Electronic supplementary material

Phylogeny and paleoecology of *Polyommatus* blue butterflies show Beringia was a climate-regulated gateway to the New World

Roger Vila, Charles D. Bell, Richard Macniven, Benjamin Goldman-Huertas, Richard H. Ree, Charles R. Marshall, Zsolt Bálint, Kurt Johnson, Dubi Benyamini, Naomi E. Pierce

Contents

List of Figures

List of Tables

Supplementary methods

Taxon sampling

To determine the phylogenetic placement of the Old and New World taxa in the section *Polyommatus*, we first performed an analysis at the tribal level, including selected ingroup taxa representing different geographic regions, and at least one representative of each section within Polyommatini *sensu* Eliot[1] as outgroups. Only the monotypic *Callictita* section, which occurs in Papua-New Guinea and is most likely not closely related to our group of interest, was not available for the study. The *Cupidopsis* section was not included in the analysis because it does not belong to the Polyommatini tribe according to our unpublished results. Four genera/subgenera were used for the *Everes* section, three for the *Glaucopsyche* section, and two each for the *Euchrysops*, *Leptotes* and *Lycaenopsis* sections, all of which are putatively closely related to the *Polyommatus* section. Two Lycaenesthini were used to root the tree. In summary, a total of 11 *Polyommatus* section taxa (4 from the Old World and 7 from the New World), 37 outgroup, and 2 root taxa were included in the analysis (Supplementary Table S1).

A total of 73 taxa were included as the ingroup in a more detailed analysis of the *Polyommatus* section (20 Old World and 53 New World taxa) (Supplementary Table S1). These consisted of at least one representative for each New World genus and subgenus within the *Polyommatus* section. For a complete list of New World taxa, see Lamas[2] and Opler & Warren[3]. At least three taxa were selected as representatives for each Old World genus *sensu* Bálint & Johnson[4]. Old World representatives of every genus or subgenus that also occurs in the New World were sampled, taking care to include all Old World genera and subgenera that had been hypothesized to be closely related to New World taxa (e.g. *Chilades* and *Aricia*, as suggested in Bálint & Johnson[4]).

Four different genera/subgenera belonging to the *Everes* section, which was the sister section of *Polyommatus* according to the results of the Polyommatini phylogeny (Supplementary Fig. S2, Supplementary Table S6), were included as outgroups, and *Leptotes trigemmatus* was used to root the tree. The specimens used in this study are listed in Supplementary Table S1. All samples are deposited in the DNA and Tissues Collection of the Museum of Comparative Zoology (Harvard University, Cambridge, MA, USA).

DNA extraction and sequencing

DNA extractions were performed using the DNeasy™ Tissue Kit (Qiagen Inc., Valencia, CA) following the manufacturer's protocols. Three mitochondrial fragments, *Cytochrome Oxidase subunit I* (*COI*), *leu-tRNA*, and *Cytochrome Oxidase subunit II* (*COII*), and six nuclear fragments, *Elongation Factor -1*α (*EF-1*α), *28S ribosome unit* (*28S*), *Histone H3* (*H3*), *wingless* (*wg*), *carbamoyl-phosphate synthetase 2/aspartate transcarbamylase/dihydroorotase* (*CAD*), and *internal transcribed spacer 2* (*ITS-2*), were used to reconstruct the phylogeny of the *Polyommatus* section. Published and/or optimized primers were used for the amplifications (Table S2). PCR was carried out in 25μL reactions using a DNA Engine™ thermal cycler (MJ Research Inc.), and typically contained 0.5μM of each primer, 0.8 mM dNTPs, 1X Qiagen PCR buffer with additional $MgCl_2$ to a final concentration of 2mM and 1.25 units Qiagen Taq DNA polymerase (Valencia, CA, USA). All reactions were initially denatured at 94°C for two minutes in a MJ Dyad Thermal Cycler (MJ Research, Waltham, MA),

then subjected to 35 cycles of 60s at 94°C denaturation, 60s at 45°C - 56°C (annealing temperature depended on marker amplified) for annealing, and 90s at 72°C extension. After amplification, the double stranded DNA was purified using QIAquick PCR purification kits (Qiagen) prior to direct sequencing in a 3100 Genetic Analyzer (Applied Biosystems/Hitachi). All sequencing was done using dye terminator cycle sequencing following the protocol specified by the ABI PRISM® Dye Terminator Cycle Sequencing Ready Reaction Kit (Revision B, August 1995, Perkin-Elmer, Norwalk, CT). Primers used for amplification served as sequencing primers. Additional internal primers were designed for sequencing purposes (Table S2) to provide overlapping sequence coverage for the entire region for selected markers. All samples were sequenced in both directions. Cycle sequencing reactions were performed in 12 μL reactions: 1.5 μL ABI PRISM® BigDye™ v3.1 (Applied Biosystems Inc., Foster City, CA), 1.0 μL 5x buffer (buffer: 400 mM Tris at pH 9.0 and 10mM $MgCl_2$), and 0.33 μL each (10 μM) primer. The remainder of the mixture was composed of ultra pure water and template to give 50-90 ng of template DNA in each reaction. Typical cycle sequence reaction parameters contained an initial denaturing step of 94°C for 2 min, followed by 25 cycles of 10s at 94°C denaturation, 5s at the annealing (temperature varied for different markers) and 4 min at 60°C (MJ Dyad Thermal Cycler, MJ Research, Waltham, MA). Annealing temperatures were: 44°C for mitochondrial markers, 51-52°C for *EF-1α*, touchdown 48°C to 38°C (20 cycles) + 50°C (20 cycles) for *28S*, touchdown 46°C to 36°C (20 cycles) + 48°C (20 cycles) for *H3*, 54-55°C for *wg*, touchdown 50°C to 40°C (20 cycles) + 48°C (20 cycles) for *CAD*, and 47°C for *ITS-2*.

Sequence alignments and characteristics

Mitochondrial and nuclear sequences were edited and aligned using Sequencher 4.2 (Genecodes Corporation, Ann Arbor, MI). Alignments were unambiguous for protein-coding genes. ClustalX (v. 1.83.1)[5] was used to align *28S* and *ITS-2*, and, in the case of the latter, ambiguous regions were excluded from the analyses, resulting in the shortening of the alignment from 685 bp to 419 bp. The *ITS-2* matrix used for phylogenetic analyses is available at http://www.ibe.upf-csic.es/ibe/_pdf/Vila_et_al_2010_ITS2_final.Nexus.txt. As already mentioned, two types of analyses were done, a tribal-level analysis (50 taxa data set), and a section-level analysis (78 taxa data set). For the tribal-level analysis, relationships were inferred using a total of approximately 5000 bp per specimen representing fragments from two mitochondrial, *Cytochrome Oxidase I* (*COI*) - (*leu-tRNA*) - *Cytochrome Oxidase II* (*COII*), and four nuclear markers, *Elongation Factor-1 α* (*EF-1α*), *28S ribosome unit* (*28S*), *Histone H3* (*H3*), and *wingless* (*wg*). For the section-level analysis, 1000 bp from two additional nuclear markers were added, *carbamoyl-phosphate synthetase 2/aspartate transcarbamylase/dihydroorotase* (*CAD*), and *internal transcribed spacer 2* (*ITS-2*). Primer sequences were cropped and missing data and ambiguities were designated by "N". All sequences were submitted to GenBank (GQ128446–GQ129111), although a few fragments were already published in GenBank from prior studies AF23356, AY496709, AY496732, AY496801, AY496805, AY496812, AY496817, AY496824, AY496827, AY496828, AY496835, AY496846, AY496849, AY675363, AY675364, AY675375, AY675410, AY675411, AY675422, DQ018884, DQ018885, DQ018913, DQ018914, DQ018946, DQ018947, DQ456536, DQ456617, EU919282, EU919287, EU919304). Separate sequences of each marker were concatenated into a single partitioned dataset in MacClade ver. 4.05[6].

Phylogenetic analyses

For both the 50 taxa and the 78 taxa data sets, we used a number of different criteria and methods to search for tree topologies. Maximum parsimony criterion searches were performed using PAUP* ver. 4.0b10[7]. Parsimony searches were conducted using heuristic search methods with tree bisection reconnection (TBR) branch swapping, collapse of zero-length branches, and equal weighting of all characters. The analyses were repeated 100 times with the "random addition" option to minimize problems of multiple islands of most parsimonious trees. Maximum parsimony searches were performed on the combined nuclear and mtDNA data partitions, as well as the complete concatenated data set. Additional parsimony searches were not performed on individual data sets that made up the separate markers (*wg*, *EF-1*α, *28s*, etc.). To investigate potential conflict between the data partitions from different genomes (i.e., mtDNA versus nuclear DNA), we performed the homogeneity partition test (i.e. ILD test) as implemented in PAUP*, using 1000 replicate searches using settings as above.

Maximum likelihood methods were also used to search for tree topologies. For each dataset, PORN*[8] was used to determine the appropriate evolutionary model for each partition based on likelihood values calculated with PAUP*. The Akaike Information Criterion (AIC) was used to evaluate the fit of competing models. In all the cases, the GTR + Γ model was selected as the most appropriate. Likelihood searches were performed using the software GARLI ver. 0.951[9] that uses a genetic algorithm to search tree space and optimize parameters. Searches were performed several times from a random starting tree. Additional searches were performed using RAXML ver. 2.0[10] in order to estimate individual marker trees for partition data in a likelihood framework. In addition to searches of each marker partition, two additional searches were performed: one that separated the data by marker (7 or 9 partitions, including one for *leu-tRNA*) and another by genome (2 partitions). Each partition was given its own GTR + Γ model of sequence evolution. For both parsimony and likelihood analyses, branch support was evaluated using 300 bootstrap replicates[11]. Search parameters for bootstrap tests were identical to those of individual likelihood searches.

In addition to maximum parsimony and maximum likelihood, tree topologies were also inferred by Bayesian methods using Markov Chain Monte Carlo (MCMC) techniques to sample the posterior distribution of trees. We employed a variety of different mixed model approaches in our Bayesian analyses: 1) we assumed a single common model across all molecular data sets (one partition), a separate model for each genome (2 partitions), and a separate model for each marker that we sequenced (7 or 9 partitions, including one for *leu-tRNA*). In each case, the underlying model was a GTR + Γ model of sequence evolution (based on AIC, see above). For parameters across partitions, we unlinked the substitution rates, character state frequencies, gamma shape parameter alpha (α), and portion of invariable sites among partitions. All other parameters (i.e, priors) were left at their default values. Posterior probabilities were calculated using the resulting trees from both runs. Bayesian analyses were conducted with the MPI-enabled version of MrBayes ver. 3.1.2[12,13], splitting runs and chains across processors. Each analysis consisted of six independent runs, with four chains (one cold and three hot) each. Initial runs were set for 10 million generations, but these indicated that five million was more than sufficient to achieve good convergence among runs (standard deviation of split values below 0.01). Subsequent runs were then done with five

million generations. Chains were sampled every 100 generations, and burnin was determined based on visual inspection of log-likelihood over time plots using Tracer ver. 1.3[14].

Ancestral area reconstruction

We used the computer software DIVA ver. 1.1[15,16] and an improved version of Lagrange[17,18] to estimate ancestral areas and dispersals within the ingroup. DIVA reconstructs ancestral areas by minimizing the number of dispersal and extinction events needed to explain a given distribution pattern. Lagrange is a biogeographical model-based ML inference method that takes into account branch lengths. In analyses of the entire clade, we coded areas as Africa (South of Sahara desert), Australia, Central America-Caribbean, East Nearctic, East Palearctic (east of the Urals and Caspian, and north of the Himalayas), Northern South America (North of the border between Peru and Ecuador), Oriental, Southern South America, West Nearctic, and West Palearctic. The biogeographic regions were coded (Table S3) based on the genus distribution range of the terminals, except in the case of genera with more than one representative in the analysis, which were coded based on the species distribution. In the case of Holarctic species (*Agriades glandon*, *Lycaeides idas* and *Vacciniina optilete*), the specimens were scored according to whether they were collected in the New or Old World. The presence of a taxon within the area was considered even when it was rather marginal (e.g. *Cyclargus* in the East Nearctic or *Lysandra* in the East Palearctic). The outgroup taxa *Leptotes* and *Everes*, which are practically cosmopolitan and are thus uninformative, were removed from the analysis. The ranges of *Freyeria trochylus* and *Tongeia*, which are distributed in more than three regions, were limited to the two areas where they were most widely distributed. We coded the distribution of the three taxa within *Hemiargus hanno* sensu lato to cover all the known distribution of the complex, even though the actual distribution limits of these taxa are unclear. DIVA analyses were performed on Bayesian and GARLI-ML trees estimated from the combined data set. The option “maxareas” was set to 2, given that when the analyses were run without constraining the maximum number of areas, all nodes had at least one most parsimonious ancestral distribution of less than three simultaneous areas.

Analyses with Lagrange were performed on a Bayesian ultrametric tree estimated from the combined data set, and on a GARLI ultrametric tree estimated from the *COI* dataset. In the *COI* tree, some polytomies existed that had to be transformed to very short branches. All possible area combinations with a maximum of two simultaneous areas were permitted. The root node was fixed at the Oriental region based on DIVA results (Supplementary Fig. S5). Since the phylogeny encompasses approximately the last 20 MY according to our molecular clock estimates, the biogeographic model used was constant through time (Supplementary Fig. S1). Dispersals between neighbouring areas were permitted bidirectionally, including dispersals through the Panama Isthmus, the North Atlantic and Beringia, even if these did not always represent a land bridge. Direct dispersals between Africa – East Palearctic and between Oriental – West Palearctic were permitted with lower probability (weight = 0.1). The following matrix of weights of dispersal events between areas was used:

Africa	[1, 0, 0, 0, 1, 1, 0.1, 0, 0, 0]
Australia	[0, 1, 0, 0, 1, 0, 0, 0, 0, 0]
West Nearctic	[0, 0, 1, 1, 0, 0, 1, 1, 0, 0]
East Nearctic	[0, 0, 1, 1, 0, 1, 0, 1, 0, 0]
Oriental	[1, 1, 0, 0, 1, 0.1, 1, 0, 0, 0]
West Palearctic	[1, 0, 0, 1, 0.1, 1, 1, 0, 0, 0]
East Palearctic	[0.1, 0, 1, 0, 1, 1, 1, 0, 0, 0]
Central America-Caribbean	[0, 0, 1, 1, 0, 0, 0, 1, 1, 0]
Northern South America	[0, 0, 0, 0, 0, 0, 0, 1, 1, 1]
Southern South America	[0, 0, 0, 0, 0, 0, 0, 0, 1, 1]

Ancestral hostplant reconstruction

Ancestral character state analyses were performed to estimate the most probable host plant of each *Polyommatus* section New World clade ancestor. Host plant families were coded as a multistate unordered character (Supplementary Table S4) and state transitions equally weighted. Larval host plant records were obtained from multiple sources[19-26], and from personal observations by RV in the case of *Paralycaeides vapa* (14.II.2003, Chucuito, Puno, Peru, 3900 m above sea level; repeated oviposition on *Trifolium* sp.) and *Pseudolucia henyah* (26.I.2003, road CH-5, Km358, 390 m above sea level, Coquimbo, Chile; repeated oviposition on *Astragalus* sp. with white flowers). Fabaceae was treated *sensu lato* in the analyses, but the results were compared to those obtained by coding Caesalpiniaceae, Fabaceae, and Mimosaceae as different states. The analyses were performed with the program Mesquite ver. 2.6[27] with MP character optimization. All the analyses were done on both Bayesian and GARLI-ML trees estimated from the 78-taxa combined data set.

Ancestral temperature tolerance reconstruction

To test the hypothesis that the changing climatic conditions from the Miocene to the Pleistocene selected the taxa capable of crossing Beringia, we estimated the temperature tolerance of the ancestors that colonized the New World. To do this, we started by selecting for each taxon in the phylogeny several of the putatively warmest localities (at lower altitudes and laltitudes) and several of the coldest localities (at higher altitudes and laltitudes). A total of 72 taxa were assessed for this analysis. When we had several representative species for a genus in the phylogeny, we included estimates for each species. However, when we only had one representative species for the genus in the phylogeny (a placeholder), we coded the characters by pooling all the species in the genus so as not to be biased by the characteristics of a single species in the phylogeny (e.g. *Agrodiaetus, Cupido, Cyclargus, Eldoradina, Everes, Talicada, Tongeia*). In a few cases, we had to limit ourselves to a group of closely related species because the genus to which the taxa belong is not clearly defined. We chose this conservative approach in order to avoid taxon sampling effects.

