Extreme High-altitude Asian and Andean Pierid Butterflies Are Not Each Others' Closest Relatives

Authors: Arthur M. Shapiro, Matthew L. Forister, and James A. Fordyce
Source: Arctic, Antarctic, and Alpine Research, 39(1) : 137-142
Published By: Institute of Arctic and Alpine Research (INSTAAR), University of Colorado
Extreme High-altitude Asian and Andean Pierid Butterflies Are Not Each Others’ Closest Relatives

Arthur M. Shapiro*
Matthew L. Forister† and
James A. Fordyce‡
*Center for Population Biology,
University of California, Davis,
California 95616, U.S.A.
amshapiro@ucdavis.edu
†Department of Ecology and Evolution,
State University of New York at Stony Brook, Stony Brook, New York 11794-5245, U.S.A.
forister@life.bio.sunysb.edu
‡Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996, U.S.A.
jfordyce@utk.edu

Abstract
Some of the world’s highest-altitude butterflies belong to the tribe Pierini of the family Pieridae. Two nominal species of Baltia occur in arid-semiarid oreal environments in Central Asia, the Himalaya, and Pamir to over 5500 m. At least 13 species currently placed in four genera (Phulia, Infraphulia, Pierphulia, Piercolias) occur in similar environments at similar altitudes in the high Andes. These genera all share numerous morphospecializations whose functional relation to the oreal environment is not understood. Their evolutionary and biogeographic relationships have been debated for over a century. We performed a phylogenetic analysis based on sequencing portions of the mitochondrial cytochrome oxidase subunit I and subunit II regions (COI and COII), incorporating all the genera but Piercolias and a variety of suspected relatives. The results from analyses of COI and COII were compared to relationships inferred from morphological and ecological characters. We conclude that Baltia is not the sister-group of the Andean genera, which are clearly nested within a Neotropical clade. The “Camelid scenario” deriving all the genera from a common ancestor no longer appears viable.

Introduction
The geological ages of mountain ranges should constrain the time frame in which their endemic high-altitude biota could have evolved. The exigencies of life in the oreal zone apparently select for characteristic phenotypic (morphological) syndromes—the giant rosette-shrub growth form of tropical oreal plants is a dramatic example—although the specific functions of the component traits may be poorly if at all understood (Descimon, 1986). This fact predisposes these organisms to convergent evolution, which may be a serious impediment to understanding their phylogeny and historical biogeography. Remarkably little attention has been paid to the question of how quickly such syndromes can arise, despite the availability of molecular tools adapted to quantifying genetic distance and reconstructing phylogeny.

Some of the world’s highest-altitude butterflies belong to the tribe Pierini of the family Pieridae, reaching 5500 m or higher in the Andes (S. Halloy, personal communication), Himalaya (Mani, 1986), Pamir and other high Central and South Asian ranges. All of these ranges, and the landscapes and semiarid to arid biotopes in which the butterflies are endemic, are believed to be quite young—no older than late Tertiary and most likely Quaternary (Andes: Lamb et al., 1997; Clapperton, 1993; Asia: Kalvoda, 1992; Shackleton and Chengfa, 1988). By most accounts the oreal climates inhabited by these butterflies are consequent to their tectonic evolution (Raymo and Ruddiman, 1992).

The true relationship between the South American and Asian species has been debated vigorously for well over a century. There are two broad historical-biogeographic scenarios for them in the literature: either the Andean and Asian genera arose independently in situ within the constraints of their respective tectonic time-lines, or they arose from a common ancestor farther in the past, arriving at their current distributions by long-distance dispersal. The latter has been compared (Shapiro, 1992) to the well-known “Camelid scenario” of Matthew (1915). The camelids apparently originated in North America, dispersed to both Asia and South America, and subsequently became extinct on their continent of origin, giving rise to an Andean-Asian disjunct distribution. In the former case the closest relatives of each should be found in its own geographic region. In the latter, the Asian oreal taxa should be most closely related to the Andean ones. We bring molecular and phylogenetic methods to bear in hopes of resolving the issue.

The butterfly genera are Baltia in Asia (two nominal species) and Phulia, Infraphulia, Pierphulia, and Piercolias (13 nominal species recognized in the last revision: Field and Herrera, 1977) in the Andes. (In older literature Infraphulia and Pierphulia were lumped into Phulia, and Piercolias is usually called Andina. Our unpublished studies indicate that cryptic species may exist on both continents.) The enigmatic, monotypic Reliquia, from the oreal of the Sierra Nevada de Santa Marta, Colombia (Ackery, 1975) may or may not be related.

