Phylogenetic relationships of the New World Troidini swallowtails (Lepidoptera: Papilionidae) based on COI, COII, and EF-1α genes

Karina Lucas Silva-Brandão a, André Victor Lucci Freitas b,*, Andrew V.Z. Brower c, Vera Nisaka Solferini a

a Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, CP 6109, CEP 13083-970, Campinas, SP, Brazil
b Museu de História Natural and Departamento de Zoologia, Instituto de Biologia, Universidade Estadual de Campinas, CP 6109, CEP 13083-970, Campinas, SP, Brazil
c Department of Zoology, Oregon State University, 97331, Corvallis, OR, USA

Received 25 May 2004; revised 6 December 2004
Available online 13 June 2005

Abstract

A phylogeny of the Neotropical members of the Tribe Troidini (Lepidoptera: Papilionidae) was obtained with sequences of three protein-coding genes: two mitochondrial (COI and COII), and one nuclear (EF-1α). Parsimony and Bayesian analyses of 33 taxa resulted in very similar trees regardless of method used with the 27 troidines always forming a monophyletic clade. Within Troidini, the genus Battus is sister group to the remaining troidines, followed by a clade formed by the Paleotropical taxa (here represented by three exemplars). The genus Euryades is the next branch, and sister group of Parides. The genus Parides is monophyletic, and is divided into four main groups by Maximum Parsimony analysis, with the most basal group composed of tailed species restricted to SE Brazil. Character optimization of ecological and morphological traits over the phylogeny proposed for troidines indicated that the use of several species of Aristolochia is ancestral over the use of few or a single host-plant. For the other three characters, the ancestral states were the absence of long tails, forest as the primary habitat and oviposition solitary or in loose group of several eggs.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Character optimization; Host-plant use; Lepidoptera; Molecular phylogeny; Parides; Swallowtail butterflies; Troidini

1. Introduction

Swallowtail butterflies (Papilionidae) are among the most popular insect taxa, and have greatly contributed to studies of ecology, behavior, and evolution in insects (Boggs et al., 2003; Scriber, 1995). Many studies have been published with this group, including on ecology (Brown et al., 1981; Morais and Brown, 1991; Spade et al., 1988), behavior (Rausher, 1978; Stamp, 1986), and chemistry (Honda and Hayashi, 1995; Klitzke and Brown, 2000; Nishida et al., 1993; Rothschild et al., 1970; Sime et al., 2000; Urzúa and Priestap, 1985).

Papilionid butterflies are divided into three subfamilies: Baroniinae, with a single genus and species occurring in Mexico (Baronia brevicornis, believed to be the most basal taxon (Caterino et al., 2001; Scriber, 1995; Tyler et al., 1994)), Parnassiinae (broad Holarctic distribution), and the cosmopolitan Papilioninae (Scriber, 1995). According to Haüser et al. (2002), the subfamily Papilioninae has 485 species, divided into three tribes: Papilionini, Graphini, and Troidini. The tribe Troidini is predominantly tropical, with most species concentrated in the lowland forests of Central and South America and in the IndoAustralian region (Weintraub, 1995). The tribe includes 130 species divided into 12 genera, three of which occur in the Neotropics: Battus (11 species), Euryades (2 species), and Parides (34 species).
(Tyler et al., 1994). The genus *Parides s. str.* is exclusively neotropical, and includes several species on official lists of endangered species (MMA, 2003). These include the Southeast Brazilian *Parides ascanius*, which is considered endangered due to the destruction of habitat and host-plants (Otero and Brown, 1986; Tyler et al., 1994), and other sensitive species such as *P. tros*, which is rare on the coastal slopes of the Atlantic Forest, and deserving of attention and monitoring now and in the future (Tyler et al., 1994).

Troidines are frequently cited in the literature as classic examples of coevolution with their host-plants *Aristolochia* (Aristolochiaceae) (Weintraub, 1995), earning them the name “Aristolochia swallowtails” (Brown et al., 1981). The features of this association agree with most of the premises of the coevolutionary hypothesis (Ehrlich and Raven, 1964). The larvae of Troidini feed almost exclusively on *Aristolochia* species, and sequester the major secondary metabolites of these plants, aristolochic acids (Klitzke and Brown, 2000). These compounds are thought to make the butterflies unpalatable to potential predators (Brower and Brower, 1964; Nishida and Fukami, 1989; Rothschild et al., 1970; Sime, 2002). Larvae and adults of many species advertise their unpalatability through aposematic coloration, making them notable in their roles as unpalatable models in mimicry rings (Sime et al., 2000; Tyler et al., 1994).

Based upon the classificatory groundwork of Haase (1892) and Rothschild and Jordan (1906), morphological studies investigating the phylogenetic relationships among troidine butterflies include those of Munroe and Ehrlich (1960), Munroe (1961), Hancock (1983), and Miller (1987). Recently, several molecular studies have been added to this list (Aubert et al., 1999; Caterino et al., 2001; Caterino and Sperling, 1999; Kondo and Shinkawa, 2003; Morinaka et al., 1999, 2000; Reed and Sperling, 1999; Zakharov et al., 2004). Most of these studies have suggested that Troidini + Papilionini form a clade, with Graphini basal to both (Caterino et al., 2001; Hancock, 1983; Kondo and Shinkawa, 2003; Miller, 1987; Zakharov et al., 2004). However, the internal relationships among members of Troidini remain controversial (Vane-Wright, 2003).

