Molecular phylogeny of Parnassiinae butterflies (Lepidoptera: Papilionidae) based on the sequences of four mitochondrial DNA segments

François Michel [1], Cécile Rebourg [2], Emmanuel Cosson [3] & Henri Descimon [4]

[1] Centre de Génétique Moléculaire du C.N.R.S, 1 Avenue de la Terrasse, F-91190 Gif-sur-Yvette, France

[2] 62 rue Rochon, F-69270 Couzon-au-Mont d'Or, France

[3] Groupe Chiroptères de Provence, rue Villeneuve, F-04230 Saint-Etienne-les-Orgues, France

[4] 2 rue Rougemont, F-13012 Marseille, France

Abstract. A molecular phylogeny of Parnassiinae (Lepidoptera: Papilionidae) was generated by combining the partial sequences of three mitochondrial genes (LSU, ND1 and CO1; 1639 aligned sites) with a somewhat enlarged version of the ND5 mitochondrial dataset of Omoto *et al.* (2004). A total of 125 individuals were sampled (109 Parnassiinae, 14 Papilioninae, two outgroups) with the emphasis being put on genus *Parnassius* (94 specimens, most of them from natural history collections). Our phylogenetic reconstructions differ in particular from recently published ones in that (i) *Baronia brevicornis* Salvin 1893, an isolated taxon from Mexico, which had generally been placed in a subfamily of its own, is suggested to belong to Parnassiini, together with *Hypermnestra* and *Parnassius*; (ii) the earliest split within *Parnassius* is shown to be between subgenus *Parnasssius* (the 'apollo' group, whose caterpillars feed on Crassulaceae, exceptionally Saxifragaceae) and the ancestor of the remaining seven subgenera whose existence is confirmed by molecular phylogenies: six of them have Fumariaceae as larval foodplant, while *Kreizbergia* uses Scrophulariaceae. Within *Parnassius*, a number of systematic rearrangements at the species level are proposed, in particular within subgenera *Parnassius* and *Koramius* (28 and 23 taxa sampled, respectively), by reanalyzing available biological information in the light of our mitochondrial phylogenies. Finally, implications of this work for the biogeography of Parnassiini and shifts in larval host plant use are briefly discussed, the evolution of other adaptive traits in Parnassiinae being the subject of a separate paper.

Résumé. Phylogénie moléculaire des Parnassiinae (Lepidoptera : Papilionidae) basée sur les séquences de quatre segments d'ADN mitochondrial. Nous avons généré une phylogénie moléculaire des Parnassiinae et, plus spécifiquement, du genre *Parnassius* en combinant les séquences partielles de trois gènes mitochondriaux ((LSU, ND1 and CO1; 1639 sites alignés au total) avec une version quelque peu élargie du tableau de séquences mitochondriales ND5 publié par Omoto *et al.* (2004). Notre échantillon comprend un total de 125 individus (109 Parnassiinae, 14 Papilioninae, deux membres d'autres familles), parmi lesquels 94 spécimens, provenant pour la plupart de collections d'histoire naturelle, appartiennent au genre *Parnassius*. Nos reconstructions phylogénétiques diffèrent en particulier de celles récemment publiées par d'autres en ce que (i) nous suggérons que *Baronia brevicornis* Salvin 1893, une espèce mexicaine qui est généralement présentée comme le seul représentant vivant d'une sous-famille distincte, pourrait appartenir en fait aux Parnassiini, à côté de *Hypermnestra* et *Parnassius*; (ii) il est démontré que la division la plus ancienne du genre *Parnassius* a séparé le sous-genre *Parnassius* (le groupe de *P. apollo* L. 1758, dont les chenilles utilisent des Crassulaceae, exceptionnellement des Saxifragaceae) des autres lignées. De fait, des sept autres sous-genres dont l'existence est confirmée par la phylogénie moléculaire, six utilisent des Fumariaceae comme plantes-hôtes larvaires, tandis que *Kreizbergia* se nourrit de Scrophulariaceae. A l'intérieur du genre *Parnassius*, de nombreux réarrangements sont proposés aux niveaux spécifique et infraspécifique après réexamen des informations biologiques disponibles à la lumière de nos phylogénies mitochondriales : c'est en particulier le cas dans les sous-genres *Parnassius* et *Koramius*, dont nous avons analysé respectivement 28 et 23 individus. Enfin, les implications de ce travail pour la biogéographie des Parnassiini et les changements de plante-hôte sont brièvement discutées, l'évolution des autres caractères adaptatifs des Parnassiinae devant faire l'objet d'une publication séparée.

Keywords: Parnassiinae systematics, molecular evolution, foodplant choice, DNA barcoding, museum DNA.

E-mail: michel@cgm.cnrs-gif.fr, cecile.rebourg@free.fr, emmanuel.cosson@gcprovence.org, descimon.henri@free.fr

Accepté le 25 octobre 2007

Among the four currently recognized families of butterflies, Papilionidae have elicited the largest number of "modern" biological and molecular studies (see *e.g.* Scriber *et al.* 1995). Parts of this family have undergone a spectacular adaptive radiation, in particular the genus *Papilio* (subfamily Papilioninae): several molecular phylogenies of *Papilio*, which sampled a reasonably representative fraction of its over 200 species worldwide, have already been published (Aubert *et al.* 1999; Caterino *et al.* 1999; Zakharov *et al.* 2004). The subject of the present work is another subfamily of the Papilionidae, the Parnassiinae, several lineages of which provide a lesser, albeit still spectacular example of bursts of speciation.

The distribution area of Parnassiinae is restricted to the Holarctis, with a centre of gravity in Asia and a few outliers in North America and Europe. Within that range, members of the Parnassiinae occupy a variety of temperate (rarely subtropical), Mediterranean, subdesertic and mountain biomes, including some extreme habitats. For instance, one species, *Parnassius hunnyngtoni* Avinoff 1916, has been claimed to reach the elevation of 6000 m in the Himalayas (Weiss 1992). The morphological and physiological adaptations that enabled Parnassiinae to colonize such an impressive diversity of environments during their extended evolutionary history have rightly fascinated a fair number of entomologists. However, far larger troops of them were drawn to these butterflies primarily by the extraordinary aesthetic appeal of many of the species when in the adult stage. This has generated a huge nomenclatural inflation, especially at the subspecies scale, with many taxa only recognizable by their own author. In contrast, studies dealing with the evolutionary processes that gave rise to this adaptive radiation are relatively few, compared to those devoted to some less attractive groups. Moreover, the evolutionary relationships of the major subdivisions of the Parnassiinae have remained controversial, so that a firm basis to build evolutionary scenarios is lacking.

About two-thirds of currently recognized species in Parnassiinae belong to genus *Parnassius*, which is the type of the subfamily and was described in 1804 by the French entomologist Latreille, founder of the Société entomologique de France, to whose memory we dedicate this work. Counting some 45 species, *Parnassius* was divided into several genera by various authors, but is still regarded as a single genus (here designated as *Parnassius s.l.*) in such recent works as Nazari *et al.* (2007), Weiss (1991–2005) and Turlin and Manil (2005). Thus, according to a classical and conservative perspective (Tab. 1), the Parnassiinae

Table 1. A list of genera and subgenera commonly included in the Parnassiinae. Numbers of species are indicated between brackets.

Genus	subgenus	type species
Hypermnestra Ménétriès 1846	(1 species)	*helios* Nickerl 1846
Parnassius Latreille 1804	(*ca* 45)	
	Parnassius	*apollo* L. 1758
	Driopa Korshunov, 1988	*mnemosyne* L. 1759
	Kailasius Moore, 1902	*charltonius* Gray 1853
	Koramius Moore, 1902	*delphius* Eversmann 1843
	Lingamius Bryk, 1935	*hardwickii* Gray 1831
	Sachaia Korshunov, 1988	*tenedius* Eversmann 1851
	Kreuzbergia Korshunov 1990	*simo* Gray 1853
	Tadumia Moore, 1902	*acco* Gray 1853
Archon Hübner 1822	(3)	*thia* Hübner 1806 = *apollinus* Herbst 1798
Luehdorfia Kruger 1878	(4)	*eximia* Crüger 1878 = *puzilo*i Erschoff 1872
Sericinus Westwood 1851	(1)	*telamon* Donovan 1798 = *montela* Gray 1853
Buthanitis Atkinson 1873	(4)	*thaidina* Blanchard 1871
Zerynthia Ochsenheimer 1816	(7)	
	Zerynthia	*hypsipyle* Fabricius 1777 = *polyxena* Denis & Schiffermüller 1775
	Allancastria Bryk 1934	*cerisyi* Godart 1824

include 7 (or 8) genera – *Hypermnestra*, *Parnassius*, *Archon*, *Luehdorfia*, *Sericinus*, *Bhutanitis*, *Zerynthia* (to include *Allancastria*) – which are distributed into Parnassiini, comprising the first three genera, and one (Luehdorfiini) or two (Luehdorfiini and Zerynthiini; e.g. Nazari *et al*. 2007) other tribes.

Even though there has long prevailed a rather general consensus about this arrangement, controversies have arisen regarding the position of some taxa. To begin with, the monophyly of the subfamily has been questioned by various authors (*e.g*. Haüser 1993b; Caterino *et al*. 2001; Stekolnikov & Kuznetsov 2003). In a rather extreme example, some peculiarities of the adult and larval morphology of *Hypermnestra* resulted in this taxon being placed in a monobasic tribe or family by a number of specialists (Dujardin 1965; Hiura 1980; Häuser 1993b; Stekolnikov & Kuznetsov 2003). *Archon* has repeatedly been dragged from Parnassiini to Zerynthiini, owing to quite contradictory morphological and biological data, and it was even moved to a tribe of its own ('Archontini'; Koçak 1989). However, recent molecular studies have revealed a quite unexpected, yet significant affinity with *Luehdorfia* (Omoto *et al*. 2004; Katoh *et al*. 2005; Nazari *et al*. 2007). The latter genus was also proposed to deserve tribe rank in a study based on male genitalia by Stekolnikov & Kuznetsov (2003). Finally, some species of *Zerynthia* were segregated in the genus *Allancastria* by Bryk (1935); such a division, as well as that previously mentioned of *Parnassius s.l*., merely reflects different conceptions – splitter or lumper – of the genus.

Another taxon of controversial status, which might be allied to the Parnassiinae, is the genus *Baronia* Salvin 1893, with one species, *B. brevicornis* Salvin 1893 from Mexico. *Baronia* has been presented as the closest living approximation to the ancestor of all Papilionidae (*e.g*. Ehrlich 1958) or even of the Papilionoidea (butterflies; Scott 1985). However, some ancient works (Jordan 1907) had it incorporated in the Parnassiinae and it has in fact been placed basal to the Parnassiinae in a quite recent study (Nazari *et al*. 2007): undoubtedly, the exact phylogenetic affinities of *Baronia* have bearings on any biogeographical scenario for the origin and diversification of the Parnassiinae.

The systematic status of the genus *Parnassius* – a set of some 45 species – is a ticklish question as well. Some systematicians, including respected ones (*e.g*. Moore 1902; Bryk 1935; Munroe 1961; Korshunov 1988, 1990; see tab. 1), have proposed to split it in various, partly incompatible ways. Irrespective of the ongoing debate about a possible objective definition of the genus concept, it is of interest to determine which of the suggested subsets are monophyletic indeed. Two recent papers (Omoto *et al*. 2004; Katoh *et al*. 2005) have used molecular data to address this question, but with only partial success, since no clear consensus could be reached regarding the order of divergence of the major subdivisions. Finally, at a fine scale, many of the uncertainties that persist about species delimitation in *Parnassius* (see Weiss 1991–2005) remain to be addressed in a systematic way by molecular tools.

The present study was undertaken in an attempt to generate a stable molecular phylogeny of Parnassiinae, with emphasis on *Parnassius*. Because the recently introduced Bayesian analyses make it straightforward to combine molecular data with morphological or physiological characters, the current trend is to present phylogenies based on 'total evidence' (*e.g.* for butterflies, Wahlberg *et al.* 2005; Nazari *et al.* 2007). Unfortunately, any subsequent analysis of the evolution of individual characters becomes circular to some extent. That is why we chose to resist the tide and use the phylograms (or, simply, trees) we generated from DNA sequences alone as a reference for discussing the evolution of other types of characters. Accordingly, the work is organized in two separate papers: in the first, present one, a thorough analysis of molecular data is carried out, the reader being provided with minimal background information about the insects, while the second one (in preparation) is specifically devoted to the evolution of adaptive traits.

Preliminary results were presented at the Fourth International Conference on the Biology of Butterflies ('Butterfly Ecology and Evolution' symposium), March 23-27, 2002, at Leyden, in the form of an abstract and poster, submitted by C. Rebourg, F. Michel, E. Cosson, H. Descimon and E. Faure. Recently, an article derived from this poster was published by three of the authors of the abstract (Rebourg *et al*. 2006). Unfortunately, this paper, which makes use of only a fraction of the sequence data that had been generated by us back in 2002, is severely flawed by factual errors and misinterpretations (see below). In contrast, we chose to analyse the much enlarged Parnasssiinae dataset we and others (Omoto *et al*. 2004; Katoh *et al*. 2005; Nazari *et al*. 2007) have accumulated since that time, and have done our best to provide the reader with a multi-faceted perspective on the evolution of this fascinating group of butterflies.

Material and Methods

DNA extraction, amplification and sequencing

Individuals analyzed in this work are listed in the Annex Table, together with their origin, when known, and voucher code.

Table 2. List of primers used for amplification and sequencing.

Segment	Designation, orientation, and location (if not terminal)	Sequence (from 5' to 3')
LSU	984 (S)	CGCCTGTTTATCAAAAACAT
	20457 (S)	TTCACTTGTTTATCAAAAACAT
	3259 (R)	CCGGTTTGAGCTCAGATCA
	22157 (R)	AAACCAACCTGGCTCACA
	20870 (S,M)	GAGAAGACCCTATAGAGTTT
	20907 (R,M)	AAACTCTATAGGGTCTTCTC
ND1	1957 (S)	CGTAAAGTCCTAGGTTATATTCAGATTCG
	3264 (R)	ATCAAAAGGAGCTCGATTAGTTTC
	21578 (S,M)	ATTTTATTTTTTTATGTTGTA
	21536 (R,M)	GTTTGTGCAACAGCTCGTAA
CO1	LCO1490[(1)] (S)	GGTCAACAAATCATAAAGATATTGG
	17379 (S)	ATTCAACAAATCATAAAGATAT
	21686 (S)	ATTCAACAAATCATAAAGATATTGG
	16736 (S)	AGCGAAGTCGACTTTAtTCWACWAATCATAAaGATATtGG
	HCO2198([1])(R)	TAAACTTCAGGGTGACCAAAAAATCA
	17378 (R)	AAACTTCTGGATGACCAAAAAATCA
	16737 (R)	TAGAATGCATGCTTCWGGRTGNCCaAAaAATCA
	22014 (S,M)	GAAAATGGRGCAGGAACT
	21930 (R,M)	AGTTCCTGCYCCATTTTC
ND5	16740 (S)	GCTGCTCGTACGCCTGTWTCWGCTTTaGTTCA
	16741 (R)	CCATAAGTCGACAAaTTHGGYATAAATCAtAT
	22158 (S,M)	ATAATAAATGATAATCAAGATATTCG
	22159 (R,M)	CTTATTCTTACTATYTCTAAAATTAAATC

Abbreviations: S, sense; R, reverse; M, middle; R, A(0.5):G(0.5); Y, T(0.5):C(0.5); W, A(0.5):T(0.5); H, A(0.33):C(0.33):T(0.33); a, A(0.9):G(0.1); t, T(0.9):C(0.1); N, A(0.25):G(0.25):T(0.25):C(0.25)
(1) Folmer et al. 1994; Hebert et al. 2003.

They are stored in Gif or Marseille, either as dried, mounted insects, or frozen material. Specific and (tentative) subspecific identification was performed by F. M. and H. D., and in case of doubt, double-checked by J.-C. Weiss.

DNA was extracted exactly as described in Aubert *et al.* (1999), usually from a single leg, occasionally from the thorax or anterior part of the abdomen. Samples were resuspended in 100 µl of 10 mM Tris-HCl pH 7.5, 1 mM Na_2-EDTA, and 0.2 µl was routinely used for amplification.

The primers used for PCR amplification of the four mitochondrial DNA segments analyzed in this work are listed in tab. 2. Amplification was performed in 50 µl of 20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 1 µM of each primer, and 0.2 mM of each dNTP. Reactions were started by addition of Taq polymerase at 90 °C. A typical PCR cycle included 10 s denaturation at 92 °C and 3 min polymerization at 62 °C preceded by 45 s at the temperature chosen for renaturation (between 38 °C and 54 °C, depending on the pair of primers). Between 20 and 40 cycles, depending on the amount and quality of the DNA, were found to be necessary to obtain a visible band at the expected distance from the origin of migration after ethidium bromide staining of agarose gels. That band was cut from the gel and 1 µl of melted agarose was used for reamplification (20 cycles). Reamplified DNA was purified with the GenElute kit from Sigma and sequenced on both strands with the primers used for amplification. Any ambiguity was resolved by examining chromatograms with BioEdit version 7.0.4.1 (Hall 1999).

A majority (69 %) of specimens were dried insects, a number of which came from old collections (up to about a century old). As has been experienced by others (Hajibabaei *et al.* 2005, and references therein), amplification from material that had been kept at room temperature for more than 15 years proved unreliable and in a number of cases, we had to resort to amplification of smaller fragments by using internal primers located about halfway between the termini of the DNA segment of interest (in the case of the CO1 and LSU segments, our two internal primers overlapped, resulting for each segment in the loss of a 20-nt stretch of sequence, which was coded as missing; recall, however, that in the case of the LSU gene, the missing stretch is invariant in all butterfly taxa that we examined so far – Aubert *et al.* 1999; Martin *et al.* 2000; F.M., unpublished data – so that the loss of information should be minimal).

We attempted to analyze the observed distribution of successes and failures as a function of the age of specimens by using a simple model according to which the probability of successful amplification and sequencing is $P = 1 - (1 - (1 - k)^{\lambda . t})^N$, where

k is the rate of decay per nucleotide and time unit; t, the time elapsed at room temperature; N, the average number of intact molecules at t_0; and λ is linearly related to the length of the segment to be amplified. Consistent with this model, the critical age at which no more than half of the samples can successfully be exploited appears inversely related to the length of the segment of interest (fig. 1); k and N were estimated from the entire dataset, whereas values of λ and their standard error were obtained by fitting the available data for individual primer combination; $t_{1/2}$ is the time at which P is calculated to be 0.5). In actual fact, the rate of successful amplification and sequencing using our internal primers was no higher than 25 % for older specimens (aged 60–112 years), which makes the use of such samples too costly for consideration, except for taxa that are extinct or extremely difficult to obtain in reasonably fresh condition.

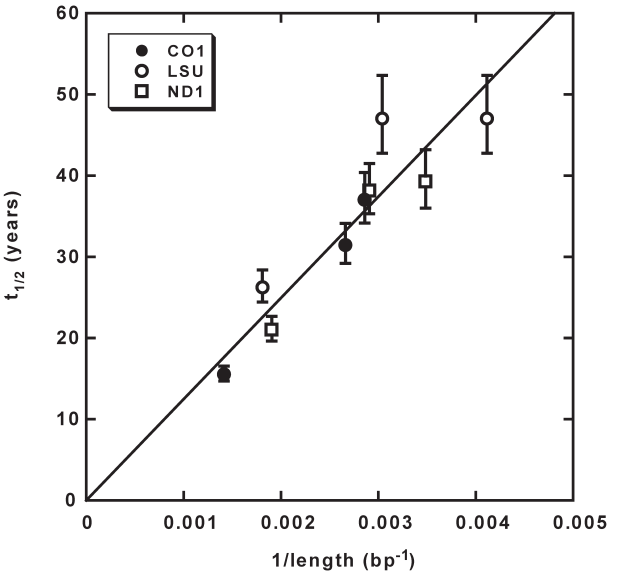


Figure 1
The critical age for a dried specimen is an inverse function of the number of nucleotides to be amplified and read. Abscissa: 1/L, where L is the length of a PCR-amplified segment – primers included – in base pairs; ordinates: $t_{1/2}$ is the duration of room-temperature storage for which half of the DNA samples could be successfully amplified and sequenced ($t_{1/2}$ and its standard error were estimated for each primer combination by fitting the distribution of successes and failure as a function of age to the equation provided in Material and Methods). Using a linear equation, A/L, to fit the data (straight line) yielded A = 12460 ± 520 bp.y (Pearson's R = 0.918; the fit was forced through the origin, since when a constant term was added, its value was smaller than its standard error).

Sequence alignment and phylogenetic analyses

While alignment of the ND1 (472 nt), CO1 (649 nt) and ND5 (807 nt) sequences was straightforward, a number of indels were detected in the LSU data. Whenever possible, the precise location of these indels was determined by taking into account both the sequence and potential secondary structure of the large ribosomal RNA (fig. 2), which we modelled after Niehuis *et al.* (2006). However, there remained a small number of subsegments that could not be unambiguously aligned and had to be left over. For analyses confined to Parnassiini, only positions 6–9 (coordinates as in fig. 2) were removed from the dataset, while in the case of Papilionidae, we chose to discard also subsegments 40–42 and 317–324, leaving 513 and 494 aligned sites, respectively (the LSU alignments used for phylogenetic analyses are available from the authors). For analysis of amino-acid substitutions in the CO1 segment (see Results) all lepidopteran sequences made available in databases by March 2007 were downloaded and aligned (a maximum of ten entries were retained for any given taxon).

Analyses based on maximum parsimony and distance were conducted with PAUP* 4.0b10 (Swofford 2002), with gaps in the LSU alignment coded as 'fifth base' and 'missing data', respectively. To calculate distances between subtrees (e.g. in order to estimate percent divergences over the CO1 segment, see Results), pairs of terminal nodes were replaced recursively by a single one, whose distances to other taxa were obtained from the arithmetic mean of the original ones. In order to assess the compatibility of any two datasets, we used the approach advocated by Delorme & Henaut (1988). Uncorrected distance matrices, generated by counting nucleotide differences, were converted into coordinates in multidimensional space by Principal Coordinate Analysis and the resulting clouds of points (each point standing for a taxon) were superimposed by minimizing the sum of squared distances between those taxa that are shared by the two datasets (Kabsch 1978). Histograms of absolute distances between shared taxa in multidimensional space were then generated and examined to detect possible outliers (fig. 4).