One of the two species of Baltia, B. butleri, is superficially extremely similar to Phulia and Infraphulia (Figs. 1c, 1d). The other, B. shawi, is equally similar to Pierphulia and Piercolias (Figs. 1a, 1b). These resemblances have occasioned much pre-cladistic and pre-molecular phylogenetic speculation. Most of the resemblances can be at least conceived as functionally related to the oreal environment (wing shape, venation, and pattern are all probably significant in aerodynamics and thermoregulation). The anatomy of the genitalia has been interpreted as sheltered from such selective influences, hence unlikely to be convergent, and useful as an indicator of true relationships (Klots, 1932). Close similarities have been alleged in the genitalia of Baltia and the Andean genera, leading Field (1958) and Field and Herrera (1977) to argue for a phylogenetic relationship. The female genitalia of Baltia and Piercolias, which are almost completely internal, are described as nearly identical.

© 2007 Regents of the University of Colorado
1523-0430/07 $7.00
The historical debate was forged within a tradition of attributing the high-Andean biota to invasion from the Northern Hemisphere (Shapiro, 1992). A variety of evolutionary hypotheses and putative ancestors, as noted above, are posited by Dixey (1894), Elwes (1895), Grote (1900), Roebel (1908–1924), Klots (1932), Forster (1958), Field (1958), Mani (1968, 1986), Field and Herrera (1977), Descimon (1986), and Shapiro (1992).

There is a fossil, *Mioperis talboti* from the Upper Miocene of Swabia, southwest Germany (Zeuner, 1942), which appears related to some or all of the genera of interest, but lacks most relevant characters except the wing venation and is not useful in resolving relationships. We thus must rely on what can be learned from extant taxa. If molecular phylogenies are more informative in general than those derived from characters subject to convergence, a molecular vs. morphological study of these genera should be helpful. If the “Camelid scenario” is valid, phylogenies based on both types of data should agree. If the Andean and Asian animals are convergent, the two phylogenies should disagree.

**Materials and Methods**

We did not seek to generate a broad phylogeny of the Pierini. Rather, we used the literature to select genera that had been suggested as relatives of the oreal pierines: *Aporia* in the Old World, *Pontia* (Synchloea) throughout the Holarctic, the Andean-Patagonian *Tatochila* and *Hypsocobia*, and the Neotropical *Ascia*. Field and Herrera (1977) considered the Andean oreal genera as specialized derivatives of *Tatochila*, with *Hypsocobia* intermediate. Several authors postulated *Baltia* as a specialized derivative of *Pontia*; some suggested *Pontia* as the ancestor of *Tatochila* and thus, indirectly, of the Andean oreal genera. For outgroups we chose *Colias eurytheme*, as well as the bizarre, taxonomically isolated *Eucheira socialis*, a Mexican montane endemic suspected of being an ancient relic.

Twenty-two recent specimens representing 15 species were obtained and sequenced (see Table 1 for collection and locality information); a *C. eurytheme* sequence was taken from GenBank (Accession #AF044024). We were unable to obtain recent material of *Piercolias* or *Reliquia*, and they are excluded. Except for a few mounted museum specimens, most of the material had been stored at −20°C until use. DNA extractions were performed with the Purgene DNA isolation kit (Gentra Systems, Minneapolis, MN, U.S.A.), and precipitated DNA was dried and resuspended in 200 µL of H2O. The basal parts of abdomens were extracted; wings, legs, and genitalia of all specimens were retained as vouchers and will be deposited in the Bohart Museum of Entomology, University of California (UC), Davis.

The following primers were used in polymerase chain reactions (PCR) to amplify a portion of the mitochondrial cytochrome oxidase subunit I region (COI): k698 (5’ TAC AAT TTA TCG CCT AAA CTT CAG CC) and k525 (5’ ACT GTA AAT ATA TGA TGA GCT CA); and for a portion of the mitochondrial cytochrome oxidase subunit II region (COII): EVA (5’ GAG ACC ATT ACT TGC TTT CAG TCA TCT 3’) and PATRICK (5’ CTA ATA TGG CAG ATT ATA TGT ATT GGA 3’) (Caterino and Sperling, 1999). PCR amplifications were run according to the following protocol: (1) 94°C for 2 min; (2) 39 cycles of 94°C for 1 min, 48°C for 45 s, and 72°C for 1 min 45 s; (3) 72°C for 10 min. PCR products were sequenced in both directions on ABI 3730 (COI) and ABI 377 (COII) automatic sequencers at the DNA Sequencing Facility, Division of Biological Sciences, UC Davis, producing 612 bp of COI and 501 bp of COII that were consistently readable. The SeqEd computer program (v.1.0.3, Applied Biosystems, Foster City, CA, U.S.A.) was used to aid in the visual alignment of sequences.