Morphological classifications (e.g., Hancock, 1983; Miller, 1987; Munroe, 1961) have divided Troidini into two subtribes: Battina, including only the genus *Battus*, and Troidina, including Southeast Asian *Cressida*, *Troides*, *Ornithoptera*, *Trogonoptera*, *Pachliopta*, *Losaria*, *Pharmacophagus*, and *Atrophaneura* (including *Panomia*) and Neotropical *Euryades* and *Parides*. The genus *Parides* is sometimes circumscribed to include both the Neotropical representatives addressed here and members of *Atrophaneura*. Morinaka et al. (1999) and Morinaka et al. (2000) studied the molecular phylogenetic relationships among Asian *Ornithoptera* butterflies, including some species in other genera of Troidini, using the mitochondrial gene ND5. Kondo and Shinkawa (2003) also used the ND5 sequences to propose a molecular phylogeny for three genera of birdwing butterflies, *Trogonoptera*, *Troides* and *Ornithoptera*, and Kato and Yagi (2004) have studied the phylogeny of geographical races of *Atrophaneura alcinous* butterflies from Asia with the same gene. Tyler et al. (1994) presented the only published phylogenetic hypothesis of the species-level relationships of the New World troidines, based on adult and larval morphological characters, adult behavior, and chemistry. In addition, the phylogeny of the genus *Battus* was studied by Racheli and Oliverio (1993) using adult morphological characters.

There are few studies focusing on the internal relationships of the genus *Parides* (see Tyler et al., 1994), and the only phylogeny published so far includes only four species in this genus. Considering the diversity and ecological importance of this group in the Neotropics, a phylogenetic hypothesis is necessary to help understand the biogeography, behavior, chemical ecology and evolution of host-plant use among *Parides* and other troidine species. The aims of this study are: (1) to infer a molecular phylogeny of the New World Troidini butterflies of the genera *Battus*, *Euryades* and especially *Parides s. str.*, based on DNA sequences of mitochondrial and nuclear genes, to propose a hypothesis about their evolutionary history; and (2) investigate the evolution of four ecological and morphological traits within the genus *Parides*.

2. Materials and methods

2.1. Specimens

Individual butterflies representing approximately half of the species of the Neotropical troidine genera, *Parides* (17 of 34 species), with representatives of all subgeneric groups recognized by Tyler et al. (1994), *Battus* (5 of 11 species), and *Euryades* (1 of 2 species) were collected in the field (Table 1). Upon collection, the wings were separated from the body and stored in glassine envelopes and the bodies were preserved in a freezer at −70 °C. In some cases, DNA was extracted from older, dried specimens from the collection of K.S. Brown. Vouchers of all samples have been deposited in the Museu de História Natural of UNICAMP. Previously published sequences of five species of Troidini, two species of Graphiini, two species of Papilionini, one of Parnassiinae, and *Baronia brevicornis* (Baroiniinae) were obtained from GenBank (Caterino et al., 2001). The final matrix has 47 terminals representing 33 species, including 27 Troidini and six non-troidine papilionids as outgroups (Table 1).
2.2. Molecular techniques

Total genomic DNA was extracted following the protocol of Genomic PrepCells and Tissue DNA Isolation Kit (Amersham Pharmacia Biotech) or DNeasy Tissue Kit (Qiagen) from thorax/abdominal tissue of frozen individuals. The DNA of three species of Troidini, *Battus crassus*, *Parides chabrias*, and *P. panthonus lysima-
chus, was extracted from the legs of dried museum specimens. For these materials, we used the protocol of DNeasy Tissue Kit modified for ancient material, that is, samples were lysed overnight and recovered with 50 μL of dilution buffer, as suggested by the manufacturer’s technical support. These protocols have been widely used by research groups studying butterfly systematics, and have been shown to be reliable (Wahlberg et al., 2003). Purified DNA was stored in TE buffer at −20 °C. For each of the specimens we sequenced the entire mitochondrial cytochrome oxidase I and II genes (COI and COII) and the nuclear gene elongation factor-1 α (EF-1α) using the primer combinations listed in Table 2. When possible, we sequenced at least two individuals of each species. In Papilionidae, these genes, and especially COI, are suitable to elucidate relationships at species and generic levels (Brower, 1994; Caterino et al., 2001; Caterino and Sperling, 1999; Zakharov et al., 2004) and Sperling (2003) has recommended the use of EF-1α in combination with COI/COII mitochondrial genes to study phylogenetic relationships among butterflies in general. Both genes have advantages and disadvantages: mtDNA is easier to amplify, does not have non-coding regions (introns), has no recombination, and evolves at higher rates. However, mtDNA may present high levels of homoplasy because of an extreme A/T bias in third positions (Harrison, 1989). Nuclear genes can be advantageous due to the less biased base composition, and generally evolve more slowly than mitochondrial genes, making them better markers for deep divergences (Lin and Danforth, 2004). The choice of these genes in our study is due mainly to their use in other studies of Papilionidae (Caterino et al., 2001; Caterino and Sperling, 1999; Reed and Sperling, 1999; Zakharov et al., 2004) since studying comparable gene regions contributes synergistically to a more comprehensive picture of the evolution of all butterfly groups (Caterino et al., 2000).

Amplification of DNA was performed using two methods. We used a direct method for COI and COII, using primers that amplified a sequence of about 500 bp (COI) or 700 bp (COII) in length. For EF-1α we used a nested method, a sequence of primers that first amplified around 1200 bp, and then amplified a smaller 500–700 bp fragment from each half of the larger piece. All fragments were amplified in a total volume of 25 μL. The following thermal cycling protocol was used for COI and COII: 96 °C for 2 min, 35 cycles of 94 °C for 1 min, 45 °C for 1:30 min, 72 °C for 1:50 min, and a final extension period of 72 °C for 4 min. The cycling profile for EF-1α was 95 °C for 2 min, 30 cycles of 95 °C for 1 min, 45 °C for 1 min, 72 °C for 1:50 min, and a final extension period of 72 °C for 4 min.