Because of the large size of our datasets, maximum likelihood analyses were carried out using the fast approach implemented in Phyml 2.4.4 (Guindon & Gascuel 2003). A GTR (general time-reversible) substitution model was used, together with a fraction of invariant sites and four rate categories. The gamma shape parameter for the distribution of rate categories and the fraction of invariant sites were optimized by maximizing the likelihood of the phylogeny prior to bootstrapping. Phylip 3.63 Consense program was used to generate a consensus tree topology and bootstrap percentages from the (at least) 200 pseudo trees generated by Phyml during bootstrapping. This consensus topology was then fed back to Phyml for parameter and branch length optimization.

Bayesian analyses were performed with program Mr Bayes 3.1.2 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). A GTR substitution model with four site-specific rates and a proportion of invariant sites was specified before running the program for 1,000,000 generations with default settings. The first 4000 trees (out of 20000) were discarded prior to computing a consensus phylogeny and posterior probabilities. For some analyses involving the LSU segment (e.g. fig. 5B) we took advantage of the diversity of models accepted by the program to code separately both the nucleotides that are part of secondary structure helices (using the 'Doublet' model option) and 18 alignment gaps (37 aligned sites, otherwise coded as 'missing data'). In the latter case, a second partition and 'restriction site' (binary) coding were used (the ascertainment bias was taken into account by setting coding=variable and the two partitions were allowed to evolve at different rates).

In order to convert the phylograms generated by maximum likelihood and Bayesian analyses into chronograms, we used penalized likelihood, as implemented in program r8s, version 1.7 (Sanderson 2002), together with the cross-validation option. To

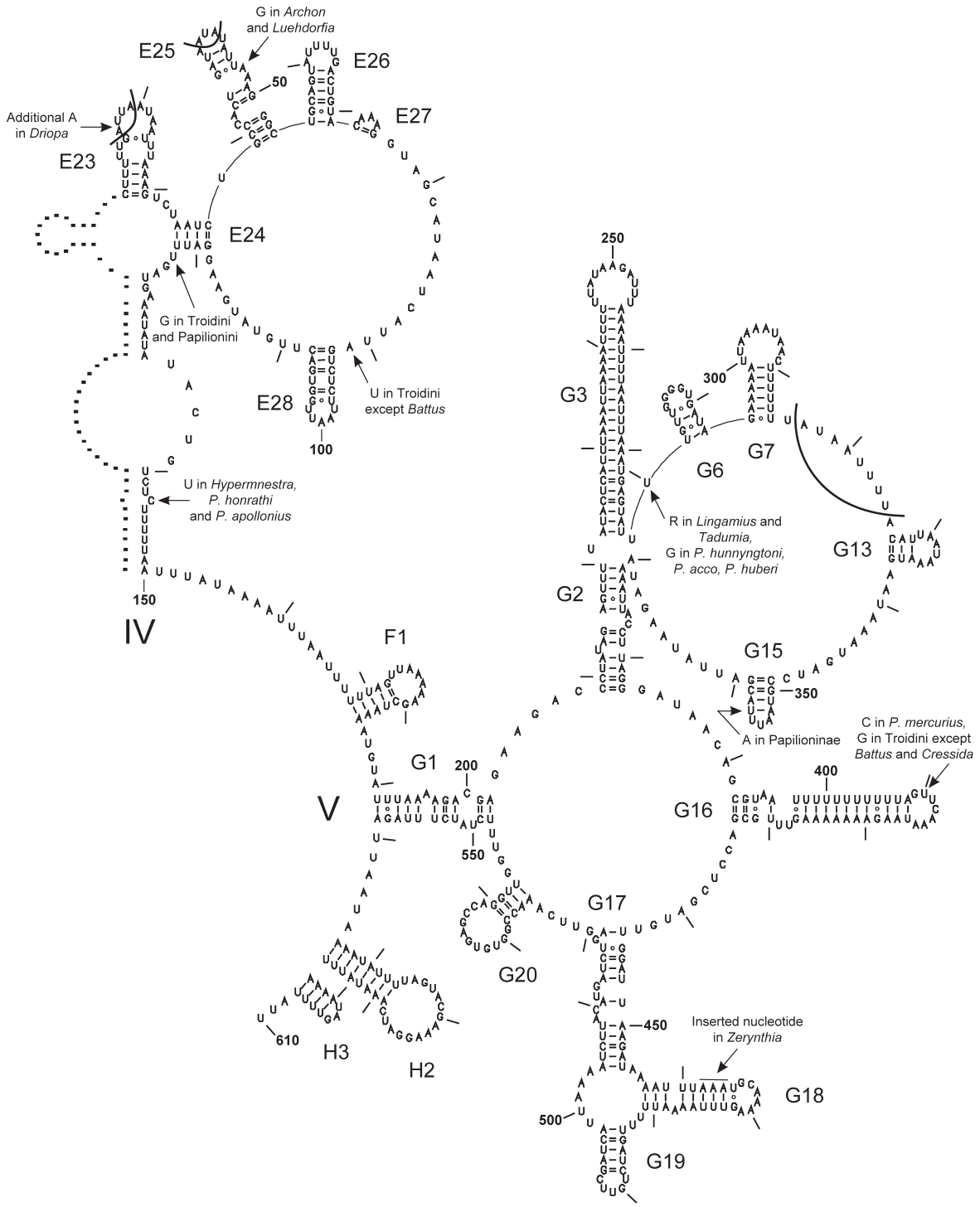


Figure 2
Predicted secondary structure of domains IV and V of the mitochondrial large ribosomal RNA of *P. apollo graslini* (AJ224055), adapted from Niehuis *et al.* (2006). The segment analyzed in this work extends from positions 0 to 484. Thick curves delimit sections that could not be unambiguously aligned throughout Papilionidae and were discarded from the analysis (for analyses limited to Parnassiini, only positions 6-9 were removed). Arrows point to rare substitution events, confined to one (exceptionally two) lineages (see Text, and for stem-loop G3, fig. 13).

generate a chronogram of the entire *Parnassius* genus (fig. 12), we proceeded as follows. First, the tree in fig. 8, which is based on all four mitochondrial DNA segments, was converted into a chronogram by program r8s and calibrated by assuming a fixed age of 38 MY (as inferred in fig. 11) for the root of *Parnassius*. The procedure was then repeated on the LSU-ND1-CO1-based Phyml (maximum likelihood) tree of the genus, some sections of which are shown in fig. 10. Finally, those subtrees with an expanded content in fig. 10 (corresponding to subgenera *Parnassius* and *Koramius* and the clade of *P. mnemosyne* L. 1759) were grafted into the chronogram derived from the tree in fig. 8 after having been brought to scale.

Results and discussion

Combining datasets

We have used four distinct sequence datasets to infer a molecular phylogeny of Parnassiinae, with emphasis on *Parnassius*. The four segments sequenced, abbreviated as LSU (518 aligned sites), CO1 (649 nt), ND1 (472 nt) and ND5 (807 nt), correspond to sections of the mitochondrial genes for the large subunit of ribosomal RNA (see fig. 2), subunit 1 of cytochrome oxidase and subunits 1 and 5 of NADH-dehydrogenase, respectively (see Material and Methods; note that the CO1 segment coincides with the section of the gene used for 'bar-coding', see Hebert *et al.* (2003)). Sequence accession numbers are provided in the Annex Table. All LSU and ND1 sequences and most of the CO1 ones were generated from DNA that was extracted either in Gif or Marseille (this work; Aubert *et al.* 1999, and Rebourg *et al.* 2006).

Insect mitochondrial DNA is maternally inherited and is not expected to undergo recombination, so that datasets consisting exclusively of mitochondrial sequences should share the same phylogenetic history, including at infraspecific levels. That the sets of sequences used in this work are mutually compatible indeed is confirmed by plotting the mean of bootstrap percentages as a function of the number of sites aligned for single and combined analyses (fig. 3). As expected, bootstrap support is found to increase steadily as a function of the combined length of sequence, whether for our 32-taxa 'Papilionidae' (fig. 5) or 65-taxa 'Parnassiinae' (fig. 8) datasets.

Overall agreement between datasets does not eliminate the possibility that a limited fraction of the data for individual taxa are actually incompatible. Three possible sources of occasional inconsistencies may be contemplated : PCR contamination, sequencing of NUMTs (nuclear-located mitochondrial pseudogenes; reviewed by Bensasson *et al.* 2001) and combination of data obtained from individuals that actually belong to different phylogenetic units. In order to try and avoid these pitfalls we systematically resorted to the approach advocated by Delorme & Henaut (1989), which consists in superimposing distance matrices generated from the datasets to be combined (see Material and Methods) and looking for inconsistently located taxa. Of special concern was the merging of our *Parnassius* LSU+ND1+CO1 dataset with the ND5 dataset since

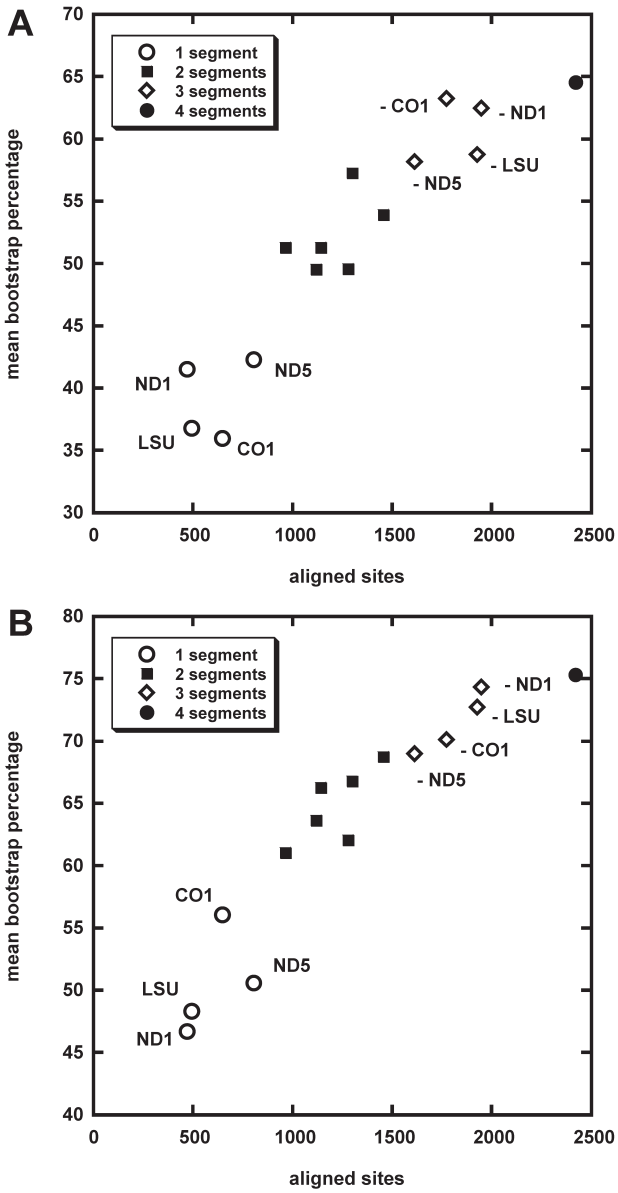


Figure 3
Mean of bootstrap percentages as a function of the number of aligned sites. Ordinates: bootstrap percentages were generated by PAUP* 4.0b10 from 200 pseudo datasets using the Neighbor-Joining algorithm and a LogDet distance measure. **A,** Papilionidae dataset (see fig. 5); **B,** Parnassiinae dataset (see fig. 8).

the vast majority of sequences in the latter had been generated in another laboratory from specimens distinct from ours (Omoto *et al.* 2004; see also Legend to the Annex Table). Nevertheless, the agreement between those two datasets is excellent, as indicated by the narrow distribution of distances between presumably homologous phylogenetic units in multidimensional space (fig. 4). *P. hardwickii* Gray 1831 is the only taxon to lie well outside the bulk of the distribution, and we believe that the likely reason for its position is not misidentification (which is quite unlikely for this very distinctive species), but, rather, that the two individuals

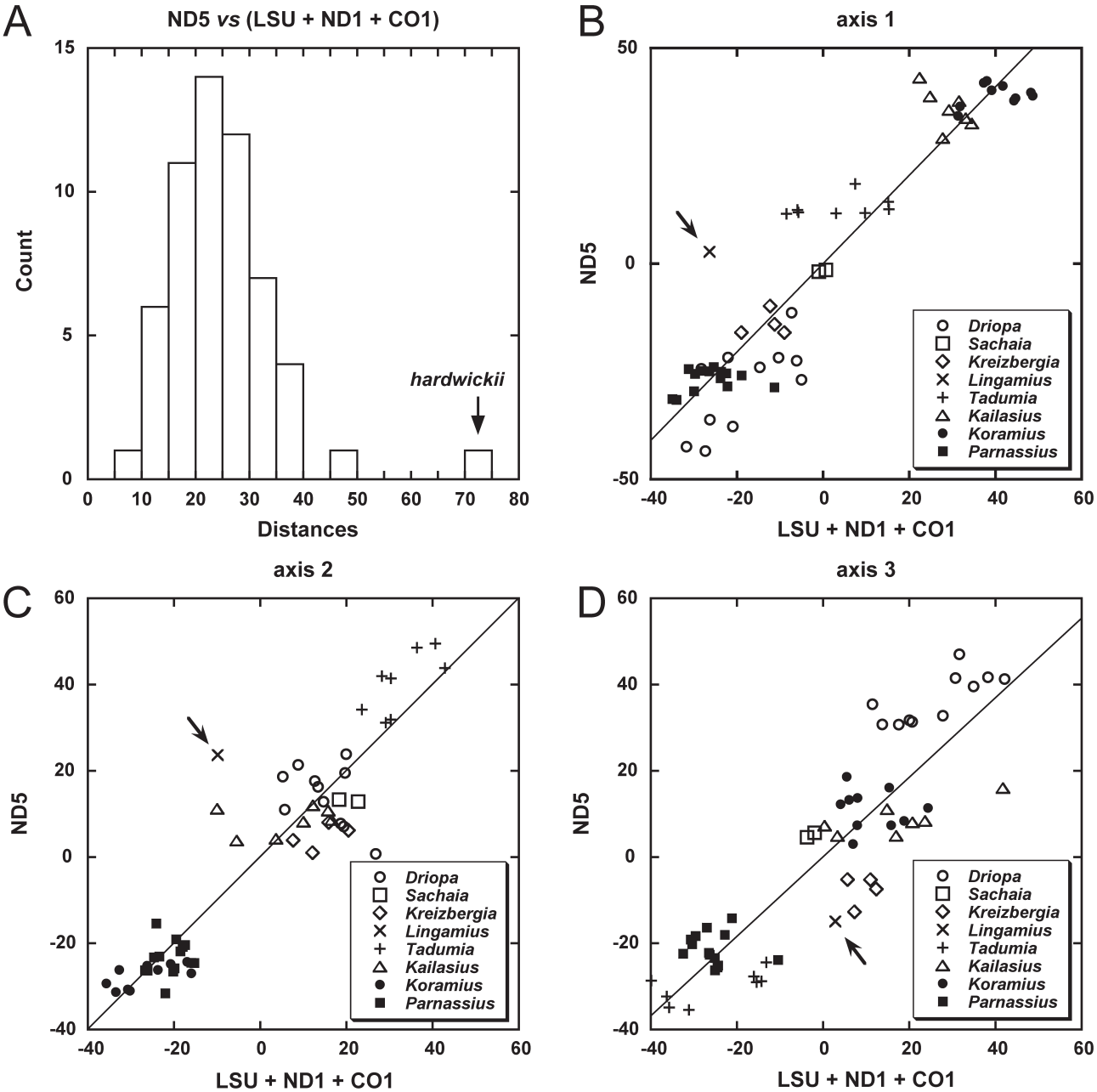


Figure 4
Superimposition of nucleotide distance matrices over segment ND5 and the combination of segments LSU, ND1 and CO1 for 57 *Parnassius* taxa (the ones in fig. 8). **A,** histogram of distances in the space defined by the first six axes. **B, C** and **D,** relative positions of individual taxa in the two matrices along the first three axes. Arrows indicate the location of *P. hardwickii*.

whose sequences were determined happened to come from the very two extremities of the species range (see Annex and further discussion below).

On the other hand, we did uncover one clear case of misidentification, by examining the data published by Rebourg *et al.* (2006). The LSU and CO1 sequences claimed by these authors to be those of *Parnassius acdestis* Grum-Grshimailo 1891 (individual UP100-22, misspelled as *'P. acdetis'* in their tab. 1; accession numbers DQ407794 and DQ407760) clearly pertain to another subgenus (*Tadumia*). We have been able to examine specimen UP100-22 and can confirm that it was misidentified and belongs instead to *P. maharaja maharaja* Avinoff 1916. By combining the two sequences determined in Marseille with the ND1 and ND5 sequences of the true *P. acdestis*, Rebourg *et al.* unwittingly created a chimera between two subgenera, thus compromising the validity of their entire phylogenetic dataset and conclusions.

Molecular phylogeny of Parnassiinae and relationships with other Papilionidae

In an attempt to derive a molecular phylogeny of Parnassiinae and test for the monophyly of the subfamily, we sampled all 15 genera and subgenera that are traditionally grouped within Parnassiinae, together with 15 species from the rest of Papilionidae and two outgroups, selected from two of the remaining three families of butterflies. The dataset we assembled (2422 aligned sites) is definitely smaller than the one recently published by Nazari *et al.* (2007), which covers 5775 nt (including all four segments we analyzed). However, the latter lacks the ND1 sequence of *Baronia* (which is generally placed in a subfamily of its own) and

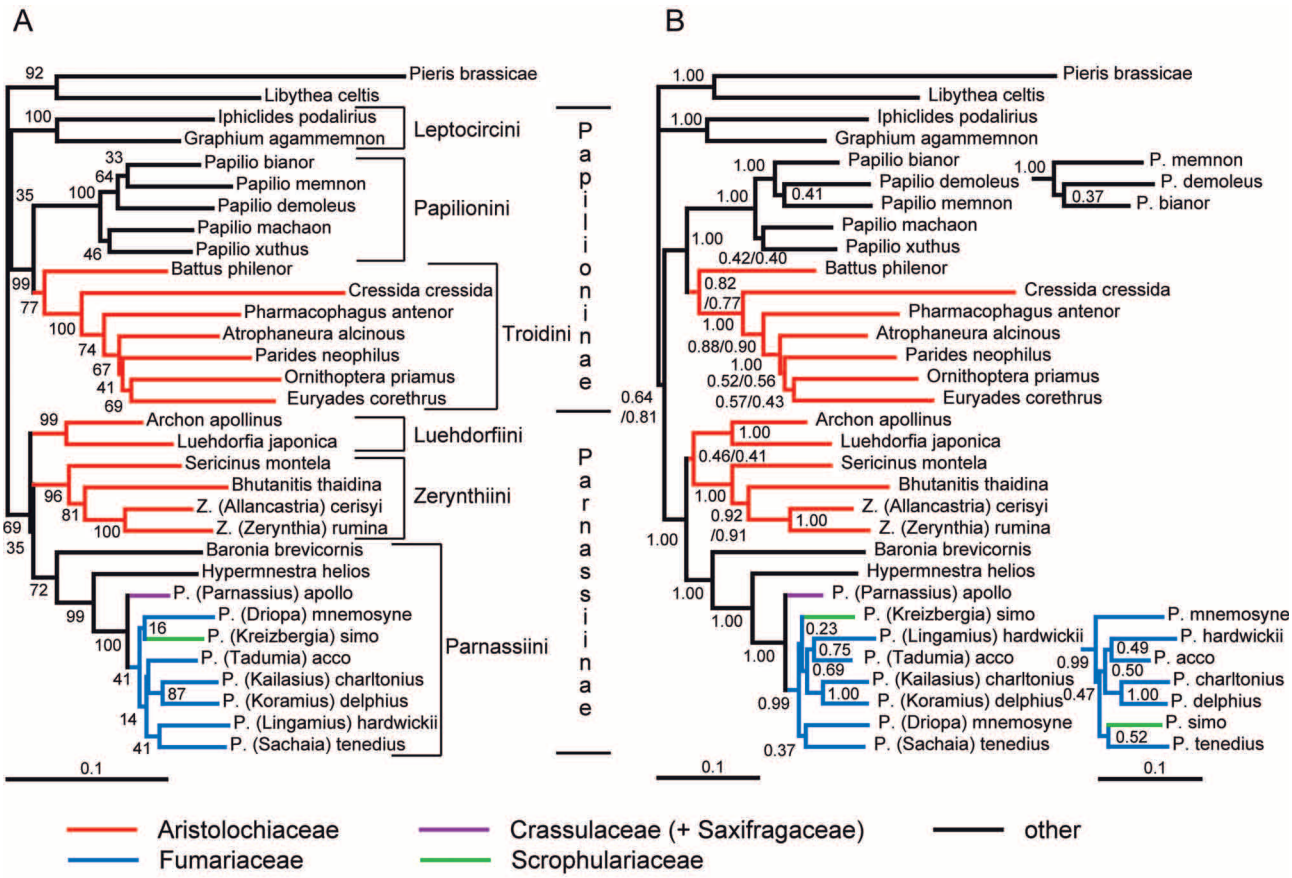


Figure 5
Molecular phylogeny of a sample of Papilionidae based on 2422 aligned sites (combined LSU, ND1, CO1 and ND5 segments). (A) consensus maximum likelihood tree (program Phyml 2.4.4, GTR substitution model, four site-specific rates plus invariant sites); numbers are bootstrap proportions. (B) consensus tree from Bayesian analysis (program mrbayes-3.1.2, GTR substitution model, four site-specific rates plus invariant sites); numbers are posterior probabilities. Subtrees at far right and numbers on the right side of slashes correspond to a variant model in which 18 indels were coded and loops were distinguished from base-paired stems in the LSU segment (see Material and Methods). Branches are coloured according to larval foodplant (indicated at bottom; see tab. 3 for a complete list).

comprises fewer Papilioninae (9 instead of 14).