For each taxon, we gathered information for as many populations as possible by consulting both literature sources and local collectors. The selection was based on a wide range of sources, including specimens from collections, papers, books and reliable web pages[19,23-25,28-42]. To the best of our knowledge, we surveyed the full distribution range for each taxon, so that the temperatures obtained represent an accurate estimate of the true tolerance ranges. From these, we selected 322 (4.5 per species on average) potentially coldest and warmest localities to include in the

analysis. In cases where the taxon of interest had an extremely restricted distribution (for example species of *Pseudolucia* and *Madeleinea* in the Andes), or the population in the most extreme environment was obvious from the distribution data, our estimates for coldest or warmest localities relied on information from only a few localities. The coldest and warmest locality for a taxon was never the same, even for taxa with very narrow distributions. In each case, the extreme localities where each taxon has been recorded are indicated by the latitude, longitude and altitude values given in Supplementary Table S5. Thus, anyone wishing to verify our estimates or recalculate values given new information regarding distributional ranges and/ or additional taxa would be able to do so using the data provided in Supplementary Table S5.

For each locality, we recorded geographic coordinates and altitude using Google Earth and exported the data in .kml format. We imported these data into ArcGIS ver. 9.3[43] and used the set of climate grids of WorldClim ver. 1.4[44] to obtain the mean annual temperature and altitude for each locality. The altitude match was used as a control, especially important in mountainous localities, and points with differences bigger than 200m where discarded. For each taxon, we selected the warmest and coldest recorded localities (with higher and lower mean annual temperature, respectively) (Supplementary Table S5). A value of 100 was added to both high and low mean annual temperatures (and substracted again after the analysis) to deal with exclusively positive values. These were coded as two ordered continuous characters and MCMC ancestral reconstructions were performed with the program BayesTraits Beta V1.1[45] on the ML-GARLI phylogram estimated from the 78-taxa combined data set. The method used is a bayesian implementation of the comparative method software Continuous[46,47]. We first tested the two available models, covariance of the two characters, and the use of *lambda*, *delta* and *kappa* parameters. We used 5 million iterations, a burn-in of 5 thousand, and a sample period of 100 in each case. The ratedev parameter was tuned in order to obtain a mean acceptance of 0.3, as suggested by the program authors. The stability of the chain was monitored by plotting the logarithm of the harmonic mean of the likelihoods with the program JMP ver. 5.1.1[48]. The significantly best model and set of parameters was used to reconstruct the ancestral character states of the nodes that involved the crossing of Beringia from the Old to the New World, according to Lagrange ancestral area reconstruction. The datadev parameter was set to 18 to obtain a "Pct Est Data taken" mean value of 0.3. The results were plotted against the mean age estimation for each of these nodes. A present-day Beringia mean annual temperature of -9°C was used, according to WorldClim data for relatively warm localities in the Beringia region.

Divergence time estimation

We used a likelihood ratio test (LRT) to test for the departure of rate constancy of molecular evolution among lineages. All likelihood values were calculated with PAUP*. In any molecular dating analysis, a calibration point, either in the form of a fossil or biogeographic event, is needed to convert inferred substitution events into absolute time. Unfortunately, neither of these kinds of external calibration points was available for the Lycaenidae used in this analysis. We were therefore forced to apply a molecular clock using published substitution rates to our inferred branch lengths (or smoothed branch lengths) to convert them to absolute time. We applied a range of substitution rates for *COI* estimated for invertebrates[49,50]: a slow rate of 6.5 x 10^{-9}substitutions/site/year, an intermediate substitution rate of 7.5 x 10^{-9} and a fast

substitution rate of 9.5 x 10^{-9}. For *COI+leu-tRNA+COII*, a substitution rate of 11.5 x 10^{-9} sub./site/year was used based on a study with heliconiine butterflies[51]. Maximum likelihood branch lengths were calculated with PAUP* for *COI* and for *COI+leu-tRNA+COII* (exact fragments as those used in the studies that estimated the substitution rates) using the maximum likelihood tree topology inferred from the combined analysis. To estimate divergence times within the ingroup, we used two different methods; a strict molecular clock and penalized likelihood (PL)[52]. Branch lengths/ substitutions per site per year under the strict molecular clock were calculated with PAUP*. The software r8s[53] was used to perform the rate smoothing procedures. An optimal smoothing parameter (λ) for penalized likelihood was determined by cross-validation. When calculating smoothing rates globally across a tree, it is usually necessary to fix the age of one of the nodes, ideally the root node. Failure to do so will often cause r8s to 'crash' during the optimization procedure. To overcome this in our data sets, we calculated a mean path length from the root node to the tips of our tree, and used the resulting mean as a fixed 'age' for the root of the tree for PL. The mean of all eight ages obtained using the different methods and rates was used as the best age estimate for each node.

Supplementary results and discussion

Phylogenetic analyses

50 taxa data set. A maximum parsimony analysis of the combined molecular data set resulted in 2 trees of 8392 steps in lengths with a CI = 0.288 a RI = 0.396 and a RC = 0.114. Maximum parsimony searches of the nuclear partition found 2 minimal length trees of 3502 steps, a CI = 0.333, a RI = 0.528, and a RC =0.176, and searches of the mitochondrial data resulted in 10 most parsimonious trees of 4768 steps and a CI = 0.263, a RI = 0.287, and a RC = 0.075. Bootstrap values for the mtDNA, nuclear, and combined data sets for various clades are presented in Table S6. For a discussion on relationships and monophyly of clades and their support please refer to "Supplementary systematic discussion".

Likelihood searches with GARLI of the combined data resulted in a single tree with a –*ln*L score of 45468.255. GARLI searches of individual data sets also recover a single tree for each marker with - *ln*L scores of 15613.143, 6826.427, 9430.5403, 2895.559, 4455.7827, and 3212.0906, for *COI*, *COII*, *EF-1*α, *H3*, *wg*, and *28S*, respectively. An analysis of all of the nuclear markers resulted in a tree with –*ln*L = 20801.106.

Searches using RAXML resulted in – *ln*L scores of 45487.418453, 44407.073237, and 44228.454 for the concatenated, 2 partition, and 7 partition data sets, respectively.

Bayesian analyses resulted in a posterior distribution of trees with harmonic means of –45552.19, -43959.97, and –44123.56 for the concatenated, 2 partition, and 7 partition data sets, respectively.

78 Taxa data set. The ILD test showed significant incongruence between the nuclear and mtDNA data partitions ($p<0.01$). A maximum parsimony analysis of the combined molecular data set resulted in 2 trees of 5693 steps in lengths with a CI = 0.360, a RI = 0.654, and a RC = 0.235. Maximum parsimony searches of the nuclear

partition found 5256 minimal length trees of 1966 steps, a CI = 0.487, a RI = 0.772, and a RC = 0.376, and searches of the mitochondrial data resulted in 16 most parsimonious trees of 3646 steps and a CI = 0.299, a RI = 0.582, and a RC = 0.174.

Likelihood searches with GARLI of the combined data resulted in a single tree with a –*ln*L score of 43322.854. GARLI searches of individual data sets also recover a single tree for each marker with - *ln*L scores of 14064.561, 6027.554, 4984.5312, 7691.5226, 1568.9735, 2214.5867, 1513.9693, and 1924.7713, for *COI*, *COII*, *CAD*, *EF-1*α, *H3*, *wg*, ITS-2 and *28S*, respectively. An analysis of all of the nuclear markers resulted in a tree with –*ln*L = 21140.911.

Searches using RAXML resulted in –*ln*L scores of 37837.968167, 36967.652627, and 36782.979750 for the concatenated, 2 partition, and 9 partition data sets, respectively.

Bayesian analyses resulted in a posterior distribution of trees with harmonic means of –37951.11, –36933.30, and –36977.26 for the concatenated, 2 partition, and 9 partition data sets, respectively.

Maximum likelihood and Bayesian support values for major clades from the 50 and 78 taxa data sets are presented in Supplementary Table S6.

Ancestral area reconstruction

DIVA[15,16] does not rely on prior assumptions about area relationships. Thus, results include all equally parsimonious possibilities, which can later be interpreted and considered based on knowledge of the system. In this case, disjunct distributions of the type Oriental-West Nearctic, Oriental-West Palearctic, West Nearctic-West Palearctic, East Nearctic-East Palearctic, East Nearctic-Northern South America, West Nearctic-Northern South America, and Central America-Southern South America were eliminated from the results when a non-disjunct distribution was equally parsimonious for a given node. Bayesian and GARLI-ML topologies are similar and the differences do not affect the DIVA reconstruction. Our analyses indicated that the most parsimonious origin of the *Polyommatus* section was in the Oriental region (Supplementary Fig. S5). The results regarding the biogeographical origin of each of the New World clades are inconclusive for the two following reasons:

1- Several equally parsimonious ancestral distributions exist.
2- Direct dispersals between unconnected Old World – New World areas are allowed. These results are sometimes the most parsimonious, but they don't provide information about the most probable steps and route followed.

Lagrange[17,18] is a biogeographical-model-based program that is more suitable to test specific biogeographical hypotheses. The conclusions obtained using both a Bayesian ultrametric tree based on the combined dataset and a GARLI-ML ultrametric tree based on the *COI* dataset are identical. Dispersal and extinction rates estimated by Lagrange are 0.09308 and 0.0147, respectively for the combined dataset. As already expected given the cosmopolitan distribution of this group, dispersal rates are high. For the *COI* dataset, these rates are higher (dispersal = 0.1607; extinction = 0.03285), probably a result of the polytomies present in this tree, which had to be transformed to very short branches to perform the analysis.

The most probable scenario according to Lagrange involves five New World colonization events through Beringia (Supplementary Fig. S6). The first one involves a long trip by the ancestor of the Neotropical group. This lineage crossed Beringia approximately 10.7 MYA (max. 15.7 MYA, min. 7.7 MYA) according to our molecular clock estimates (Supplementary Table S7), and from the Western Nearctic to Central America – Caribbean and radiating in all the Neotropics. Almost no historical trace of this dispersal remains today, as most of the ancestors went extinct in North America. Only the lineage that eventually produced *Echinargus isola* might have survived in the Western Nearctic as well as the Caribbean-Central America region. The other two species of this group that also reach the Southern Nearctic region (*Hemiargus hanno* and *Cyclargus ammon*) seem to have secondarily colonized this region from the Caribbean-Central America. The other four colonization events are the following: *Icaricia-Plebulina* clade ca. 9.3 MYA, *Lycaeides* clade ca. 2.4 MY, *Agriades glandon* ca. 1.1 MYA, *Vacciniina optilete* ca. 1.0 MYA. These lineages also followed the Beringia gateway, but they didn't extend south into the Neotropics, only to the East Nearctic in some cases. For *Lycaeides,* Lagrange results suggest at least one recent crossing of Beringia back from the New World to the Old World. The case of *Lycaeides* is quite complex[54,55] and the clarification of the clade internal relationships and biogeographic dispersals will require a specific and detailed analysis. Therefore our results indicate that taxa of the *Polyommatus* section crossed from Asia to Alaska not only before, but also well after, the formation of the Bering Strait[56] and show that the Northern Atlantic passage was not a route of colonization for these butterflies. These results are entirely consistent with Nabokov's hypothesis, and agree with a scenario in which the Northern Atlantic bridge had already disappeared when the *Polyommatus* blues colonized the New World[57].

Ancestral hostplant reconstruction

Ancestral hostplant reconstruction (Supplementary Fig. S7) based on current hostplant data (Supplementary Table S4) shows that the ancestors of the Neotropical clade, the *Icaricia-Plebulina* clade, and the *Lycaeides* clade all used species of Fabaceae as hostplants. When coding Caesalpiniaceae, Fabaceae, and Mimosaceae as different states, the ancestor of these clades is always Fabaceae *sensu stricto*. In contrast, the ancestors of New World *Vacciniina* and *Agriades* were feeding on Ericaceae and Primulaceae, respectively. Maximum Parsimony analyses on both Bayesian and GARLI-ML phylogenies of the *Polyommatus* section lead to the same conclusions. These results are in good agreement to what is known about Fabaceae paleobiogeography[58-61].

Ancestral temperature tolerance reconstruction

We tested all possible combinations of models and parameters of BayesTraits MCMC analysis to determine which one best fits the phylogram topology and temperature data. This led us to select Model B with *lambda* parameter estimation allowing for the "Test for trait correlation". The logarithm of the harmonic mean of the likelihoods under these conditions (-503.8) was significantly better than any other. This result indicates that the two studied traits (mean temperature at coldest and mean temperature at warmest locations) covary, as expected. The better fit of Model B (directional) over model A (random walk) demonstrates directionality in the evolution of temperature tolerance. This corresponds well with results obtained for ancestors that crossed Beringia, as well as for known changes in climate from the Miocene to the Pleistocene. Since estimating *delta* and *kappa* scaling parameters

does not significantly affect the fit of the model to the data, the tempo of evolution (branch lengths and overall path lengths) of the tree agrees well with the data. However, an estimated *lambda* value of 0.925 clearly improves the fit of the model, indicating that the tree topology slightly overestimates the covariance among species. The estimated value for *lambda* is close to one, which indicates that the evolution of thermal tolerances has a strong phylogenetic signal. The reconstruction analysis (Supplementary Fig. S8, Supplementary Table S8) shows that the more recent the colonization events, the more cold-adapted the ancestors that crossed Beringia. This result strongly suggests that the Beringia route has been progressively difficult for warm-adapted taxa over the last 11 million years, closely matching both paleoclimate estimates and the fossil record[62-65]. These results also indicate that Neotropical species appear to have more restricted ranges than their Nearctic relatives. It is possible that the colonizers that crossed Beringia were rather widely distributed species, able to adapt to new conditions and cross areas of relatively unsuitable habitat. Once in the New World, however, these lineages had the opportunity to diversify and produce taxa with more specialized niches, as reflected by current Neotropical descendants.

Supplementary systematic discussion

The six-marker tribal-level phylogeny is the first detailed hypothesis published for relationships in the Polyommatini (Supplementary Fig. S2, Supplementary Table S6). One of the most significant and unexpected systematic results is the strongly supported sister relationship between the *Polyommatus* section and the *Everes* section, both sister to *Leptotes* section. The close relationship between these three sections has never been proposed before using traditional morphological characters. For example, Eliot placed the sections *Euchrysops*, *Glaucopsyche* and *Lycaenopsis* as those closest to *Polyommatus*[1]. Like *Polyommatus*, the center of diversity for the *Everes* section is in the Old World, and only a few species occur in the Nearctic region, extending south into the Neotropics. It is interesting to note that both *Everes* and *Leptotes* sections include taxa with hind wing tails at vein terminus CuA2. This trait is shared by *Chilades* and *Edales*, which are sister to the rest of the *Polyommatus* section. This strongly suggests that ancestral *Polyommatus* blues were tailed, a character that has been lost in most but not all of the taxa in the section.

Within the *Polyommatus* section (Supplementary Table S6, Article Fig. 3), *Chilades* Moore, [1881] and *Edales* Swinhoe, [1910] form a clade that is sister to the rest. *Freyeria* Courvoisier, 1920 is frequently treated as a subgenus of *Chilades* by modern authors[4,19]. Our results show that *Freyeria*, which is sister to all the Holarctic taxa, cannot possibly be subsumed within *Chilades,* and deserves separate generic status. Our analysis includes one specimen of *Freyeria* from Turkey (taxon *trochylus* Freyer, 1845), and one from Australia (taxon *putli* Kollar, [1844]). The taxon *putli* has until recently been considered a subspecies of *F. trochylus*[66,67], but now most authors treat it as a good species[4,68]. In our analysis, *trochylus* and *putli* appear as sister taxa, and we estimate that they diverged ca. 6.4 MYA. This is a surprisingly old divergence, comparable to the degree of divergence estimated between some genera, and indicates that *putli* should be treated as a species. It is interesting to note that Nabokov correctly considered *Freyeria* a good genus, and *putli* a good species. He even pointed out that "*Freyeria* is less close to *Chilades* than to *Lycaeides*, its nearest ally"[69].

Our analyses show that all Neotropical taxa belong to the *Polyommatus* section, and assertions that *Itylos* Draudt, 1921 and other genera belong to other sections[4,70,71] are not supported. On the contrary, all Neotropical taxa form a well-supported monophyletic clade that is sister to the rest of the *Polyommatus* section except for *Chilades* and *Edales*. Bálint and Johnson's taxonomic hypothesis[4], which considered this group as polyphyletic, is likewise not supported. The Neotropical taxa are divided into four well-supported clades. Two of them, very probably sister clades, are formed by Andean, typically high-altitude adapted taxa that occur south of Central Colombia. These are *Eldoradina* Balletto, 1993, *Nabokovia* Hemming, 1960 and *Pseudolucia* Nabokov, 1945 on the one hand, and *Itylos*, *Madeleinea* Bálint, 1993 and *Paralycaeides* Nabokov, 1945 on the other. The other two clades are formed by lowland taxa, including all the Caribbean representatives and species occurring north of Central Colombia, plus a few with more southern distributions. One clade is formed by *Cylargus* Nabokov, 1945, *Echinargus* Nabokov, 1945 and *Hemiargus* Hübner, 1818, and the other by *Pseudochrysops* Nabokov, 1945. The position of *Pseudochrysops* with respect to the other three clades is unresolved, probably due to its early divergence and very long branch.