PCR products were cleaned by using a GFXPCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech) or a Gel Purification Kit (Qiagen), and then amplified for sequencing using the protocol of ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit. PCR fragments were sequenced in ABI 373 or ABI 377 automated sequencers. All fragments were sequenced in both directions, using the same primers listed in Table 2. Sequences were analyzed with the program SeqEd version 1.0.3 (Applied Biosystems), and aligned manually by the Se.Al program (Rambaut, 1996) using the translated amino acid sequences and the Drosophila yakuba sequence for COI and COII (Clary and Wolstenholme, 1985). All sequences were deposited in GenBank (Accession Nos. in Table 1).

Table 2

<table>
<thead>
<tr>
<th>Gene</th>
<th>Name</th>
<th>F/R</th>
<th>Locationa,b(3’end)</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COI</td>
<td>K698</td>
<td>F</td>
<td>1460</td>
<td>TACAATTTATCGCTAAACTTCAGCC</td>
</tr>
<tr>
<td></td>
<td>K699</td>
<td>R</td>
<td>1840</td>
<td>AGGGAGTAAAACAGTTCACCC</td>
</tr>
<tr>
<td></td>
<td>Ron-mod</td>
<td>F</td>
<td>1751</td>
<td>GGTTCACCTGATATAGCATTCCC</td>
</tr>
<tr>
<td></td>
<td>Nancy-mod</td>
<td>R</td>
<td>2192</td>
<td>CCTGTTAAttaaaatataaaaacttc</td>
</tr>
<tr>
<td></td>
<td>Jerry</td>
<td>F</td>
<td>2183</td>
<td>CAACATTATTTGATTTTTG</td>
</tr>
<tr>
<td></td>
<td>Mila</td>
<td>R</td>
<td>2659</td>
<td>GCTAATCCAGTGAATAATG</td>
</tr>
<tr>
<td></td>
<td>BrianXV</td>
<td>F</td>
<td>2495</td>
<td>CATCAATCTCATGAAGATTAGG</td>
</tr>
<tr>
<td></td>
<td>PatII</td>
<td>R</td>
<td>3014</td>
<td>TCCATCATACATAACTCCAGCTATTAG</td>
</tr>
<tr>
<td>COII</td>
<td>Patrick</td>
<td>F</td>
<td>3038</td>
<td>CTAATATGCCAGATTATATGTAATGGA</td>
</tr>
<tr>
<td></td>
<td>Eva</td>
<td>R</td>
<td>3782</td>
<td>GACACGATCTACATTTGG</td>
</tr>
<tr>
<td>EF-1α</td>
<td>Hillary</td>
<td>F</td>
<td>2103</td>
<td>CACATYACATTGTGTSATYG</td>
</tr>
<tr>
<td></td>
<td>Monica</td>
<td>R</td>
<td>2645</td>
<td>CATRTTGTCKCCTGTCCTC</td>
</tr>
<tr>
<td></td>
<td>Al</td>
<td>F</td>
<td>2582</td>
<td>GAGGAAATYYARAAAGAAG</td>
</tr>
<tr>
<td></td>
<td>Tipper</td>
<td>R</td>
<td>3344</td>
<td>ACAGCVACKGTYTGYCTCARTTC</td>
</tr>
</tbody>
</table>

Y = C/T; S = C/G; R = G/A; K = G/T; V = C/G/A.

a Position relative to Drosophila yakuba (Clary and Wolstenholme, 1985) for COI and COII primers. Primers obtained from Caterino and Sperling (1999).
b Position relative to Drosophila melanogaster to EF-1α primers. Primers obtained from Cho et al. (1995).
2.3. Phylogenetic analyses

The phylogenetic analyses were performed with PAUP* 4.0b10 (Swofford, 2002), using Maximum Parsimony. Bayesian analysis was carried out with MrBayes 3.08v (Huelsenbeck and Ronquist, 2001). The purpose of doing Bayesian analysis was to investigate the effects on the results under the most restrictive assumptions of data analysis.

The Partition Homogeneity Test of PAUP* 4.0b10 (Swofford, 2002) was used to assess congruence among molecular data sets. This test is equivalent to the ILD test of Farris et al. (1994), which has been employed as a method for determining whether separated data sets should be combined in a single parsimony analysis (Bull et al., 1993; Yoder et al., 2001). In the present study, this test was used as a measure of heterogeneity among the data sets (as in Freitas and Brown, 2004), and not as a way to validate or invalidate the combined analysis (see also Brower et al., 1996; DeSalle and Brower, 1997). We performed the test under parsimony, using the following parameters: heuristic search, TBR branch-swapping, with 100 random addition sequences, and 500 replicates to generate the null hypothesis. The transition/transversion ratio was estimated in MEGA, version 2.1 (Kumar et al., 2001).

Maximum Parsimony analyses (MP) were performed on the entire data set, as well as for each gene separately, using heuristic search with 500 random taxon addition replicates, TBR branch-swapping, gaps scored as missing data, and all characters equally weighted. A strict consensus tree was computed whenever multiple equally parsimonious trees were obtained. The consistency index (CI) and the retention index (RI) were calculated by the PAUP “tree scores” option. The robustness of each branch was determined using the non-parametric bootstrap test (Felsenstein, 1985), with 1000 replicates and 10 random taxon additions. Bremer support and Partitioned Bremer support values (to obtain the contribution of each data set to the Bremer support values of the combined analysis) (Baker and DeSalle, 1997; Baker et al., 1998; Bremer, 1988, 1994) were calculated using TreeRot (Sorensen, 1999), in conjunction with PAUP* 4.0b10 (Swofford, 2002). The analysis was conducted with 100 random taxon addition replicates, TBR branch-swapping and 100 trees held in each replicate. Following Wahlberg and Nylin (2003) and Wahlberg et al. (2003), we will refer to the support values as either giving weak, moderate, good or strong support when discussing our results. We define “weak support” as Bremer support values of 1–2 (mostly corresponding to bootstrap values of 50–61%), “moderate support” as values between 3 and 4 (bootstrap values 62–74%), “good support” as values between 5 and 8 (bootstrap values 75–88%) and “strong support” as values >8 (bootstrap values 89–100%).