Maximum Parsimony (MP), Maximum Likelihood (ML) and Bayesian Maximum Likelihood (BML) were used with the COI and COII sequences to examine the relationships among taxa. Using hierarchical likelihood ratio tests (α < 0.01) in the program Modeltest v.3 (Posada and Crandall, 1988), 56 models of evolution were evaluated for the portions of COI and COII sequenced (Table 2).

MP (1000 bootstrap pseudoreplicates) and ML (100 pseudoreplicates) analyses were done with the program PAUP* v.4.0b10.
TABLE 1
Locality and collection information for specimens studied. Numbers in parentheses indicate the number of individuals sampled of each species. Collectors as follows: all specimens from Chile, California, and Texas (A. M. Shapiro); A. crataegi (M. Jagelka); B. butleri (M. Zubrik); B. shawiti (R. Westphal); E. socialis (D. Underwood). One *Colias eurytheme* (not listed) sequence was obtained from GenBank.

<table>
<thead>
<tr>
<th>Species</th>
<th>Location</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aporia crataegi (1)</td>
<td>Tolfa, Italy</td>
<td>2001</td>
</tr>
<tr>
<td>Ascia monuste (1)</td>
<td>Welder Ranch, San Patricio Co., TX, U.S.A.</td>
<td>iv.23.05</td>
</tr>
<tr>
<td>Baltia shawiti (2)</td>
<td>3900 m, Koh-I-Parandeh, Panshir Province, Afganistan</td>
<td>vii.10–20.02</td>
</tr>
<tr>
<td>Baltia butleri (2)</td>
<td>Gyatisola Pass, Himalaya Mts., southwest Tibet</td>
<td>v.27.96</td>
</tr>
<tr>
<td>Eucheria socialis (1)</td>
<td>Valle de Bravo, Estado de Mexico, Mexico</td>
<td>1990*</td>
</tr>
<tr>
<td>Tatochila blanchardii (1)</td>
<td>Conaripe, Prov. Valdivia, Chile</td>
<td>i.8.96</td>
</tr>
<tr>
<td>Tatochila macroleuca (1)</td>
<td>Putre, Prov. Parinacota, Chile</td>
<td>xi.20.94</td>
</tr>
<tr>
<td>Tatochila mercedis (1)</td>
<td>Conaripe, Prov. Valdivia, Chile</td>
<td>i.8.96</td>
</tr>
<tr>
<td>Tatochila theodice (2)</td>
<td>Laguna Solis, Prov. Ultima Esperanza, Chile</td>
<td>i.23.94</td>
</tr>
<tr>
<td>Infracalidia illyoides (1)</td>
<td>Salar de Suriri, Prov. Parinacota, Chile</td>
<td>xi.22.94</td>
</tr>
<tr>
<td>Phalaena nymphula (2)</td>
<td>Salar de Suriri, Prov. Parinacota, Chile</td>
<td>xi.22.94</td>
</tr>
<tr>
<td>Pierphalia rosea (2)</td>
<td>Parque Nacional Lauca, Prov. Parinacota, Chile</td>
<td>xi.20.94</td>
</tr>
<tr>
<td>Hypochoila microdice (1)</td>
<td>Cerro Guado, Prov. Ultima Esperanza, Chile</td>
<td>xii.2.94</td>
</tr>
<tr>
<td>Pontia proteodice (2)</td>
<td>Sierra Valley, Sierra Co., California, U.S.A.</td>
<td>x.12.00</td>
</tr>
<tr>
<td>Pontia occidentalis (2)</td>
<td>Sierra Valley, Sierra Co., California, U.S.A.</td>
<td>ix.12.00</td>
</tr>
</tbody>
</table>

* Collected as larvae in 1989, emerged in 1990.

(Swofford, 2002). Bayesian analysis was performed with the program MRBAYES (Huelsenbeck and Ronquist, 2001; available online at http://morphbank.ubc.ca.mrbayes/), which uses a metropolis-coupled, Markov-chain Monte Carlo method to estimate posterior probabilities associated with nodes. A total of 2 million pseudoreplicates were run in MRBAYES; trees were sampled every 100 generations. A “burn-in” value was estimated by plotting log-likelihood values against generation time, and a consensus tree with posterior probabilities was then generated in PAUP* for all sampled trees posterior to the “burn-in” generation.