Distance methods fail to take all information into account, while maximum parsimony is hardly appropriate to establish deep relationships with fast-evolving sequences (that is, unless sites are weighted according to their rates of evolution – Goloboff 1993). This is why we chose to resort to maximum likelihood and Bayesian analyses in order to derive the phylogenetic information presented in fig. 5. Both approaches allow site-specific rate variation to be taken into account, something which largely obviates the need for data partitioning provided the individual subsets to be combined have comparable base composition and base-specific patterns of substitution. Still, an attempt was made to take advantage of the variety of models that can be integrated into Bayesian analysis by coding the indels in the LSU segment (otherwise interpreted as 'missing data') and treating separately the nucleotides that are included in secondary structure pairings, whose evolution is subject to (partly) explicit evolutionary constraints (see Material and Methods; it should, however, be stressed that the continued use of such naive evolutionary models as the 'Doublet' option of MrBayes is quite questionable now that the publication of atomic resolution structures of entire ribosomes with their ligands – *e.g.* Selmer *et al.* 2006 – has made it possible to gauge the complexity and diversity of constraints to which the evolution of individual nucleotides is subjected). These two refinements were expected to change in opposite ways the relative weight of the LSU segment in the complete dataset and in fact, they can be seen to have pretty little effect on the resulting tree topology and posterior probabilities: the only nodes to be affected are the most poorly supported ones (see fig. 5B).

Very similar phylograms were generated by maximum likelihood and Bayesian analysis of the Papilionidae dataset (fig. 5): only the statistical support associated with individual nodes differs significantly, the bootstrap test associated with the former method being definitely more conservative than Bayesian posterior probabilities. Monophyly of the three traditionally recognized tribes of Papilioninae is well supported in both trees, as is the sister-clade status of Papilionini and Troidini, and the basal position within the latter of *Battus*. These features were already apparent in the morphology-based cladistic analyses of Hancock (1983) and Miller (1987) and in at least some molecular studies (e.g. Caterino *et al.* 2001; Braby *et al.* 2005; Nazari *et al.* 2007). Within Parnassiinae, the sister-clade relationships of *Archon* and *Luehdorfia* on the one hand, and *Hypermnestra* and *Parnassius* on the other (both of them already recognized by Katoh *et al.* 2005, and Nazari *et al.* 2007) are all the more compelling that they are supported by separate analyses of each of the four sequenced segments (not shown). The former was totally unanticipated on morphological grounds: 'common sense', based on superficial analysis of wing patterns, had *Luehdorfia* grouped with *Bhutanitis*, despite the latter sharing a divided hind-wing cell pattern with *Zerynthia* and *Sericinus*, and larval morphology (Igarashi 1984) with *Zerynthia*. As for grouping of *Parnassius* with *Hypermnestra* rather than *Archon*, it has repeatedly been debated (Hiura 1980; Haüser 1993), despite the traditional designation of *Hypermnestra helios* Nickerl 1846 as 'Desert apollo'. A third, well-supported Parnassiinae clade consists of *Sericinus*, *Bhutanitis* and *Zerynthia*. Its phylogenetic relationship with the other two Parnassiinae subdivisions remains uncertain, which justifies the recognition of three tribes (fig. 5), as proposed by Nazari *et al.* (2007).

As expected from the redundancy of datasets, the trees in fig. 5 are very similar indeed to those recently presented by Nazari *et al.* (2007). Still, our analysis differs significantly from theirs in the placement of *Baronia*, presumably because of the inclusion of the ND1 sequence of this insect in our dataset. *Baronia brevicornis*, an isolated taxon confined to Mexico and a relatively late discovery (1893), has been regarded as a 'living fossil' and is traditionally placed in a separate subfamily of its own, at the root of Papilionidae. It was therefore somewhat unexpected to find it located well within the Papilionidae consensus tree, next to the root of the Parnassiinae, in the maximum likelihood analysis of Nazari *et al.* (2007). Its position in our own phylogenetic trees is even more extreme, since it is part of the clade labelled Parnassiini, together with *Parnassius* and *Hypermnestra*.

Such an unconventional definition of Parnassiini is reasonably well supported by statistical analyses,

Figure 6
Diversity of wing patterns in subgenus *Koramius*. Scale: x 0.96. Voucher codes (Annex Table) are provided for individuals whose mitochondrial DNA was analyzed or that came from the same population as individuals analyzed (in the latter case, the code is between square brackets). From left to right, 1st row: *P. delphius* male [UP100-12], *P. patricius uzyngyrus* D. Weiss 1979 male (W231), *P. hide meveli* Weiss & Michel 1989 female (W225); 2nd row: *P. staudingeri infernalis* male (Taldyk Pass, Kirghizstan), *P. staudingeri darvasicus* male (W310), *P. staudingeri kiritshenkoi* female (W243); 3rd row: *P.* (*staudingeri*) *jacobsoni* male (W278), *P. cardinal* male [W246], *P.* (*staudingeri*) *ruth* male (W334); 4th row: *P.* (*staudingeri*) *kohibaba* male (W323), *P.* (*staudingeri*) *affinis* female [UP100-10], *P. stoliczkanus nobuko* male (W256); 5th row: *P. stoliczkanus zojilaica* male [W267], *P.* (*stoliczkanus*) *stenosemus* male [W224], *P. stoliczkanus stoliczkanus* male [W261].

with a Bayesian posterior probability of 1.0 and a bootstrap percentage of 72. In order to assess the precise significance of this finding, the 200 trees generated from pseudo-datasets and used to compute bootstrap percentages were sorted according to the position of *Baronia* and the root of Papilionidae (fig. 7). Two topologies dominate among unrooted Papilionidae trees: in the prevailing one (Type I, 131 trees), *Baronia* is next to *Parnassius* and *Hypermnestra*, like in fig. 5, whereas in the other one (Type II, 29 trees), it is located as proposed by Nazari *et al.* (2007), at the root of Parnassiinae (27 of the remaining 40 trees were discarded because at least one traditionally recognized phylogenetic unit was not monophyletic). Once the position of the root is taken into account, *Baronia* is found to be the sister group of *Parnassius* + *Hypermnestra* in 122 trees, against 29 in which it lies at the root of Parnassiinae. It should be added that a maximum likelihood consensus tree derived from two nuclear genes that were not sampled in our study (Nazari *et al.* 2007) also places *Baronia* at the root of Parnassiinae.

Molecular phylogeny of *Parnassius*

The molecular phylogeny of *Parnassius* shown in fig. 8 is the most complete one to date: we have sampled essentially all taxa generally assumed or suspected to constitute species (with the exception of the very recently described *P. davydovi* Churkin 2006), as well as some major subspecies. It is also the one that rests on the largest number of aligned sites (2422).

Eight outgroups, whose phylogenetic relationships can be verified to be the same as in unrooted versions of the Parnassiinae subtrees of fig. 5, were used to root our *Parnassius* trees. Inclusion of a larger number of Parnassiinae outgroup taxa, whose sequences were available for all but the ND5 segment, did not change significantly either the relationships between outgroups or those within *Parnassius* (fig. 9). And just as was the case for the phylogeny of Papilionidae, consensus trees obtained by Bayesian and maximum likelihood analyses were found to differ only by a few exchanges at some poorly supported nodes (Legend to fig. 8).

The genus *Parnassius* is currently subdivided into some eight subgenera (tab. 1) that correspond closely to Munroe's ten 'groups' (1961), which were based primarily on venational and genitalic characters (the close relationship of the *acco* Gray 1853 and *szechenyii* Frivaldszky 1886 groups, which constitute together subgenus *Tadumia*, was already recognized by Munroe). As seen in fig. 8, all these subgenera receive strong statistical support from molecular data (Bayesian posterior probability 1.0; bootstrap percentages ranging from 97 to 100); only the position of *acdestis*, within *Kailasius* rather than the related *Koramius*, is incongruent with the traditional classification based on morphology. The same conclusion was already reached by Omoto *et al.* (2004) and Katoh *et al.* (2005), based on smaller datasets. However, our phylogeny of *Parnassius* differs significantly from theirs, and also from the one recently published by Nazari *et al.* (2007), which was based on an incomplete set of six segments and eight taxa, in that there is a major split at the base of the genus between subgenus *Parnassius* (the 'apollo' group) and the other species.

The early branching position of *P. apollo* L. 1758 and its relatives has rather good statistical backing (Bayes posterior probability 1.0, bootstrap percentage 73) and is further supported by a rare amino acid substitution in the CO1 segment we sequenced, at codon 159 (this is codon 174 in the reference *Drosophila yakuba* Burla 1954 sequence of Clary & Wolstenholme 1985; accession number NC_001322). At that position, which is part of the variable loop between transmembrane helices IV and V (Tsukihara *et al.* 1996), asparagine (AAY) predominates in Lepidoptera,

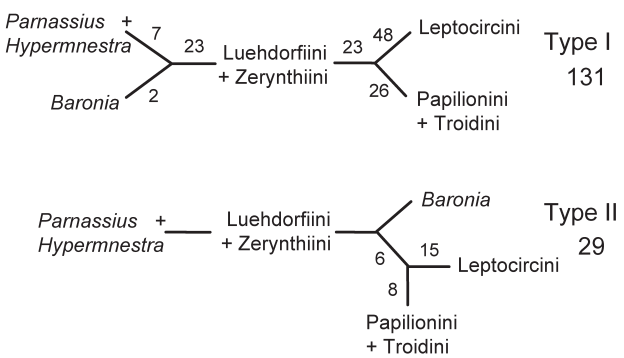
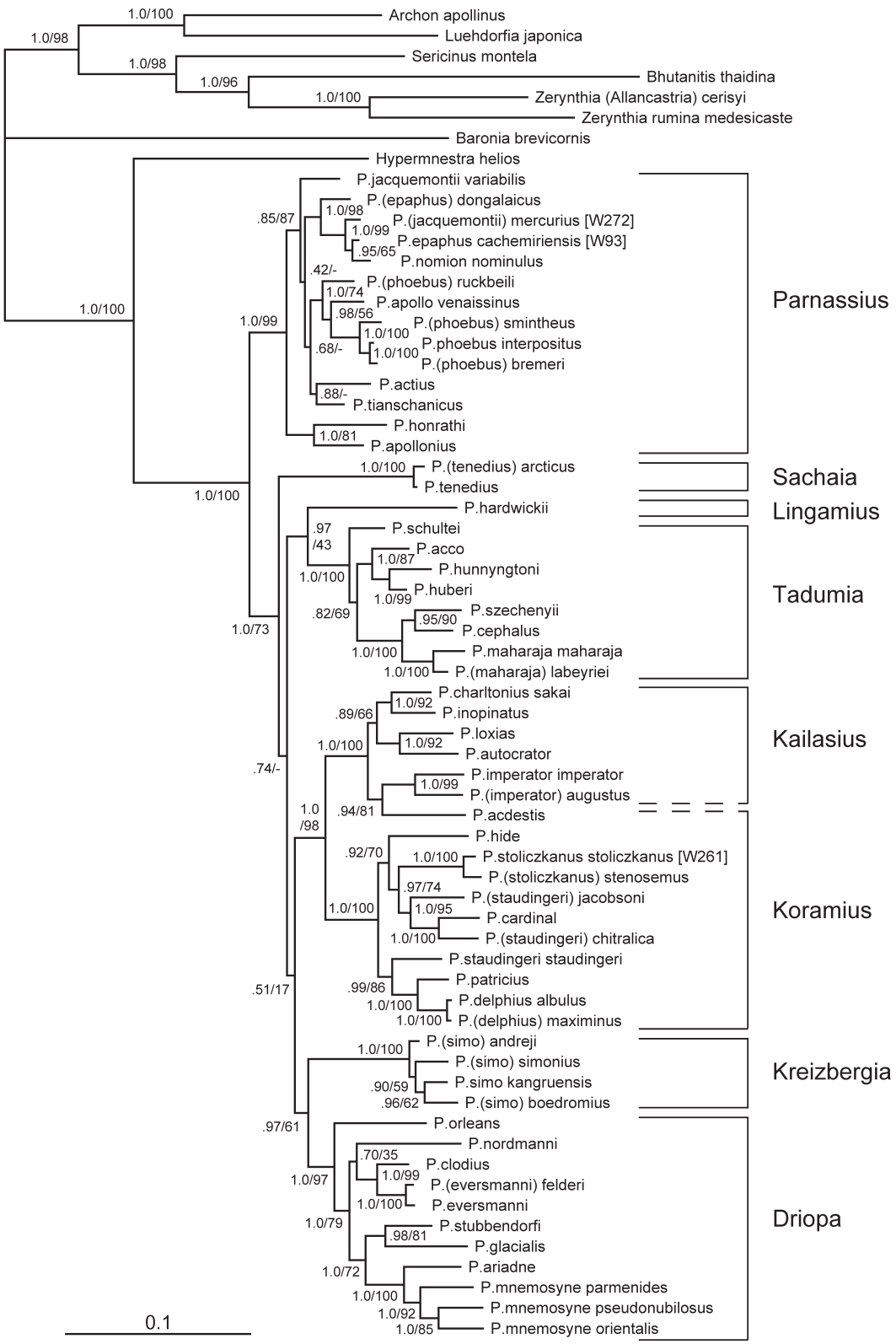


Figure 7
Two possible topologies for the unrooted phylogenetic tree of Papilionidae. The 200 bootstrap trees used to generate the consensus maximum likelihood tree in fig. 5A were sorted according to the position of *Baronia* with respect to the six major lineages of Papilionidae (27 trees were rejected because one major subdivision was poly- or paraphyletic and 13 trees had topologies incompatible with types I and II). Figures next to branches are numbers of trees in which the root lay within that branch.

Figure 8
Phylogeny of Parnassiinae with emphasis on *Parnassius* according to Bayesian and maximum likelihood analysis of 2422 aligned sites. Numbers next to nodes are Bayesian posterior probabilities (left of slashes) and bootstrap percentages (right of slashes). The consensus tree from Bayesian analysis is shown, the maximum likelihood consensus tree has the same topology except within subgenus *Parnassius* (see also fig. 10) and with respect to *Sachaia* and *Lingamius*, which form a distinct, poorly supported (43 percent) clade branching between *Tadumia* and the rest of *Parnassius*.


Archon apollinus
Luehdorfia japonica
Sericinus montela
Bhutanitis thaidina
Zerynthia (Allancastria) cerisyi
Zerynthia rumina medesicaste
Baronia brevicornis
Hypermnestra helios
P.jacquemontii variabilis
P.(epaphus) dongalaicus
P.(jacquemontii) mercurius [W272]
P.epaphus cachemiriensis [W93]
P.nomion nominulus
P.(phoebus) ruckbeili
P.apollo venaissinus
P.(phoebus) smintheus
P.phoebus interpositus
P.(phoebus) bremeri
P.actius
P.tianschanicus
P.honrathi
P.apollonius
P.(tenedius) arcticus
P.tenedius
P.hardwickii
P.schultei
P.acco
P.hunnyngtoni
P.huberi
P.szechenyii
P.cephalus
P.maharaja maharaja
P.(maharaja) labeyriei
P.charltonius sakai
P.inopinatus
P.loxias
P.autocrator
P.imperator imperator
P.(imperator) augustus
P.acdestis
P.hide
P.stoliczkanus stoliczkanus [W261]
P.(stoliczkanus) stenosemus
P.(staudingeri) jacobsoni
P.cardinal
P.(staudingeri) chitralica
P.staudingeri staudingeri
P.patricius
P.delphius albulus
P.(delphius) maximinus
P.(simo) andreji
P.(simo) simonius
P.simo kangruensis
P.(simo) boedromius
P.orleans
P.nordmanni
P.clodius
P.(eversmanni) felderi
P.eversmanni
P.stubbendorfi
P.glacialis
P.ariadne
P.mnemosyne parmenides
P.mnemosyne pseudonubilosus
P.mnemosyne orientalis
Parnassius
Sachaia
Lingamius
Tadumia
Kailasius
Koramius
Kreizbergia
Driopa
0.1

being present in 85 % of Rhopalocera and 95 % of a sample of other families (Material and Methods). In addition to asparagine, two amino-acids, serine (AGN) and lysine (AAR), appear to constitute well tolerated, yet rather short-lived variants at position 159, with scattered occurrences in all, or nearly all, families (overall frequencies 6.5 and 2.7 % of Rhopalocera, respectively). In contrast, glycine (GGN) and histidine (CAY), the other two significant variants in terms of frequency (2.1 and 3.5 %, respectively), are confined to a small number of subtrees. Nearly all occurrences of the former are in Nymphalidae (in *Coenonympha* Hübner 1819, *Junonia* Hübner 1819, and the American section of *Limenitis* Fabricius 1807, in particular), while the latter is restricted almost exclusively to *Parnassius* (only three additional occurrences in 2238 *Rhopalocera* sequences): whereas subgenus *Parnassius* (30 sequences) has asparagine (or serine, in *P. apollo graslini* Oberthür 1891), the remaining seven subgenera (80 sequences) have histidine (with three exceptions; secondary mutations leading to tyrosine (UAY) occurred in *P. nordmanni* Nordmann 1851 and

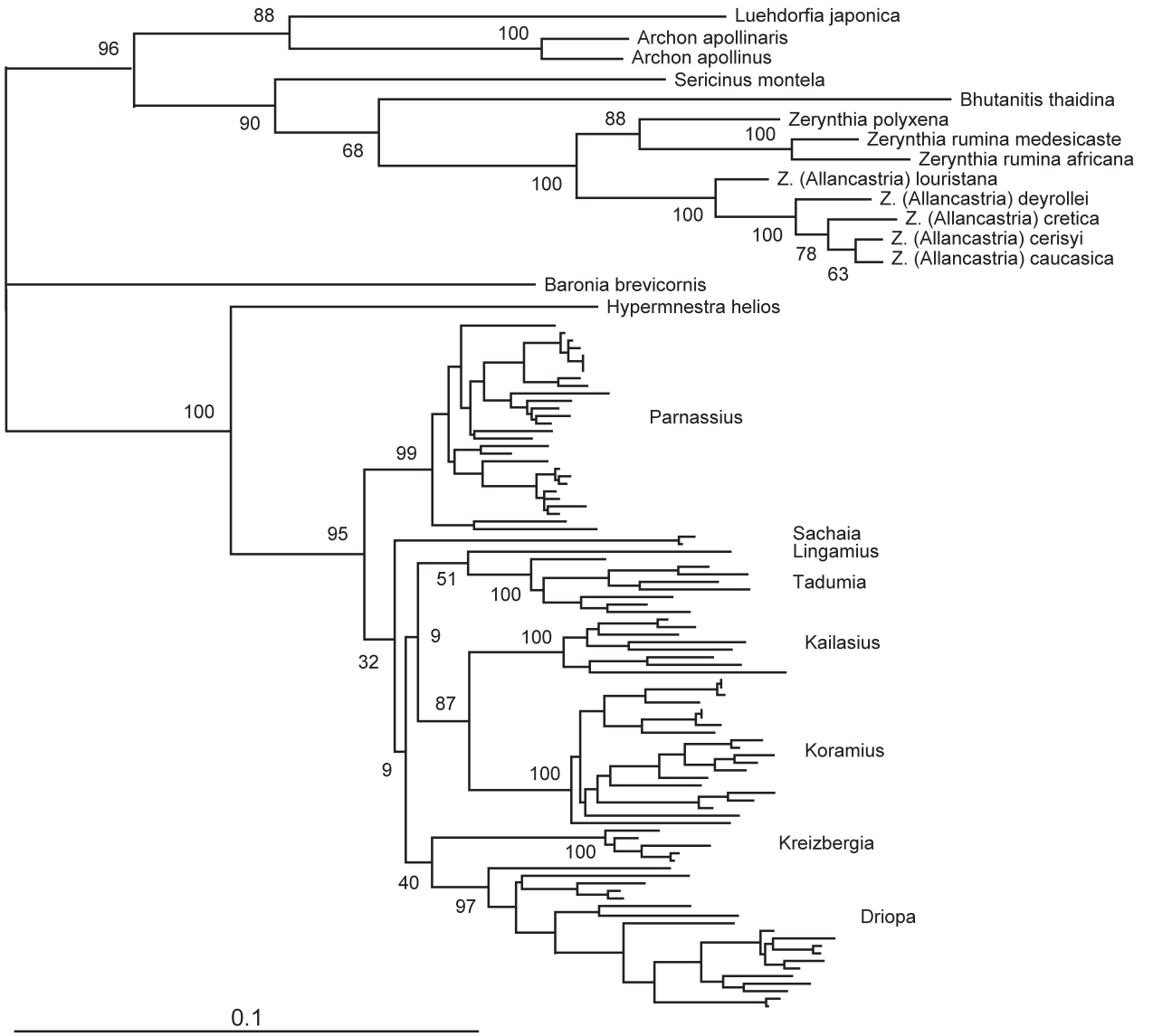


Figure 9
Unrooted maximum likelihood consensus tree of Parnassiinae for the combined LSU, ND1 and CO1 segments (1615 aligned sites). Numbers next to nodes are bootstrap proportions.