We used the program MODELTEST 3.06 (Posada and Crandall, 1998) to determine the available substitution model with the best fit to each partitioned data set. Bayesian analyses (Huelsenbeck and Ronquist, 2001; Huelsenbeck et al., 2001, 2002) were carried out for the combined data set under the model GTR + G + I (General Time-Reversible model (Rodrigues et al., 1990), with gamma distribution (Γ) and with proportion of invariable sites (I)). According to Nylander et al. (2004), analysis of combined data by Bayesian methods permits partition-specific substitution models and parameters. For that reason, all substitution model parameters (gamma shape parameter, proportion of invariable sites, character state frequencies, substitution rates of GTR model) were allowed to vary across partitions (=genes). We conducted six simultaneous chains for 1.0 × 10^6 generations, sampling trees every 100 cycles. Stability of the process was assessed plotting the likelihood scores against generation time (Lin and Danforth, 2004). The 1000 first trees were discarded as “burn in.” For all analyses, Baronia brevicornis (Baroiniinae) was used as outgroup to root the tree.

2.4. Analyses of character evolution

We investigated the evolution of some ecological and morphological traits superimposed onto the phylogenetic hypothesis proposed for Troidini butterflies.

Character 1. “Use of host-plants,” using data obtained from several literature and field sources (Brown et al., 1981, 1995; DeVries, 1987; Freitas and Ramos, 2001; Klitzke and Brown, 2000; Morais and Brown, 1991; Otero and Brown, 1986; Papaj, 1986; Rausher, 1980; Rausher and Odendaal, 1987; Spade et al., 1988; Stamp, 1986; Tyler et al., 1994; Weintraub, 1995; AVLF, K.S. Brown, unpublished data), Janz et al. (2001) and Wahlberg (2001) discuss the difficulty in coding the use of host-plants in analyses of character optimization. Here we chose to use the number of Aristolochia species used as host by each Troidini species as a multistate character (Table 3); the results were the same if compared with a binary character coded as “generalist” or “specialist” (using the same definition of Janz et al. (2001)). We followed the approach of Janz et al. (2001) to test the hypothesis of specialization as a “dead end” (Futuyma and Moreno, 1988), i.e., we compared the number of host-plant gains (colonizations) and losses (specializations) in our analyses of character evolution. According to Janz et al. (2001), more host-plant losses than gains indicate a trend toward increasing specialization.

Character 2. “Presence or absence of long tail on the hindwing” (we considered a “long” tail a hindwing projection three times longer than any other projection).

Character 3. “Primary habitat”: forest (dominant tall trees with closed canopy); scrub (open woody forest,
### Table 3

Host-plants used by Troidini butterflies and the outgroups used in phylogenetic analyses

<table>
<thead>
<tr>
<th>Species</th>
<th>Host-plant family</th>
<th>Host-plant species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baronia brevicornis</td>
<td>Leguminosae</td>
<td>A. macroura, A. veraguensis</td>
<td>11</td>
</tr>
<tr>
<td>Parides acraea</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. acraea</td>
<td>2, 3, 6, 8, 9, 10, 11, 13</td>
</tr>
<tr>
<td>Parides ascanius</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. elegans, A. esperanzae, A. littoralis, A. macroura</td>
<td>10, 11</td>
</tr>
<tr>
<td>Parides burchius</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. elegans, A. esperanzae, A. littoralis, A. macroura</td>
<td>3, 2, 9, 13</td>
</tr>
<tr>
<td>Parides chabrias</td>
<td>Aristolochiaceae</td>
<td>A. acutifolia, A. barbarata, A. bicolor, A. burchelli</td>
<td>10</td>
</tr>
<tr>
<td>Parides儿童ae</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides eurimeades</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides foetida</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides neophila</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>9, 10, 11</td>
</tr>
<tr>
<td>Parides panthonus</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides photinus</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides proeus</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>9, 13</td>
</tr>
<tr>
<td>Parides sesostris</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides tros</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
<tr>
<td>Parides vertumnus catora</td>
<td>Aristolochiaceae</td>
<td>A. arcuaea, A. esperanzae, A. melastoma, A. triangularis</td>
<td>10</td>
</tr>
</tbody>
</table>

including chaparral, semi arid formations, and Brazilian “cerrado” and “restinga”); and open (dominant open, mainly grassland or small herbs, including south Brazilian “pampas”).

Character 4. “Gregarious immatures,” based mainly on oviposition patterns (following Tyler et al. (1994), defining two kinds of oviposition: tight bunch of many ordered eggs vs. solitary or loose group of several eggs).

Character states were treated as unordered and of equal weights, and were optimized on the MP phylogeny, using MacClade 3.08 program (Maddison and Maddison, 1999). We performed the analysis over the MP tree due to the nature of optimizations algorithms, which are based on parsimony. “Host-plant use” could be analyzed asymmetrically, with gains costing more than losses (as in Wahlberg, 2001), but we considered that within a single genus, the costs of gains and losses of a host-plant are similar, since the biological, behavioral, and ecological apparatus of oviposition and larval feeding do not have to experience significant changes. Analyses of character evolution used the same outgroups as the phylogenetic analyses (the results were the same if only Papilioninae outgroups are used).