All the genera for which we have more than one representative taxon are monophyletic in our analyses, except for *Echinargus*, which is paraphyletic with respect to *Hemiargus*. The validity of the genera *Cyclargus*, *Echinargus* and *Hemiargus* has been debated[72]. Our results cast doubt on the validity of *Echinargus* from a phylogenetic point of view, which should probably be considered a junior subjective synonym of *Hemiargus*. A deeper study including a yet-undescribed new species of *Echinargus*[2] and more *Cyclargus* taxa would be advisable to complete the picture. The taxonomy of the *Hemiargus hanno* (Stoll, [1790]) complex is at present unclear and many taxa with uncertain status have been described. For example, many authors treat *ceraunus* as a different species than *hanno*[3,72,73]. In this study, we have followed the nomenclature of the checklist prepared by Gerardo Lamas[2]. *H. h. ceraunus* (Fabricius, 1793) (from Puerto Rico) and *H. h. bogotana* Draudt, 1921 (from Colombia) form a clade that is sister to *H. ramon* (Dognin, 1887), which has always been considered a good species, while *H. h. gyas* (Edwards, 1871) (from California and Arizona, USA) is sister to all them. Since the morphological and geographical limits of these and other taxa are not clearly defined, it would be necessary to study many samples to define the real number of species, their distibutions and natural histories. At present, we can conclude that at least three species exist within the genus *Hemiargus, sensu stricto*.

Since Nabokov's first description[74], the genus *Icaricia* Nabokov, [1945] has been frequently treated as a junior subjective synonym or as a subgenus of either *Aricia* Reichenbach, 1817[4] or *Plebejus* Kluk, 1780[3,75-77]. The situation of the monotypic genus *Plebulina* Nabokov, [1945][74] is similar, and its single species *emigdionis* Grinnel, 1905 is usually considered to belong to the genus *Plebejus*[3,4,75-77]. In all our analyses, the taxa within *Icaricia* and *Plebulina*, plus the taxon *saepiolus* Boisduval, 1852, form a Nearctic clade that is sister to all the rest of Holarctic taxa. This strongly supported result indicates that this clade is the result of a relatively old colonization of the New World that occurred ca. 9.3 MYA. Such a topology in the phylogeny is unexpected given modern taxonomic treatments of these groups, and implies that *Icarica* and *Plebulina* cannot possibly be included in *Plebejus*. Indeed, *Plebejus* is more closely related to, for example, *Polyommatus* Latreille, 1804 than it is to *Icaricia* and *Plebulina*.

Within the *Icaricia-Plebulina* clade, *Plebulina emigdionis* is sister to the rest, and *saepiolus* is either sister to the *Icaricia* clade or included within it. The taxon *saepiolus* has been almost invariably been considered, even by Nabokov, to belong to the genus *Plebejus*, although no closely allied species have ever been convincingly pointed out. Its close relatedness to *Icaricia* has only previously been recognized by Bálint and Johnson, who considered it to belong to the *icarioides*-group within their *Aricia sensu lato*[4]. Given our phylogenetic results, we propose here to reinstate the genus names *Plebulina* and *Icaricia*, and include within the latter the taxon *saepiolus*. Thus, we use in this paper the terminology *Icaricia saepiolus* **comb. nov.** Indeed, Ballmer and Pratt indicate, "In many respects, the larvae of this species [*P. saepiolus*] are similar to those of *Icaricia*"[22]. Our reasons to keep *Plebulina* as the monotypic sister genus of *Icaricia* include the fact that the larvae of the taxon *emigdionis* have five to seven instars as opposed to only four in *Icaricia*, and also feed on a different plant family, the Chenopodiaceae, as opposed to Fabaceae or Polygonaceae. As pointed out by Ballmer and Pratt[22], "This [*P. emigdionis*] is the most distinctive California member of the Polyommatinae in terms of biology and larval morphology". They also note, "it is the only one whose larvae lack a spatulate lobe on the prolegs". The position of *emigdionis* as sister to all the *Icaricia* species –for which we have a complete sampling[3]– is very well supported. Finally, the estimated age for the splitting of *emigdionis* is ca. 7.6 MY, well within the range of ages that we estimate for other widely accepted genera.

In all our analyses, species of *Lycaeides* Kluk, 1802 form a monophyletic clade sister to *Plebejus argus*. However, Nearctic *Lycaeides* appear as polyphyletic, with unexpected, yet strongly supported, sister relationships between Old and New World taxa. This result is similar to that obtained independently by other researchers[54,55] and deserves further study with more specimens. The well supported position of the clade *Plebejus* + *Lycaeides* as sister to the rest of Holarctic taxa except for *Icaricia*, restricts the use of *Plebejus* as supergenus. A number of authors consider *Agriades*, *Vacciniina*, *Plebejides* and *Plebejidea* as synonyms or subgenera of *Plebejus*[77], but our results show that all these taxa are more closely related to *Aricia* and *Polyommatus* than they are to *Plebejus*. Some examples of supports for this result: 100% Bayesian posterior probability, 80% bootstrap support for GARLI-ML, 80% for RAxML partitioned by gene, 78% for non-partitioned RAxML.

Interesting Nabokov citations

"I find it easier to give a friendly little push to some of the forms and hang my distributional horseshoes on the nail of Nome rather than postulate transoceanic land-bridges in other parts of the world."[69]

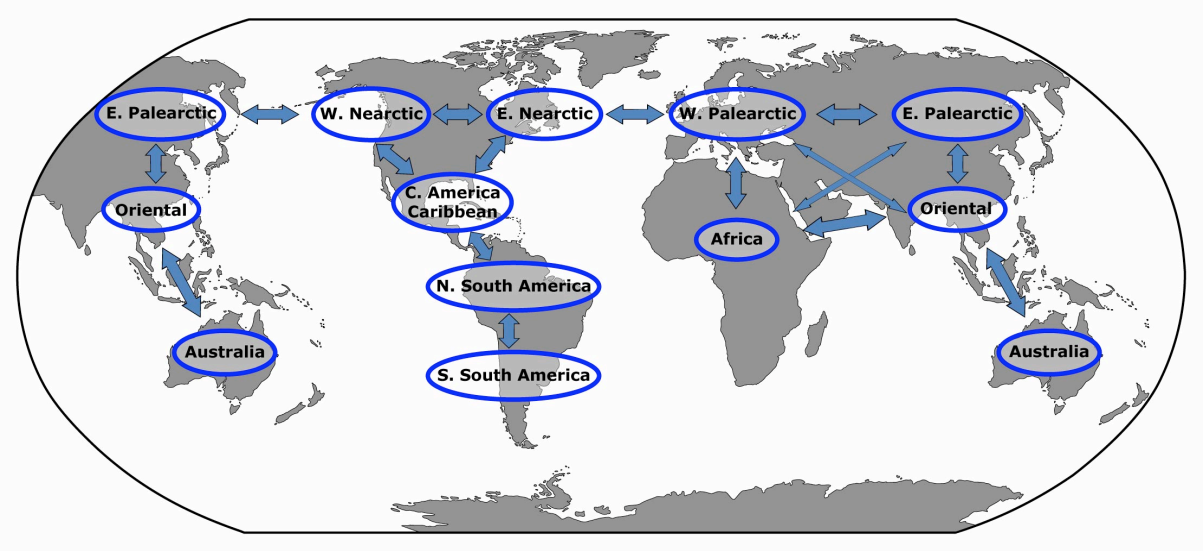
"One can assume, I think, that there was a certain point in time when both Americas were entirely devoid of *Plebejinae* but were on the very eve of receiving an invasion of them from Asia where they had been already evolved. Going back still further, a modern taxonomist straddling a Wellsian time machine with the purpose of exploring the Cenozoic era in a "downward" direction would reach a point –presumably in the early Miocene—where he still might find Asiatic butterflies classifiable on modern structural grounds as Lycaenids, but would not be able to discover among them anything definitely referable to the structural group he now diagnoses as *Plebejinae*. On his return journey, however, he would notice at some point a confuse adumbration, then a tentative "fade-in" of familiar shapes (among other, gradually

vanishing ones) and at last would find *Chilades*-like and *Aricia*-like and *Lycaeides*-like structures in the Palaearctic region.

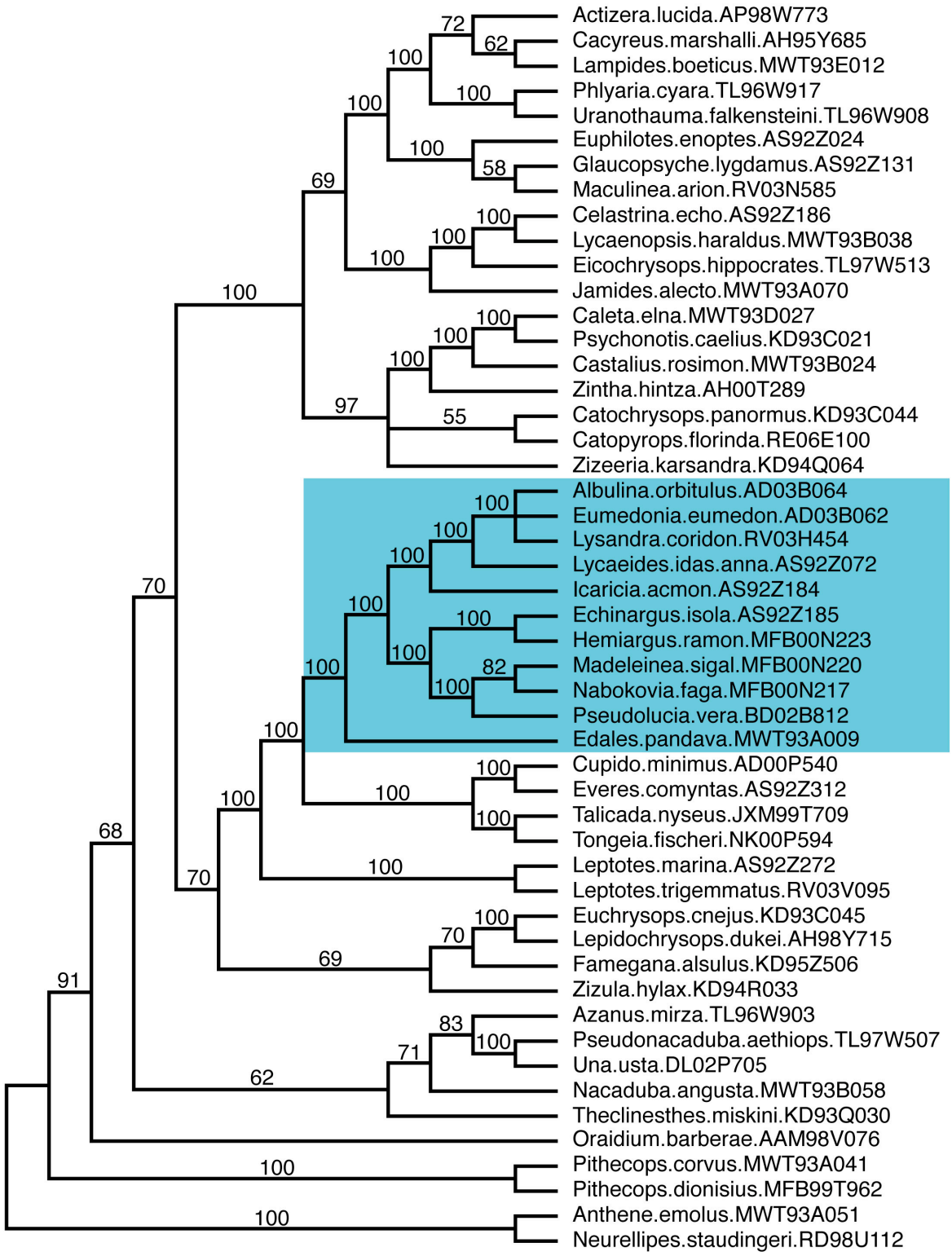
It is impossible to imagine the exact routes these forms took to reach Chile, and I have no wish to speculate on the details of their progress, beyond suggesting that throughout the evolution of *Lycaenidae* no two species ever became differentiated from each other at the same time in the same habitat (*sensu stricto*), and that the arrival of *Plebejinae* in South America preceded the arrival in North America (and differentiation from Old World ancestors) of the genera *Icaricia* and *Plebulina* (and of the species *Plebejus saepiolus*) while the latter event in its turn preceded the invasion of North America by holarctic species which came in the following sequence: *Lycaeides argyrognomon* (subsequently split), *Agriades glandon*, *Vacciniina optilete*."[69]

"Few things indeed have I known in the way of emotion or appetite, ambition or achievement, that could surpass in richness and strength the excitement of entomological exploration"[78]

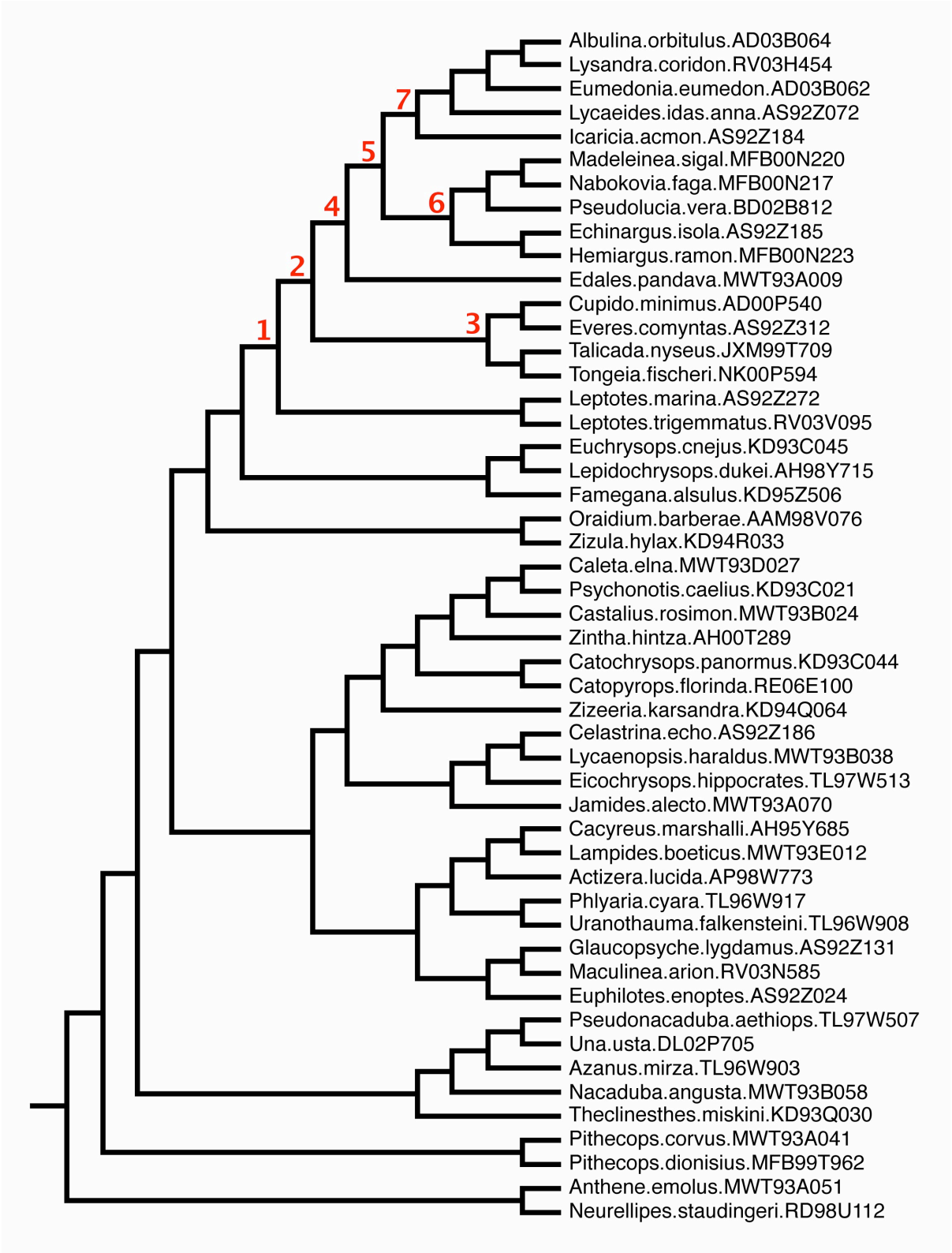
"I have hunted butterflies in various climes and disguises: as a pretty boy in knickerbockers and sailor cap; as a lanky cosmopolitan expatriate in flannel bags and beret; as a fat hatless old man in shorts."[78]