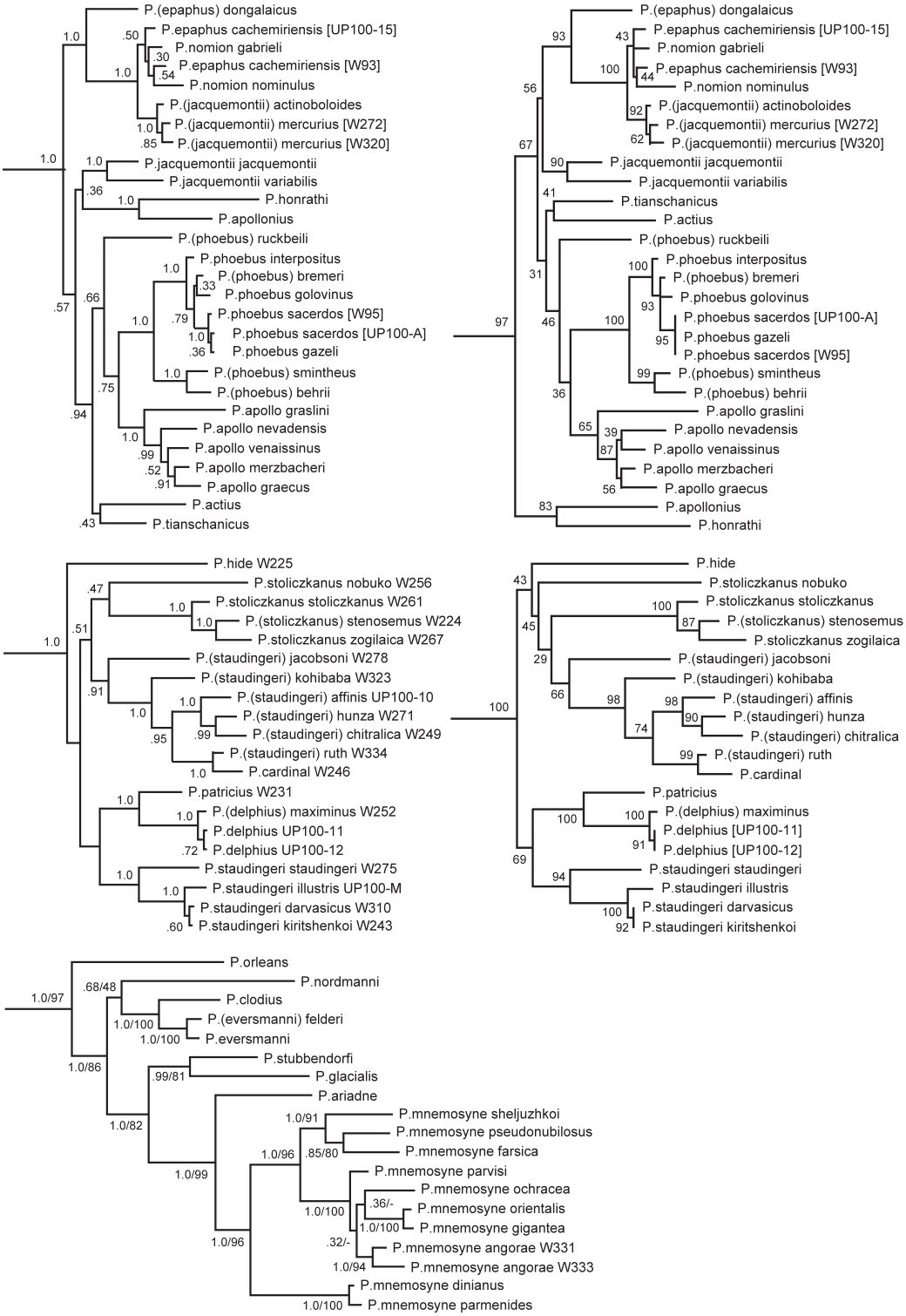


Figure 10
Internal phylogenetic structure of sugenera *Parnassius*, *Koramius* and *Driopa* according to Bayesian analysis (left) and maximum likelihood analysis (right) of the combined LSU, ND1 and CO1 segments (1634 aligned sites). Selected sections of consensus trees of *Parnassius* + *Hypermnestra* are shown, drawn to the same scale. Numbers next to nodes are Bayesian posterior probabilities (left) and bootstrap percentages (right; see fig. 12 for the complete topology of the maximum likelihood tree). The two *Driopa* consensus trees were combined into a single one since their topology was the same except within the (poorly resolved) eurasiatic clade of *P. mnemosyne*.

P. mnemosyne farsica Bang-Haas 1938, and there was a reversion to asparagine in *P. mnemosyne shelzhujkoi* Bryk 1912). Since all other Parnassiinae (49 sequencs) have asparagine at position 159, by far the most likely interpretation of the pattern observed in *Parnassius* is that (i) the asparagine codon mutated to a histidine codon after subgenus *Parnassius* had separated from the rest; (ii) histidine confers some sort of selective advantage, which accounts for its evolutionary stability in the lineage that acquired it.

Aside from the *apollo* group, the relationships between *Parnassius* subgenera remain mostly obscure. *Kailasius* and *Koramius* clearly constitute sister clades, as again anticipated by Munroe, and it is rather likely (Bayesian posterior probabilities 0.97) that *Kreizbergia* and *Driopa* on the one hand, and *Lingamius* and *Tadumia* on the other, are related (the latter share a rare transversion at position 280 of the LSU segment, see fig. 2). However, the relative arrangement of the four resulting subdivisions – *Sachaia*, *Lingamius* + *Tadumia*, *Kailasius* + *Koramius*, *Kreizbergia* + *Driopa* – is largely uncertain, which suggests that divergence occurred over a short period of time, possibly as a result of the colonization of the new ecological niche constituted by Fumariaceae (see tab. 3 and General Discussion).

Within subgenera, phylogenetic relationships tend to be better recovered. In *Driopa* for instance, *P. orleans* Oberthür 1890, which is the only species to have retained a primitive pattern of blue marginal spots, is located basally and the special relatedness of *P. clodius* Ménétriés 1855 and *P. eversmanni* Ménétriés 1855 (the two species that coexist in North America) on the one hand, and *P. stubbendorfi* Ménétriés 1849 and *P. glacialis* Butler 1866 on the other, is confirmed. The fine structure of subgenera *Parnassius* and *Koramius* (fig. 10) is discussed in a subsequent subsection, together with the possible bearings of molecular data on species delimitation.

A temporal frame for the evolution of Parnassiinae and *Parnassius*

While the phylograms in figs. 5 and 8 were generated without any assumption about the pace of evolution, hence their unequal branch lengths, all terminal nodes can be forced to stand an equal distance from the root (*P. brassicae* L. 1758 + *L. celtis* Laicharting 1782 in fig. 5) by assuming a molecular clock to exist (clock option in Paup*). This, however, is at the expense of likelihood, which jumps from –21552.1 to –21633.0 for the tree topology in fig. 5A, indicating that individual lineages experienced significantly different rates of molecular evolution over the DNA segments analyzed (Felsenstein 1988).

Different rates of mitochondrial evolution can result from some localized sections of mitochondrial DNA experiencing markedly slower or faster evolution in some lineages or else, from the evolution of the entire genome changing pace, *e.g.* following some mutation(s) in the replication and reparation apparatus: in the former, but not the latter case, the magnitude of the effect should decrease as the number of sampled genome segments grows (the undetected presence of NUMTs among sequence data could also

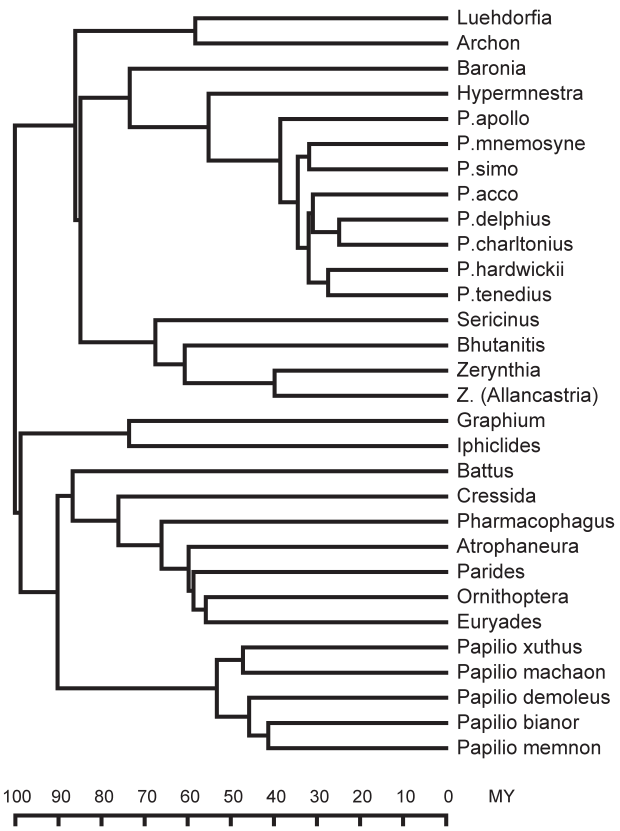
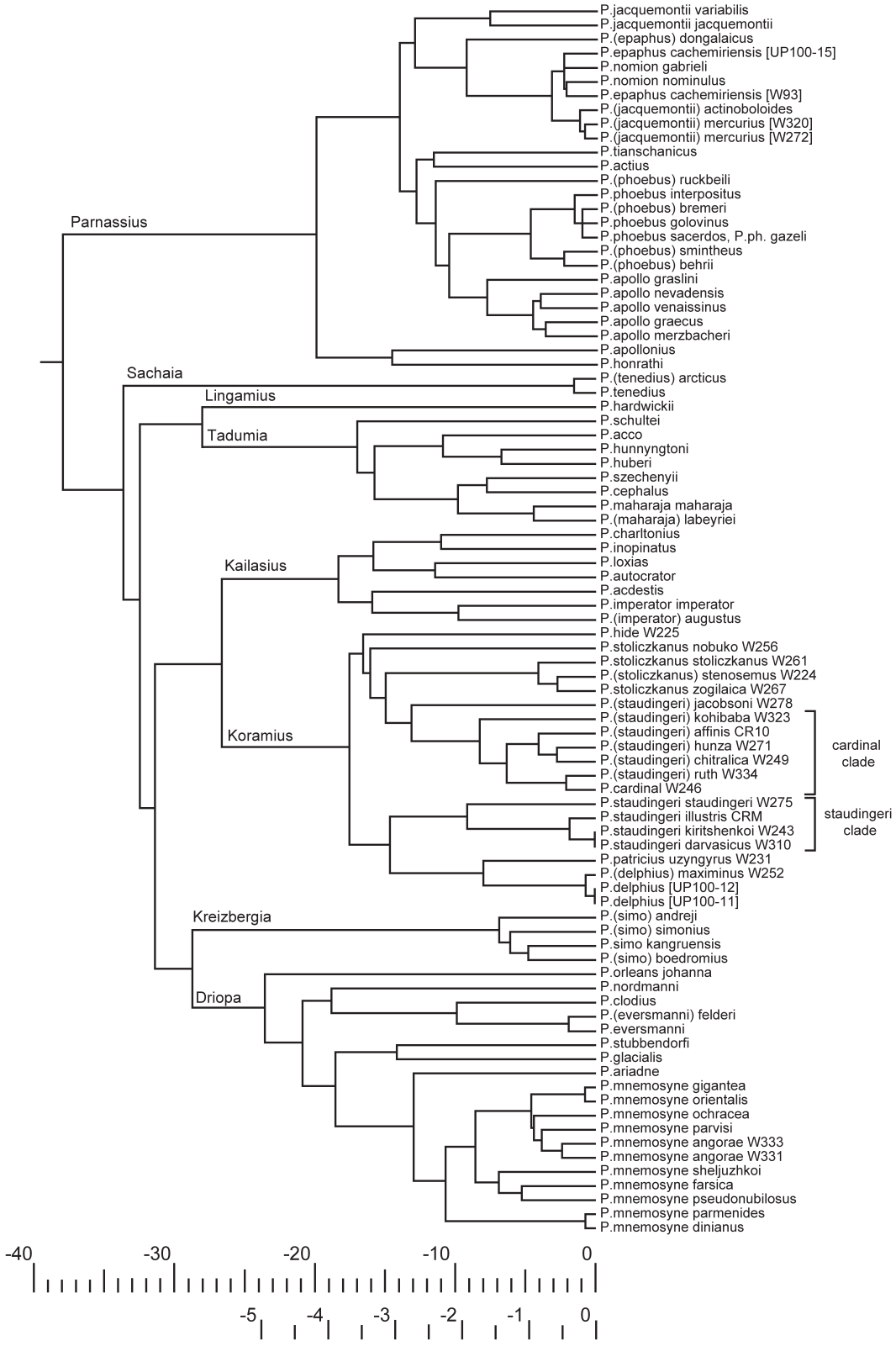


Figure 11
A possible temporal frame for the evolution of Papilionidae. The maximum likelihood consensus tree of fig. 5A was processed by program r8s (Sanderson 2002). The solution shown is the one generated by the cross-validation test. Calibration was generated by placing the root at -100 million years (see Text).

Figure 12
Possible time frames for the evolution of *Parnassius*. The maximum likelihood consensus tree of fig. 10 (rooted with *Hypermnestra helios*) was processed by program r8s (Sanderson 2002). The solution shown is the one generated by the cross-validation test. Two putative calibrations were generated (i) by placing the root of the genus at -38 million years, as in fig. 11; (ii) by assuming an uncorrected rate of 1.5 substitutions per site per million year (see Text).


P.jacquemontii variabilis
P.jacquemontii jacquemontii
P.(epaphus) dongalaicus
P.epaphus cachemiriensis [UP100-15]
P.nomion gabrieli
P.nomion nominulus
P.epaphus cachemiriensis [W93]
P.(jacquemontii) actinoboloides
P.(jacquemontii) mercurius [W320]
P.(jacquemontii) mercurius [W272]
P.tianschanicus
P.actius
P.(phoebus) ruckbeili
P.phoebus interpositus
P.(phoebus) bremeri
P.phoebus golovinus
P.phoebus sacerdos, P.ph. gazeli
P.(phoebus) smintheus
P.(phoebus) behrii
P.apollo graslini
P.apollo nevadensis
P.apollo venaissinus
P.apollo graecus
P.apollo merzbacheri
P.apollonius
P.honrathi
P.(tenedius) arcticus
P.tenedius
P.hardwickii
P.schultei
P.acco
P.hunnyngtoni
P.huberi
P.szechenyii
P.cephalus
P.maharaja maharaja
P.(maharaja) labeyriei
P.charltonius
P.inopinatus
P.loxias
P.autocrator
P.acdestis
P.imperator imperator
P.(imperator) augustus
P.hide W225
P.stoliczkanus nobuko W256
P.stoliczkanus stoliczkanus W261
P.(stoliczkanus) stenosemus W224
P.stoliczkanus zogilaica W267
P.(staudingeri) jacobsoni W278
P.(staudingeri) kohibaba W323
P.(staudingeri) affinis CR10
P.(staudingeri) hunza W271
P.(staudingeri) chitralica W249
P.(staudingeri) ruth W334
P.cardinal W246
P.staudingeri staudingeri W275
P.staudingeri illustris CRM
P.staudingeri kiritshenkoi W243
P.staudingeri darvasicus W310
P.patricius uzyngyrus W231
P.(delphius) maximinus W252
P.delphius [UP100-12]
P.delphius [UP100-11]
P.(simo) andreji
P.(simo) simonius
P.simo kangruensis
P.(simo) boedromius
P.orleans johanna
P.nordmanni
P.clodius
P.(eversmanni) felderi
P.eversmanni
P.stubbendorfi
P.glacialis
P.ariadne
P.mnemosyne gigantea
P.mnemosyne orientalis
P.mnemosyne ochracea
P.mnemosyne parvisi
P.mnemosyne angorae W333
P.mnemosyne angorae W331
P.mnemosyne sheljuzhkoi
P.mnemosyne farsica
P.mnemosyne pseudonubilosus
P.mnemosyne parmenides
P.mnemosyne dinianus
Parnassius
Sachaia
Lingamius
Tadumia
Kailasius
Koramius
Kreizbergia
Driopa
cardinal clade
staudingeri clade
-40
-30
-20
-10
0
-5
-4
-3
-2
-1
0

be responsible for a seemingly higher rate of evolution for particular segments and taxa). In order to try and distinguish between these possibilities we compared distances to outgroups for each of the four segments sequenced. When uncorrected, absolute numbers of substitutions relative to *P. brassicae* and *L. celtis* (fig. 5) are tabulated (not shown; 3rd codon positions were excluded from protein-coding sequences), *Cressida* is found to rank second out of 30 ingroups for the LSU segment, first for ND1, first (by far) for CO1 and fourth for ND5, which leaves little doubt that all or most of the mitochondrial DNA of that taxon has experienced a marked increase in its rate of evolution compared to the other Papilionidae we sampled. On the slow side, however, there is no undisputed winner among the taxa connected by much shorter branches than their closest relatives: *P. apollo* ranks 26/30 for LSU, 26/30 for ND1, 23/30 for ND5, but only 14/30 for CO1; similarly, *B.phileno*r L. 1771 (a Troidini, like *Cressida*) is 26/30, 25/30 and 27/30 for LSU, CO1 and ND5, respectively, but only 13/30 for ND1.

In order to convert our phylograms into chronograms, we resorted to the 'r8s' program of Sanderson (2002) which performs semi-parametric rate smoothing under a penalized likelihood framework and provides the user with a cross-validation procedure to adjust the extent of smoothing to a presumably optimal value. The outcome is shown in fig. 11 for the tree of fig. 5A. The root of the tree was put somewhat arbitrarily at –100 MY (million years), in a compromise attempt to match as closely as possible both the date of divergence of Troidini (–90 MY) in Braby *et al.* (2005) – –87 MY in fig. 11 – and the slightly conflicting date of divergence of Troidini from Papilionini (between –82.5 and –89.1 MY) in Zakharov *et al.* (2004) – –90 MY in fig. 11.

Strikingly, all six Papilionidae tribes are seen to antedate the Cretaceous-Tertiary boundary (at –65 MY), as seems also to be the case for some 43 lineages of placental mammals (Bininda-Emonds *et al.*, 2007) and 40 bird lineages (Brown *et al.*, 2007). Parnassiinae (including *Baronia*) appears somewhat older than proposed by Nazari *et al.* (2007) – *ca* 86 MY instead of *ca* 77 MY. Finally, despite our internal phylogeny of *Parnassius* being distinct from that of Nazari *et al.* (2007), the age we inferred for the genus (38 MY) is very similar to the one estimated by these authors.

Just as is the case for Papilionidae, some lineages within *Parnassius* (*e.g.* subgenus *Parnassius*) evolved at a distinctly slower pace than the rest, while others (*e.*g. the clade of *P. mnemosyne*) underwent accelerated evolution. Nevertheless, a chronogram may be generated for the entire genus (fig. 12; see Material and Methods) by resorting again to program r8s, together with a fixed age of –38 MY for the root of *Parnassius*. As immediately apparent from that chronogram, once subgenus *Parnassius* had separated from the rest, the successive splits that gave rise to the other seven subgenera occurred in rather quick succession – between *ca* –34 and *ca* –27 MY according to fig. 12, that is, within a lapse of time of less than one-fifth the age of the genus. It is tempting to ascribe this rapid diversification, which makes it difficult to recover the precise phylogenetic relationships of the major *Parnassius* subdivisions (fig. 8), to a newly available ecological niche, which we suggest was created by the initial colonization of Fumariaceae (see General Discussion).

In a second phase, from ca –24 to –17 MY, subgenera that currently include multiple species (*Parnassius*, *Tadumia*, *Kailasiu*s, *Koramius* and *Driopa*) started in turn to diversify. *Kreizbergia* is a special case, because the extant taxa in this phylogenetic unit appear to have diverged from one another much later than is the case for other subgenera – posterior to –7 MY according to the temporal calibration used in this subsection, and possibly much more recently (see next subsection). While a tardy process (fig. 12), the diversification of *Kreizbergia* was rapid and allowed its member taxa to occupy nearly all of the central Asian mountains. A likely possibility is that this rapid expansion resulted from the colonization of a new niche (tab. 3): all *P. simo* Gray 1853 relatives appear to use *Lagotis* (Scrophulariaceae) rather than *Corydalis* as a larval foodplant (*Veronica*, another genus from Scrophulariaceae, was also recorded by Kreuzberg 1987, as an alternative larval foodplant for *P. simonius* Staudinger 1889).

An alternative time frame for relatively recent events

Extending our temporal calibration of *Parnassius* evolution to the specific and infra-specific levels (fig. 12) results in unexpectedly elevated estimates for the ages of mitochondrial clades associated with species. For instance, the ancestral mitochondrial genomes of *P. apollo* and *P. mnemosyne* are calculated to be 8.0 and 10.5 MY old, respectively. Just the same, the estimated dates of divergence of many pairs of sister species are surprisingly large. Thus, *P. apollonius* Eversmann 1847 and *P. honrathi* Staudinger 1882, which occasionally generate viable hybrids in nature (one of them is illustrated in Turlin & Manil 2005), are found to have separated some 15.0 MY ago according to fig. 12. The problem is not specific to this work; similar figures were proposed by Nazari *et al.* (2007), who estimated for instance the date of divergence of *Zerynthia polyxena*

and *Z. rumina* at some –15 MY (much the same value was obtained by calibrating the tree in fig. 9 – data not shown). This seems exaggerate, given the fact that the two species, whose ranges overlap only in parts of South-Eastern France, yield hybrids that, although very rare in nature, are not only viable, but proved quite fertile when backcrossed to the parental stocks (Descimon & Michel 1989).

Should then the assumptions used by Zakharov *et al.* (2004), Braby *et al.* (2005), Nazari *et al.* (2007) or ourselves to generate a time frame for the diversification of Papilionidae be challenged? All these works have in common that the divergence of major Papilionidae subdivisions is assumed to antedate the final break-up of Gondwanaland, a process which, as thoroughly discussed by Braby *et al.* (2005), is held responsible for the present distribution of Troidini by most authors in the field. The problem is, that as already noted by Zakharov *et al.* (2004), these assumptions result in estimates of numbers of substitutions per site and million years that tend to be much smaller than those published for other insects. However, the latter are mostly based on relatively recent events, so that this apparent discrepancy, rather than being specific to butterflies or insects, is more likely to illustrate the now well-recognized transition in metazoans between a high mutation rate on the short-term and a much lower substitution rate on the long term (Ho *et al.* 2005).

Unfortunately, the calibration curves that would enable one to convert small percentages of nucleotide substitutions into times of divergence remain to be generated for butterflies. Nevertheless, it should remain meaningful to compare relative estimates of dates as long as they are within the same range: in order to try and provide the reader with an alternate, tentative time scale for relatively recent events, we have added at the bottom of fig. 12 a second ruler, which corresponds to an uncorrected substitution rate of 1.5 substitutions per site per million year, as estimated by Farrell (2001) for CO1, and adopted by Kandul *et al.* (2004) to calibrate their phylogenetic tree of *Agrodiaetus* Hübner 1822, a butterfly genus estimated to be about 3 MY old (see Material and Methods). All times of divergence quoted in the next subsection, which discusses primarily the relationships between mitochondrial clades and species trees, are based on the latter calibration; they are indicated in italics and differ from the ones used heretofore by nearly 5-fold.

Mitochondrial lineages and species delimitation in *Parnassius*

There is considerable controversy currently regarding the extent to which intra- and interspecific levels of mitochondrial DNA divergence overlap. Whereas advocates of the identification of species by DNA 'barcoding' (using the CO1 segment of mitochondrial DNA we sequenced) claim that interspecific nucleotide divergence is as a rule far greater than intraspecific variation (Hebert *et al.* 2003), no such gap is apparent in a comprehensive analysis of available barcode sequences of Lycaenid butterflies (Wiemers & Fiedler 2007). Admittedly, there appears to be an upper bound for intraspecific divergence in the *Lycaenidae* that were sampled – it is less than 3.2 % in at least 95 percent of species – but there exists no lower limit to interspecific distance, which is often found to be less than 1 % (fig. 5 in Wiemers & Fiedler) and can be nil.