To test whether there is a phylogenetic signal in the characters traced, we used the methodology proposed by Wahlberg (2001), modified from the PTP test described by Faith and Cranston (1991). The test consists in comparing the number of steps of the tree constructed with the actual data with the number of steps obtained for each random reshuffling of the states of each separated character. We performed 300 random reshufflings of character states among the fixed terminal taxa, with the equally weighted data set, using the option “shuffle” from the utilities menu in the MacClade program (Maddison and Maddison, 1999). The probability (P) that the observed pattern does not differ randomly is given by the number of replications as short or shorter than the tree obtained with the actual data, plus one, divided by the number of replications. Following Faith and Cranston (1991), a significant phylogenetic signal is observed when P is less than 0.05, and here, the minimal value should be 0.003 (number of trees as short or shorter than the original tree + 1/300).

3. Results

The full data set contained 3330 nucleotides, 2169 from the mitochondrial DNA and 1161 from EF-1α. Single-codon gaps were found both in COI and COII. Two taxa, Parides chabrias and P. photinus, had gaps at the position 1984–1986 of COI in relation to Drosophila yakuba sequence, and Battus crassus and P. chabrias had gaps at the position 3423–3425 of COII. All Parides species showed a gap at the position 3458–3460, while only one Battus species (B. polydamas) showed a gap at that position. The alignment of EF-1α did not show indels. No differences were found in base composition among sequences within each of the partitioned genes (Table 4). However, the transition/transversion ratio among the three genes was quite different among codon positions (Table 4), and in the third codon position of the mitochondrial genes COI and COII this ratio was 0.7 and 0.9, respectively, suggesting transition saturation at this position. EF-1α sequence did not show this problem of saturation at third positions, although the difference among codons was strong (Table 4).

3.1. Phylogenetic analyses

3.1.1. Partitioned data

Parsimony analyses of the three genes separately resulted in different topologies (Fig. 1). The COI sequences resulted in five equally parsimonious trees, with 2416 steps (CI = 0.348; RI = 0.598). The strict consensus tree is shown in Fig. 1A. The analyses of COII resulted in one most parsimonious tree (Fig. 1B), with 978 steps (CI = 0.425; RI = 0.667), and EF-1α analyses resulted

---

**Table 4**

Summary of the sequence statistics over gene partitions

<table>
<thead>
<tr>
<th></th>
<th>All genes</th>
<th>COI</th>
<th>COII</th>
<th>EF-1α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of analyzed characters</td>
<td>3330</td>
<td>1527</td>
<td>642</td>
<td>1161</td>
</tr>
<tr>
<td>Number of invariant characters</td>
<td>2072</td>
<td>928</td>
<td>336</td>
<td>802</td>
</tr>
<tr>
<td>Number of variable characters</td>
<td>1258</td>
<td>599</td>
<td>306</td>
<td>359</td>
</tr>
<tr>
<td>Parsimony-informative sites</td>
<td>1000</td>
<td>480</td>
<td>233</td>
<td>287</td>
</tr>
<tr>
<td>CI</td>
<td>0.4</td>
<td>0.348</td>
<td>0.425</td>
<td>0.521</td>
</tr>
<tr>
<td>RI</td>
<td>0.655</td>
<td>0.598</td>
<td>0.667</td>
<td>0.773</td>
</tr>
<tr>
<td>Ti/Tv ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All positions</td>
<td>1.2</td>
<td>0.9</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>1st codon position</td>
<td>2.3</td>
<td>2.5</td>
<td>2.4</td>
<td>1.4</td>
</tr>
<tr>
<td>2nd codon position</td>
<td>1.2</td>
<td>1.6</td>
<td>1.1</td>
<td>0.7</td>
</tr>
<tr>
<td>3rd codon position</td>
<td>1.0</td>
<td>0.7</td>
<td>0.9</td>
<td>2.5</td>
</tr>
<tr>
<td>Frequencies of A, C, G, T</td>
<td>0.30, 0.15, 0.13, 0.42</td>
<td>0.34, 0.13, 0.10, 0.42</td>
<td>0.27, 0.26, 0.24, 0.23</td>
<td></td>
</tr>
<tr>
<td>Base frequencies</td>
<td>$\chi^2_{138} = 44.69; P = 1.0$</td>
<td>$\chi^2_{138} = 55.44; P = 0.99$</td>
<td>$\chi^2_{138} = 36.94; P = 1.0$</td>
<td></td>
</tr>
</tbody>
</table>
in 144 equally parsimonious trees (941 steps, CI = 0.521; RI = 0.773), with strict consensus shown in Fig. 1C. The homoplasy in the partitions was greater in the COI sequence (CI = 0.348) than in the COII sequence, and EF-1α showed the smallest homoplasy. Both COI and COII gave well-resolved trees with strong support at the tips of the tree (Figs. 1A and B), and EF-1α presented strong support in the internal nodes also (Fig. 1C).

3.1.2. Combined data

The ILD test results suggest that partitions of the data into COI + COII (mtDNA) and EF-1α were not incongruent ($P = 0.708$).

The total number of parsimony-informative sites in the full combined data set was 1000 (30%). COII showed the highest proportion of parsimony-informative sites, followed by COI and EF-1α (Table 4).
Parsimony searches over the equally weighted combined data set resulted in one most parsimonious tree, with 4368 steps (CI = 0.4; RI = 0.655) (Fig. 2). The tribe Troidini appeared as a monophyletic group. The genus Battus appeared as monophyletic and sister group to all remaining Troidini, supported by strong bootstrap and good Bremer values. Battus is divided in two groups (with strong bootstrap and Bremer values), one of them containing B. polydactylus, B. belus, and B. crassus, and the other containing B. polydactylus and B. philenor. The clade with the three paleotropical genera Troides + Byasa + Losaria is the sister of the remaining neotropical taxa (with strong bootstrap and Bremer support), but the relationships among these genera are still unclear. The genus Euryades is sister to the monophyletic genus Parides (with strong bootstrap and Bremer support). The Parides clade is in turn divided in four groups: group 1. ascanius + bunichus, with strong bootstrap and Bremer support; group 2. agavus + peneus, with weak bootstrap and Bremer support; group 3. chabrias + childrenae + photinus + sesostris + anchises + vertumnus, with no bootstrap support and weak Bremer support; and group 4. aeneas + tros + euryides + neophilus + zacynthus + lysander + panthonus, supported by strong bootstrap and Bremer values. Based on our sampling, groups 1 and 2 comprise species of Parides with tails on the hindwing that are
restricted to Southern South America (the only exception is the tailed P. tros, that appears in group 4). In addition, all species represented by more than one exemplar appeared as monophyletic entities.