For MP and ML analyses, the portions of COI and COII sequenced for each individual were analyzed as contiguous sequences following evaluation with a partition-homogeneity test in PAUP* (1000 replicates), which returned P-values near or below 0.05, suggesting acceptable congruence of the phylogenetic information in each region (one run of the test produced P = 0.053, and a second run produced P = 0.042). Bayesian analyses were implemented with six partitions, such that each codon position of each gene was modeled separately (see Table 2 for the models used for each data partition).

A morphological matrix (Table 3) was generated for 13 taxa using data already available from published descriptions or our own work. Characters were derived from both adults and immatures. Host plant, which is often phylogenetically informative in butterflies (Braby et al., 2006), was also used as a character. A number of potentially useful characters were omitted because relevant information was lacking or difficult to interpret for some of the genera. The data were used in the computer program PAUP* to identify the most parsimonious relationships among taxa. The Shimodaira-Hasegawa (SH) Test was then used to compare the topology of the trees generated with parsimony analysis of morphological and ecological characters to the best maximum-likelihood sequence tree (combined COI and COII), with similarity rejected at the level p < 0.05 (1000 RELL bootstrap pseudoreplicates were used).

**Results and Discussion**

Analysis of the portions of COI and COII sequenced produced 187 variable sites with 140 informative sites from COI, and 132 variable sites with 103 informative sites from COII.

| TABLE 2 DNA substitution models selected using maximum likelihood ratio tests. |
|---------------------------------|-----------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Data partition                  | DNA substitution model      | Number substitution types | Invariant sites? | Substitution rates | Shape (±) (e2, 95% credibility interval) |
| COI 1st Codon                   | TrN*                        | 6                | No              | Gamma distributed | 0.303 (0.019, 0.137–0.609) |
| COI 2nd Codon                   | F81*                        | 1                | Yes             | Equal            | 0.739 (0.017, 0.528–1.043) |
| COI 3rd Codon                   | HKY*                        | 2                | No              | Gamma distributed | 0.188 (0.005, 0.057–0.338) |
| COII 1st Codon                  | TrN                         | 6                | No              | Gamma distributed | 0.569 (0.018, 0.377–0.911) |
| COII 2nd Codon                  | F81                         | 1                | No              | Equal            | 1.026 (0.063, 0.593–1.571) |
| COII 3rd Codon                  | TrN                         | 6                | No              | Gamma distributed | 0.303 (0.019, 0.137–0.609) |
| COI + COII                      | GTRd                        | 6                | Yes             | Gamma distributed | 0.303 (0.019, 0.137–0.609) |

* Hasegawa et al. (1985) model.
* General time reversible model.
TABLE 3

Morphological character matrix. Characters and codes are as follows: (1) number of radial veins: 0 = 2, 1 = 3, 2 = 4; (2) size of pulvilli: 0 = large, 1 = reduced; (3) size of paronychia: 0 = large, 1 = reduced; (4) M1 stalked on forewing: 0 = no, 1 = yes; (5) M2 stalked on forewing: 0 = no, 1 = yes; (6) length of third segment of pulpus: 0 = very short, 1 = short, 2 = “normal”; (7) tegumen size: 0 = very short, 1 = large, 2 = “normal”; (8) apex of forewing rounded: 0 = rounded, 1 = angled; (9) larval host: 0 = no, 1 = yes; (10) conical development of lower juxta: 0 = not, 1 = somewhat, 2 = pronounced; (11) tibial spurs on middle and hind legs: 0 = no, 1 = yes; (12) harpe simple, rounded: 0 = no, 1 = yes; (13) overall VHW “thermal” melanization: 0 = no, 1 = yes; (14) wing vein darkening: 0 = veins not melanized relative to ground color, 1 = thin black line directly on veins, 2 = dark edging on veins, veins not themselves darkened, 3 = broad dark lines including the veins; (15) chevrons concolorous with vein-lines on ventral hindwing: 0 = no chevrons, 1 = no, 2 = yes (Hypsochilla variable); (16) direction of “arrow” of ventral hindwing chevrons: 0 = no chevrons, 1 = point marginad, 2 = no point, 3 = point basad; (17) pupal attachment: 0 = no, 1 = very stalked, veins themselves short; 0 = no, 1 = yes.