A deeper issue at stake is whether or not mitochondrial phylogenetic trees are likely to be congruent with species trees for recently diverged taxa. Of course, recently isolated populations are likely to share some of the mitochondrial DNA diversity that existed prior to their separation, so that young species may not appear reciprocally monophyletic in mitochondrial phylogenies until their mitochondrial genetic pools have been purified by lineage sorting. Of more concern is the possibility of phylogenetic relationships being confounded by the repeated introgression of mitochondrial DNA. In the case of Lepidoptera, since the female, through which mitochondrial DNA is inherited, also happens to be the heterogametic sex, it is a consequence of Haldane's rule (*i.e.* the much higher susceptibility of heterogametic hybrids to inviability and sterility), that introgression of mitochondrial DNA should become negligible once postzygotic incompatibility is established (except, perhaps, if transfection by male-killing *Wolbachia* bacteria, as observed in *Acraea*, were to prove a widespread phenomenon; Jiggins 2003; Hurst & Jiggins 2005). On the whole then, mitochondrial clades should generally agree with other criteria used to delimit lepidopteran species, provided the latter are defined according to their ability to maintain genetic integrity when in contact and have not diverged too recently (reviewed by Sperling 2003). Thus, we regard the figure reported by Wiemers & Fiedler (2007), that at most 43 % of *Agrodiaetus* species appear monophyletic in a mitochondrial phylogeny of the genus, as likely to reflect primarily an explosive speciation rate – 1.6 species per MY according to Kandul *et al.* (2007) – and possibly also, some problems in the delimitation of species.

In the case of *Parnassius*, we have systematically confronted our inferred mitochondrial clades and estimated relative dates of divergence with currently available information regarding species and their

delimitation, using the latest version (2005) of C. Häuser's Papilionidae checklist as a primary reference (in order to allow comparison with Wiemers & Fiedler, 2007, percent divergence over CO1, calculated according to Kimura's two-parameter model, is quoted next to the date of divergence, estimated from all sequence data over the LSU, ND1 and CO1 segments). As is to be seen hereunder, the main lesson of this exercise is that despite the extraordinary attractiveness of *Parnassius* both to amateurs and professionals, much remains to be learnt about the biological systematics of its member species.

(i) Subgenus *Parnassius s. str.* Latreille 1804

In all our trees but the one based on Bayesian analysis of the LSU+ND1+CO1 segments (fig. 10), the first lineage to split from the main trunk of subgenus *Parnassius* consists of *P. apollonius* and *P. honrathi*. The two taxa, whose sister-species status has long been recognized, have overlapping ranges in the mountains of western Central Asia. On the other hand, we were unable to come up with a stable phylogeny for the rest of the subgenus. However, a number of significant facts emerge from the inspection of available data and trees.

In our sampling of subgenus *Parnassius*, we focused on *P. apollo* and *P. phoebus* Fabricius 1793 on the one hand, and taxa traditionally grouped into *P. epaphus* Oberthür 1879, *P. nomion* Fischer De Waldheim 1823 and *P. jacquemontii* Boisduval 1836 on the other. *P. apollo*, which is widely distributed from southern Spain to Mongolia, has attracted much interest because of conservation issues: a number of local European populations, many of which had been granted subspecific status by aficionados based on minor deviations in wing pattern, have gone extinct or are endangered. While we sampled only five individuals from the entire range of the species, two points are already apparent: *P. apollo* displays considerable mitochondrial polymorphism (figs. 12–13), pointing to its relative antiquity (estimated at *–1.7* MY; 2.6 % nucleotide divergence between *graslini* and other subspecies); however, as far as can be judged from such limited data, there does not seem to be a major geographical component to that variability.

In the Alps, the distribution of *P. apollo* broadly overlaps with that of *P. phoebus sacerdos* Stichel 1806 and hybrids are not infrequent in some of the localities where the two species meet (Deschamps-Cottin *et al.* 2000). Interestingly, these hybrids are fertile, pointing to a closer relationship between *P. apollo* and *P. phoebus* than with the rest of the genus, something that is also apparent from the trees in figs. 8 and 10. In contrast to *P. apollo*, there has been much controversy regarding the taxa that should be included into *P. phoebus*, which tends currently to be split into at least three, and more frequently four or five distinct morphological 'species'. The distribution of P. *phoebus* in Eurasia is a highly disjoint one, with the morphologically distinct *P. phoebus sacerdos* from the European Alps being quite isolated from the main nucleus, which extends from the Urals to far eastern Siberia. Even in the relatively limited territory of the Alps, two ecologically distinct populations coexist without overlapping: *gazeli* Praviel 1936 in the extreme Southwest and *styriacus* Fruhstörfer 1907 in the East share with other *phoebus* subspecies the use of *Crassulaceae* as larval foodplant, whereas the remaining populations have shifted to *Saxifraga aizoides*.

At the southern edge of the Asian range of *phoebus*, the status of *ruckbeili* Deckert 1909 has remained controversial. Initially described as a subspecies of *phoebus*, it has alternatively been grouped with *actius* Eversmann 1843 or regarded as a distinct species (Häuser 2005, and references therein). Nor is the situation clearer in North America. The asiatic *P. phoebus* reaches western Yukon as the weakly differentiated subspecies *golovinus* Holland 1930. In southern Yukon, it is replaced by populations of the *smintheus* Doubleday 1847 group, with a different wing pattern, body vestiture and egg microsculpture (Shepard & Manley 1998). The latter character has been used to raise not only *smintheus*, but the allied *behrii* Edwards 1870 from the Sierra Nevada of California, to species rank. A final potential point of controversy is the status of *P. bremeri* Bremer 1864, which replaces *P. phoebus* in far eastern Russia, Manchuria and Korea. The wing pattern and antennae of typical *bremeri* are quite different from those of typical *phoebus*, and it has long been regarded as a different species. However, there exist populations – *e.g. amgunensis* Sheljuzhko 1928, from the lower Amur basin – with characters intermediate between those of *phoebus* and *bremeri*, and their existence raises the question of the actual identity of the latter taxon (Korshunov & Gorbunov 1995; Gorbunov 2001).

We sampled all aforementioned taxa of the *phoebus* complex and used the resulting data to generate the mitochondrial phylogenies and chronogram shown in figs. 10 and 12. Our results confirm and extend those of Omoto *et al.* (2004) who published ND5 sequences of *phoebus*, *bremeri* and *smintheus*. Mitochondrial DNA variation in the Alps is nil: there is not a single nucleotide difference between *gazeli* and neighbouring *sacerdos* populations over the LSU, ND1 and CO1 segments. *P. bremeri* from Korea stands halfway

between *P. phoebus interpositus* Herz 1903 and *P. p. golovinus* – four nucleotide substitutions from each; the distance between *interpositus* and *golovinus* is 6 substitutions – which leaves little room for doubt that instead of being a distinct species, *bremeri* is just another Asian form of *phoebus*. As expected, *smintheus* and *behrii* are closely allied, but the two are rather well separated from the rest of *phoebus*, with an estimated date of divergence (*–1.0* MY; 1.8 % nucleotide divergence) which is smaller than the age of *P. apollo*, for instance, but larger than that at which *mercurius* Grum-Grshimailo 1890 separated from the *nomion-epaphus* lineage (see below). We believe that unless localities in which the two taxa coexist happened to be found in Yukon, it is only by examining the viability and fertility of hybrids between the *phoebus* and *smintheus* lineages that it might be possible to decide whether they should be regarded as constituting distinct species.

The argument for elevating *ruckbeili* to species rank would seem *a priori* much stronger: in all but the tree based on maximum parsimony (fig. 13), the branch leading to *ruckbeili* lies basally with respect to the point of divergence of *apollo* from *phoebus* and the date estimated for the separation from the *apollo-phoebus* lineage is a relatively elevated one (*–2.3* MY; 2.4 %). Still, *ruckbeili* shares with all the other taxa potentially grouped with *phoebus* a unique substitution (an A to C transversion) at position 265 of the LSU segment. This mutation creates a U:C mismatch in the middle of ribosomal RNA helix G3 (fig. 13), and it may have been positively selected, since at least in the basal part of that helix, most substitutions are compensatory (i.e. they restore base-pairing), even at the intraspecific level (compare the sequences of *P. apollo graslini*, *venaissinus* Fruhstörfer 1921, *graecus* Ziegler 1901 and *nevadensis* Oberthür 1891), which indicates that on an evolutionary time scale, mismatches tend to be short-lived. However, the evidence from another rare event, at codon 151 of the ND1 segment, is contradictory and pleads strongly in favour of *ruckbeili* being a distinct entity: whereas the *apollo* and *phoebus* lineages show a serine codon at that position, *ruckbeili* has a cysteine codon, which it shares with all other Parnassiinae, including *Baronia*. Either the latter substitution or the one in the gene coding for the large ribosomal RNA molecule must have become fixed more than once, possibly in response to the selection pressure created by a mutation in a nuclear gene, whose history need not be congruent with that of the rest of the genome.

The other section of subgenus *Parnassius* whose systematics are clearly addressed by sequence data comprises the taxa generally grouped into *jacquemontii*, *epaphus* and *nomion*. *P. jacquemontii* is traditionally presented as a species with marked geographical variation and a curiously disjoint distribution. Our data indicate that it corresponds to the artificial grouping of two essentially unrelated lineages, *jacquemontii* in western Central Asia and *mercurius* in central China. The latter includes populations with lightly marked wings (*actinoboloides* Bang-Haas 1928) or with an intermediate pattern (*mercurius*), as well as others with a dark grey background (*tibetanus* Rühl 1893), which we were unable to sample but that most likely belong here. As seen in figs. 8, 10 and 13, *mercurius* is part of a well-supported clade, whose members share a unique A to C transversion in the terminal loop of helix G3 (the C subsequently mutated to U in *nomion nominulus* Staudinger 1895 and *nomion mandschuricus* Oberthür 1891 – see Katoh *et al.*, 2005, for the latter sequence). The sister lineage of *mercurius* consists of the *epaphus-nomion* taxon group: the two lines diverged rather recently, around *–0.7* MY (1.3 % nucleotide divergence). Despite this relatively late separation, *mercurius* is readily distinguished on morphological grounds from those members of the *epaphus-nomion* complex with which it cohabits. In contrast, the small, high-altitude *epaphus* and the large, low-altitude *nomion*, which had always been regarded as distinct species, cannot be distinguished from one another by their mitochondrial DNA sequences, nor can all individuals from those localities where the distributions of *epaphus* and *nomion* meet be confidently sorted out based on morphological criteria (H. Descimon, unpublished observations). Finally, *dongalaicus* Tytler 1926, which has a very restricted distribution in southern Tibet and had often been grouped with *epaphus* despite its (slightly) keeled sphragis and its sympatry with some populations of the latter, does possess the A to C transversion in helix G3. However, it diverged from the *epaphus-nomion-mercurius* clade some *2.0* MY ago (3.3 % divergence), that is, long before the split between *mercurius* and its sister taxa.

(ii) Subgenus *Sachaia* Korshunov 1988

As already noted by Omoto *et al.* (2004), the extreme similarity of the ND5 sequences of *tenedius* Eversmann 1851 and *arcticus* Eisner 1968 was somewhat unexpected given the divergent wing patterns and ecology of the two taxa. The lack of genetic differentiation between *tenedius* and *arcticus* has been confirmed by Chichvarkhin (2004), who sequenced sections of two nuclear genes in addition to other mitochondrial segments from a larger number of individuals and observed minimal variation. The

small number of substitutions we observed over the LSU, ND1 and CO1 segments (zero, three and two, respectively) provide additional evidence of the recent divergence of *tenedius* and *arcticus* (estimated at *–0.3* MY; 0.3 % nucleotide divergence).

(iii) Subgenus *Lingamius* Bryk 1935

When our LSU + ND1 + CO1 distance matrix was superimposed with the ND5 matrix of Omoto *et al.* (2004), it was noted that the distance between the two *P. hardwickii* samples, which came from quite distant localities, exceeded any other one in that dataset to such an extent that it stood well apart from the rest of the distribution (fig. 4A). The LSU and ND1 sequences of the individual extracted by Omoto *et al.* (2004) were subsequently determined by Katoh *et al.* (2005) and comparison with our own data does reveal a large number of nucleotide substitutions – 13 and 18, respectively. While those are unexpectedly large figures for a presumably intraspecific comparison, they look no longer exceptional when compared with the observed distances between alpine and non-alpine *P. mnemosyne* (between 15 and 23 nucleotide substitutions for the ND1 segment; up to 28 substitutions for the LSU segment), which are estimated to have separated *–2.1* MY ago (3.2 % nucleotide divergence; fig. 12 and discussion below). It should be of interest to sample additional individuals of *P. hardwickii* from intermediate localities.

(iv) Subgenus *Tadumia* Moore 1902

In subgenus *Tadumia*, the traditionally recognized *szechenyii* group (*P. szechenyii*, *P. cephalus* Grum-Grshimailo 1891 and *P. maharaja*) forms a well-supported clade and the recently discovered *P. huberi* Paulus 1999 is clearly related to *P. hunnyngtoni*: both are confined to high altitudes, the former at the northern edge, and the latter at the southern edge, of the Tibetan Plateau. Another late discovery with a restricted range at very high altitudes in southern Tibet, *P. schultei* Weiss & Michel 1989, could occupy a basal position.

The wing pattern of *P. labeyriei* Weiss & Michel 1989 is somewhat different from that of *P. maharaja* and its sphragis is markedly larger: it was accordingly described as a separate species rather than as a subspecies of *maharaja* when it was discovered in south-central Tibet by Weiss & Michel (1989). We estimated the date at which the individuals that we were able to examine diverged at *–0.9* MY (1.7 % nucleotide divergence). This is larger than some interspecific distances, but much smaller than a number of presumably intraspecific ones, so that no decision can be drawn on the exact status of the two taxa from those data alone.

(v) Subgenus *Kailasius* Moore 1902

As already noted by Omoto *et al.* (2004), *P. acdestis* clearly belongs to subgenus *Kailasius*, and not to *Koramius*, where it had traditionally been placed based on the relatively small size and simple structure of its sphragis. It is interesting to note that the wing pattern of some subspecies of *acdestis* (*e.g. imperatoides* Weiss & Michel 1989) is quite similar to that of *imperator* Oberthür 1883 and *augustus* Fruhstorfer 1903, which suggests that the two lineages could be sister clades, as is the case indeed in fig. 8.

In the other subsection of *Kailasius*, the close relationship of *P. loxias* Püngeler 1901 and *P. autocrator* Avinov 1913 is confirmed, despite the latter having originally been described as a form of the somewhat more distantly related *P. charltonius* Gray 1853. Like *P. davydovi*, a quite recent discovery from Kyrgyzstan (Churkin 2006), these insects are cliff-dwellers and their larvae feed on members of the 'Strictae' section of *Corydalis* (Kreuzberg 1987), many of which are chasmophytes indeed (Lidén & Zetterlund 1997).

P. augustus had been regarded as a subspecies of *P. imperator* until 1989, when Weiss & Michel reported that (i) the geographical ranges of the two taxa overlapped in the mountains to the North-West of Lhassa, (ii) their early stages differed and (iii) they had seemingly distinct life cycles (as was to be confirmed by subsequent breeding; F. Michel, unpublished observations). A small, yet apparently constant difference in male genitalia was subsequently discovered by Sugisawa & Kawasaki (1997): the base of the uncus protrudes dorsally in *P. imperator*, but not in *P. augustus*. It should be added that the genetic distance between *imperator* and *augustus* (estimated time of divergence *–2.1* MY; 3.3 % nucleotide divergence) corresponds to the upper intraspecific limit and is well within the range for sister species, which they most probably are.

(vi) Subgenus *Koramius* Moore 1902

Subgenus *Koramius* constitutes a major challenge for taxonomists, with the estimated number of species after removal of *acdestis* ranging from two – *delphius* Eversmann 1843 and *patricius* Niepelt 1911 (Eisner 1966; Ackery 1975) – to close to ten, depending on authors. The subgenus has a nearly continuous distribution in the high mountains of central Asia (fig. 14) and adults present a bewildering variety of wing patterns (fig. 6), which resulted in the naming of many local taxa (as well as a disproportionate number of individual forms). Strikingly enough, it is

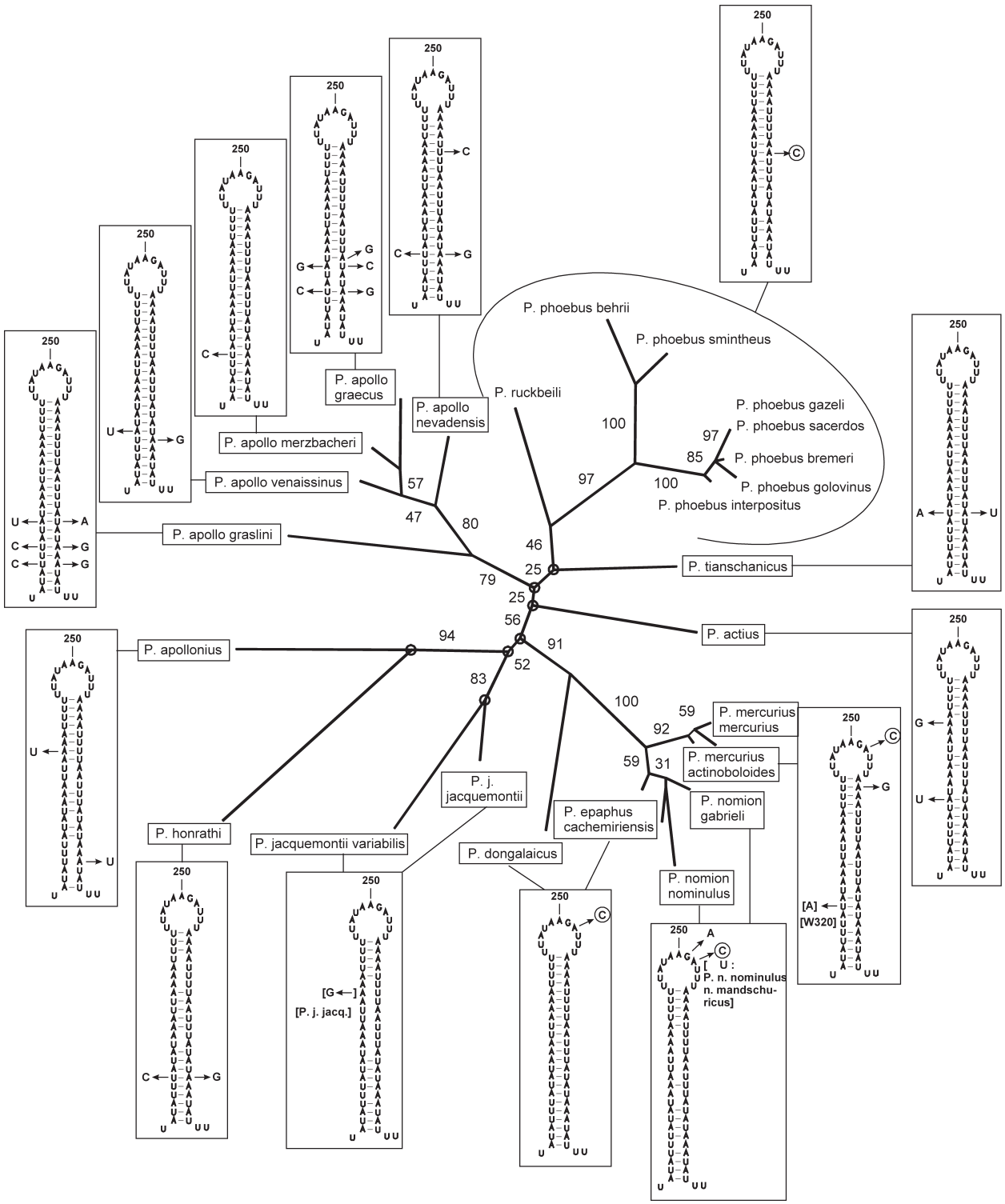


Figure 13
Maximum parsimony analysis of subgenus *Parnassius* over the combined LSU, ND1 and CO1 segments (1606 aligned sites). A consensus unrooted tree is shown with numbers next to nodes corresponding to bootstrap percentages (200 pseudoreplicates; branch lengths were generated by using the consensus topology as input tree). Inserts show changes in stem-loop G3 (sequence numbering as in fig. 2), relative to its proposed ancestral sequence – the same reconstructed state was recovered at all circled nodes. The circled C's at positions 252 of the *epaphus-nomion-mercurius* clade and 265 of the *phoebus* clade are unique so far among Papilionidae (this work, Aubert *et al.* 1999, and F.M., unpublished data).

seldom the case that more than one of those taxa will be found in the same locality and biotope. Exceptions to the rule are few: despite having different altitudinal preferences, *P. patricius* and *P. delphius* fly occasionally together in Kyrgyzstan and Xinjiang and the same appears to be true of *stoliczkanus* Felder & Felder 1865 and *stenosemus* Honrath 1890 in Indian Zanskar (see below). Admittedly also, some morphologically quite distinct entities that are not known to coexist in space and time, nevertheless have partly overlapping ranges, within which they occupy markedly different biotopes. Thus, *cardinal* Grum-Grshimailo 1887 and (*staudingeri*) *infernalis* Elwes 1886 both occur in Central Tajikistan, the latter flying at higher altitudes than the former. And there is evidence as well of some overlap between *stoliczkanus* and forms traditionally grouped with *staudingeri* Bang-Haas 1882: in Ladakh, the two taxa appear to have completely separate ranges on either side of the valley of the Indus, but to the West, in Pakistani Kashmir, *affinis* Bang-Haas 1915 definitely exists south of that river (*e.g.* individual UP100-10) while *stoliczkanus* has been reported from the Haramosh mountains, north of the Indus (fig. 14). As could have been expected, both contacts, between *cardinal* and the northern *staudingeri* subspecies on the one hand, and between *stoliczkanus* and the southern '*staudingeri*' populations on the other, coincide with major genetic discontinuities (fig. 12).

Such occasional cohabitations and range overlaps as well as supposedly constant differences in adult wing pattern, larval morphology, life cycle and, in some cases, minor differences in genitalic structure, have been used indeed to delimit a number of potential species. *P. stoliczkanus*, which had generally been kept separate from the rest of the *delphius* complex based on its distinctive wing pattern (fig. 6) and slightly different sphragis, now tends to be split into two entities, *stoliczkanus* and *stenosemus*. *P. staudingeri*, which was raised to specific rank by Kreuzberg (1985) based on constant differences in male genitalia and wing pattern, was meant to include all taxa in-between the ranges of *delphius* and *stoliczkanus* with the exception of *cardinal*, which was set apart from the populations surrounding its restricted range owing to its striking hindwing eyespots (fig. 6) and distinctive ecology. A further split suggested by Kreuzberg was between *delphius* and its westernmost subspecies: the latter, designated as *maximinus* Staudinger 1891, was stated to differ from the higher altitude, neighbouring *delphius* populations to the East in terms of larval coloration and voltinism. Even so, the status of a number of taxa, especially within and around the poorly accessible Pamir region (*jacobsoni* Avinoff 1913, *kiritshenkoi* Avinoff 1910, *hunza* Grum-Grshimailo 1888), has remained uncertain, with occasional reports of cohabitations between distinct forms (e.g. Tuzov *et al.* 1997).