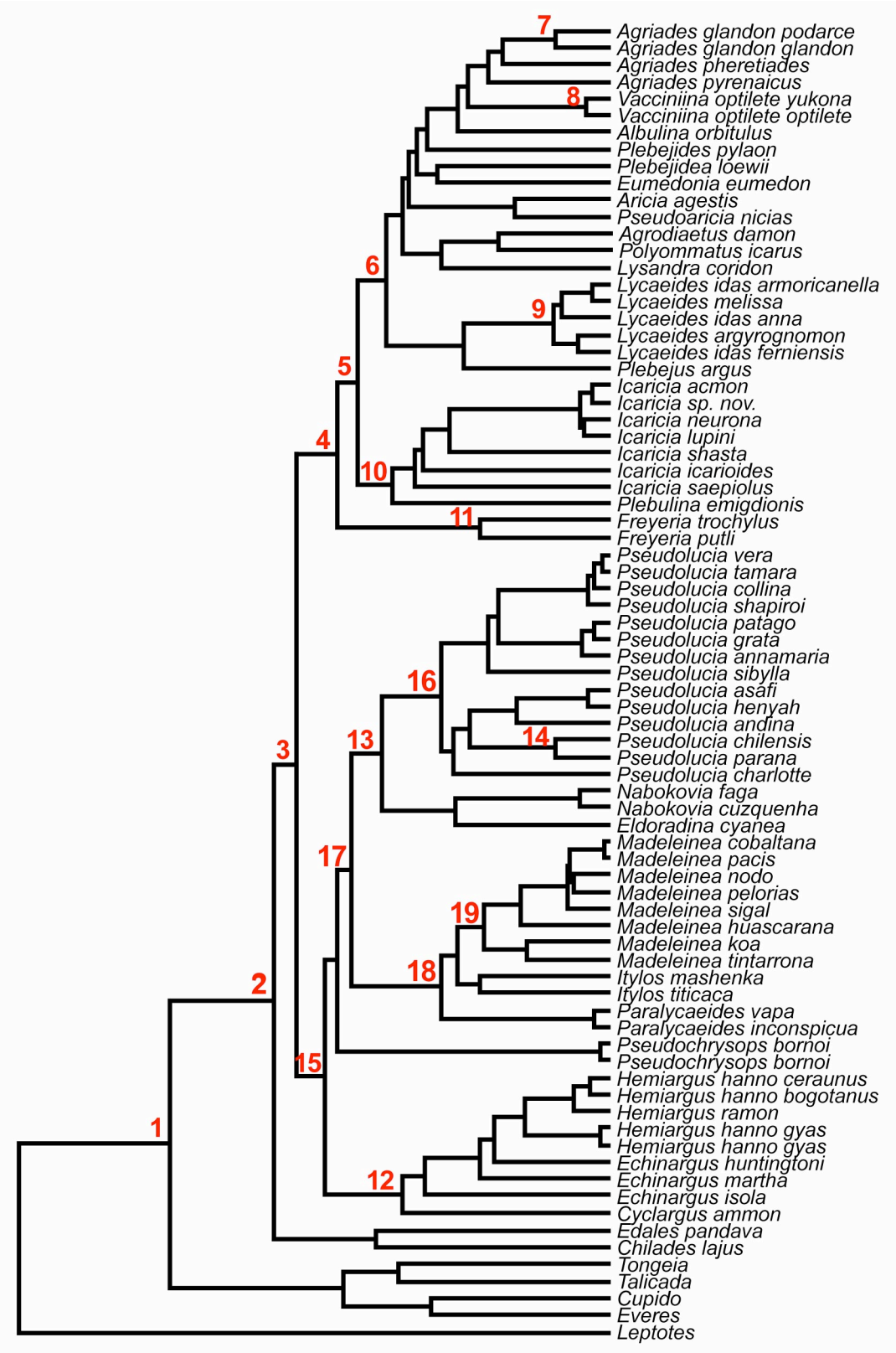
Supplementary Figure S1. Biogeographical model. Model used to infer ancestral areas with the program Lagrange[17,18]. Permitted dispersals are shown by arrows between areas. All dispersals are bidirectional.



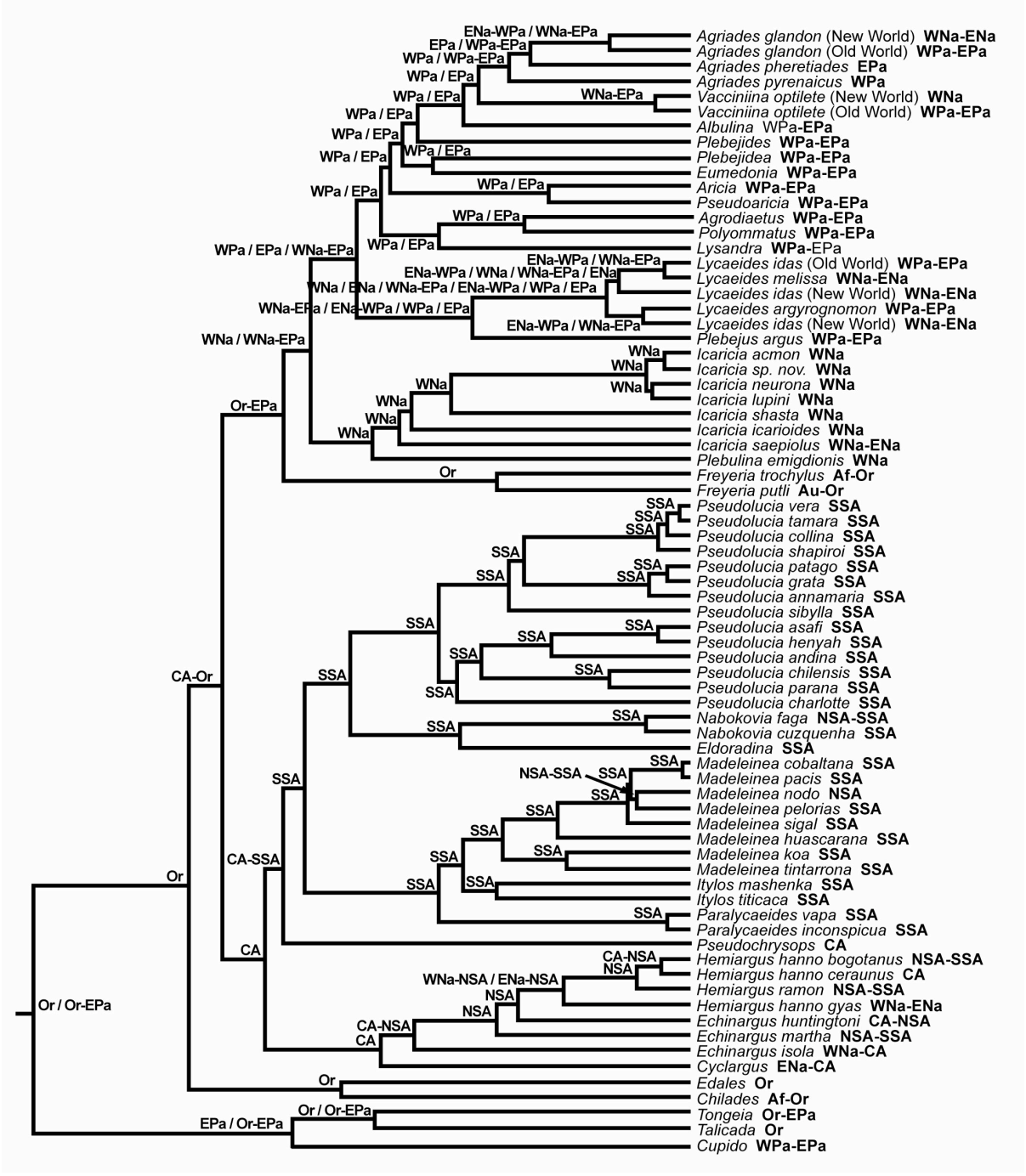
Supplementary Figure S2. Bayesian cladogram of the Polyommatini tribe. Cladogram inferred from 4939 bp of the markers *COI*-(*leu-tRNA*)-*COII*, *EF-1*α, *28S*, *H3*, and *wg* for 50 taxa. The *Polyommatus* section is highlighted in blue. The dataset was partitioned by marker and the GTR+Γ model was used, -lnL = 44123.56. Posterior probabilities (pP) are shown above recovered branches.



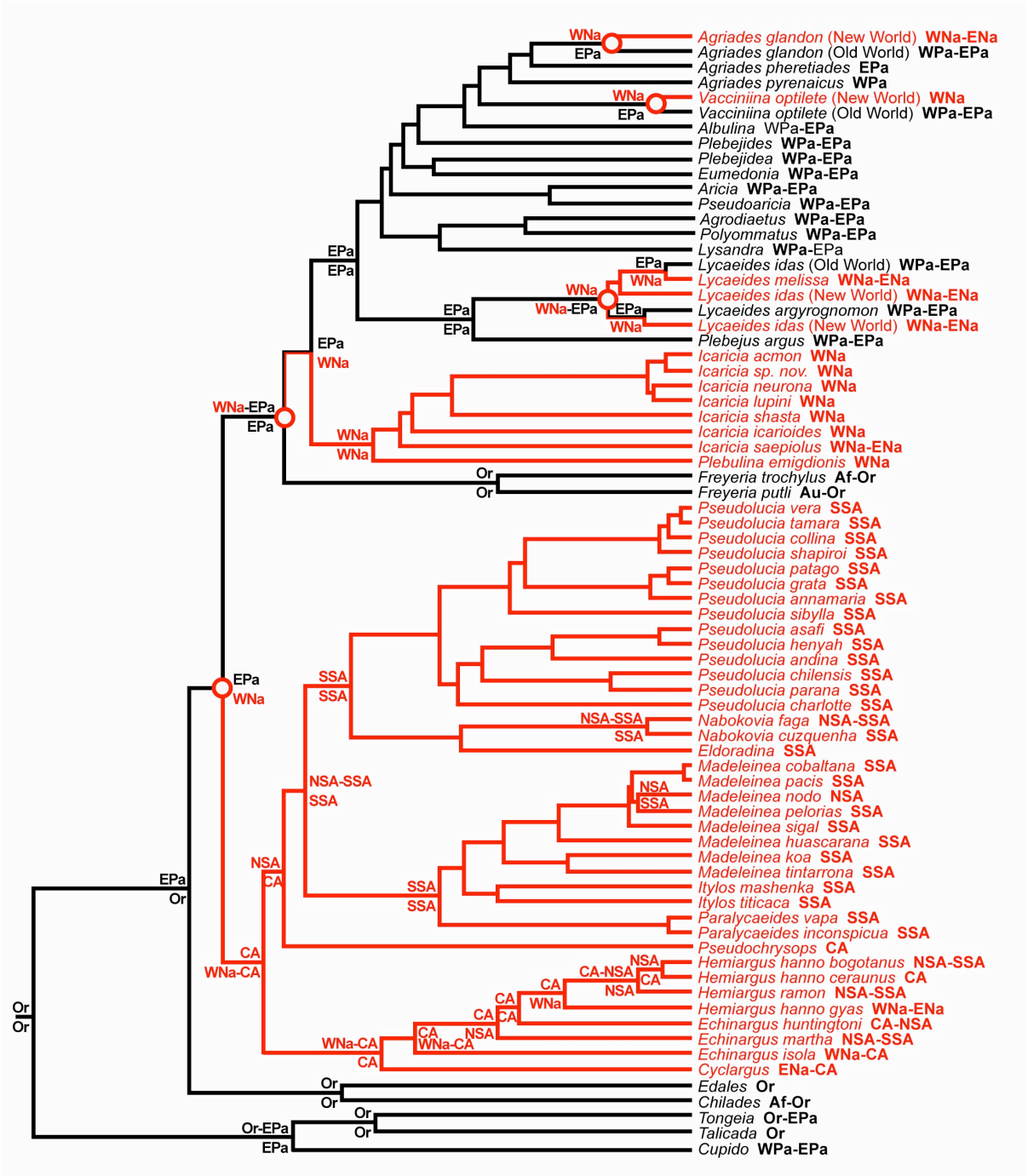
Supplementary Figure S3. Polyommatini tribe node numbers. GARLI-ML cladogram of the Polyommatini tribe with interesting nodes numbered in red corresponding to those in Supplementary Table S6 for the 50-taxa dataset.



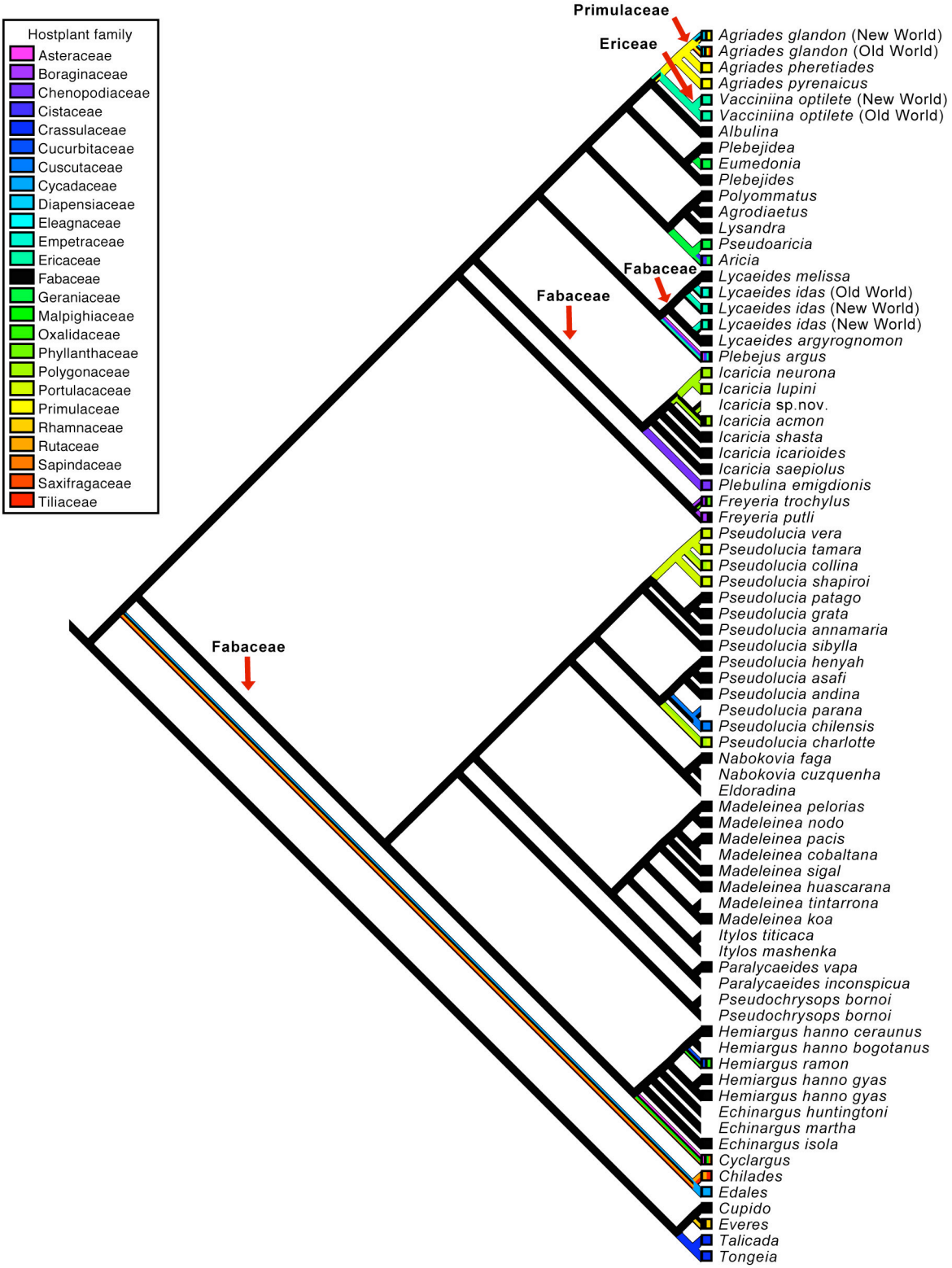
Supplementary Figure S4. *Polyommatus* section node numbers. Bayesian cladogram of the *Polyommatus* section with nodes numbered in red corresponding to those in Supplementary Tables S7 & S8 for the 78-taxa dataset.