Sequences from all specimens are available on GenBank (Accession nos. DQ463374 to DQ463417). Saturation plots examined for each codon position in both COI and COII gave no indication that any of the sites were saturated with multiple substitutions.

The ML, BML, and MP phylogenies put the Andean oreal genera into a monophyletic group with Andean non-oreal genera (as predicted by Field and Herrera, 1977), but group no Andean genus with the Northern Hemisphere (Fig. 2). Baltia is also monophyletic in these phylogenies, as is Pontia (Synchloe). The relationships described in Figure 2 falsify the hypothesis, which Field (1958) had posited, that Baltia and the Andean oreal genera are sister-groups.

Nine equally parsimonious trees were produced by analyses of morphological and ecological characters. The topologies of these trees are similar in that a clade composed of Baltia shawi,
Baltia butleri, Phulia, Infraphulia, and Pierphulia was recovered with 80% bootstrap support. An exemplar of the nine most parsimonious trees is shown in Figure 3 (the clade comprised of the five focal taxa is also present in the consensus tree, not shown). The results from parsimony analyses of morphological and ecological data unsurprisingly reflect the historic ambiguity found in narrative phylogenies—embedding both Baltia species among Andean taxa but not as each other’s sister-species. These results underscore that, at least at the coarse level of resolution used to date in scoring morphological characters, convergence and possibly reversals run rampant in this group. The nine equally parsimonious trees are all statistically different from the combined ML COI and COII tree at the level of \( P < 0.0001 \). If we place greater confidence in molecules than morphology, both sets of morphologically similar taxa (B. shawi / P. rosea; B. butleri / P. nymphula, I. illydes) are convergent.

The internal structure of the Andean clade is complicated by problems with the existing taxonomy. Field and Herrera (1977), as earlier noted, saw Hypsochila as an intermediate step in the evolution of the oreal genera from a generalized, lowland Tatochila ancestor. At that time little was known about these butterflies except their adult morphology. Subsequent studies of the immature stages, biology, and host plants have suggested that both Tatochila and Hypsochila are almost certainly polyphyletic (Shapiro, 1991 and unpublished). The species we sequenced were selected to represent the perceived austral diversity in Tatochila. T. macrodice (central Andes, Puna, and Altiplano) and T. mercidea (lowland Mediterranean Chile) are fully interfertile and biologically conspecific, part of a large polytypic complex ranging from Colombia to Tierra del Fuego (Shapiro, 1991), but they did not fall out as sister-taxa here. Hypsochila microdice fell squarely within this cluster, despite the fact that on morphological grounds it belongs to a different genus! Tatochila theodice belongs in its own monotypic genus, based on both adult and early-stage characters. Whatever the proper relationships at this level and the implied insensitivity of the COI and COII sequences to them, they do not affect the higher-level question this study addresses. The position of Ascia, as recovered by ML, suggests that the entire Andean group of genera may be derived from the lowland tropics rather than from Holarctic ancestors, as commonly supposed (however, MP and BML failed to support this node).

Independently of our efforts, Braby et al. (2006), in a sweeping phylogenetic survey of the Pieridae involving some of the same taxa but different molecular markers, find that Baltia (they had butleri only) clusters with Pontia and the Andean oreal Phulia, Pierphulia, and Infraphulia cluster with Tatochila, Hypsochila, and Theochila (which we did not have). Like us, they had neither Reliquia, which is morphologically very similar to Pontia, nor Piercolias. Despite differences in detail, largely due to having used several different taxa, their results agree with ours in rejecting the “Camelid scenario” for the oreal genera.

Despite persistent ambiguities in the overall phylogeny, we conclude that Baltia and the Andean oreal genera are not sister-taxis and do not form a clade antedating the current geographic pattern, as in the Camelidae. Their remarkable similarities are convergent (extending even to the genitalia), and must be accommodated within the Quaternary time frame of their oreal biomes—suggesting quite rapid morphological evolution and, in the Andes, taxonomic differentiation.

Acknowledgments

Collection of material from the Andes was supported by NSF grant BSR-83-06922 and two travel grants from the Committee on Research of the Academic Senate, UC Davis, all to Shapiro. The Committee on Research also supported sequencing. We thank Brad Shaffer for the use of his facilities, Tom Near for help with the analyses, and Michael Braby and Naomi Pierce for sharing their unpublished results with us.

References Cited


Klots, A. B., 1932: A generic revision of the Pieridae (Lepidoptera), together with a study of the male genitalia. Entomologica Americana, 12: 139–204.


Ms accepted April 2006