We have done our best to sample much of the western part of the range of *Koramius* (fig. 14). Strikingly, the resulting trees (figs. 8 and 10) support a model in which the present phylogeographic structure of *Koramius* was generated by the progressive differentiation of populations within an essentially static spatial distribution. Thus, the deepest split (dated around –17.6 MY) was between the North-West populations (*delphius*, *staudingeri* s.str., *patricius*) and the rest of the range. *P. staudingeri* – here meant as including *staudingeri*, *illustris* Grum-Grshimailo 1888, *darvasicus* Avinoff 1916 and, somewhat unexpectedly given its distinctive wing pattern (fig. 6), *kiritshenko*i – then separated from the *delphius-patricius* lineage. The range of *staudingeri* appears to extend north of the Karakoram range at least as far as Xaidulla into Xinjiang: while only a fragment of the ND1 segment of a quite old specimen from the type locality of *abramovi* Bang-Haas 1915 (extract W327) could be successfully amplified, its sequence reveals a close relationship (only one nucleotide substitution) with other members of the *staudingeri* clade.

P. delphius and *P. patricius* were recognized as distinct species ever since the description of the latter nearly one century ago and have largely overlapping ranges, within which they are readily distinguishable from one another on a morphological basis, whether as larvae (F. Michel & E. Zinszner, unpublished observations) or adults (fig. 6). It is therefore somewhat surprising that these sister taxa should have separated relatively recently compared to some other *Koramius* lineages (*–1.6* MY; 2.1 % nucleotide divergence; this date is actually somewhat younger than the separation of *staudingeri staudingeri* from *illustris*, *kiritshenkoi* and *darvasicus*, which suggests that the latter lineage might comprise more than one specific entity). An interesting possibility is that *P. patricius* was the product of a parapatric speciation event, having adapted to a different foodplant (*Cysticorydalis feldtshenkoana*) with a restricted, high-altitude ecological niche within the range of *P. delphius*. One would then expect the genetic diversity of *P. delphius* to be significantly larger than that of its offspring. Whether or not this expectation is borne out by far larger a sample than we were able to examine, it must be emphasized that as already noted by Omoto *et al.* (2004), the taxon *maximinus*, despite claims to its possible specific status, appears very closely related genetically to the rest of *delphius*: only three nucleotide differences over segments LSU+ND1+CO1 separate sample W252 from the two

delphius individuals from central Kyrgyzstan.

South of the range of the *staudingeri* clade as defined in fig. 12, the picture suggested by the phylogenetic trees in figs. 10 and 12 differs deeply from the one that prevails in the current literature. Far from being an isolated entity, *P. cardinal* is closely related to *ruth* Kotzch 1915 (1.1 % nucleotide divergence), its immediate neighbour to the South-East, and the two are part in turn of a statistically well-supported clade whose range extends from Central Afghanistan to Ladakh and includes not only *kohibaba* Clench & Shoumatoff 1956, *chitralica* Verity 1911, *hunza* and *affinis* Pesche & Eisner 1934, but also *tytlerianus* Bryk & Eisner 1932 and *mamaievi* Bang-Haas 1915: the LSU sequence of the former, a 1935 specimen from Ishkuman, Pakistan (W282 in fig. 14), showed it to be closely related to its *chitralica* neighbour, rather than to *stoliczkanus*, as claimed by some (see Ackery 1975), while that of the latter (W288, from Khalsi, Ladakh) is the same as *hunza* (W271), with a single difference from *affinis* (UP100-10). To summarize, *P. staudingeri*, as defined by Kreuzberg (1985), consists of two separate mitochondrial lineages, *staudingeri* to the North and *cardinal* to the South (the names correspond to the earliest described taxon in each entity, see Annex Table), that actually correspond to distinct species, whose ranges overlap in central Tajikistan. As for *jacobsoni*, which separates the *staudingeri* clade from the *cardinal* clade in southern Pamir, its genetic distance from the latter is sufficient to suggest yet another possible specific separation.

Typical *stoliczkanus* is readily distinguished from *stenosemus* by its wing pattern (fig. 6) and most of those who experienced collecting the two taxa at

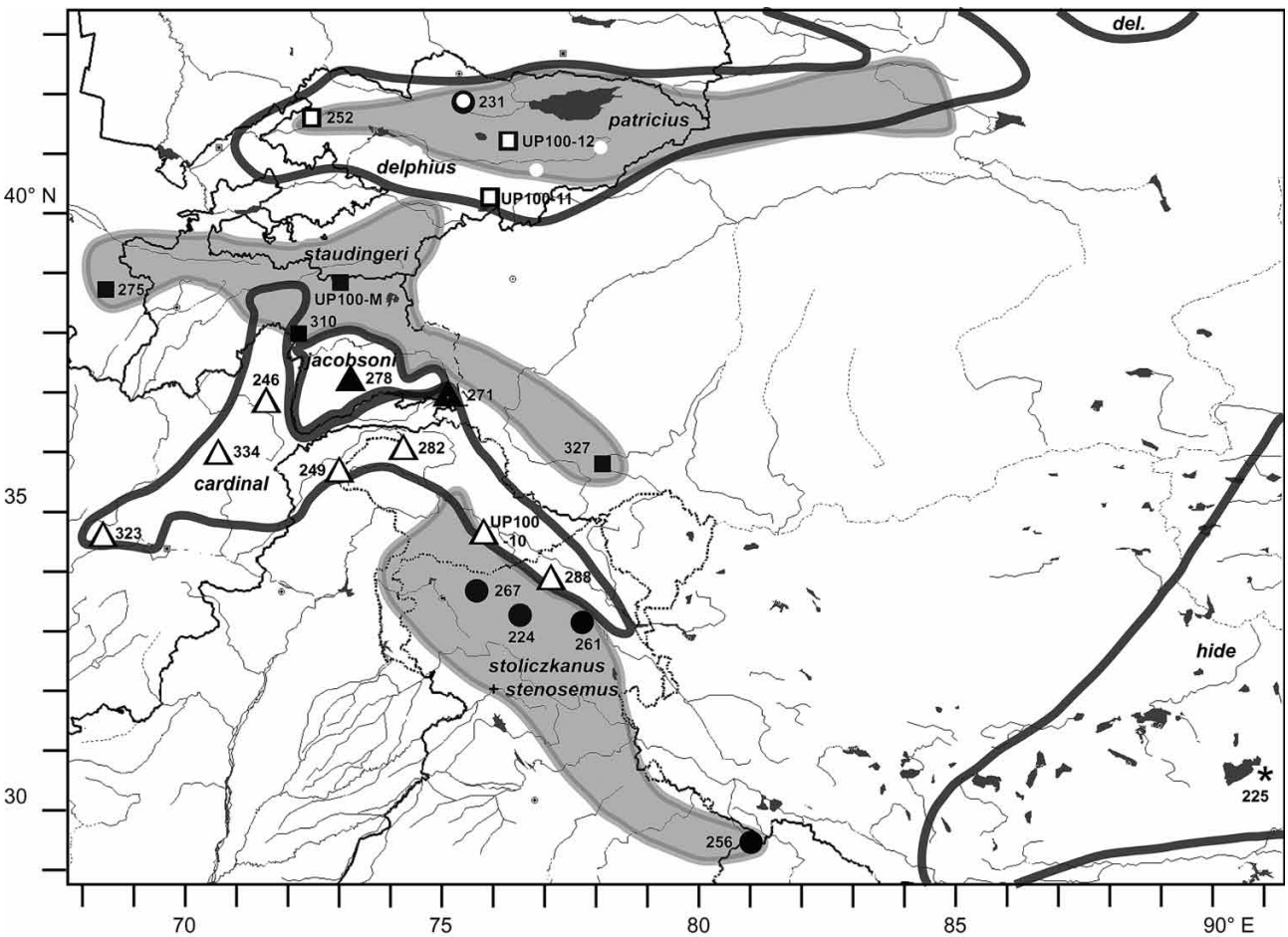


Figure 14
Geographic distribution of *Koramius* mitochondrial lineages. Numbers next to symbols are voucher codes for individuals analyzed (see Annex Table); empty squares, *delphius* lineage; empty circle, *P. patricius*; filled squares, *staudingeri* (*sensu stricto*) lineage; filled triangle, *jacobsoni*; empty triangles, *cardinal* lineage; filled circles, *stoliczkanus* + *stenosemus* lineage; asterisk, *P. hide* Koiwaya 1987. Putative ranges of taxa were inferred primarily from Weiss (1992); Tshikolovets (2004); Tshikolovets (2005a); Tshikolovets (2005b); Turlin & Manil (2005).

the same place in Zanskar (e.g. at Rohtang Pass; at Baralacha Pass; in Nira) are convinced those are distinct species. Still, the genetic distance between *stoliczkanus* and *stenosemus* is relatively low (estimated time of divergence according to fig. 12: *–0.8* MY; 1.1 % nucleotide divergence) and more significantly, there exist many populations with intermediate characters, which have been treated in different ways depending on authors. Thus, *zogilaica* Tytler 1926 (fig. 6; and more generally, *atkinsoni* Moore 1902 from Kashmir) was regarded as belonging to *stoliczkanus* by Weiss (1992), whereas it was grouped together with *stenosemus* by Tshikolovets (2005a): the former arrangement would result in a paraphyletic *stoliczkanus* according to the mitochondrial phylogeny in fig. 10. Given this lack of agreement over the systematics of the *stoliczkanus* complex, it should be of interest to sample a far larger number of populations and forms. However, our finding that *nobuko* Ohya 1996 (fig. 6), from the far eastern end of the *stoliczkanus-stenosemus* range, is so divergent in terms of mitochondrial sequences from the west-central populations of Kashmir and Zanskar that it is excluded from the *stoliczkanus-stenosemus* clade in all but the maximum likelihood tree of fig. 10, suggests that the actual genetic diversity of the *stoliczkanus* complex is far higher than could have been anticipated from wing patterns alone.

(vii) Subgenus *Kreizbergia* Korshunov 1990

The formerly monotypic *P. simo* Gray 1853 group is currently regarded by most authors as being constituted of four species. As emphasized by Kreuzberg (1985), the wing pattern, uncus and shape of the seventh tergite of *simo*, *simonius* Staudinger 1889 and *boedromius* Püngeler 1901 are (moderately) distinct. However, these three taxa are allopatric and their mitochondrial DNAs (as well as that of *andreji* Eisner 1930) are only moderately divergent (1.7–2.2 %), so that they should probably best be regarded as conspecific pending investigations of their genetic compatibility. On the other hand, Koiwaya (1995) reported that *simo* and *andreji* are sympatric, which would justify of course the specific separation of the two taxa if confirmed.

(viii) Subgenus *Driopa* Korshunov 1988

Several of the members of *Driopa* have wide geographical ranges, but *mnemosyne* is particularly noteworthy for the number and diversity of biogeographical provinces it has populated from the Spanish Pyrenees to the Tian-Shan mountains of Central Asia and from coastal Norway (63° N) to the central Zagros mountains of Iran (30° N). As might have been expected, the species displays quite significant variation both in wing shape and wing pattern within that range, but the analysis of the latter is made difficult by the constant absence of both the blue and red eyespots and a trend towards reduction of even the black markings in some populations.

Our sampling of mitochondrial genotypes reveals the presence of three well-defined clades within *P. mnemosyne*. The first one groups the southernmost populations from Iran and Turkey, which are readily distinguished morphologically by their angular forewing and the presence of a well-developed series of white markings within the submarginal area of the forewing. These populations give way in Central Turkey to the main eurosiberian stock, which is shown by molecular data to range from Central Asia at least to Greece and, possibly, even further west. Less expectedly, the genetic discontinuity between eurosiberian and western alpine subspecies (3.2 % nucleotide divergence over CO1) was found to be even larger than the one between northern and southern Turkish subspecies (2.6 %). While the former value corresponds to the upper limit for intraspecific variation, it could be argued that the *mnemosyne* subtree having definitely longer branches compared to the rest of the *Driopa* tree (fig. 10), genetic divergence between *mnemosyne* populations is likely to have been overestimated. However, examination of available data (not shown) reveals no detectable excess of substitutions over the ND1 and CO1 segments in the *mnemosyne* clade with respect to the rest of *Driopa* when other subgenera are used as outgroups; the LSU segment was the only one to be affected by an unusually elevated number of events (up to 28 – out of 518 aligned sites – between *parmenides* Fruhstörfer 1908 and *gigantea* Staudinger 1886).

To summarize, individuals from the Western Alps and from the southern edge of the range are so different in terms of their mitochondrial DNA from other *mnemosyne* samples that they may belong to different species. It should be of particular interest to locate the contact between alpine and Eurasiatic populations in Central Europe and also, to evaluate the viability and fertility of hybrids between individuals belonging to the three major mitochondrial clades.

General discussion

Molecular versus traditional taxonomy

In an ideal world, not only would taxonomic designations represent accurately the phylogenetic relations of the organisms, only truly monophyletic groups being judged worth of consideration, but the use of hierarchical categories (genus, tribe, family) should be

strictly codified according to levels of genetic divergence, as estimated by DNA sequencing. Unfortunately – or, perhaps, fortunately – that is still a remote goal. To take an example from this work, whether the initial split within Parnassiinae was between (Zerynthiini + Luehdorfiini) and Parnassiini (Fig. 5B; fig. 10 of Nazari *et al.*, 2007), in which case Zerynthiini should be dropped as a tribe; or, rather, between Luehdorfiini and (Zerynthiini + Parnassiini), as weakly suggested by fig. 5A, remains statistically indeterminable at present. In the same vein, whether *Parnassius* s.l. and *Zerynthia* s.l., which are about the same age (fig. 11), should both be divided into several genera is still very much a matter of taste (except that, as quite appropriately stated back in 1975 by Ackery, 'If *Allancastria* Bryk is to be recognized as a valid genus it would seem to me that there is equal justification for raising the status of the species groups of *Parnassius* to genera').

However premature it would then be to bring down taxonomy to DNA sequencing, it is just as impossible to ignore the potential impact of molecular phylogenies. Among the novel associations that distinguish our phylogenies of Parnassiinae and *Parnassius* from recently published ones (Omoto *et al.* 2004; Katoh *et al.* 2005; Nazari *et al.* 2007), two are particularly noteworthy: we provide rather strong evidence that the earliest split within *Parnassius* was between subgenus *Parnassius* – the 'P. apollo' group – and the rest of the genus, and also suggest that *Baronia*, rather than being a sister lineage to all Papilionidae, as generally claimed, not only belongs in fact to Parnassiinae, but separated from the ancestor of *Parnassius* and *Hypermnestra* after the divergence of Zerynthiini and Luehdorfiini (fig. 5).

As noted by Vane–Wright (2003) regarding the classification of the Papilionidae as a whole, 'Schemes abound, but we remain far from any consensus'. Still, of the two unrooted topologies in fig. 7, type II is essentially identical to the one proposed in a majority of recent works, whether based on cladistic analyses of morphological characters (Hancock 1983; Miller 1987) or molecular sequences (Nazari et al. 2007), while type I differs from the former by a single exchange between neigbouring nodes. The real source of conflict, then, is about how to root these trees. Whereas *Baronia* had almost universally been assumed to be the first offshoot in the family tree, neither our data, nor the combined molecular phylogeny in Fig. 10 of Nazari et al. (2007) agree and the question now is, which conclusion will turn out to be supported by the sampling of additional genes and outgroups.

Assuming *Baronia* is a Parnassiinae indeed and, possibly, a Parnassiini, why should it have consistently been misplaced by morphologists for the last 100 years – Jordan (1907–1908) was the last one to state that it 'belongs in the neighbourhood of *Parnassius*' ? Part of the answer lies in the unusual wing pattern of the adult (there is evidence that at least some female forms are mimetic; Tyler *et al.* 1994) and the existence of two hindwing anal veins (this, however, has been a much debated character, due to the presence of a vestigial second hindwing vein in many Papilionidae). Still, it is striking to observe the extent to which subjective weighting and polarization of characters

Table 3. Larval foodplants of the Parnassiinae.

Genus	Subgenus	Plant family	Genus (species)
Baronia		Mimosaceae	*Acacia cymbispina*
Hypermnestra		Zygophyllaceae	*Zygophyllum*
Parnassius	*Parnassius*	Crassulaceae	*Sedum, Sempervivum, Rhodiola*
		Saxifragaceae	*Saxifraga aizoides*
	Driopa	Fumariaceae	*Corydalis, Dicentra*
	Kailasius	Fumariaceae	*Corydalis*
	Koramius	Fumariaceae	*Corydalis*
	Lingamius	Fumariaceae	*Corydalis*
	Sachaia	Fumariaceae	*Corydalis*
	Kreuzbergia	Scrophulariaceae	*Lagotis, Veronica*
	Tadumia	Fumariaceae	*Corydalis*
Archon		Aristolochiaceae	*Aristolochia*
Luehdorfia		Aristolochiaceae	*Heterotropa, Asiasarum (=Asarum)*
Sericinus		Aristolochiaceae	*Aristolochia*
Buthanitis		Aristolochiaceae	*Aristolochia*
Zerynthia	*Zerynthia*	Aristolochiaceae	*Aristolochia*
	Allancastria	Aristolochiaceae	*Aristolochia*

have been indulged in so as to reinforce the assumption that *Baronia* is 'primitive'. A perfect example of circular reasoning is provided by the treatment given to a venational character that *Baronia* shares uniquely with *Parnassius* and *Hypermnestra* among Papilionidae: 'If we suppose that the loss of [forewing vein R4] is homologous in the two groups, we should be driven to consider that *Baronia* had evolved between *Archon* (with all 12 veins) and *Parnassius*, in which vein 9 is lost. It seems scarcely credible that a distinct subfamily should have arisen within so homogenous a group as the existing Parnassiinae.' (Ford 1944, cited and approved by most subsequent authors). Similarly, the underground, apparently girdle-less pupa of *Baronia* (Vazquez & Perez 1961; Igarashi 1984) has typically been interpreted as primitive (Tyler *et al.* 1994; de Jong *et al.* 1996) despite the fact that in addition to a majority of Papilionidae, a majority of subdivisions in two of the remaining three butterfly families (Pieridae and Lycaenidae), some skippers (Hesperiidae, the sister group of Papilionoidea, *i.e.* butterflies) and the Hedylidae (the proposed sister group of Papilionoidea + Hesperidae – Scoble 1986) all have girdled pupae (de Jong *et al.* 1996)! Actually, the pupae of *Parnassius* and *Hypermnestra* also lack a girdle and they lie under or on the ground. However, this character appears homoplasious in Papilionidae, for it is also present in *Archon*.

It is also worth noting that the early split between subgenus *Parnassius* and the other subgenera was already proposed by Munroe (1961), who stated 'The *apollo* group and the *mnemosyne* group [*Driopa*] appear to represent divergent primitive strains; the remaining groups can be regarded as direct or indirect derivatives of the *mnemosyne* group.'; and also, 'It is probably useful to recognize two subgenera: *Parnassius*, with a spatulate process (uncus?) between the two lateral processes (socii?) of the male tenth abdominal segment, and *Doritis*, without such a central process'.

Larval foodplants and cladogenesis in the Parnassiinae

As seen at a glance from tab. 3 and fig. 5, cladogenesis is closely correlated with foodplant change in the Parnassiinae. While the physiological and phytochemical aspects of this phenomenon will be considered in detail in a separate paper, two key questions, the answers to which remain particularly elusive, are worth recalling here.

What was the foodplant of the ancestral Parnassiinae?

Neither *Baronia*, whose caterpillars feed on *Acacia cymbispina*, nor *Hypermnestra* and *Parnassius* use Aristolochiaceae as larval foodplants. In contrast, all species of the other two lineages – Zerynthiini and Luehdorfiini – that meet at the root of Parnassiinae are exclusive Aristolochiaceae feeders. Admittedly, the latter two tribes may have split from one another after the initial diversification of Parnassiinae – the possibility is left open by the trees in fig. 5 – but even if that were the case, their common history was clearly a short one. It is therefore a plausible hypothesis that Aristolochiaceae were the ancestral foodplant of all Parnassiinae, rather than having been secondarily colonized by one or two lineages within the subfamily. Interestingly, the estimated age of Parnassiinae meets closely that assumed in fig. 11 for Troidini – a lineage that also feeds exclusively on Aristolochiaceae (Berenbaum 1995). This coincidence, and the fact that the age assumed for the root of Papilionidae, 100 MY, is somewhat younger than the date (*ca* 110 MY) at which Aristolochiaceae have been estimated to have separated from their sister lineage (Magallon & Sanderson 2005) leaves once again open the vexing question of whether the similar larval feeding habits and larval morphology of Zerynthiini and Troidini result from convergence or, rather, from common descent, Aristolochiaceae feeding having subsequently been lost by Papilionini, Leptocircini (=Graphiini) and Parnassiini. Paradoxically, the fact that Aristolochiaceae feeding is 'addictive' – not a single species within the three tribes that share the habit shifted to another foodplant – pleads in favour of convergence. Moreover, convergence of both host plant use and larval habitus is not without precedent in Papilionidae: mature larvae of *Papilio alexanor* Esper 1799 and *Papilio machaon* L. 1758 look very similar and share umbelliferous foodplants (in contrast to a majority of other *Papilio* species, which use Rutaceae), yet belong to quite distant lineages (Zakharov *et al.* 2004).

What was the foodplant of the ancestral Parnassiini?

As already underscored, the divergence of the lineage leading to *P. apollo* and its relatives was followed by the rapid diversification of the other branch, possibly in response to a shift to Fumariaceae as a host plant (of the remaining subgenera, only *Kreizbergia* uses instead Scrophulariaceae and the trait is clearly a late acquisition, see fig. 12). Such a scenario implies that the ancestral *Parnassius* larva did not feed on Fumariaceae and raises the question of what might have been the original foodplant of the genus. Crassulaceae, the larval host of subgenus *Parnassius* (with the exception of some populations of *P. phoebus*, which have shifted to Saxifragaceae) is a possibility, but we regard it as more likely that colonization of that plant family was also a

relatively late event, that ensured the success of subgenus *Parnassius*. Yet another possibility is *Zygophyllum*, the foodplant of *Hypermnestra helios*. However, feeding on *Zygophyllum* may just as well be one of the many specializations of *Hypermnestra*, which would leave forever open the question of the nature of the foodplant that the ancestor of *Parnassius* may have used in the semi-arid environment it possibly shared with the forbearers of its *Hypermnestra* sister lineage.