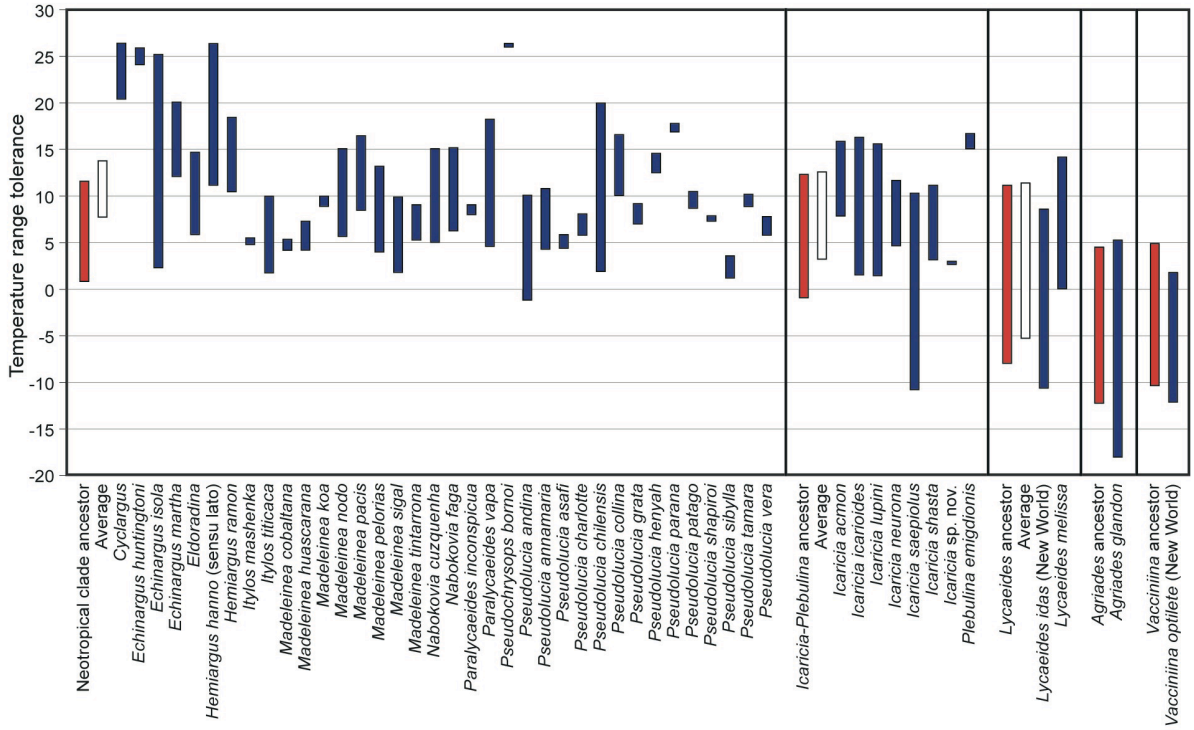
Supplementary Figure S5. DIVA ancestral area reconstruction. The most parsimonious ancestral distributions for the *Polyommatus* section are shown, except for disjunct distributions which were removed provided other non-disjunct distributions were equally parsimonious. The analysis was done with DIVA ver. 1.1[15,16] on the Bayesian tree estimated from the 78-taxa combined dataset. Using the slightly different GARLI-ML topology resulted in identical conclusions. The distribution character states used in the analysis are indicated after the taxa names. Af = Africa, Au = Australia, CA = Central America-Caribbean, ENa = East Nearctic, EPa = East Palearctic, NSA = Northern South America, Or = Oriental, SSA = Southern South America, WNa = West Nearctic, WPa = West Palearctic.



Supplementary Figure S6. Lagrange ancestral area reconstruction. Most likely ancestral range subdivision/ inheritance scenarios ('splits') at internal nodes are shown, except for invariable distributions. The analysis was done with a modified version of Lagrange[17,18] on the Bayesian tree estimated from the 78-taxa combined dataset. Using the slightly different GARLI-ML tree estimated from the 78-taxa *COI* dataset resulted in identical conclusions. Black: Old World lineages; red: New World lineages. The five New World colonization events through Beringia estimated by Lagrange are indicated with red circles. The distribution character states used in the analysis are indicated after the taxa names. Af = Africa, Au = Australia, CA = Central America-Caribbean, ENa = East Nearctic, EPa = East Palearctic, NSA = Northern South America, Or = Oriental, SSA = Southern South America, WNa = West Nearctic, WPa = West Palearctic.



Supplementary Figure S7. Ancestral hostplant reconstruction. Ancestral character state reconstruction for larval host plant families in the *Polyommatus* section. The most parsimonious scenario, requiring 41 steps, is shown. The analysis was done with Mesquite ver. 2.6 on the Bayesian tree estimated from the 78-taxa combined dataset. Host plant family was treated as a multistate unordered character, and state transitions equally weighted. Using the slightly different GARLI-ML topology resulted in identical conclusions. The host plant families of terminals and branches are indicated by colour. Estimated character states for New World clade ancestors are shown.



Supplementary Figure S8. Ancestral and current thermal range tolerances. Current mean annual temperature tolerances for all New World taxa studied (blue), as well as the average for each clade (white) and the range for their colonizing ancestor estimated using the program BayesTraits (red).

Supplementary Table S1. Samples used in this study. Taxon name, specimen label, sample accession number at MCZ and sample collection locality used in the analysis.

Section	Genus	Species (& ssp.)	Sample code	Collection locality
Ingroup (*Polyommatus* section)				
Polyommatus	*Agriades*	*glandon*	VL-05-Z994	Russia, Altai, Sailugem Range
Polyommatus	*Agriades*	*pheretiades*	NK-00-P690	Kazakhstan, Kayandy
Polyommatus	*Agriades*	*glandon podarce*	AS-92-Z130	USA, California, Leek Spring
Polyommatus	*Agriades*	*pyrenaicus dardanus*	AD-00-P259	Armenia, Gnishyk, Aiodzor Mts.
Polyommatus	*Agrodiaetus*	*damon*	MAT-99-Q841	Spain, Pyrenees, Font Llebrera
Polyommatus	*Albulina*	*orbitulus*	AD-03-B064	Russia, Altai, Aktash
Polyommatus	*Aricia*	*agestis*	NK-00-P712	Kazakhstan, Kayandy
Polyommatus	*Chilades*	*lajus*	DL-99-T242	Thailand, Thap Sakae
Polyommatus	*Cyclargus*	*ammon*	JE-01-C283	USA, Florida, Big Pine Key
Polyommatus	*Echinargus*	*huntingtoni*	RE-01-H234	Costa Rica, P.N. Santa Rosa, Guanacaste
Polyommatus	*Echinargus*	*isola*	AS-92-Z185	USA, California, Alpine, Carson River
Polyommatus	*Echinargus*	*martha*	RV-04-I212	Peru, Huánuco
Polyommatus	*Edales*	*pandava*	MWT-93-A009	Malaysia, Kepong
Polyommatus	*Eldoradina*	*cyanea*	RV-05-M735	Peru, Lima, Oyón
Polyommatus	*Eumedonia*	*eumedon*	AD-03-B062	Russia, Altai, Aktash
Polyommatus	*Freyeria*	*putli*	RE-02-A007	Australia, Queensland, Trinity Beach
Polyommatus	*Freyeria*	*trochylus*	VL-01-L462	Turkey, Artvin, Kiliçkaya
Polyommatus	*Hemiargus*	*hanno bogotanus*	SR-03-K069	Colombia, Caldas, Chinchina
Polyommatus	*Hemiargus*	*hanno ceraunus*	MH-01-I001	Puerto Rico, Culebra Island, Flamenco Beach
Polyommatus	*Hemiargus*	*hanno gyas*	AS-92-Z255	USA, California, Los Angeles, Pyramid Lake
Polyommatus	*Hemiargus*	*hanno gyas*	DL-02-P801	USA, Arizona, Chiricahua Mts.
Polyommatus	*Hemiargus*	*ramon*	MFB-00-N223	Chile, Arica, Molino
Polyommatus	*Icaricia*	*acmon*	AS-92-Z184	USA, California, Alpine, Carson River
Polyommatus	*Icaricia*	*icarioides*	AS-92-Z065	USA, California, Nevada, Donner Pass
Polyommatus	*Icaricia*	*lupini*	AS-92-Z098	USA, California, Nevada, Lang Crossing
Polyommatus	*Icaricia*	*neurona*	CCN-05-I855	USA, California, Kern, Wofford Hghts.
Polyommatus	*Icaricia*	*saepiolus*	AS-92-Z069	USA, California, Nevada, Donner Pass
Polyommatus	*Icaricia*	*shasta*	AS-92-Z465	USA, California, Nevada, Castle Peak
Polyommatus	*Icaricia*	sp. nov.	ADW-05-I828	USA, Oregon, Deschutes, Dutchman Flat
Polyommatus	*Itylos*	*mashenka*	MFB-00-N166	Peru, Junín
Polyommatus	*Itylos*	*titicaca*	MFB-00-N206	Chile, P.N. Lanca, Las Cuevas
Polyommatus	*Lycaeides*	*argyrognomon*	AD-00-P560	Russia, Tula, Tatinki
Polyommatus	*Lycaeides*	*idas anna*	AS-92-Z072	USA, California, Nevada, Donner Pass
Polyommatus	*Lycaeides*	*idas ferniensis*	NGK-02-C411	Canada, British Columbia, Castlegar
Polyommatus	*Lycaeides*	*idas armoricanella*	NK-00-P165	Russia, St. Petersburg, Luga
Polyommatus	*Lycaeides*	*melissa*	AS-92-Z005	USA, California, Nevada, Verdi
Polyommatus	*Lysandra*	*coridon*	RV-03-H454	Spain, Barcelona, El Brull
Polyommatus	*Madeleinea*	*cobaltana*	RV-03-V314	Peru, Junín, La Oroya
Polyommatus	*Madeleinea*	*huascarana*	RV-04-I403	Peru, Ancash, Pitec

Polyommatus	*Madeleinea*	*koa*	RV-03-V327	Peru, Junín, Huasahuasi
Polyommatus	*Madeleinea*	*nodo*	RV-04-I789	Ecuador, Cotopaxi, Quilotoa Lake
Polyommatus	*Madeleinea*	*pacis*	RV-03-V194	Peru, Puno, Chucuito
Polyommatus	*Madeleinea*	*pelorias*	MFB-00-N221	Chile, Socoroma
Polyommatus	*Madeleinea*	*sigal*	MFB-00-N220	Chile, Socoroma
Polyommatus	*Madeleinea*	*tintarrona*	RV-03-V182	Peru, Arequipa, Cañón del Colca
Polyommatus	*Nabokovia*	*cuzquenha*	RV-03-V234	Peru, Cuzco, Pisac
Polyommatus	*Nabokovia*	*faga*	MFB-00-N217	Chile, Socoroma
Polyommatus	*Paralycaeides*	*inconspicua*	RV-03-V188	Peru, Arequipa, Cañón del Colca
Polyommatus	*Paralycaeides*	*vapa*	RV-03-V198	Peru, Puno, Chucuito
Polyommatus	*Plebejidea*	*loewii*	AD-00-P266	Armenia, Gnishyk, Aiodzor Mts.
Polyommatus	*Plebejides*	*pylaon*	AD-00-P066	Russia, Volgograd, Kamyshinsky
Polyommatus	*Plebejus*	*argus*	NK-00-P135	Ukraine, Krim, Ai-Petri Mt.
Polyommatus	*Plebulina*	*emigdionis*	CCN-05-I856	USA, California, Kern, W. Onyx
Polyommatus	*Polyommatus*	*icarus*	NK-00-P562	Kazakhstan, Altai, Oktyabrsk
Polyommatus	*Pseudoaricia*	*nicias*	AD-03-B041	Russia, Altai, Aktash
Polyommatus	*Pseudochrysops*	*bornoi*	MAC-04-Z109	Dominican Republic, Punta Cana
Polyommatus	*Pseudochrysops*	*bornoi*	MAC-04-Z114	Dominican Republic, Punta Cana
Polyommatus	*Pseudolucia*	*andina*	BD-02-B788	Argentina, Mendoza, Valle de las Lenas
Polyommatus	*Pseudolucia*	*annamaria*	RV-03-V101	Chile, Coquimbo, Alcohuas
Polyommatus	*Pseudolucia*	*asafi*	RV-03-V020	Chile, Céspedes, Illapel
Polyommatus	*Pseudolucia*	*charlotte*	BD-02-B813	Chile, Temuco
Polyommatus	*Pseudolucia*	*chilensis*	MFB-00-N227	Chile, Farellones
Polyommatus	*Pseudolucia*	*collina*	BD-02-B796	Argentina, Neuquén, Lago Alumine
Polyommatus	*Pseudolucia*	*grata*	BD-02-B797	Argentina, Neuquén, Lago Alumine
Polyommatus	*Pseudolucia*	*henyah*	RV-03-V073	Chile, Coquimbo, Fray Jorge
Polyommatus	*Pseudolucia*	*parana*	OM-05-G417	Brazil, Parana
Polyommatus	*Pseudolucia*	*patago*	BD-02-B807	Chile, Aisen, Chile Chico
Polyommatus	*Pseudolucia*	*shapiroi*	BD-02-B792	Argentina, Mendoza, Valle de las Lenas
Polyommatus	*Pseudolucia*	*sibylla*	RV-03-V112	Chile, Coquimbo, Río La Laguna
Polyommatus	*Pseudolucia*	*tamara*	BD-02-B801	Argentina, Neuquén, Río Trafal
Polyommatus	*Pseudolucia*	*vera*	BD-02-B812	Chile, Temuco, Volcán Villarica
Polyommatus	*Vacciniina*	*optilete optilete*	VL-01-B424	Russia, St. Petersburg, Tamengont
Polyommatus	*Vacciniina*	*optilete yukona*	JB-05-I879	Canada, Yukon, km 359 Dempster Hwy.
Outgroup (other sections)				
Actizera	*Actizera*	*lucida*	AP-98-W773	Republic of South Africa, Kwazulu Natal, Hillcrest
Azanus	*Azanus*	*mirza*	TL-96-W903	Ghana, Kibi, Atewa
Brephidium	*Oraidium*	*barberae*	AAM-98-V076	Republic of South Africa, Springbok
Cacryeus	*Cacyreus*	*marshalli*	AH-95-Y685	Republic of South Africa, Capetown, Pinelands
Castalius	*Castalius*	*rosimon*	MWT-93-B024	Malaysia, Selangar
Catochrysops	*Catochrysops*	*panormus*	KD-93-C044	Australia, Queensland, Pialba
Danis	*Psychonotis*	*caelius*	KD-93-C021	Australia, Queensland, Nathan
Eicochrysops	*Eicochrysops*	*hippocrates*	TL-97-W513	Cameroon, Korup N.P.
Euchrysops	*Euchrysops*	*cnejus*	KD-93-C045	Australia, Queensland, Pialba
Euchrysops	*Lepidochrysops*	*dukei*	AH-98-Y715	Republic of South Africa, Eastern Swartberg, Blesberg Mt.
Everes	*Cupido*	*minimus*	AD-00-P540	Russia, Tula, Tatinki
Everes	*Everes*	*comyntas*	AS-92-Z312	USA, California, Davis
Everes	*Talicada*	*nyseus*	JXM-99-T709	India, Karala, Trivandrum

Everes	*Tongeia*	*fischeri*	NK-00-P594	Russia, Altai, Oktyabrsk
Famegana	*Famegana*	*alsulus*	KD-95-Z506	Australia, Queensland, Wulguru
Glaucopsyche	*Euphilotes*	*enoptes*	AS-92-Z024	USA, California, Nevada, Donner Pass
Glaucopsyche	*Glaucopsyche*	*lygdamus*	AS-92-Z131	USA, California, Leek Spring
Glaucopsyche	*Maculinea*	*arion*	RV-03-N585	Spain, Barcelona, Gombrèn
Jamides	*Jamides*	*alecto*	MWT-93-A070	Malaysia, Kepong
Lampides	*Lampides*	*boeticus*	MWT-93-E012	Malaysia, Poring Hot Spring
Leptotes	*Leptotes*	*marina*	AS-92-Z272	USA, California, Santa Barbara
Leptotes	*Leptotes*	*trigemmatus*	RV-03-V095	Chile, Coquimbo, Alcohuas
Lycaenopsis	*Celastrina*	*echo*	AS-92-Z186	USA, California, Alpine, Carson River
Lycaenopsis	*Lycaenopsis*	*haraldus*	MWT-93-B038	Malaysia, Pelindung
Nacaduba	*Nacaduba*	*angusta*	MWT-93-B058	Malaysia, Kepong
Petrelaea	*Pseudonacaduba*	*aethiops*	TL-97-W507	Cameroon, Korup N.P.
Phlyaria	*Phlyaria*	*cyara*	TL-96-W917	Ghana, Kibi, Atewa
Pithecops	*Pithecops*	*corvus*	MWT-93-A041	Malaysia, Kepong
Pithecops	*Pithecops*	*dionisius*	MFB-99-T962	Australia, Morobe, Lae, Marabi
Theclinesthes	*Theclinesthes*	*miskini*	KD-93-Q030	Australia, Queensland, Mt. Gammie
Una	*Una*	*usta*	DL-02-P705	Thailand, Chiang Mai, Doi Suthep-Pui N.P.
Upolampes	*Caleta*	*elna*	MWT-93-D027	Malaysia, Sabah, Kokol
Uranothauma	*Uranothauma*	*falkensteini*	TL-96-W908	Ghana, Kibi, Atewa
Zintha	*Zintha*	*hintza*	AH-00-T289	Republic of South Africa, Gautony, Helepoort
Zizeeria	*Zizeeria*	*karsandra*	KD-94-Q064	Australia, Queensland, Townsville, Hermit Park
Zizula	*Zizula*	*hylax*	KD-94-R033	Australia, Queensland, Inglewood
Root (Lycaenesthini)				
	Anthene	*emolus*	MWT-93-A051	Malaysia, Kepong
	Neurellipes	*staudingeri*	RD-98-U112	D. R. of the Congo, Beni

Supplementary Table S2. Primer sequences. mt: mitochondrial, n: nuclear. T = thymine, A = adenine, G = guanine, C = cytosine, K = G+T, W = A+T, M = A+C, Y = C+T, R = A+G, S = G+C, V = G+A+C, I = Inosine, N = A+C+G+T.