The contribution of each gene to the combined tree, assessed by Partitioned Bremer Support, shows that there are few conflicting nodes. The COI data provide the greatest source of conflict, as found by Wahlberg et al. (2003) for the family Nymphalidae (Lepidoptera). COI sequences showed conflict in seven of 44 nodes, COII in five, and EF-1α in only one node.

Bayesian analyses for the combined data set became stationary well before generation 100,000. The topology of the tree was quite similar to that obtained by MP (Fig. 3), with the Bayesian tree differing in the relationships mainly among the species in group 4. In the Bayesian analysis, P. eurimedes is the sister taxon to P. zacynthus, with a low posterior probability (PP) (67%), conflicting with the sister relationship of P. neophilus + P. zacynthus found in the MP analysis. In both MP and Bayesian analyses, these three species formed a monophyletic clade with strong support (100% of bootstrap and PP support). Also different from MP, Bayesian analyses implied three groups within Parides, joining the groups 1 and 2 in a single clade (88% PP support). Clades which did not agree with MP trees showed weak or moderate PP support in the Bayesian analysis, such as P. proneus as the sister species of P. ascanius + P. bunichus (51% PP support) and Troides helena as the sister group of Euryades + Parides + Battus (59% of PP).

3.2. Analyses of character evolution

Character optimization of troidine ecological and morphological traits over the inferred MP molecular

---


---

Fig. 3. Phylogenetic tree inferred by Bayesian analysis from combined data set. Values above the branches indicate Bayesian posterior probability.
The ancestral states for the Aris-tolochia species were: 1, the use of many Aristolochia species as host-plant; 2, absence of long tail; 3, forest as main habitat; and 4, oviposition solitary or in loose group of several eggs (Fig. 4). The ancestral states for the Par-ides clade were similar, except for character 2—“presence or absence of long tail,” which showed an equivocal result (Fig. 4).

We found that the number of host-plant losses along the evolution of Troidini was higher than the number of host-plant gains, with four unambiguous events of host-plant losses and none of host-plant gain (Fig. 4). DELTRAN (which maximizes parallel changes) and ACCTRAN (which maximizes early gains) tracings were used to resolve the only ambiguity found (Maddison and Maddison, 1999). Both reconstructions resulted in five events of host-plant losses.

The presence of a long tail has unambiguously arisen four times along the evolutionary history of the Troidini. DELTRAN tracing indicated five gains of a long tail, and ACCTRAN tracing pointed out four gains and one loss. The shift of primary habitat from forest to scrub or to open habitat has unambiguously occurred three times and once, respectively (Fig. 4), and the oviposition of tight bunches of many ordered eggs is inferred to have arisen once along the evolution of troidine butterflies.

The phylogenetic signals for “use of host-plants” and “gregariousness of the immatures” were both strong (P = 0.003 for both characters), suggesting that the distribution of these traits among taxa can be explained by their phylogenetic relationships. The phylogenetic signals of the characters “presence or absence of long tail,” and “primary habitat,” were both marginally significant (P = 0.053 and P = 0.056, respectively).
4. Discussion

4.1. The basal Troidini

Our results for the intergeneric relationships among the Troidini largely corroborate those obtained by prior morphological researchers (Hancock, 1983; Miller, 1987; Munroe, 1961) and the molecular hypothesis of Caterino et al. (2001). The general topology is maintained even with the inclusion of additional taxa in some clades (in our case, Euryades, Parides, and Battus). However, this topology has some weakly supported branches in both this study and in Caterino et al. (2001), especially the branch of the IndoAustralian species, which is poorly sampled in both studies. A complete sampling of all Troidini genera, including representatives of all major lineages, will be necessary before we can fully understand the relationships among the troidine genera, considered to be of the greatest uncertainty within the Papilionidae (Vane-Wright, 2003).

The position of Battus as basal to all remaining Troidini agrees with Munroe (1961), Hancock (1983), and Miller (1987), each of whom placed this genus in the subtribe Battina, sister of the subtribe Troidina (which includes all troidine genera except Battus), corresponding also with Caterino et al.’s (2001) hypothesis. Munroe and Ehrlich (1960), in addition, considered Battus the most distinctive genus in the tribe based on several singular morphological features, but their study leaves unclear if Battus or Euryades + Cressida should be the most basal troidines (neither included in Caterino et al., 2001). Those results disagree with those of Morinaka et al. (1999), based on the ND5 gene, that showed Battus closer to the Graphiini, far from all remaining Troidini. That result is weakly supported, however, and is likely due to insufficient taxonomic and character sampling, given that they sequenced only a short mitochondrial region. The internal relationships among Battus reported here agree completely with Racheli and Oliverio (1993) (including all species in the genus), who also found the genus Battus divided into two major groups.

We found that Euryades occupies a sister group position with Parides, differing from Miller (1987), who placed it with a group of Paleotropical genera. This close association agrees with Hancock’s (1983) hypothesis, who suggested a sister relationship among (Cressida + Euryades) + Parides. The support in the clade Euryades + Parides could eventually be improved by the inclusion of Cressida and the other species of Euryades, as well as better representation of other Old World troidine genera.