Finally, it is interesting to note that at least in *Parnassius*, which has by far the largest number of species among Parnassiinae, adaptation to new foodplants and conditions is an ongoing process. In the Alps, a majority of *P. phoebus sacerdos* populations recently left Crassulaceae in favour of *Saxifraga aizoides.* In Kyrgyzstan, parapatric speciation linked both to habitat and foodplant differentiation is likely to be responsible for the relatively recent divergence between *P. delphius* and *P. patricius*. And the recent success of subgenus *Kreizbergia* shows that the conquest of a new ecological niche can mean also areal expansion.

Biogeographical origins of *Parnassius*

Considerations on foodplant change bring us to the question of the geographical origin of *Parnassius* and its major subdivisions. Biogeographical hypotheses for the diversification of Parnassiinae were discussed in detail by Nazari *et al.* (2007), who suggested that the group may have originated in the Irano-Touranian region (which happens to be the first author's homeland). Facts are, that *Parnassius* has a wide distribution nowadays, ranging from southern Spain to New Mexico. However, five of the eight extant subgenera are strictly confined to the mountains of Central Asia. Moreover, in subgenus *Parnassius*, only two lineages (*nomion* and *apollo* + *phoebus*) managed to expand out of Central Asia and both have as closest relatives taxa from the Central Asian mountains (*ruckbeili* is confined to a rather small area of southern Mongolia and *epaphus* is a dweller of Tibet and surrounding mountains). This leaves only *P. tenedius*, whose home range covers parts of Siberia in addition to Mongolia and *Driopa*, which is well distributed over Eurasia and North America, to plead in favour of a non-centralasian origin. However, the first offshoot from the main *Driopa* trunk, *P. orleans*, also happens to be confined to Central Asia, so that the balance is definitely in favour of the latter region being not only the region of highest *Parnassius* diversity, but also the genus homeland. It should be added first, that the age we inferred for *Parnassius* – 38 MY – is clearly posterior to the onset of the Indian-Asian plate collision (currently estimated at ca –50 MY; e.g. Zhu *et al.* 2005), which gave rise to the mountain system of Central Asia, and second, that the sister taxon of *Parnassius*, *Hypermnestra*, is a dweller of the semi-desert areas that lie to the West and Southwest of the Central Asian mountains.

Mitochondrial clades and species delimitation

Coming now to the species level, our data confirm and extend the conclusions of Omoto *et al.* (2004), which rested on the sequencing of a single gene (ND5). Our suggestions for nomenclatural changes within *Parnassius* are provided in the Annex Table and will be briefly commented here. As anticipated from field studies which had revealed the existence of morphologically intermediate populations, *bremeri* and *felderi* Bremer 1861 display weak genetic differentiation from *phoebus* and *eversmanni*, respectively, and are undoubtedly subspecies of the latter. Despite their ecological isolation from *tenedius* and *delphius*, respectively, the same is probably true of *arcticus* and *maximinus*. In contrast to these 'bad species' (i.e. entities that do not comply with the criteria of species concepts; Descimon & Mallet, in press), *augustus* and *ruckbeili* are highly distinct genetically from *imperator* and *phoebus* and should be regarded as sister species of the latter. Still, it is specifically in subgenera *Parnassius* and *Koramius*, which were significantly better sampled in the present work (28 and 23 individuals, respectively, as against 13 and 9 in Omoto *et al.*), that the trees of figs. 8, 10 and 12 suggest the largest number of changes to previous systematic arrangements (Ackery 1975; Weiss 1991–2005). True to say, in *Parnassius*, *mercurius* and *dongalaicus* were already treated as separate from *jacquemontii* and *epaphus*, respectively, by the latter author, but we now show that the first pair lacks any close relationship in terms of mitochondrial genotypes, while the second one could constitute a monophyletic entity only if *nomion* and *epaphus* were regarded as conspecific, contrary to a tradition that is more than a century-old. Similarly, in *Koramius*, the many taxa that had been grouped under the label '*staudingeri*' are now found to be distributed into two separate lineages that lack any close affinity. The first of these lineages includes nominate *staudingeri* while the second one should be called *cardinal*: the latter taxon had been regarded as isolated in Northern Afghanistan and Central Tajikistan, but is now shown to be closely related genetically to its immediate neighbours to the south.

Since all our inferences regarding species delimitation are based exclusively on mitochondrial genotypes, the priority now should be to sample nuclear genes and check the extent to which their phylogeny agrees with the picture drawn by the trees in fig. 10. That, however,

will require going back to the field in a quest for fresh samples, from which not only mitochondrial, but also nuclear segments of DNA may readily be amplified. In any case, finer geographic sampling and larger numbers of individuals are necessary to corroborate and extend our conclusions. We hope that the data presented here will serve as an useful guide for further work on speciation in *Parnassius*, which should encompass both some additional phylogeography and biological investigations of the viability and fertility of the offspring of crosses. Except in subgenera *Parnassius* and *Driopa*, a majority of the known species of *Parnassius* have yet to be successfully reared and in fact, early stages are still unknown for many of the taxa that inhabit poorly accessible districts and biotopes. This is another area in which our database should prove useful. Now that most of the major taxa in *Parnassius* have been 'barcoded' by their mitochondrial genotype (Hebert *et al*. 2003), it should be possible to identify any larva, pupa or egg collected and photographed in the wild by PCR amplification of its mitochondrial DNA, without the need to rear it successfully to the adult stage.

Acknowledgements. We are grateful to F. Carbonell, M. Gaboly, H.-J. Geiger, J. Michel, B. Mollet, V. Pierrat and above all, to J.-C. Weiss, for generously providing us with material for DNA extractions. Special thanks are also due to J.-C. Weiss for checking the identification of some specimens and providing us with rare pieces of entomological literature. Constructive comments by Richard Vane-Wright, an anonymous referee and Pierre Rasmont helped us to improve this article.

References

Ackery P.R. 1975. A guide to the genera and species of Parnassiinae (Lepidoptera: Papilionidae). *Bulletin of the British Museum (Natural History), Entomology* **31**: 71-105.

Aubert J., Legal, L., Descimon, H., Michel, F. 1999. Molecular phylogeny of swallowtail butterflies of the tribe Papilionini (Papilionidae, Lepidoptera). *Molecular Phylogenetics and Evolution* **12**: 156–167.

Bensasson D., Zhang D.-X., Hartl D.L., Hewitt G.M. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *TRENDS in Ecology* & *Evolution* **16**: 314-321.

Berenbaum M.R. 1995. Chemistry and oligophagy in the Papilionidae, p. 27-38 *in*: **Scriber J.M., Tsubaki Y., Lederhouse R.C. (eds.),** *Swallowtail Butterflies, Their Ecology and Evolutionary Biology*. Scientific Publishers, Gainesville.

Bininda-Emonds O.R., Cardillo M., Jones K.E., MacPhee R.D., Beck R.M., Grenyer R., Price S.A., Vos R.A., Gittleman J.L., Purvis A. 2007. The delayed rise of present-day mammals. *Nature* **446**: 507-12.

Braby M.F., Trueman J.W.H., Eastwood R. 2005. When and where did troidine butterflies (Lepidoptera : Papilionidae) evolve? Phylogenetic and biogeographic evidence suggests an origin in remnant Gondwana in the Late Cretaceous. *Invertebrate Systematics* **19**: 113-143.

Brown J.W., Payne R.B., Mindell D.P. 2007. Comment. Nuclear DNA does not reconcile 'rocks' and 'clocks' in Neoaves: a comment on Ericson *et al*. *Biological Letters* **3**: 257-259.

Bryk F. 1935. *Lepidoptera, Parnassiidae. pars II (Subfam. Parnassiinae)*. Tierreich 65, Walter de Gruyter et C°, Berlin, li + 790 p.

Caterino M.S., Reed R.D., Kuo M.M., Sperling F.A.H. 2001. A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). *Systematic Biology* **50**: 106–127.

Caterino M.S., Sperling F.A.H. 1999. *Papilio* Phylogeny Based on Mitochondrial Cytochrome Oxidase I and II Genes. *Molecular Phylogenetics and Evolution* **11**: 122-137.

Chichvarkhin A. 2004. Molecular phylogeny of Parnassian butterflies of the (sub) genus *Sachaia* Korshunov, 1988: Unexpectedly low genetic divergence sheds light to their systematics and evolution. *Wallace* **9**: 25-29.

Churkin, S. 2006. A new species of *Parnassius* Latreille, 1804 from Kyrgyzstan (Lepidoptera, Papilionidae) *Helios* **7**: 142-158.

Clary D.O., Wolstenholme D.R. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *Journal of Molecular Evolution* **22**: 252-271

de Jong R., Vane-Wright R.I., Ackery P.R. 1996. The higher classification of butterflies (Lepidoptera): problems and prospects. *Entomologia Scandinavica* **27**: 65–102.

Delorme M.O., Henaut A. 1988. Merging of distance matrices and classification by dynamic clustering. *Computer Applied Biosciences* **4**: 453-8.

Deschamps-Cottin M., Aubert J., Barascud B., Descimon H. 2000. Hybridation et introgression entre « bonnes espèces » : le cas de *Parnassius apollo* et *P. phoebus* (Lepidoptera : Papilionidae). *Comptes Rendus de l'Académie des Sciences de Paris, Life Sciences* **323**: 327-337.

Descimon H., Michel F. 1989. Expériences d'hybridation intra- et interspécifique dans le genre *Zerynthia* (Papilionidae). Relativité des critères mixiologiques de l'espèce. *Nota lepidopterologica* **12** (Suppl. 1): 28-31.

Dujardin F. 1965. Papilionidae. Espèces de France et des Alpes-Maritimes. *Entomops* **1**: 77-89.

Ehrlich P.R. 1958. The comparative morphology, phylogeny and higher classification of butterflies (Lepidoptera: Papilionoidea). *University of Kansas Science Bulletin* **34**: 305-370.

Eisner C. 1966. Parnassiidae-Typen in der Sammlung J.C. Eisner. *Zoölogische Verhandlungen, Leiden* **81**: 1-190.

Farrell B.D. 2001. Evolutionary assembly of the milkweed fauna: Cytochrome oxydase I and the age of *Tetraopes* beetles. *Molecular Phylogenetics and Evolution* **18**: 467-478.

Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* **22**: 521–565

Folmer O., Black M., Hoeh W., Lutz R., Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**: 294–299.

Ford E.B. 1944. Studies on the chemistry of pigments in the Lepidoptera, with reference to their bearing on systematics. 4. The classification of the Papilionidae. *Transactions of the Royal entomological Society of London* **94**: 201-223.

Goloboff P.A. 1993. Estimating character weights during tree search. *Cladistics* **9**: 83-91.

Gorbunov P.Y. 2001. *The butterflies of Russia: classification, genitalia, keys for identification*. Thesis. Russian Academy of Sciences, Institute of Plant and Animal Ecology, Ekaterinenburg, 320 p. [in English].

Guindon S., Gascuel O. 2003. PhyML - A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**: 696-704.

Hajibabaei M., deWaard J.R., Ivanova N.V., Ratnasingham S., Dooh R.T., Kirk S.L., Mackie P.M., Hebert P.D. 2005. Critical factors for assembling a high volume of DNA barcodes. *Philosophical Transactions of the Royal Society London* B **360**: 1959-67.

Hall T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids. Symposia Ser.* **41**: 95-98.

Hancock D.L. 1983. Classification of the Papilionidae: a phylogenetic approach. *Smithersia* **2**: 1-48.

Häuser C. L. 1993. Critical comments on the phylogenetic relationships within the family Papilionidae (Lepidoptera). *Nota Lepidopterologica* **16**: 34–43.

Häuser C. L. 2005 (in cooperation with **de Jong R., Lamas G., Robbins R. K., Smith C., Vane-Wright R.I.**). Papilionidae *revised GloBIS/GART species checklist* (2nd draft). http://www.insects-online.de/frames/papilio.htm

Hebert P. D., Cywinska A., Ball S. L., deWaard J. R. 2003. Biological identifications through DNA barcodes. *Proceedings of Biological Sciences* **270**: 313-321.

Hiura L. 1980. A phylogeny of the genera of Parnassiinae based on analysis of wing pattern, with description of a new genus (Lepidoptera: Papilionidae). *Bulletin of the Osaka Museum of Natural History* **33**: 71-85.

Ho S. Y., Phillips M. J., Cooper A., Drummond A. J. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* **22**:1561-1568.

Huelsenbeck J. P., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754-755.

Hurst G. D. D., Jiggins F. M. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B* **272**: 1525–1534.

Igarashi S. 1984. The classification of the Papilionidae mainly based on the morphology of their immature stages. *Tyô to Ga* **34**: 41-96.

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

Jordan, K. (1907) 1926. Papilionidae, p. 45 *in* **Seitz A. (ed.)**, *Les Macrolépidoptères du Globe.* French edition, Le Moult, Paris.

Kabsch W. 1978. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica* **A34**: 827-828

Kandul N. P., Lukhtanov V. A., Dantchenko A. V., Coleman J.W., Sekercioglu C. H., Haig D., Pierce N.E. 2004. Phylogeny of *Agrodiaetus* Hubner 1822 (Lepidoptera: Lycaenidae) inferred from mtDNA sequences of COI and COII and nuclear sequences of EF1-alpha: karyotype diversification and species radiation. *Systematic Biology* **53**: 278-98.

Kandul N. P., Lukhtanov V. A., Pierce N.E. 2007. Karyotypic diversity and speciation in *Agrodiaetus* butterflies. *Evolution* **61**: 546-59.

Katoh T., Chichvarkin A., Yagi T., Omoto K. 2005. Phylogeny and evolution of butterflies of the genus *Parnassius*: inferences from mitochondrial 16*S* and *ND1* sequences. *Zoological Science* **22**: 343–351.

Koçak A. O. 1989. Description of the genus *Adoritis* (*gen. nov.*) with notes on closely related groups in Parnassiinae (Papilionidae, Lepidoptera). *Priamus* **4**: 163-170.

Koiwaya S. 1995. Re-assortment of *Parnassius simo* and *andreji* - On the discovery of sympatric habitat of *P. simo* and *P. andreji*, with the description of new ssp. of *andreji*. *Gekkan-Mushi* **297**: 8-13.

Korshunov Y. P. 1988. Novye bulavusye tscheshuekrylye (Lepidoptera, Rhopalocera iz Khakasii, Tuvy I Jakutii. *Novye I Maloizvestiye Vidy Fauny Sibiri* **20**: 65-80. (in Russian).

Korshunov Y. P. 1990. New genera of the subfamily Parnassiinae Swainson 1840, p. 99-105 *in*: **Zolotarenko G. S. (ed.)**, *Arthropods and Helminths: collected scientific works*. Nauka, Novosibirsk.

Korshunov Y. P., Gorbunov P. Y., 1995. *Dnevnye babochki aziatskoi chasti Rossii. Spravochnik.* [*Butterflies of the Asian part of Russia. A handbook*]. Ural University Press. Ekaterinenburg, 202 p. (in Russian)

Kreuzberg A. V.-A. 1985. Parusniki grupp *delphius*, *charltonius*, *simo* (Lepidoptera, Papilionidae) fauny SSSR, p. 25-68 *in*: **Ministerstvo kultury Uzbekskoj SSR (ed.)**, *Issledovanija flory i fauny srednej Azii*. Tashkent.

Kreuzberg A. V.-A. 1987. Stenophagy in *Parnassius* (Lepidoptera: Papilionidae) of Central Asia and Altai. *Entomologist's Gazette* **38:** 95-102.

Lidén M., Zetterlund H. 1997. *Corydalis, a gardener's guide and a monograph of tuberous species*, *in*: **C. Grey-Wilson (ed.)**. AGS Publications, Avon Bank, U.K. 146 p.

Magallon S. A., Sanderson M. J. 2005. Angiosperm divergence times: the effect of genes, codon positions, and time constraints. *Evolution* **59**:1653-70.

Martin J. F., Gilles A., Lörtscher M., Descimon H. 2002. Phylogenetics and differentiation among the western taxa of the *Erebia tyndarus* group (Lepidoptera: Satyrinae*). Biological Journal of the Linnean Society* **75***:* 319-332.

Miller J. S. 1987. Phylogenetic studies in the Papilioninae (Lepidoptera: Papilionidae). *Bulletin of the American Museum of Natural History* **186:** 365–512.

Moore F. 1902. *Lepidoptera Indica. Rhopalocera: family Papilionidae, sub-families Parnassiinae, Thaidinae, Leptocircinae and Papilioninae*. Vol. **5**, p. 97-225. Lovell Reeve & Co., London.

Munroe E. 1961. The classification of the Papilionidae (Lepidoptera). *Canadian Entomologist* Suppl. **17**: 1–51.

Nazari V., Zakharov E. V., Sperling F. A. 2007. Phylogeny, historical biogeography, and taxonomic ranking of Parnassiinae (Lepidoptera, Papilionidae) based on morphology and seven genes. *Molecular Phylogenetics and Evolution* **42**:131-56.

Niehuis O., Yen S-H., Naumann C. M., Misof B. 2006. Higher phylogeny of zygaenid moths (Insecta: Lepidoptera) inferred from nuclear and mitochondrial sequence data and the evolution of larval cuticular cavities for chemical defence. *Molecular Phylogenetics and Evolution* **39**: 812-829.

Omoto K., Katoh T., Chichvarkin A., Yagi T. 2004. Molecular evolution of the 'apollo' butterflies of the genus *Parnassius* (Lepidoptera: Papilionidae). *Gene* **326**: 141-147.

Rebourg C., Péténian F., Cosson E., Faure E. 2006. Patterns of speciation and adaptive radiation in *Parnassius* butterflies. *Journal of Entomology* **3**: 204-215.

Ronquist F., Huelsenbeck J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574.

Sanderson M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**: 101-109.

Scoble M. J. 1986. The structure and affinities of the Hedyloidea: a new concept of butterflies. *Bulletin of the British Museum Natural History (Entomology)* **53**: 251–286.

Scott J. A. 1985. The phylogeny of butterflies (Papilionoidea and Hesperioidea). *J. Res. Lepid.* **23**: 241-281.

Scriber J. M., Tsubaki Y., Lederhouse R. C. 1995. *Swallowtail Butterflies: their Ecology and Evolutionary Biology*. Scientific Publishers, Gainesville, 459 p.

Selmer M., Dunham C. M., Murphy F. V. 4th, Weixlbaumer A, Petry S., Kelley A. C., Weir J. R., Ramakrishnan V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**:1935-42.

Shepard J. H., Manley T. R. 1998. A species revision of the *Parnassius phoebus* complex in North America (Lepidoptera: Papilionidae), p. 717-726 *in*: **Emmel, T. C. (ed.)**, *Systematics of Western North American Butterflies*. Mariposa Press, Gainesville, Florida.

Sperling F. A. H., 2003. Butterfly molecular systematics: from species definitions to higher level phylogenies, p. 431-458 *in***: Boggs C. L., Watt W. B., Ehrlich P. R. (eds**.), *Ecology and Evolution Taking Flight: Butterflies as Model Study Systems*. University of Chicago Press, Chicago.

Stekolnikov A. A., Kuznezov V. I. 2003. Evolution of the male genitalia, phylogenesis and systematic position of the subfamilies Baroniinae Salvin 1893, Luehdorfiinae Tutt 1896 *stat. nov.* and Zerynthiinae Grote 1899 in the family Papilionidae (Lepidoptera) *Entomological Review* **83**: 350-436.

Sugisawa S., Kawasaki Y. 1997. Re-classification of *Parnassius imperator* from central and southern parts of Tibet. *Gekkan-Mushi*, **318**: 4-9.

Swofford D. L. 2002. PAUP*: Phylogenetic Analysis Using Parsimony [and Other Methods]. 4.0 Beta. Sinauer Associates, Sunderland, Massachusetts.

Tshikolovets V. 2004. *The Butterflies of Tajikistan***.** 500 p., 80 colour plates. Tshikolovets, Pardubice, Czech Republic.

Tshikolovets V. 2005a. *The Butterflies of Ladakh (N.-W. India)***.** Tshikolovets, Pardubice, Czech Republic, 176 p., 30 colour pls.

Tshikolovets V. 2005b. *The Butterflies of Kyrgyzstan***.** 511 p., 108 colour plates. Tshikolovets, Pardubice, Czech Republic.

Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H., Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S. 1996. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å . *Science* **272**:1136-44.

Turlin B., Manil L. 2005. *Etude synoptique et répartition mondiale des espèces du genre Parnassius Latreille 1804 (Lepidoptera Papilionidae).* Lépidoptéristes Parisiens, Paris, 80 p.

Tuzov V. K., Bogdanov P. V., Devyatkin A. L., Kaabak L. V., Korolev V. A., Murzin V. S., Samodurov G.D., Tarasov E. A. 1997. *Guide to the Butterflies of Russia and adjacent territories (Lepidoptera, Rhopalocera).* Vol. 1, Hesperiidae, Papilionidae, Pieridae, Satyridae. Pensoft, Sofia, Moscow, 480 p.

Tyler H. A., Brown K. S., Wilson K. 1994. *Swallowtail Butterflies of the Americas: A Study in Biological Dynamics, Ecological Diversity, Biosystematics and Conservation*. Scientific Publishers, Gainsville, 377 p.

Vane-Wright R. I. 2003. Evidence and identity in butterfly systematics, p. 477-513 *in*: **Boggs C. L., Watt W. B., Ehrlich P. R. (eds.)**, *Butterflies: ecology and evolution taking flight*. University of Chicago Press, Chicago.

Vasquez L., Perez H. 1962. Observaciones sobre la biologia de *Baronia brevicornis* Salv. (Lepidoptera: Papilionidae – Baroniinae). *Annales del Instituto de Biologia de la Universidad Nacional Autonoma de Mexico* **32**: 295-311.

Wahlberg N., Braby M. F., Brower A. V. Z., de Jong R., Lee M. M., Nylin S., Pierce N., Sperling F. A. H., Vila R., Warren A. D., Zakharov E. 2005. Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proceedings of the Royal entomological Society B* **272**: 1577-1586.