Primer location	Primer name	Direction	Sequence (5' to 3')
mt *COI*	LCO1490[79]	forward	GGTCAACAAATCATAAAGATATTGG
mt *COI*	Ron[80,81]	forward	GGATCACCTGATATAGCATTCCC
mt *COI*	Nancy[81]	reverse	CCCGGTAAAATTAAAATATAAACTTC
mt *COI*	Tonya[81]	forward	GAAGTTTATATTTTAATTTTACCGGG
mt *COI*	Hobbes[81]	reverse	AAATGTTGNGGRAAAAATGTTA
mt *COII*	George[81,82]	forward	ATACCTCGACGTTATTCAGA
mt *COII*	Phyllis[81,82]	reverse	GTAATAGCIGGTAARATAGTTCA
mt *COII*	Strom[81,82]	forward	TAATTTGAACTATYTTACCIGC
mt *COII*	Eva[81,82]	reverse	GAGACCATTACTTGCTTTCAGTCATCT
n *CAD*	CAD787F[83]	forward	GGDGTNACNACNGCNTGYTTYGARCC
n *CAD*	CADFa	forward	GDATGGTYGATGAAAATGTTAA
n *CAD*	CADRa	reverse	CTCATRTCGTAATCYGTRCT
n *EF-1α*	ef135[84,85]	forward	CAAATGYGGTGGTATYGACAAACG
n *EF-1α*	ef684[84,85]	reverse	TCCTTRCGCTCCACSTGCCAYCC
n *EF-1α*	ef531[84,85]	forward	TACAGYGAGCSCCGTTTYGAGGA
n *EF-1α*	ef929[84,85]	reverse	GCCTCTTGGAGAGCTTCGTGGTG
n *EF-1α*	ef51.9[84,85]	forward	CARGACGTATACAAAATCGG
n *EF-1α*	efrcM4R[84,85]	reverse	ACAGCVACKGTYTGYCTCATRTC
n *H3*	H3F[86]	forward	ATGGCTCGTACCAAGCAGACVGC
n *H3*	H3R[86]	reverse	ATATCCTTRGGCATRATRGTGAC
n *ITS-2*	ITS-3[87]	forward	GCATCGATGAAGAACGCAGC
n *ITS-2*	ITS-4[87]	reverse	TCCTCCGCTTATTGATATGC
n *wg*	LepWg1[88]	forward	GARTGYAARTGYCAYGGYATGTCTGG
n *wg*	LepWg2E	reverse	ACNACGAACATGGTCTGCGT
n *28S*	S3660[89]	forward	GAGAGTTMAASAGTACGTGAAAC
n *28S*	A335[89]	reverse	TCGGARGGAACCAGCTACTA

Supplementary Table S3. Character states for DIVA and Lagrange analyses. Taxon name and biogeographic regions in which it occurs (two regions at most). Af = Africa, Au = Australia, CA = Central America-Caribbean, ENa = East Nearctic, EPa = East Palearctic, NSA = Northern South America, Or = Oriental, SSA = Southern South America, WNa = West Nearctic, WPa = West Palearctic.

Taxon	Biogeographic Regions
Ingroup (*Polyommatus* section)	
Agriades glandon (New World)	WNa-ENa
Agriades glandon (Old World)	WPa-EPa
Agriades pheretiades	EPa
Agriades pyrenaicus	WPa
Agrodiaetus	WPa-EPa
Albulina	WPa-EPa
Aricia	WPa-EPa
Chilades	Af-Or
Cyclargus	ENa-Ca
Echinargus huntingtoni	CA-NSA
Echinargus isola	WNa-CA
Echinargus martha	NSA-SSA
Edales	Or
Eldoradina	SSA
Eumedonia	WPa-EPa
Freyeria putli	Au-Or
Freyeria trochylus	Af-Or
Hemiargus hanno bogotanus	NSA-SSA
Hemiargus hanno ceraunus	CA
Hemiargus hanno gyas	WNa-ENa
Hemiargus ramon	NSA-SSA
Icaricia acmon	WNa
Icaricia icarioides	WNa
Icaricia lupini	WNa
Icaricia neurona	WNa
Icaricia saepiolus	WNa-ENa
Icaricia shasta	WNa
Icaricia sp. nov.	WNa
Itylos mashenka	SSA
Itylos titicaca	SSA
Lycaeides argyrognomon	WPa-EPa
Lycaeides idas (New World)	WNa-Ena
Lycaeides idas (Old World)	WPa-EPa
Lycaeides melissa	WNa-ENa
Lysandra	WPa-EPa
Madeleinea cobaltana	SSA
Madeleinea huascarana	SSA
Madeleinea koa	SSA
Madeleinea nodo	NSA
Madeleinea pacis	SSA
Madeleinea pelorias	SSA
Madeleinea sigal	SSA
Madeleinea tintarrona	SSA
Nabokovia cuzquenha	SSA
Nabokovia faga	NSA-SSA
Paralycaeides inconspicua	SSA
Paralycaeides vapa	SSA

Plebejidea	WPa-EPa
Plebejides	WPa-EPa
Plebejus argus	WPa-EPa
Plebulina emigdionis	WNa
Polyommatus	WPa-EPa
Pseudoaricia	WPa-EPa
Pseudochrysops	CA
Pseudolucia andina	SSA
Pseudolucia annamaria	SSA
Pseudolucia asafi	SSA
Pseudolucia charlotte	SSA
Pseudolucia chilensis	SSA
Pseudolucia collina	SSA
Pseudolucia grata	SSA
Pseudolucia henyah	SSA
Pseudolucia parana	SSA
Pseudolucia patago	SSA
Pseudolucia shapiroi	SSA
Pseudolucia sibylla	SSA
Pseudolucia tamara	SSA
Pseudolucia vera	SSA
Vacciniina optilete (New World)	WNa
Vacciniina optilete (Old World)	WPa-EPa
Outgroup	
Cupido	WPa-EPa
Talicada	Or
Tongeia	Or-EPa

Supplementary Table S4. Character states for ancestral hostplant reconstruction. Taxon name and its larval host plant family.

Taxon	Hostplant family
Ingroup (*Polyommatus* section)	
Agriades glandon (New World)	Diapensiaceae&Fabaceae&Primulaceae
Agriades glandon (Old World)	Diapensiaceae&Fabaceae&Primulaceae&Saxifragaceae
Agriades pheretiades	Primulaceae
Agriades pyrenaicus	Primulaceae
Agrodiaetus	Fabaceae
Albulina	Fabaceae
Aricia	Cistaceae&Geraniaceae
Chilades lajus	Rutaceae&Tiliaceae
Cyclargus	Asteraceae&Fabaceae&Malpighiaceae&Sapindaceae
Echinargus huntingtoni	Unknown
Echinargus isola	Fabaceae
Echinargus martha	Unknown
Edales pandava	Cycadaceae
Eldoradina	Unknown
Eumedonia	Geraniaceae
Freyeria putli	Boraginaceae&Fabaceae
Freyeria trochylus	Boraginaceae&Fabaceae&Phyllanthaceae
Hemiargus hanno bogotana	Unknown
Hemiargus hanno ceraunus	Fabaceae
Hemiargus hanno gyas	Fabaceae
Hemiargus ramon	Cucurbitaceae&Fabaceae&Oxalidaceae
Icaricia acmon	Polygonaceae&Fabaceae
Icaricia icarioides	Fabaceae
Icaricia lupini	Polygonaceae
Icaricia neurona	Polygonaceae
Icaricia saepiolus	Fabaceae
Icaricia shasta	Fabaceae
Icaricia sp. nov.	Unknown
Itylos mashenka	Unknown
Itylos titicaca	Unknown
Lycaeides argyrognomon	Fabaceae
Lycaeides idas (New World)	Empetraceae&Ericaceae&Fabaceae
Lycaeides idas (Old World)	Eleagnaceae&Empetraceae&Ericaceae&Fabaceae
Lycaeides melissa	Fabaceae
Lysandra	Fabaceae
Madeleinea cobaltana	Unknown
Madeleinea huascarana	Fabaceae
Madeleinea koa	Fabaceae
Madeleinea nodo	Fabaceae
Madeleinea pacis	Fabaceae
Madeleinea pelorias	Fabaceae
Madeleinea sigal	Fabaceae
Madeleinea tintarrona	Unknown
Nabokovia cuzquenha	Unknown
Nabokovia faga	Fabaceae
Paralycaeides inconspicua	Unknown
Paralycaeides vapa	Fabaceae
Plebejidea	Fabaceae
Plebejides	Fabaceae
Plebejus argus	Asteraceae&Cistaceae&Ericaceae&Fabaceae

Plebulina emigdionis	Chenopodiaceae
Polyommatus	Fabaceae
Pseudoaricia	Geraniaceae
Pseudochrysops	Unknown
Pseudolucia andina	Fabaceae
Pseudolucia annamaria	Fabaceae
Pseudolucia asafi	Fabaceae
Pseudolucia charlotte	Polygonaceae&Portulacaceae
Pseudolucia chilensis	Cuscutaceae
Pseudolucia collina	Polygonaceae&Portulacaceae
Pseudolucia grata	Fabaceae
Pseudolucia henyah	Fabaceae
Pseudolucia parana	Unknown
Pseudolucia patago	Fabaceae
Pseudolucia shapiroi	Portulacaceae
Pseudolucia sibylla	Fabaceae
Pseudolucia tamara	Portulacaceae
Pseudolucia vera	Portulacaceae
Vacciniina optilete (New World)	Ericaceae
Vacciniina optilete (Old World)	Ericaceae
Outgroup (*Everes* section)	
Cupido	Fabaceae
Everes	Fabaceae&Rhamnaceae
Talicada	Crassulaceae
Tongeia	Crassulaceae

Supplementary Table S5. Character states for ancestral temperature tolerance reconstruction. Mean annual temperature, latitude, longitude and altitude of the coldest and warmest localities where each taxon occurs. Temperatures were obtained from WorldClim ver. 1.4[44].

Taxon	Locality	Mean annual temp (°C)	Latitude	Longitude	Altitude (m)
Agriades glandon (New World)	Coldest	-18.0	81°49'58"N	70°25'1"W	375
Agriades glandon (New World)	Warmest	5.3	34°00'46"N	109°30'18"W	2803
Agriades glandon (Old World)	Coldest	-13.3	70°58'01"N	179°36'41"E	115
Agriades glandon (Old World)	Warmest	2.8	42°46'41.81"N	0°25'30.30"W	1960
Agriades pheretiades	Coldest	-6.0	38°33'16"N	73°37'52"E	4600
Agriades pheretiades	Warmest	5.3	39°44'52.96"N	69°51'42.41"E	2095
Agriades pyrenaicus	Coldest	2.8	42°46'41.81"N	0°25'30.30"W	1960
Agriades pyrenaicus	Warmest	7.8	43°36'22.51"N	18°3'57.75"E	1101
Agrodiaetus	Coldest	-4.8	49°40'14"N	88°19'43"E	2610
Agrodiaetus	Warmest	16.0	29°04'14"N	56°55'32"E	1963
Albulina (orbitulus sp. group)	Coldest	-2.0	53°0'19.63"N	106°42'46.81"E	553
Albulina (orbitulus sp. group)	Warmest	3.3	44°08'25"N	07°40'29"E	2122
Aricia (agestis sp. group)	Coldest	-7.7	54°22'55"N	119°26'27"E	995
Aricia (agestis sp. group)	Warmest	20.6	33°10'30"N	35°34'30"E	80
Chilades lajus	Coldest	19.1	28°15'34"N	84°04'06"E	1200
Chilades lajus	Warmest	27.0	11°29'46"N	99°36'51"E	16
Cupido	Coldest	-13.0	62°51'56"N	155°09'46"E	776
Cupido	Warmest	17.8	36°46'54.33"N	15°2'22.24"E	24
Cyclargus	Coldest	20.4	29°34'47"N	82°10'58"W	25
Cyclargus	Warmest	26.4	18°33'01"N	68°23'10"W	11
Echinargus huntingtoni	Coldest	24.1	11°08'47"N	74°07'05"W	600
Echinargus huntingtoni	Warmest	25.9	20°41'25"N	88°36'15"W	30
Echinargus isola	Coldest	2.3	50°31'58"N	101°49'58"W	526
Echinargus isola	Warmest	25.2	18°25'51"N	99°00'33"W	960
Echinargus martha	Coldest	12.1	06°04'58"S	77°38'20"W	2980
Echinargus martha	Warmest	20.1	02°17'39"S	78°59'19"W	1200
Edales pandava	Coldest	19.1	28°15'34"N	84°04'06"E	1200
Edales pandava	Warmest	26.9	06°30'53"N	126°06'17"E	30
Eldoradina	Coldest	5.9	10°42'05"S	76°43'45"W	3857
Eldoradina	Warmest	14.7	11°54'05"S	76°43'12"W	2000
Eumedonia (*eumedon* sp. group)	Coldest	-4.1	53°26'17"N	121°57'46"E	379
Eumedonia (*eumedon* sp. group)	Warmest	11.1	36°12'23.02"N	50°45'11.64"E	2000
Everes	Coldest	-9.0	62°49'51"N	162°10'47"E	217
Everes	Warmest	26.9	06°30'53"N	126°06'17"E	30
Freyeria putli	Coldest	16.6	30°26'50"N	78°04'08"E	1600
Freyeria putli	Warmest	26.9	06°30'53"N	126°06'17"E	30
Freyeria trochylus	Coldest	9.3	40°30'11"N	73°01'50"E	1377
Freyeria trochylus	Warmest	27.1	27°27'07"N	56°33'42"E	245
Hemiargus hanno (sensu lato)	Coldest	11.2	37°47'40"N	115°19'01"W	1490
Hemiargus hanno (sensu lato)	Warmest	26.4	18°33'01"N	68°23'10"W	11
Hemiargus ramon	Coldest	10.5	18°15'14.81"S	69°40'43.93"W	3200
Hemiargus ramon	Warmest	18.5	18°29'38.03"S	70°16'33.65"W	90
Icaricia acmon	Coldest	7.9	38°43'08"N	119°44'50"W	1905
Icaricia acmon	Warmest	15.9	31°22'23"N	115°41'02"W	950
Icaricia icarioides	Coldest	1.6	51°16'29.89"N	121°55'2.11"W	1559
Icaricia icarioides	Warmest	16.3	35°46'48"N	118°26'35"W	875
Icaricia lupini	Coldest	1.5	52°27'02"N	109°04'46"W	663