4.2. The genus Parides

This paper presents the first detailed phylogenetic hypothesis for the genus Parides, including 17 of the 34 recognized species representing all of the subgeneric groups recognized by Tyler et al. (1994) (all other studies published to date have examined only 1–4 species (Caterino et al., 2001; Morinaka et al., 1999, 2000; Tyler et al., 1994)). Our more intensive sampling makes possible the identification of the major lineages (groups 1–4) within the genus. The basal position of group 1 (P. bunichus + P. ascanius) was in part suggested by Tyler et al. (1994, p. 179), who argue that the restriction to a single host-plant species and the presence of tails in the hindwings of Parides ascanius (a species not included in the cladograms in that work) suggest that this is the most basal species in the genus. The strongly supported sister relationships between P. ascanius and P. bunichus is sustained both by their preference for scrub habitats over forest (Tyler et al., 1994), and the possibility of hybridization in the field and laboratory (Otero and Brown, 1986). Parides ascanius is very similar morphologically in adult features to P. bunichus, especially P. bunichus chamissonia, including genitalia and minor elements of wing color-pattern (Otero and Brown, 1986; Tyler et al., 1994). A complete study on the relationships of all subspecies of the P. bunichus–P. ascanius group could help to clarify the patterns of diversity within this clade.

Group 2 is now composed of P. proneus and P. aga- tus, two species with tails on the hindwings also restricted to the southern Neotropics. The results of Bayesian analysis show groups 1 and 2 forming a single clade basal to all remaining Parides, with all species sharing the presence of tails on the hindwings and geographic distribution restricted to the southern Neotropics. This group could also include some additional tailed species such as P. phalaecus (Ecuador), P. montezuma (Mexico), P. gundlachianus (Cuba), and P. alopius (NW Mexico), according to Tyler et al. (1994).

Groups 3 and 4 include all remaining Parides, many with wide distributions in Amazonia and Central America. The clade eurimedes + zacynthus + neophilus in group 4 is the only point where different analyses disagreed. This still unresolved clade includes one broadly distributed species (P. neophilus) and two restricted species; P. zacynthus from the coastal Atlantic sand forests, and P. eurimedes with Transandean distribution. Future work with multiple populations of these three species could add important information about the patterns of colonization of different habitats by Parides in the Neotropics, and the relevance of this to formation of new subspecies.

In our results, P. chabrias is included in group 3, with no bootstrap support and just a weak Bremer support in MP analyses, and a strong PP value in Bayesian analyses. Tyler et al. (1994) included P. chabrias in a group of species with unusual wing-shape in males, including also P. halmelii (tailed), P. quadratus, P. pizarro, P. vercingetorix, and P. klagesi. In our results, P. chabrias could be
placed in group 3, but a more basal position, forming a separate species group, cannot be ruled out. It is intriguing that *P. chabrias* and putative relatives belong to a separate mimicry complex that includes the ithomines *Methona* and *Thyridia*, rather than the black with pink-and-green or white-spots-pattern typical of most of the “derived” *Parides* species.

### 4.3. Analyses of character evolution

This section is mostly based on the two characters with a strong phylogenetic signal, “use of host-plants” and “gregariousness of the immatures.” The characters “presence or absence of long tail,” and “primary habitat,” both with *P* values above 0.05, are only briefly discussed.

#### 4.3.1. Use of host-plants

The concept that specialization can be an evolutionary dead end has been central in host-plant/herbivore evolutionary studies. Many factors can constrain the use of plants as food, such as secondary compounds found in them (Bernays, 1998; Futuyma et al., 1993; Jai-nike, 1990), female oviposition preferences (Ronquist and Nylin, 1990), and the geographical distribution of species (Dobler et al., 1996; Kelley and Farrell, 1998; Pasteels and Rowell-Rahier, 1991). However, it seems that over evolutionary time, diet breadth in insects (Colwell and Futuyma, 1971) has both increased and decreased (Bernays, 1998), sometimes leading to specialization (Kelley and Farrell, 1998; Moran, 1988; Nosil, 2002; Ronquist and Nylin, 1990), and others leading to generalism (Armbuster and Baldwin, 1998; Janz et al., 2001; Scheffer and Wiegmann, 2000).

Based upon the available data for host-plant use, the ancestral state of host-plant use for both Troidini and *Parides* is the use of many *Aristolochia* species (Fig. 4), with a tendency to advance towards increased specialization. The present results show that terminal taxa usually feed on fewer *Aristolochia* species compared with basal taxa, and this pattern agrees with that found by Kelley and Farrell (1998). *Parides ascanius* could be suggested as an exception to this pattern, since it belongs to a basal clade of *Parides* and is a specialist in *Aristolochia macroura* (a fact confirmed through extensive field and laboratory data, including experiments with other available species of potential host-plants (Otero and Brown, 1986, L.S. Otero, pers. comm.). These conclusions could be tested with the addition of more species of tailed *Parides* (see above) to confirm if they are part of this clade making *bunichus-ascanius* a terminal clade in group 1. In addition, the hypothesis that specialists are more sensitive than generalists to changes in abundance of their host-plants (Futuyma and Moreno, 1988) could also be tested for *P. ascanius*. Otero and Brown (1986) argue that the main factor threatening this species is habitat destruction rather than host-plant availability (*A. macroura* is common in most swampy coastal habitats in SE Brazil). The rarity of *P. ascanius* could be an example of habitat fidelity rather than host fidelity (Dobler et al., 1996), since this species is specialized not only to its host-plant, but also to the physical and biotic environment where this plant grows (Bernays, 1996).

The trend toward specialization among the troidine butterflies examined here could be a result of the geographic distribution of the species, reflecting the pattern proposed by Weintraub (1995): species with restricted geographical ranges tend to be specialists, while those with broad geographical ranges are usually generalists, and in this case, the diet breadth is mirroring the host-plant availability and abundance (Pasteels and Rowell-Rahier, 1991). This could explain monophagy of *Parides ascanius* and *Parides panthonus jaguarae*, two species with very restricted distributions (the first in swampy areas of the coastal plain in Rio de Janeiro and the second in narrow galley forests in central Minas Gerais). It is also interesting to note that, even if we are considering a probable subspecies here (*P. panthonus jaguarae*), this is not a factor of bias in our results, since another subspecies of this taxa, *P. panthonus aglaope*, is known to feed on up to five species of *Aristolochia* through its geographic range (Moss, 1920; Tyler et al., 1994). Moreover, additional data on host-plant use in *Parides* could bear upon the hypothesis of Fox and Morrow (1981), who argued that a species that uses many hosts over its geographical range could in fact be using only one or a few host-plants in each site. Gomez-Zurita et al. (2000) found that specialization in *Timarcha* beetles is dependent on geographical distribution and thus on the availability of host-plants. Again, *Parides panthonus*, with seven known subspecies, could be a good taxon for testing this hypothesis. Within this species, five subspecies have “broad” geographic distributions (*P. panthonus panthonus*, *P. p. aglaope*, *P. p. ecaudatus*, *P. p. callicles*, and *P. p. lysimachus*) and two are very restricted (*P. panthonus jaguarae* and *P. panthonus castilhoi*). Knowledge about host-plant use within these seven geographic entities could help to understand the relationships between distribution and host associations.

#### 4.3.2. Hindwing tails

The significance of particular characters of butterfly wings, including tails, has been widely discussed, but no solid conclusions have been attained (Wootton, 1992). Our objective in the present paper is to map the presence of long tails in species of Troidini, especially *Parides*, to check an idea proposed in Tyler et al. (1994, p. 179), that the tailed *Parides* are the basal species of the genus. Even though the species of tailed *Parides* appear to belong mainly to basal clades (with the exception of *P. tros*, that is tailed and not basal), the
absence of phylogenetic signal in this character makes it difficult to draw additional inferences. Again, additional species of tailed *Parides* could help to clarify the evolution of this morphological feature. Among Papilionidae as a whole, tails appear to have evolved and disappeared on multiple occasions, with the most basal taxa (e.g., *Baronia*, Paprassiiinae) lacking them, and all of the more derived tribes containing both tailed and tailless species.

4.3.3. Habitat

There are no a priori hypotheses about ancestral habitats of Troidini. Based on our results, the Troidini is a group that originated from forest ancestors; colonization of open habitats and scrub forest was secondary in the group. This pattern is the same for all analyzed genera except *Euryades*, whose two known species are specialized on open habitats. The Australian genus *Cressida*, a proposed sister group of *Euryades* (see above) occurs in a variety of habitats including grassland and open forests. If this sister group relationship is confirmed, it could be reasonable to infer that the *Cressida Euryades* clade had its origins from open forest/grassland ancestors with Austral origin.

4.3.4. Gregariousness of immatures

According to our results, “immatures solitary or in loose groups” is the plesiomorphic state for this character, and since all known Troidini immatures are considered aposemetic (and probably all are chemically defended, Klitzke and Brown, 2000; Tyler et al., 1994), it is likely that, at least in Troidini, evolution of aposematism has occurred prior to the evolution of gregariousness (as proposed by Sillen-Tullberg, 1988). Gregariousness thus appeared in the ancestor of the genus *Battus*, as this character is shared by all known species in this genus (Tyler et al., 1994). Also, the results show that gregariousness has arisen only once in Troidini, a conclusion different from that stated by Sillen-Tullberg (1988), who proposed a minimum of two (or three) events of evolution of gregariousness in this tribe. Additionally, host-plant features cannot be neglected when discussing the evolution of this trait, strongly related with advantages of group feeding behavior in some species of *Battus* (Fordyce and Agrawal, 2001; Fordyce and Nicsce, 2004).

4.3.5. Concluding remarks

The present study clarifies the major internal relationships of the genus *Parides* and provides a useful hypothesis to test ecological and biogeographical theories about the evolution of this group. There are many questions still open, however, that could guide future investigation in this group of butterflies.

Future work to help reveal the internal relationships within the genus *Parides* could include more species of *Parides* in the data set (covering some of the lacunae indicated above, especially additional taxa in the *chabrias* group of Tyler et al. (1994) and “tailed *Parides*” from N and S sectors), and bring in new sources of information, including different genes (such as wingless and 28S) and morphological characters of a wide range of adults and immatures. The available phylogeny could be useful in mapping host plant chemistry, helping to understand host plant use in different lineages within this genus.

The Troidini are highly diversified, easy to study, and part of various mimetic rings in tropical regions of the world. Even though much research has already been done with troidines, there are several interesting aspects to be still investigated in the future, such as age of the group in relation to its host plants and diversification in the different continents, detailed ecology of immatures and adults and the internal relationships in polytypic species (usually as numerous distinct geographical populations).

Acknowledgments

We gratefully acknowledge the invaluable assistance of the following agencies: Instituto Florestal, Fundação Florestal, Fundação José Pedro de Oliveira, IBAMA, and Hemocentro-UNICAMP. This work was supported in part by the Harold and Leona Rice Endowment for Systematic Entomology at the Oregon State University, by Fapesp (BIOTA-FAPESP Program, Grants 98/05101-8, 00/01484-1, and 04/05269-9 to AVLF), by the National Science Foundation (DEB-0089870 to AVZB; DEB-0316505 to AVLF), and by CAPES (fellowship to KLBS). We are grateful to Debra Murray for her assistance in the lab. Felix Sperling greatly encouraged our work since 2000, made many helpful suggestions all along the project and provided a supply of important specimens to conclude this work. We also thank Keith S. Brown Jr. and Niklas Wallberg for comments on the manuscript.

References


Haae, E., 1892. Untersuchungen über die Mimikry auf Grundlage eines natürlichen systems der Papilioniden. 2. Untersuchung über die Mimikry. