in these two areas is estimated to have occurred at the end of the Pliocene epoch, at a time when the temperature dropped dramatically for a short period of time (Veith, Kosuch & Vences, 2003). This temperature drop is estimated to have been considerably lower than during the Quaternary glacial cycles, and was accompanied by a probable wide-scale extent of continental ice sheets and a decrease of sea level, leading to shrinkage of the Straits of Gibraltar. This may have allowed some butterflies to cross and settle on the opposite shore. The final case of a split between Eurasia and North Africa is *M. phoebe* and *M. punica*, for which divergence is estimated to be in between the previous two scenarios, although the credibility intervals do not exclude either a Messinian origin or a Pliocene origin for the divergence.

The use of multiple calibration points spread widely in a phylogeny is known to make the credibility intervals narrower, and this is what happened in the present study. By being able to use the precisely dated geological events as references in multiples nodes, the analysis becomes much more powerful because variations in the rhythm of speciation over time are likely to be taken into consideration more extensively. On the one hand, it is clear that our last analysis relies completely on a single and debatable hypothesis that is based on an approximation given by the first one. On the other hand, the reasoning is not completely circular because the genetic data are not the only arguments that support this hypothesis. The populations of species present on both sides of the Straits of Gibraltar show a clear pattern of genetic differentiation and, in addition, some extant species are widely spread in Spain but do not appear at all on the other side of the Straits: *M. celadussa*, *M. trivia* and *M. parthenoides* are 'unexpectedly absent, although the conditions of life should be quite suitable' (Higgins, 1941). All these elements are evidence demonstrating that the Mediterranean Sea currently represents a physical barrier between Western Europe and North African for the genus *Melitaea*, which is quite surprising because the Straits of Gibraltar is currently less than 20 km wide, an observation that has also been made for other butterflies (Weingartner *et al.*, 2006; Wiemers & Fiedler, 2007; Kodandaramaiah & Wahlberg, 2009). However, the winds in the Straits are usually strong and predominantly orientated East/West in the area. We have found evidence that this barrier disappeared at some point during the recent geological ages and that these migration events between the two continents are likely to have occurred at the origin of diversification in the lineages after the barrier reappeared, which is a classic case of speciation through geographical isolation (allopatry) (Mayr, 1970).

Conclusions

Changes in the climate or in the morphology of the Earth have often been described as causes that could explain the break-up and subsequent divergence of isolated taxonomic clusters; and this is especially well-documented for recent geologic times. However, the present study is particularly innovative in the sense that it combines relatively precise information about the dating of the diversification event of an extensively sampled group at the lowest taxonomic level together with data about phylogeography, as well as data about paleoenvironmental changes in the concerned areas. First, the present study shows a spatiotemporal correlation between the estimated speciation event and the major paleoenvironmental changes in the Palaearctic region. We believe that this gives more credibility to the results concerning the phylogeny reconstruction, and particularly to the dating of the nodes, as well as to our inference about the evolutionary history of the genus *Melitaea* by answering the seminal questions for evolutionary biologists: when and where (i.e. in what circumstances, with which patterns) does speciation occur.

Concerning the factors implied in the speciation, changes in the environment, such as major climatic

modifications, have been highlighted as being the most important in the diversification of the genus *Melitaea*. The rate of speciation has increased during wide-scale paleoenvironmental events, giving information about the mechanisms likely to be involved in the divergence events: opening of new niches and consequent relocation of populations followed by geographical isolation. This is particularly well symbolized by the impact of the relatively recent Messinian Age on the populations of several species of *Melitaea*, a pattern that is most likely more widely spread among the Palaearctic organisms than previously described. The present study also shows that it is likely that interactions between closely-related lineages at the genus level have shaped the diversity of extant lineages; patterns of competitive exclusion in specific areas such as mountain ranges are visible in the inferences.

Another aspect highlighted in the present study is the question of the duration of speciation. This requires more extensive experimental work (because it implies testing the definition of biological species) but our study shows that the timing of speciation is far from being constant, even when comparing closely-related sister species. Morphologically recognizable sister species sometimes have been estimated by our analysis to have diverged approximately 1 million years ago, although there are cases of extant phylogroup pairs that appear to have been reproductively isolated for more than 5 million years. As shown in other studies (Smith, Lafay & Christen, 1992), it appears that there is a poor correlation between morphological and molecular evolution, which implies that time is acting together with other causes (such as demographical and ecological factors) on the divergence between isolated populations.

ACKNOWLEDGEMENTS

We are very grateful to all those who have provided us with specimens for this study, including Alexei Belik, Sergei Churkin, Willy De Prins, Henri Descimon, Wolfgang Eckweiler, Pavel Gorbunov, Ilkka Hanski, Jaakko Kullberg, Yuri Marusik, Kari Nupponen, Ilya Osipov, Constanti Stefanescu, and Michel Tarrier. We also thank Luc Legal and an anonymous reviewer for comments on the manuscript. This study was funded by grants from the Academy of Finland (grant number 118369) and the Swedish Research Council to N.W.