Icaricia lupini	Warmest	15.6	30°38'57"N	108°29'23"W	1636
Icaricia neurona	Coldest	4.7	34°48'43"N	119°08'44"W	2680
Icaricia neurona	Warmest	11.7	35°33'17"N	118°26'24"W	1400
Icaricia saepiolus	Coldest	-10.8	69°22'58"N	132°10'1"W	33
Icaricia saepiolus	Warmest	10.3	39°18'51"N	120°39'39"W	1409
Icaricia shasta	Coldest	3.2	39°22'02"N	120°21'09"W	2700
Icaricia shasta	Warmest	11.2	36°25'04"N	115°45'53"W	1820
Icaricia sp. nov.	Coldest	2.7	44°00'09"N	121°40'04"W	1930
Icaricia sp. nov.	Warmest	3.0	42°52'38"N	122°09'33"W	2000
Itylos mashenka	Coldest	4.8	11°19'29"S	75°53'28"W	4150
Itylos mashenka	Warmest	5.5	11°22'32.19"S	75°52'56.29"W	4019
Itylos titicaca	Coldest	1.8	17°40'30"S	69°45'30"W	4500
Itylos titicaca	Warmest	10.0	8°42'23.79"S	77°52'6.53"W	3700
Lycaeides argyrognomon	Coldest	-11.4	62°28'52"N	136°30'28"E	266
Lycaeides argyrognomon	Warmest	17.1	37°19'48"N	67°13'28"E	309
Lycaeides idas (New World)	Coldest	-10.6	67°49'58"N	115°05'59"W	30
Lycaeides idas (New World)	Warmest	8.6	39°21'42"N	120°40'11"W	1687
Lycaeides idas (Old World)	Coldest	-1.9	70°01'33"N	25°02'38"E	92
Lycaeides idas (Old World)	Warmest	11.5	39°05'41"N	67°04'33"E	1181
Lycaeides melissa	Coldest	0.1	54°16'01"N	101°49'01"W	280
Lycaeides melissa	Warmest	14.2	35°21'29"N	118°13'12"W	1170
Lysandra	Coldest	3.0	54°43'56"N	56°46'38"E	133
Lysandra	Warmest	17.1	36°3'56.59"N	5°29'49.97"W	145
Madeleinea cobaltana	Coldest	4.2	11°29'6.19"S	75°53'59.55"W	4200
Madeleinea cobaltana	Warmest	5.4	11°21'53"S	75°53'05"W	4100
Madeleinea huascarana	Coldest	4.2	09°00'11"S	77°41'00"W	4273
Madeleinea huascarana	Warmest	7.3	09°30'10"S	77°26'07"W	4000
Madeleinea koa	Coldest	8.9	11°19'43.83"S	75°37'7.08"W	3590
Madeleinea koa	Warmest	10.0	13°30'15"S	71°59'51"W	3600
Madeleinea nodo	Coldest	5.7	00°37'42"S	78°41'06"W	3850
Madeleinea nodo	Warmest	15.1	00°14'25"S	78°20'28"W	2639
Madeleinea pacis	Coldest	8.5	15°27'55.23"S	69°7'27.06"W	3922
Madeleinea pacis	Warmest	16.5	16°25'56.58"S	67°37'8.56"W	2609
Madeleinea pelorias	Coldest	4.0	17°45'52"S	69°45'59"W	4300
Madeleinea pelorias	Warmest	13.2	20°11'38"S	69°17'11"W	2173
Madeleinea sigal	Coldest	1.8	17°40'30"S	69°45'30"W	4500
Madeleinea sigal	Warmest	9.9	18°12'6.66"S	69°34'31.61"W	3400
Madeleinea tintarrona	Coldest	5.3	11°32'09"S	75°54'12"W	4180
Madeleinea tintarrona	Warmest	9.1	15°36'29"S	71°52'32"W	3680
Nabokovia cuzquenha	Coldest	5.0	12°59'51"S	75°33'24"W	4063
Nabokovia cuzquenha	Warmest	15.1	14°1'20.83"S	73°12'38.86"W	2500
Nabokovia faga	Coldest	6.3	18°28'45.02"S	69°29'26.55"W	3850
Nabokovia faga	Warmest	15.2	18°31'58.03"S	69°56'54.85"W	1600
Paralycaeides inconspicua	Coldest	8.0	13°16'35.50"S	72°15'45.19"W	3830
Paralycaeides inconspicua	Warmest	9.1	11°25'26"S	75°45'27"W	3465
Paralycaeides vapa	Coldest	4.6	23°38'8.32"S	65°19'38.46"W	4419
Paralycaeides vapa	Warmest	18.2	26°53'30.09"S	65°29'54.51"W	938
Plebejidea (only *loewii*)	Coldest	5.1	43°24'27"N	41°44'05"E	1698
Plebejidea (only *loewii*)	Warmest	26.5	26°52'10"N	56°03'00"E	74
Plebejides (*pylaon* sp. group)	Coldest	-2.3	50°10'12.02"N	87°44'33.95"E	1892
Plebejides (*pylaon* sp. group)	Warmest	18.7	36°11'43.60"N	28°4'28.69"E	105
Plebejus argus	Coldest	-4.8	64°50'56.68"N	60°26'14.38"E	344
Plebejus argus	Warmest	17.8	36°51'12.78"N	6°21'6.39"W	11
Plebulina emigdionis	Coldest	15.1	36°08'03"N	117°56'59"W	1135
Plebulina emigdionis	Warmest	16.7	34°25'26"N	118°32'35"W	348

Polyommatus (*icarus* sp. group)	Coldest	-6.8	66°56'50.73"N	65°36'57.68"E	288
Polyommatus (*icarus* sp. group)	Warmest	25.6	30°53'34"N	49°25'41"E	35
Pseudoaricia (*nicias* sp. group)	Coldest	-3.4	49°54'53"N	108°14'23"E	1287
Pseudoaricia (*nicias* sp. group)	Warmest	5.3	42°46'35.34"N	0°50'15.63"E	1600
Pseudochrysops bornoi	Coldest	26.0	17°57'47"N	66°54'05"W	70
Pseudochrysops bornoi	Warmest	26.4	18°33'01"N	68°23'10"W	11
Pseudolucia andina	Coldest	-1.2	32°51'17.51"S	70°5'54.57"W	3800
Pseudolucia andina	Warmest	10.1	33°45'1.94"S	70°9'5.43"W	1800
Pseudolucia annamaria	Coldest	4.3	32°5'18.28"S	70°40'8.13"W	2850
Pseudolucia annamaria	Warmest	10.8	30°10'47.99"S	70°29'11.33"W	1450
Pseudolucia asafi	Coldest	4.4	31°15'14.52"S	70°50'18.37"W	2800
Pseudolucia asafi	Warmest	5.9	31°15'7.74"S	70°50'47.48"W	2500
Pseudolucia charlotte	Coldest	5.8	39°23'18.32"S	71°56'35.54"W	1432
Pseudolucia charlotte	Warmest	8.1	38°30'35.81"S	71°30'6.37"W	1400
Pseudolucia chilensis	Coldest	1.9	31°29'42.60"S	70°39'40.72"W	3300
Pseudolucia chilensis	Warmest	20.0	26°20'12.90"S	70°36'15.64"W	33
Pseudolucia collina	Coldest	10.1	33°45'1.94"S	70°9'5.43"W	1800
Pseudolucia collina	Warmest	16.6	34°1'49.99"S	71°5'6.35"W	200
Pseudolucia grata	Coldest	7	38°39'10.92"S	71°45'32.43"W	1400
Pseudolucia grata	Warmest	9.2	38°55'30"S	71°03'30"W	1250
Pseudolucia henyah	Coldest	12.5	33°10'30"S	71°02'30"W	1150
Pseudolucia henyah	Warmest	14.6	30°48'04"S	71°34'13"W	390
Pseudolucia parana	Coldest	16.9	25°22'52"S	51°24'40"W	1050
Pseudolucia parana	Warmest	17.8	25°14'58"S	50°00'05"W	880
Pseudolucia patago	Coldest	8.7	46°33'6.66"S	71°43'29.42"W	250
Pseudolucia patago	Warmest	10.5	46°35'36.21"S	70°19'42.41"W	400
Pseudolucia shapiroi	Coldest	7.3	35°9'31.60"S	70°3'50.78"W	2150
Pseudolucia shapiroi	Warmest	7.9	35°11'41.35"S	70°3'42.62"W	2068
Pseudolucia sibylla	Coldest	1.2	30°13'4.25"S	69°55'11.33"W	3700
Pseudolucia sibylla	Warmest	3.6	30°15'30.39"S	70°0'38.81"W	3200
Pseudolucia tamara	Coldest	8.9	40°30'24.40"S	71°10'33.96"W	847
Pseudolucia tamara	Warmest	10.2	40°37'58.67"S	70°40'13.93"W	665
Pseudolucia vera	Coldest	5.8	39°23'18.32"S	71°56'35.54"W	1432
Pseudolucia vera	Warmest	7.8	37°45'30"S	72°55'30"W	800
Talicada	Coldest	22.5	30°24'41"N	77°50'08"E	550
Talicada	Warmest	26.6	08°30'04"N	76°56'53"E	40
Tongeia	Coldest	0.0	57°10'21"N	85°04'48"E	131
Tongeia	Warmest	25.4	06°27'21"N	101°17'38"E	400
Vacciniina optilete (New World)	Coldest	-12.1	69°24'35"N	140°04'42"W	315
Vacciniina optilete (New World)	Warmest	1.8	48°37'24"N	90°52'32"W	450
Vacciniina optilete (Old World)	Coldest	-9.3	70°24'26"N	70°08'21"E	50
Vacciniina optilete (Old World)	Warmest	7.5	49°16'16"N	43°29'34"E	150

Supplementary Table S6. Support values for major clades. Maximum likelihood and parsimony bootstrap, and Bayesian posterior probability (x100) greater than 50% for major clades as calculated with GARLI (Ga), RAXML (Rx), PAUP* (MP) and MrBayes (BI). Node numbers refer to those in Supplementary Figs. S3 & S4 for the 50-taxa dataset (Polyommatini phylogeny) and the 78-taxa dataset (*Polyommatus* section phylogeny), respectively.

Node	*COI* Ga/Rx	*COII* Ga/Rx	*mtDNA* Ga/Rx/MP	*28s* Ga/Rx	*EF-1a* Ga/Rx	*H3* Ga/Rx	*wg* Ga/Rx	*ITS-2* Ga/Rx	*CAD* Ga/Rx
50-taxa dataset									
1	ns/ns	ns/ns	ns/ns/ns	ns/ns	77/84	ns/ns	ns/ns	------	------
2	ns/ns	ns/ns	ns/ns/ns	ns/ns	99/100	ns/ns	ns/ns	------	------
3	ns/61	ns/ns	80/88/86	ns/ns	100/100	62/80	ns/ns	------	------
4	ns/51	ns/ns	61/ns//ns	ns/ns	ns/ns	51/53	ns/ns	------	------
5	ns/ns	ns/ns	ns/ns/59	ns/ns	ns/ns	55/61	ns/ns	------	------
6	ns/ns	ns/ns	58/53/ns	ns/ns	ns/ns	52/56	ns/ns	------	------
7	70/80	ns/ns	80/91/73	ns/53	ns/ns	90/95	ns/ns	------	------
78-taxa dataset									
2	93/86	ns/ns	93/97/96	6969/	63/63	66/68	66/85	52/83	100/100
3	ns/ns	ns/ns	ns/ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns
4	ns/ns	ns/ns	ns/ns/ns	ns/ns	ns/ns	55/69	ns/ns	ns/ns	94/86
5	ns/ns	ns/ns	ns/ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns
6	100/56	51/58	97/100/81	77/72	62/ns	ns/ns	ns/ns	ns/ns	ns/ns
7	100/99	ns/ ns	94/98//96	81/92	99/99	95/97	96/99	85/86	ns/ns
8	ns/ns	ns/ ns	ns/ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	99/99
9	100/98	88/95	100/99/100	ns/ns	97/100	ns/ns	94/100	ns/ns	99/98
10	100/100	97/98	100/100/100	ns/ns	98/100	ns/ns	ns/ns	69/92	ns/ns
11	96/96	92/ns	100/100/100	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	100/100
12	ns/ns	ns/ns	ns/ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	89/98
13	ns/ns	ns/ns	ns/ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	90/72
14	80/88	ns/ns	79/90/82	ns/ns	82/98	56/67	52/95	ns/ns	ns/ns
15	ns/ns	ns/ns	ns/ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	57/ns
16	96/94	ns/ns	99/99/100	ns/ns	90/55	ns/ns	64/78	ns/ns	96/93
17	59/54	ns/ns	63/59/59	ns/ns	ns/ns	ns/54	ns/ns	ns/ns	ns/ns
18	98/95	ns/ns	99/100/95	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	100/100
19	100/99	97/97	100/100/100	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns	ns/ns

Node	Nuclear Ga/Rx/MP	Combined Single model Ga/Rx/MP/BI	Combined By marker Rx/BI	Combined By genome Rx/BI
50-taxa				
1	86/89/99	82/90/80/100	94/99	83/94
2	100/100/100	99/100/100/100	100/100	100/100
3	100/100/100	100/100/100/100	100/100	100/100
4	89/89/93	100/100/100/100	100/100	100/100
5	92/96/86	93/100/83/100	99/100	100/100
6	99/96/90	98/99/98/100	99/100	99/100
7	95/96/95	100/100/99/100	100/100	100/100
78 taxa				
2	100/100/100	100/100/100/100	100/100	100/100
3	72/79/63	68/67/55/99	76/100	70/99
4	ns/51/60	81/78/82/100	77/99	72/99
5	67/82/81	65/65/57/90	66/98	70/85
6	99/100/100	100/100/100/100	100/100	100/100
7	100/100/100	100/100/100/100	100/100	100/100
8	57/ns/ns	65/65/57/97	100/100	99/99
9	89/100/100	100/100/100/100	100/100	100/100
10	100/100/100	100/100/100/100	100/100	100/100
11	70/73/70	100/100/100/100	100/100	100/100
12	ns/80/72	77/74/66/99	ns/90	70/95
13	ns/ns/ns	70/57/57/94	ns/76	53/62
14	100/100/100	100/100/100/100	100/100	100/100
15	ns/65/75	52/53/56/81	ns/87	ns/92
16	100/100/100	100/100/100/100	100/100	100/100
17	ns/ns/53	82/87/72/99	84/99	84/100
18	63/68/68	99/100/100/100	100/100	100/100
19	56/53/ns	100/100/100/100	100/100	100/100

Supplementary Table S7. Divergence time estimates. Tests for molecular clock and divergence times (in million years) based on three substitution rates for *COI* (intermediate rate on top, minimum and maximum rates in parentheses below) and one substitution rate for *COI+leu-tRNA+COII*. The mean of the eight estimated ages obtained using different methods and rates for each node is shown and is taken as the best estimation. The inferred chronological order of the colonization events is also indicated. χ^2 = chi square test statistic for rate constancy among lineages, *p*= probability associated with χ^2 statistic and n-2 degrees of freedom (where n = number of taxa), PL = penalized likelihood, λ = smoothing parameter determined through cross validation. Node numbers refer to those in Supplementary Fig. S4 for the *Polyommatus* section phylogeny.

		COI	***COI + COII***		***COI***	***COI + COII***	**Mean**	**Colonization event**
Clock	χ^2	134.14		*PL*				
	p	<0.001		λ	3200			
node1		18.8 (21.7-14.8)	13.3		17.7 (20.5-14.0)	9.0	16.2	
node2		14.2 (16.4-11.2)	9.5		13.9 (16.1-11.0)	8.4	12.6	
node3		13.6 (15.7-10.7)	9.3		9.7 (11.2-7.7)	7.7	10.7	1st colonization event
node4		11.4 (13.1-9.0)	7.1		9.3 (10.7-7.3)	6.3	9.3	2nd colonization event
node5		10.2 (11.7-8.0)	7.1		8.4 (9.7-6.7)	6.3	8.5	
node6		9.6 (11.1-7.6)	7.1		6.2 (7.1-4.6)	5.9	7.4	
node7		2.1 (2.4-1.6)	1.5		0.19 (0.22-0.15)	0.83	1.1	4th colonization event
node8		1.2 (1.4-1.0)	0.55		0.62 (0.72-0.49)	2.0	1.0	5th colonization event
node9		2.4 (2.7-1.9)	1.5		1.9 (2.2-1.5)	4.7	2.4	3rd colonization event
node10		9.3 (10.7-7.3)	5.9		7.8 (9.0-6.1)	4.7	7.6	
node11		6.7 (7.7-5.3)	5.6		6.9 (8.0-5.5)	5.6	6.4	
node12		9.8 (11.3-7.7)	8.7		4.4 (5.0-3.4)	5.1	6.9	
node13		9.9 (11.4-7.8)	6.9		4.7 (5.4-3.7)	2.9	6.6	
node14		1.9 (2.2-1.5)	1.3		1.9 (2.1-1.5)	2.1	1.8	
node15		12.7 (14.7-10.0)	9.0		7.5 (8.6-5.9)	5.4	9.2	
node16		6.6 (7.6-5.2)	4.3		4.3 (4.9-3.4)	2.5	4.9	
node17		11.8 (13.6-9.3)	8.3		5.9 (6.8-4.7)	3.2	8.0	
node18		8.5 (9.8-6.7)	4.5		3.4 (3.9-2.7)	2.2	5.2	
node19		5.9 (6.9 –4.7)	4.5		2.7 (3.1-2.1)	2.2	4.0	

Supplementary Table S8. Results of ancestral temperature tolerance reconstruction. The table shows the reconstructed range of mean annual temperatures tolerated by the ancestors that crossed from the Old World to the New World and the estimated age of colonization.