Weiss J. C. 1991-2005. ***The Parnassiinae of the world***. Part I 1991, II 1992, III 1999, IV 2005. Sciences Nat, Venette, 400 p.

Weiss J.-C., Michel F. 1989. Description de nouveaux taxa du genre *Parnassius* provenant du Tibet (Chine). *Bulletin de la Société Sciences Nat* **61**: 5-19.

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

Zakharov E. V., Caterino, M. S, Sperling F. 2004. Molecular phylogeny, historical biogeography and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera : Papilionidae). *Systematic Biology* **53**: 193-215.

Zhu B., Kidd W. S. F., Rowley D. B., Currie B. S., Shafique N. 2005. Age of Initiation of the India-Asia Collision in the East-Central Himalaya. *Journal of Geology* **113**: 265–285.

Appendix

Taxa, with locality, identification code and sequence accession numbers of specimens analyzed. Taxa with a bracketed species name are those whose specific affiliation has been in dispute in the recent literature; for these combinations, we have indicated in bold type the species name whose use we suggest based on our and other data (see Text). Bracketed accession numbers are from other, published works. When combining the ND5 data, most of which come from Omoto *et al.* (2004), with our LSU, ND1 and CO1 data, care was taken to match individuals from closely located geographic sites, whenever possible.

Taxon	Locality	Code	LSU	ND1	CO1	ND5
genus *Parnassius* subgenus *Parnassius*						
P. apollo venaissinus Frühstörfer 1921	Col du Rousset, Drôme, France	W265	AJ972017	AJ972114	AM231489	(AB095636) Mt Ventoux, France
P. apollo nevadensis Oberthür 1891	Sierra Nevada, Spain	W309	AJ972018	AJ972115	AM231490	
P. apollo graecus Ziegler 1901	Katara pass, Veluchi, Greece	UP100-C	(DQ407778)	(DQ407805)	(DQ407763)	
P. apollo graslini Oberthür 1891	Dödegül Dag, Isparta, Turkey	W35	(AJ224055)	(AJ224093)	AM231488	
P. apollo merzbacheri Frühstörfer 1906	N. Tian-Shan, Kyrgyzstan	W286	AJ972019	AJ972116	AM231491	
*P. (**phoebus**) bremeri* Bremer 1864	South Korea	UP100-21	(DQ407787)	AJ972126	AM231501	(AB095611) Middle Amur, Russia
*P. (**phoebus**) smintheus* Doubleday 1847	Wind River Range, Wyoming, USA	W91	AJ972028	AJ972125	AM231495	AB095653) Colorado, USA
*P. (**phoebus**) behrii* Edwards 1870	Mono Co, California, USA	W222	AJ972027	AJ972124	AM231494	
P. phoebus golovinus Holland 1930	Nome, Alaska, USA	W240	AJ972026	AJ972123	AM231500	

Taxon	Locality	Code	LSU	ND1	CO1	ND5
P. phoebus sacerdos Stichel 1906	Les Ayes, Hautes Alpes, France	UP100-A	(DQ407782)	(DQ407804)	(DQ407764)	
P. phoebus sacerdos	Cervières, Hautes Alpes	W95	AJ972022	AJ972119	AM231496	
P. phoebus gazeli Praviel 1936	Boréon, Alpes Maritimes, France	W100	AJ972024	AJ972121	AM231498	
P. phoebus interpositus Herz 1903	Tomtor, Yakutia, Russia	W85	AJ972025	AJ972122	AM231499	(AB095654) Magadan, Russia
P. (phoebus) ***ruckbeili*** Deckert 1909	Barkul, Xinjiang, China	W233	AJ972021	AJ972118	AM231493	(AB095638) N. Hami, Xinjiang, China
P. jacquemontii jacquemontii Boisduval 1836	Katoh, Spiti, India	UP100-3	AJ972014	AJ972111	AM231486	
P. jacquemontii variabilis Stichel 1906	Tenghisbay Pass, Alai, Kyrgyzstan	UP100-J	(DQ407784)	(DQ407810)	(DQ407766)	(AB095647) Alai, Kyrgyzstan
P. (jacquemontii) ***mercurius*** Grum-Grshimailo 1890	Daban Shan, Qinghai, China	W320	AM28309	AM28308	AM231484	
P. (jacquemontii) ***mercurius***	Songpan, Sichuan, China	W272	AJ972010	AJ972107	AM231483	AM283089
P. (jacquemontii) ***mercurius*** *actinoboloides* Bang-Haas 1928	N. Çaka, Qinghai, China	W313	AJ972015	AJ972112	AM231485	
P. nomion nominulus Staudinger 1895	Buriatia, Russia	UP100-R	(DQ407779)	AJ972108	(DQ407759)	(AB095609) Primorye, Russia
P. nomion gabrieli Bryk 1934	Datong Shan, Qinghai, China	UP100-4	(DQ407780)	AJ972109	AM231480	
P. epaphus cachemiriensis Oberthür 1891	Satpara pass, Deosai, Pakistan	UP100-15	(DQ407786)	AJ972105	(DQ407761)	
P. epaphus cachemiriensis	Hankar, Ladakh, India	W93	AJ972007	AJ972104	AM231478	(AB095610) Qilianshan, Gansu, China
P. (epaphus) ***dongalaicus*** *rikihiroi* Tytler 1926	Monda La, Tibet, China	W244	AJ972009	AJ972106	AM231482	AM283088
P. actius minutus Vérity 1910	Dolon Pass, Kyrgyzstan	UP100-E	(DQ407783)	(DQ407807)	(DQ407765)	(AB095622) Tianshan, Xinjiang, China
P. tianschanicus alexander Bryk & Eisner 1935	Dolon Pass, Kyrgyzstan	UP100-D	(DQ407785)	(DQ407806)	(DQ407767)	(AB095648) Alai, Kyrgyzstan
P. honrathi Staudinger 1882	Ghissarski Mts, Uzbekistan	UP100-Q	(DQ407788)	AJ972129	(DQ407772)	(AB096091) Uzbekistan
P. apollonius Eversmann 1847	Susamyr valley, Kyrgyzstan	UP100-G	(DQ407789)	(DQ407808)	(DQ407768)	(AB095646) Kazakhstan
subgenus *Driopa*						
P. mnemosyne dinianus Frühstörfer 1908	Montagne de Lachens, Alpes Maritimes, France	UP100-F	(DQ407790)	AJ972048	(DQ407769)	
P. mnemosyne parmenides Frühstörfer 1908	Boréon, Alpes Maritimes, France	W78	AJ971952	AJ972049	AM231417	AM283081
P. mnemosyne ochracea Austaut 1891	Tachtakaratscha Pass, Zerafschanskyi Mts, Tadjikistan	W292	AJ971953	AJ972050	AM231421	
P. mnemosyne sheljuzhkoi Bryk 1912	Hasanbeyli, Adana Prov., Turkey	W335	AM283041	AM283060	AM231418	
P. mnemosyne orientalis Vérity 1907	S. Zailijsky Mts., Kyrgyzstan	W330	AM283042	AM283061	AM231422	AM283083
P. mnemosyne gigantea Staudinger 1886	Chirchik, Chatkalski Mts., Uzbekistan	W329	AM283043	AM283062	AM231423	
P. mnemosyne angorae Bryk & Eisner 1931	Kizilcahamam, Ankara Prov., Turkey	W331	AM283044	AM283063	AM231424	
P. mnemosyne angorae	Ilgazdagi Geçidi, Kastamonu Prov., Turkey	W333	AM283045	AM283064	AM231425	
P. mnemosyne parvisi Turati 1919	Pelkofito, Grammos, Greece	W336	AM283046	AM283065	AM231426	
P. mnemosyne farsica Bang-Haas 1938	N-O Ardakan, Fars, Iran	W280	AJ971955	AJ972052	AM231420	
P. mnemosyne pseudonubilosus Vérity 1909	Likbin, Ourmia Lake, Iran	W311	AJ971954	AJ972051	AM231419	AM283082
P. ariadne ariadne Lederer 1889	Shebalino, Altai, Russia	W237	AJ971958	AJ972055	AM231429	(AB09497) Altai, Russia
*P. (**eversmanni**) felderi* Ménétriés 1855	Amur District, Russia	W251	AJ971959	AJ972056	AM231430	(AB095608) Middle Amur, Russia

Taxon	Locality	Code	LSU	ND1	CO1	ND5
P. eversmanni ssp.	unknown	W263	AJ971960	AJ972057	AM231431	(AB094971) Hokkaido, Japan
P. orleans johanna Bryk 1932	Mt Tai Bei-Shan, Shaanxi, China	W212	AJ971964	AJ972061	AM231433	(AB095623) Qinghai, China
P. clodius Ménétriés 1855	Mineral, Tehama Co., California, USA	W257	AJ971961	AJ972058	(AF170871)	(AB095624) Montana, USA
P. glacialis mikado Bryk & Eisner 1932	W. Matsumoto, Nagano Pref., Japan	W216	AJ971957	AJ972054	AM231428	AM283085
P. stubbendorfi hoenei Schweitzer 1912	Shikaoi, Hokkaido, Japan	W218	AJ971956	AJ972053	AM231427	AM283084
P. nordmanni Ménétriés 1850	Dzhubga, Krasnodar Reg.,W. Caucasus, Russia	W226	AJ971962	AJ972059	AM231432	(AB094968) Dombay, N Caucasus, Russia
subgenus *Sachaia*						
*P. (**tenedius**) arcticus* Eisner 1968	Jakutaka, Chajota, Yakutia, Russia	W217	AJ971965	AJ972062	AM231434	(AB095639) N.E.Yakutia
P. tenedius Eversmann 1851	unknown	W228	AJ971966	AJ972063	AM231435	(AB095658) Artyk, NE Yakutia, Russia
subgenus *Lingamius*						
P. hardwickii Gray 1831	Deosai, Pakistan	UP100-5	(DQ407791)	AJ972069	(DQ407770)	AB094969) E. Nepal
subgenus *Tadumia*						
P. szechenyii frivaldszkyi Bang-Haas 1928	Qilianshan, Gansu, China	UP100-7	(DQ407792)	AJ972074	(DQ407771)	(AB095642) Qinghai, China
P. acco tagalangi Bang-Haas 1927	Taglang La, Ladakh, India	UP100-25	(DQ407793)	AJ972070	AM231442	(AB095617) Karo-la, S. Tibet, China
P. hunnyngtoni Avinoff 1916	Mt. Cho-Oyu, Tibet, China	W235	AJ971974	AJ972071	AM231443	(AB095633) Karo-la, S. Tibet, China
P. huberi Paulus 1999	Tanggula pass, Tibet, China	W247	AJ971975	AJ972072	AM231444	(AB095631) Tanggulashan, C. Tibet, China
P. schultei Weiss & Michel 1989	Karo La, S. Tibet, China	W242	AJ971976	AJ972073	AM231445	(AB095619) Karo-la, Tibet, China
P. maharaja maharaja Avinoff 1916	Taglang La, Ladakh, India	UP100-22	AJ971979	AJ972076	AM231448	(AB095615) Ladakh, India
*P. (**maharaja**) labeyriei* Weiss & Michel 1989	Largeh La, Damxung, Tibet, China	W214	AJ971980	AJ972077	AM231449	(AB095614) Qinghai, China
P. cephalus irene Bryk & Eisner 1937	150 km N. Golmud, Kun Lun Shan, China	W219	AJ971978	AJ972075	AM231447	(AB095616) Qamdo, Tibet, China
subgenus *Koramius*						
*P. (**stoliczkanus**) stenosemus* Honrath 1890	Rangdum, Zanskar, India	W224	AJ971992	AJ972089	AM231461	(AB095656) Baralacha Pass, Ladakh, India
P. stoliczkanus stoliczkanus Felder & Felder 1864	Stok, Ladakh, India	W261	AJ971990	AJ972087	AM231459	(AB095650) Ladakh, India
P. stoliczkanus zojilaica Tytler 1926	Sonamarg, Kashmir, India	W267	AJ971991	AJ972088	AM231460	
P. stoliczkanus nobuko Ohya 1996	Mahakali Tata, Nepal	W256	AJ971989	AJ972086	AM231458	
P. hide meveli Weiss & Michel 1989	Largeh La, Tibet, China	W225	AJ971993	AJ972090	AM231462	(AB095613) Wenquan, Qinghai, China
P. delphius Eversmann 1843	Tash Rabat, Tian-Shan, Kyrgyzstan	UP100-11	(DQ407795)	AJ972092	DQ407762	(AB095632) Kyrgyzstan
P. delphius	Dolon Pass, Kyrgyzstan	UP100-12	(DQ407796)	AJ972093	AM231465	
*P. (**delphius**) maximinus* Staudinger 1891	Chandalash Mt., Kyrgyzstan	W252	AJ971997	AJ972094	AM231466	(AB095651) Tianshan, Xinjiang, China
P. patricius uzyngyrus D. Weiss 1979	Uzun-Gur, Kirghiz Ala-Too, Kyrgyzstan	W231	AJ971994	AJ972091	AM231463	(AB095620) Tianshan, Xinjiang, China
P. cardinal Grum-Grshimailo 1887	Baharak, Badakhshan Prov., Afghanistan	W246	AJ971998	AJ972095	AM231467	(AB095644) Tajikistan
*P. (staudingeri) **cardinal** ruth* Kotzsch 1936	Pushta-e-Daraz, S. Warsadj, Takhar Prov., Afghanistan	W334	AM283047	AM283066	AM231468	

Taxon	Locality	Code	LSU	ND1	CO1	ND5
P. (staudingeri) ***cardinal*** *kohibaba* Bang-Haas 1915	Mt. Shah Fuladi, Koh-i-Baba Mts., Afghanistan	W323	AM283048	AM283067	AM231469	
P. (staudingeri) ***cardinal*** *affinis* Peschke & Eisner 1934	Satpara Pass, Deosai, Pakistan	UP100-10	AJ971999	AJ972096	AM231472	
P. (staudingeri) ***cardinal*** *mamaievi* Bang-Haas 1915	Chalsi, Sham, Ladakh, India	W288	incomplete			
P. (staudingeri) ***cardinal*** *chitralica* Vérity 1911	Shandur Pass, Chitral, Pakistan	W249	AJ972000	AJ972097	AM231470	(AB095637) Shandur Pass, NW Pakistan
P. (staudingeri) ***cardinal*** *tytlerianus* Bryk & Eisner 1932	Ishkuman, Jasin, Chitral, Pakistan	W282	incomplete			
P. (staudingeri) ***cardinal*** *hunza* Grum-Grshimailo 1888	Kamjut Mts., Tadjikistan	W271	AJ972001	AJ972098	AM231471	
P. (staudingeri) ***jacobsoni*** Avinoff 1913	Dzelanby, S.Pamir, Tadjikistan	W278	AJ972002	AJ972099	AM231473	AM283086
P. staudingeri abramovi Bang-Haas 1915	Xaidulla, Xinjiang, China	W327		incomplete		
P. staudingeri darvasicus Avinoff 1916	Gyshkhun valley, Vanch Mts, W. Pamir, Tadjikistan	W310	AJ972003	AJ972100	AM231474	
P. staudingeri illustris Grum-Grshimailo 1888	Aram Kunghei,Transalai, Kyrgystan	UP100-M	(DQ407797)	(DQ407812)	(DQ407773)	
P. staudingeri kiritshenkoi Avinoff 1910	Petrai, Pamir, Tadjikistan	W243	AJ972005	AJ972102	AM231476	
P. staudingeri staudingeri Bang-Haas 1882	Kaltakol, W. Gissar, Uzbekistan	W275	AJ972006	AJ972103	AM231477	AM283087
subgenus *Kailasius*						
P. charltonius ella Bryk 1932	Satpara pass, Deosai, Pakistan	UP100-L	(DQ407800)	(DQ407811)	(DQ407774)	
P. charltonius sakai Eisner 1978	Sonamarg, Kashmir, India	W32	AJ971982	AJ972079	AM231451	(AB095630) Kyrgyzstan
P. loxias tashkorensis Kreuzberg 1984	Kaindy-Ketta Mt., E. Kyrgyzstan	W250	AJ971983	AJ972080	AM231452	(AB096090) Kyrgyzstan
P. inopinatus Kotzsch 1940	Djamak, N. Mardes, Afghanistan	W255	AJ971984	AJ972081	AM231453	(AB095641) Afghanistan
P. autocrator Avinoff 1913	Sarez Lake, Pamir, Tadjikistan	W245	AJ971985	AJ972082	AM231454	(AB095634) Tajikistan
P. imperator Oberthür 1883	Qilianshan, Gansu, China	UP100-6	(DQ407801)	AJ972083	(DQ407775)	(AB095612) Qinghai, China
P. (imperator) ***augustus*** Frühstörfer 1903	Shigatse, Tibet, China	W266	AJ971987	AJ972084	AM231456	(AB095645) Tibet, China
P. acdestis limitis Weiss & Michel 1989	Nyalam, Tibet, China	W220	AJ971988	AJ972085	AM231457	(AB095621) Karo-la, S. Tibet, China
subgenus *Kreizbergia*						
*P. (**simo**) simonius* Staudinger 1889	Tenghisbay Pass, Alai, Kyrgyzstan	UP100-I	(DQ407798)	(DQ407809)	(DQ407758)	(AB095649) Kyrgyzstan
P. simo Gray 1852	Anju La, S. Tibet, China	UP100-24	(DQ407799)	AJ972065	AM231437	
P. simo kangruensis Eisner & Weiss 1990	Stok, Ladakh, India	W28	AJ971969	AJ972066	AM231438	(AB095640) Ladakh, India
*P. (**simo**) boedromius* Püngeler 1901	Sary-Djaz riv., E. Kyrgyzstan	W238	AJ971970	AJ972067	AM231439	(AB095629) Tianshan, Xinjiang, China
*P. (**simo**) andreji* Eisner 1930	Daban pass, 100 km N. Xining, Qinghai, China	W253	AJ971971	AJ972068	AM231440	(AB095643) Sichuan, China
other Parnassiinae						
Hypermnestra helios Nickerl 1846	Chinaz, Uzbekistan	W64	AJ972034	AJ972131	AM231506	(AB095659) Uzbekistan
Archon apollinus apollinus Herbst 1798	unknown	W3	AJ972035	AJ972132	AM231507	(AB095661) Izmir, Turkey
Archon apollinaris Staudinger 1892	Kermanshah, Iran	W279	AJ972036	AJ972133	AM231508	
Zerynthia rumina medesicaste Hoffmannsegg 1803	Chabrières, Alpes de Haute Provence, France	UP100-S'	(DQ407802)	AJ972143	(DQ407777)	(AB095660) Aix-en-Provence, France
Zerynthia rumina africana Stichel 1907	El Ksiba, Morocco	W264	AJ972045	AJ972142	AM231514	

Taxon	Locality	Code	LSU	ND1	CO1	ND5
Zerynthia polyxena cassandra Geyer 1828	Gard, France	W31	AJ972047	AJ972144	AM231516	
Zerynthia (Allancastria) cerisyi Godart 1824	Bafa, Turkey	W36	AJ972040	AJ972137	(AF170869)	(AB095662) Izmir, Turkey
Zerynthia (Allancastria) caucasica Lederer 1864	Lake Abant, Turkey	W268	AJ972041	AJ972138	AM231510	
Zerynthia (Allancastria) cretica Rebel 1904	Crete, Greece	W262	AJ972042	AJ972139	AM231511	
Zerynthia (Allancastria) deyrollei Oberthür 1869	Road Aksaray-Golcük, Turkey	W258	AJ972043	AJ972140	AM231512	
Zerynthia (Allancastria) louristana Le Cerf 1908	Khorramabad, SW Iran	W276	AJ972044	AJ972141	AM231513	
Bhutanitis thaidina Blanchard 1871	Konkashan (Kanting), Sichuan, China	W89	AJ972038	AJ972135	AM231509	(AB026728)
Luehdorfia japonica Leech 1889	Ena, Gifu Pref., Japan	W30	AJ972037	AJ972134	(AF170867)	(AB095663) Kyoto, Japan
Sericinus montela Gray 1853	Fuchu, Tokyo Pref., Japan	W33	AJ972039	AJ972136	(AF170868)	(AB095665) Kyoto, Japan
Baronia brevicornis Salvin 1893	Morelos, Mexico	W42	AM283057	AM283076	(AF170866)	AM283090
other Papilionidae						
Iphiclides podalirius L. 1758	Drôme, France	W46	(AJ224049)	(AJ224087)	(AF170873)	(AB059546) Czech Republic
Graphium agammemnon L. 1758	Butterfly farm	W2	AM283050	AM283069	(AF170874)	(AB059511)
Papilio demoleus L. 1758	Butterfly farm	W108	(AJ224061)	(AJ224099)	(AF044000)	(AB013159)
Papilio machaon hippocrates Felder & Felder 1864	Oiso, Japan	W107	AM283051	AM283070	(AY457593)	(AB095666)
Papilio xuthus L. 1767	Uwajima, Ehime Pref., Japan	W75	(AJ224067)	(AJ224105)	(AF043999)	(AB013149)
Papilio bianor Cramer 1777	Butterfly farm	W77	(AJ224056)	(AJ224094)	(AY457572)	(AB013156)
Papilio memnon L. 1758	Butterfly farm	W83	(AJ224064)	(AJ224102)	(AY457578)	(AB013155)
Battus philenor L. 1771	Butterfly farm	W6	(AJ224048)	(AJ224086)	(AF170875)	(AB027573)
Parides neophilus olivencius Bates 1861	Satipo, Peru	W18	(AJ224065)	(AJ224103)	(AY804373)	(AB027581)
Ornithoptera priamus L. 1758	Bulolo, Papua New Guinea	W122	AM283052	AM283071	(AY919291)	(AB044656)
Euryades corethrus Boisduval 1836	unknown	W39	AM283053	AM283072	(AY804356)	(AB027576)
Pharmacophagus antenor Drury 1773	Butterfly farm	W1	AM283054	AM283073	(AY919288)	(AB027582)
Atrophaneura alcinous Klug 1736	Japan	W45	AM283055	AM283074	(AF170876)	(AB013145)
Cressida cressida Fabricius 1775	Brisbane, Australia	W67	AM283056	AM283075	AM283079	(AB027577)
non-Papilionidae						
Pieris brassicae L. 1758	Laboratory strain	WE	AM283058	AM283077	AM283080	AM283091
Libythea celtis Laicharting 1782	Blida, Algeria	W49	AM283059	AM283078	(AY090198)	(AB013163)