REFERENCES

Agassiz L, Gould AA. 1860. *Principle of Zoölogy: touching the structure, development, distribution, and natural arrangement of the races of animals, living and extinct*. Boston, MA: Gould & Lincoln.

Agusti J, Garces M, Krijgsman W. 2006. Evidence for African-Iberian exchanges during the Messinian in the Spanish mammalian record. *Palaeogeography, Palaeoclimatology, Palaeoecology* **238:** 5–14.

Agusti J, Oms O, Meulenkamp JE. 2006. Introduction to the Late Miocene to Early Pliocene environment and climate change in the Mediterranean area. *Palaeogeography, Palaeoclimatology, Palaeoecology* **238:** 1–4.

Albre J, Gers C, Legal L. 2008. Molecular Phylogeny of the Erebia Tyndarus (Lepidoptera, Rhopalocera, Nymphalidae, Satyrinae) complex of species combining CoxII and ND5 mitochondrial genes: a case study of a recent radiation. *Molecular Phylogenetics and Evolution* **47:** 196–210.

Avise JC, Walker D, Johns GC. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society B: Biological Sciences* **265:** 1707–1712.

Barraclough TG, Nee S. 2001. Phylogenetics and speciation. *Trends in Ecology and Evolution* **16:** 391–399.

Bohme J. 2003. The Miocene climatic optimum: evidence from ectothermic vertebrates of Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology* **195:** 389–401.

Braby MF, Pierce NE. 2007. Systematics, biogeography and diversification of the Indo-Australian genus *Delias* Hübner (Lepidoptera: Pieridae): phylogenetic evidence supports an 'out-of-Australia' origin. *Systematic Entomology* **32:** 2–25.

Coleman M, Hodges K. 1995. Evidence for Tibetan plateau uplift before 14 Myr ago from a new minimum age for east-west extension. *Nature* **374:** 49–52.

Cuvier G. 1812. *Recherche sur les ossements fossiles de quadrupèdes où l'on établit les caractères de plusieurs espèces d'animaux que les révolutions du globe paraissent avoir détruites*. Paris: Deterville.

Donoghue PCJ, Benton MJ. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology and Evolution* **22:** 424–431.

Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJP. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution* **20:** 248–254.

Drummond AJ, Ho SYW, Rawlence N, Rambaut A. 2007. A rough guide to BEAST 1.4. Available at: http://beast.bio.ed.ac.uk/

Drummond AJ, Rambaut A. 2004. BEAST v1.2.2. Latest version of BEAST available at: http://beast.bio.ed.ac.uk/

Fang X, Lü L, Yang S, Li J, An Z, Pa J, Chen X. 2002. Loess in Kunlun Mountains and its implications on desert development and Tibetan Plateau uplift in west China. *Science in China Series D: Earth Sciences* **45:** 289–299.

Francois L, Ghislain M, Otto D, Micheels A. 2006. Late Miocene vegetation reconstruction with the CARAIB model. *Palaeogeography, Palaeoclimatology, Palaeoecology* **238:** 302–320.

Glazko GV, Koonin EV, Rogozin IB. 2005. Molecular dating: ape bones agree with chicken entrails. *Trends in Genetics* **21:** 89–92.

Goloboff PA, Farris JS, Kallersjo M, Oxelman B, Ramirez MJ, Szumik CA. 2003. Improvements to resampling measures of group support. *Cladistics* **19:** 324–332.

Goloboff PA, Farris JS, Nixon KC. 2008. TNT: a free program for phylogenetic analysis. *Cladistics* **24:** 774–786.

Goloboff PA. 1999. Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics* **15:** 415–428.

Hall T. 1999. Bioedit: a user-friendly biological alignement editor and Sequence Alignement Editor and analysis program for Windows 95/98/NT. 7.0.5.3. *Nucleic Acids Symposium Series* **41:** 95–98.

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

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B, Biological Sciences* **359:** 183–195.

Higgins LG. 1941. An illustrated catalogue of the Palaearctic Melitaea (Lep. Rhopalocera). *Transactions of the Royal Entomological Society of London* **91:** 175–365.

Higgins LG. 1955. A descriptive catalogue of the genus *Mellicta* Billberg (Lepidoptera: Nymphalidae) and its species, with supplementary notes on the genera Melitaea and Euphydryas. *Transactions of the Royal Entomological Society of London* **106:** 1–127.

Higgins LG. 1981. A revision of *Phyciodes* Hübner and related genera, with a review of the classification of the Melitaeinae (Lepidoptera: Nymphalidae). *Bulletin of the British Museum (Natural History)* **43:** 77–243.

Hsu KJ, Ryan WBF, Cita MB. 1973. Late Miocene desiccation of the Mediterranean. *Nature* **242:** 240–244.

Janz N, Nylin S, Wahlberg N. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology* **6:** 4.

Kodandaramaiah U, Wahlberg N. 2007. Out-of-Africa origin and dispersal mediated diversification of the butterfly genus Junonia (Nymphalidae: Nymphalinae). *Journal of Evolutionary Biology* **20:** 2181–2191.

Kodandaramaiah U, Wahlberg N. 2009. Phylogeny and biogeography of *Coenonympha* butterflies (Nymphalidae: Satyrinae) – patterns of colonization in the Holarctic. *Systematic Entomology* **34:** 315–323.

Kovac M, Barath I, Fordinal K, Grigorovich AS, Halasova E, Hudackova N, Joniak P, Sabol M, Slamkova M, Sliva L, Vojtko R. 2006. Late Miocene to Early Pliocene sedimentary environments and climatic changes in the Alpine–Carpathian–Pannonian junction area: a case study from the Danube Basin northern margin (Slovakia). *Palaeogeography, Palaeoclimatology, Palaeoecology* **238:** 32–52.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and progression of the Messinian salinity crisis. *Nature* **400:** 652–655.

Lafranchis T. 2000. *Les papillons de jour de France, Belgique et Luxembourg et leurs chenilles*. Mèze, France: Biotope.

Laming DJC. 1965. Age of the New Red Sandstone in South Devonshire. *Nature* **207:** 624–625.

Lavé J, Avouac JP. 2001. Fluvial incision and tectonic uplift across the Himalayas of central Nepal. *Journal of Geophysical Research* **106:** 26561–26591.

Losos JB, Glor RE. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution* **18:** 220–227.

Lunter G, Miklos I, Drummond A, Jensen J, Hein J. 2005. Bayesian coestimation of phylogeny and sequence alignment. *BMC Bioinformatics* **6:** 83.

McKenna DD, Farrell BD. 2006. Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. *Proceedings of the National Academy of Sciences* **103:** 10947–10951.

Maki T, Hase Y, Kawamuro K, Minoura K, Oda T, Miyoshi N. 2003. Vegetation changes in the Baikal region during the Late Miocene based on pollen analysis of BDP-98-2 Core. *Long Continental Records from Lake Baikal* Springer 123–137.

Mayr E. 1970. *Population, species and evolution: an abridgment of animal species and evolution*. Cambridge, MA: Harvard University Press.

Nee S, Holmes EC, May RM, Harvey PH. 1994. Extinction rates can be estimated from molecular phylogenies. *Philosophical Transactions of the Royal Society B, Biological Sciences* **344:** 77–82.

Nylin S, Wahlberg N. 2008. Does plasticity drive speciation? Host-plant shifts and diversification in nymphaline butterflies (Lepidoptera: Nymphalidae) during the tertiary. *Biological Journal of the Linnean Society* **94:** 115–130.

Pagani M, Freeman KH, Arthur MA. 1999. Late miocene atmospheric CO_2 concentrations and the expansion of C4 grasses. *Science* **285:** 876–879.

Peña C, Wahlberg N. 2008. Prehistorical climate change increased diversification of a group of butterflies. *Biology Letters* **4:** 274–278.

Pérez Tris J, Bensch S, Carbonell R, Helbig A, Tellería JL. 2004. Historical diversification of migration patterns in a passerine bird. *Evolution* **58:** 1819–1832.

Quade J, Cater JML, Ojha TP, Adam J, Harrison TM. 1995. Late Miocene environmental change in Nepal and the northern Indian subcontinent; stable isotopic evidence from Paleosols. *Geological Society of America Bulletin* **107:** 1381–1397.

Ronquist F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* **46:** 195–203.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian pylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

Santucci F, Emerson BC, Hewitt GM. 1998. Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology* **7:** 1163–1172.

Sharma M, Wasserburg GJ, Hofmann AW, Chakrapani GJ. 1999. Himalayan uplift and osmium isotopes in oceans and rivers. *Geochimica et Cosmochimica Acta* **63:** 4005–4012.

Smith AB, Lafay B, Christen R. 1992. Comparative variation of morphological and molecular evolution through geologic time: 28S ribosomal RNA versus morphology in

echinoids. *Philosophical Transactions of the Royal Society B, Biological Sciences* **338:** 365–382.

Veith M, Kosuch J, Vences M. 2003. Climatic oscillations triggered post-Messinian speciation of Western Palearctic brown frogs (Amphibia, Ranidae). *Molecular Phylogenetics and Evolution* **26:** 310–327.

Wahlberg N. 2001. The phylogenetics and biochemistry of host-plant specialization in Melitaeine butterflies (Lepidoptera: Nymphalidae). *Evolution* **55:** 522–537.

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

Wahlberg N, Brower AVZ, Nylin S. 2005. Phylogenetic relationships and historical biogeography of tribes and genera in the subfamily Nymphalinae (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society* **86:** 227–251.

Wahlberg N, Freitas AVL. 2007. Colonization of and radiation in South America by butterflies in the subtribe Phyciodina (Lepidoptera: Nymphalidae). *Molecular Phylogenetics and Evolution* **44:** 1257–1272.

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

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

Wahlberg N, Zimmermann M. 2000. Pattern of phylogenetic relationships among members of the tribe Melitaeini (Lepidoptera: Nymphalidae) inferred from mitochondrial DNA sequences. *Cladistics* **16:** 347–363.

Weingartner E, Wahlberg N, Nylin S. 2006. North Africa is the source of ancestral populations of all Pararge species. *Systematic Entomology* **31:** 621–632.

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

Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292:** 686–693.