Ancestor	Age of colonization (MYA)	Mean temperature at warmest location (°C)	Mean temperature at coldest location (°C)
Neotropical clade	10.7	11.6	0.8
Icaricia-Plebulina	9.3	12.3	-0.9
Lycaeides	2.4	11.2	-8.0
Agriades	1.1	4.5	-12.2
Vacciniina	1.0	4.9	-10.3
		Corr. coef.=0.790	Corr. coef.=0.985 $P<.01$

References

1 Eliot, J.N., The higher classification of the Lycaenidae (Lepidoptera): a tentative arrangement. *Bulletin of the British Museum of Natural History* 28, 371-505 (1973).

2 Lamas, G. ed., Checklist: Part 4A. Hesperioidea - Papilionoidea. (Association for Tropical Lepidoptera; Scientific Publishers, Gainesville, 2004).

3 Opler, P.A. & Warren, A.D., Butterflies of North America. 2. Scientific Names List for Butterfly Species of North America, north of Mexico. *Contributions of the C. P. Gillette Museum of Arthropod Diversity, Colorado State University*, 83 pp. (2004).

4 Bálint, Z. & Johnson, K., Reformation of the *Polyommatus* section with a taxonomic and biogeographic overview (Lepidoptera, Lycaenidae, Polyommatini). *Neue Entomologische Nachrichten* 40, 1-68 (1997).

5 Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., & Higgins, D.G., The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25, 4876-4882 (1997).

6 Maddison, W.P. & Maddison, D.R., *MacClade: Analysis of phylogeny and character evolution*. (Sinauer Associates, Inc., Sunderland, MA. USA., 1992).

7 Swofford, D.L., PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4 (Sinauer Associates, Sunderland, Massachusetts, USA, 2002).

8 Bell, C.D., PORN*: a hierarchical likelihood ratio calculator for LINUX, http://www.phylodiversity.net/pornstar. (Yale University, New Haven, Connecticut, USA, 2001).

9 Zwickl, D.J., Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation. *The University of Texas at Austin* (2006).

10 Stamatakis, A., Ludwig, T., & Meier, H., Raxml-iii: A fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* 21, 456-463 (2005).

11 Felsenstein, J., Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783-791 (1985).

12 Huelsenbeck, J.P. & Ronquist, F., MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17 (8), 754–755 (2001).

13 Ronquist, F. & Huelsenbeck, J.P., MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572-1574 (2003).

[14] Rambaut, A. & Drummond, A.J., *Tracer v1.3.* (University of Oxford, Oxford, UK, 2004).
[15] Ronquist, F., *DIVA (Dispersal-Vicariance Analysis)*, version 1.1. ed. (Uppsala University, Uppsala, 1996).
[16] Ronquist, F., Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* 46 (1), 195-203 (1997).
[17] Ree, R.H., Moore, B.R., Webb, C.O., & Donoghue, M.J., A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59, 2299-2311 (2005).
[18] Ree, R.H. & Smith, S.A., Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57 (1), 4-14 (2008).
[19] Tolman, T. & Lewington, R., *Collins Field Guide, Butterflies of Britain & Europe*. (HarperCollins Publishers Ltd, London, 1997).
[20] Benyamini, D., Synopsis of biological studies of the Chilean Polyommatini (Lepidoptera, Lycaenidae). *Reports of the Museum of Natural History. University of Wisconsin (Stevens Point)* 52, 1-51 (1995).
[21] Roque-Albelo, L., The butterflies (Papilionoidea, Hesperioidea) of the Galápagos Islands, Ecuador: Distribution, hostplants and biology. *Journal of the Lepidopterists' Society* 58 (1), 33-43 (2004).
[22] Ballmer, G.R. & Pratt, G.F., A survey of the last instar larvae of the Lycaenidae (Lepidoptera) of California. *Journal of Research on the Lepidoptera* 27 (1), 1-81 (1988).
[23] Henriksen, H.J. & Kreutzer, I., *The Butterflies of Scandinavia in Nature*. (Skandinavisk Bogforlag, Odense, Denmark, 1982).
[24] Warren, A.D., Llorente-Bousquets, J.E., Luis-Martínez, A., & Vargas-Fernández, I., Interactive Listing of Mexican Butterflies http://www.mariposasmexicanas.com, (2006).
[25] Opler, P.A., Pavulaan, H., Stanford, R.E., & Pogue, M., Butterflies and Moths of North America, http://www.butterfliesandmoths.org, (2006).
[26] Lafranchis, T., *Papillons d'Europe*. (Diatheo, Paris, France, 2007).
[27] Maddison, W.P. & Maddison, D.R., Mesquite: a modular system for evolutionary analysis. Version 2.01, http://mesquiteproject.org. (2007).
[28] Lukhtanov, V.A. & Lukhtanov, A.G., *Die Tagfalter Nordwestasiens*. (Marktleuthen, Germany, 1994).
[29] Tshikolovets, V.V., *Butterflies of Kyrgyzstan*. (Pensoft, Sofia & Moscow, 2005).
[30] Kudrna, O., The Distribution Atlas of European Butterflies. *Oedippus* 20, 1-342 (2002).
[31] Tshikolovets, V.V., Bidzilya, O.V., & Golovoskin, M.I., *The Butterflies of Transbaikal Siberia*. (Nacional Na Akademija Nank Ukrainy, Brno-Kyev, 2002).
[32] Korshunov, Y. & Gorbunov, P., *[Butterflies of the Asian Russia. A Guide]*. (Ural State University Press, Ekaterinburg, 1995).
[33] Lee, S.-M., *Butterflies of Korea*. (Commitee of Insecta Koreana, Seoul, 1982).
[34] García-Barros, E. *et al.*, *Atlas de las mariposas diurnas de la Península Ibérica e islas Baleares (Lepidoptera: Papilionoidea & Hesperioidea)*. (SEA, UAM & MEC, Zaragoza, Spain, 2004).
[35] ZipcodeZoo.com, http://zipcodezoo.com.
[36] Global Biodiversity Information Facility, http://www.gbif.org.
[37] Benyamini, D., Bálint, Z., & Johnson, K., Additions to the Diversity of the Polyommatine Genus *Madeleina* Bálint (Lepidoptera, Lycaenidae). *Reports of the*

Museum of Natural History, University of Wisconsin (Stevens Point) 47, 1-5 (1995).
[38] Benyamini, D., Bálint, Z., & Johnson, K., Two new Pseudolucia species from the High Andean Region of temperate South America with revision of the status of P. andina neuqueniensis Balint and Johnson. *Reports of the Museum of Natural History, University of Wisconsin (Stevens Point)* 48, 1-9 (1995).
[39] Benyamini, D., Bálint, Z., & Johnson, K., Recently discovered new species of Pseudolucia Nabokov (Lepidoptera, Lycaenidae) from austral South America. *Reports of the Museum of Natural History, University of Wisconsin (Stevens Point)* 53, 1-2 (1995).
[40] Benyamini, D. & Bálint, Z., Studies of life history and myrmecophily in certain Chilean Pseudolucia Nabokov (Lycaenidae). *Reports of the Museum of Natural History, University of Wisconsin (Stevens Point)* 51, 1-7 (1995).
[41] Bálint, Z. & Benyamini, D., Taxonomic notes, faunistics and species descriptions of the austral South American polyommatine lycaenid genus Pseudolucia (Lepidoptera: Lycaenidae): the chilensis and collina species-groups. *Annales Historico-Naturales Musei Nationalis Hungarici* 93, 107-149 (2001).
[42] Bálint, Z., Benyamini, D., & Johnson, K., Species descriptions and miscellaneous notes on the genus Pseudolucia (Lepidoptera: Lycaenidae). *Folia Entomologica Hungarica Rovartani Kozlemenyek* 62, 151-165 (2001).
[43] Environmental Research Institute, *ArcGIS 9: ArcGIS Spatial Analyst Tutorial*. (ESRI Press, Redlands, USA, 2006).
[44] Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., & Jarvis, A., Very high resolution interpolated climate surfaces for global land areas, http://www.worldclim.org. *International Journal of Climatology* 25, 1965-1978 (2005).
[45] Pagel, M. & Meade, A., BayesTraits, http://www.evolution.rdg.ac.uk. (2008).
[46] Pagel, M., Inferring the historical patterns of biological evolution. *Nature* 401, 877-884 (1999).
[47] Organ, C.L., Shedlock, A.M., Meade, A., Pagel, M., & Edwards, S.V., Origin of avian genome size and structure in non-avian dinosaurs. *Nature* 446, 180-184 (2007).
[48] JMP (SAS Institute Inc., Cary, North Carolina, USA, 2009).
[49] Quek, S.P., Davies, S.J., Itino, T., & Pierce, N.E., Codiversification in an ant-plant mutualism: the phylogeny of host use in *Crematogaster* (Formicidae) associates of *Macaranga* (Euphorbiaceae). *Evolution* 58 (3), 554–570 (2004).
[50] Schubart, C.D., Diesel, R., & Hedges, S.B., Rapid evolution to terrestrial life in Jamaican crabs. *Nature* 393, 363–365 (1998).
[51] Brower, A.V.Z., Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Science, USA* 85 (23), 6491–6495 (1994).
[52] Sanderson, M.J., Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution* 19 (1), 101-109 (2002).
[53] Sanderson, M.J., r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19 (2), 301-302 (2003).
[54] Gompert, Z., Fordyce, J.A., Forister, M.L., & Nice, C.C., Recent colonization and radiation of North American *Lycaeides* (*Plebejus*) inferred from mtDNA. *Molecular Phylogenetics and Evolution* 48, 461-490 (2008).

[55] Nice, C.C., Anthony, N., Gelembiuk, G., Raterman, D., & Ffrench-Constant, R., The history and geography of diversification within the butterfly genus *Lycaeides* in North America. *Mol. Ecol.* 14 (6), 1741-1754 (2005).
[56] Marincovich Jr., L. & Gladenkov, A.Y., Evidence for an early opening of the Bering Strait. *Nature* 397, 149-151 (1999).
[57] Tiffney, B.H. & Manchester, S.R., The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences* 162 (6 Suppl.), S3-S17 (2001).
[58] Lavin, M. *et al.*, Identifying tertiary radiations of Fabaceae in the Greater Antilles: Alternatives to cladistic vicariance analysis. *International Journal of Plant Sciences* 162 (6 Suppl.), S53-S76 (2001).
[59] Lavin, M., Herendeen, P.S., & Wojciechowski, M.F., Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology* 54 (4), 575-594 (2005).
[60] Ellison, N.W., Liston, A., Steiner, J.J., Williams, W.M., & Taylor, N.L., Molecular phylogenetics of the clover genus (*Trifolium*—Leguminosae). *Molecular Phylogenetics and Evolution* 39, 688-705 (2006).
[61] Wojciechowski, M.F., *Astragalus* (Fabaceae): A molecular phylogenetic perspective. *Brittonia* 57 (4), 382-396 (2005).
[62] Matthews, J.V. & Ovenden, L.E., Late Tertiary plant macrofossils from localities in Arctic/Subarctic North America: A review of the data. *Arctic* 43 (4), 364-392 (1990).
[63] Micheels, A., Bruch, A.A., Uhl, D., Utescher, T., & Mosbrugger, V., A Late Miocene climate model simulation with ECHAM4/ML and its quantitative validation with terrestrial proxy data. *Palaeogeography, Palaeoclimatology, Palaeoecology* 253, 251-270 (2007).
[64] Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K., Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292, 686-693 (2001).
[65] Vincent, J.-S., Late Tertiary and early Pleistocene deposits and history of Banks Island, southwestern Canadian arctic archipelago. *Arctic* 43 (4), 339-363 (1990).
[66] Common, I.F.B. & Waterhouse, D.F., *Butterflies of Australia*. (Angus and Robertson, 1981).
[67] Parsons, M., *The butterflies of Papua New Guinea. Their systematics and biology*. (Academic Press, San Diego, USA, 1999).
[68] Braby, M.F., *Butterflies of Australia: Their Identification, Biology, and Distribution*. (CSIRO Publishing, Collingwood, 2000).
[69] Nabokov, V., Notes on Neotropical Plebejinae (Lycaenidae, Lepidoptera). *Psyche* 52 (1-2), 1-61 (1945).
[70] Bálint, Z. & Johnson, K., Neotropical polyommatine diversity and affinities, I. Relationships of the higher taxa (Lepidoptera: Lycaenidae). *Acta Zoologica Academiae Scientiarum Hungaricae* 41, 211-235 (1995).
[71] Bálint, Z. & Johnson, K., Polyommatine lycaenids of the Oreal biome in the Neotropics, Part I: The *Itylos* section (Lepidoptera: Lycaenidae, Polyommatini). *Annales Historico-Naturales Musei Nationalis Hungarici* 86, 53-77 (1994).
[72] Johnson, K. & Bálint, Z., Distinction of *Pseudochrysops*, *Cyclargus*, *Echinargus* and *Hemiargus* in the Neotropical Polyommatini (Lepidotera, Lycaenidae). *Reports of the Museum of Natural History. University of Wisconsin (Stevens Point)* 54, 1-14 (1995).
[73] Schwartz, A., *The Butterflies of Hispaniola*. (University of Florida Press, Gainesville, Florida, USA, 1989).

[74] Nabokov, V., Notes on the morphology of the genus *Lycaeides* (Lycaenidae, Lepidoptera). *Psyche* 50, 104-138 (1944).
[75] Scott, J.A., *The Butterflies of North America: A Natural History and Field Guide*. (Stanford University Press, Stanford, 1986).
[76] Brock, J.P. & Kaufman, K., *Butterflies of North America*. (Houghton Mifflin, New York, USA, 2003).
[77] Gorbunov, P.Y., *The butterflies of Russia: classification, genitalia, keys for identification (Lepidoptera, Hesperioidea and Papilionoidea).* (Thesis. Russian Academy of Sciences, Institute of Plant and Animal Ecology, Ekaterinburg, Russia, 2001).
[78] Nabokov, V., *Speak, Memory: An Autobiography Revisited*. (Putnam, New York, 1966).
[79] Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R.C., DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *MOl. Marine Biol. Biotech.* 3, 294–299 (1994).
[80] Simon, C. *et al.*, Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* 87 (6), 651-701 (1994).
[81] Monteiro, A. & Pierce, N.E., Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII, and EF-1alpha gene sequences. *Molecular Phylogenetics and Evolution* 18, 264-281 (2001).
[82] Brower, A.V.Z., Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). *Molecular Phylogenetics and Evolution* 3 (2), 159–174 (1994).
[83] Moulton, J.K. & Wiegmann, B.M., Evolution and phylogenetic utility of cad (rudimentary) among Mesozoic-aged eremoneuran Diptera (Insecta). *Molecular Phylogenetics and Evolution* 31, 363-378 (2004).
[84] Cho, S. *et al.*, A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1alpha recovers morphology-based tree for heliothine moths. *Mol. Biol. Evol.* 12 (4), 650-656 (1995).
[85] Kandul, N.P. *et al.*, Phylogeny of *Agrodiaetus* Hübner 1822 (Lepidoptera: Lycaenidae) inferred from mtDNA sequences of *COI* and *COII*, and nuclear sequences of *EF1-*α: karyotype diversification and species radiation. *Systematic Biology* 53 (2), 278–298 (2004).
[86] Colgan, D.J. *et al.*, Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* 46, 419-437 (1998).
[87] White, T.J., Bruns, S., Lee, S., & Taylor, J., Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics in *PCR protocols: a guide to methods and applications*, edited by M. A. Innis, Gelfandm D. H., J. J. Snisky, & T. J. White (Academic Press, New York, 1990), pp. 315–322.
[88] Brower, A.V.Z. & DeSalle, R., Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Molecular Biology* 7 (1), 73-82 (1998).
[89] Sequiera, A.S., Normark, B.B., & Farrell, B., Evolutionary assembly of the conifer fauna: Distinguishin ancient from recent associations in bark beetles. *Proceedings of the Royal Entomological Society (London) B* 267, 2359-2366 (2000).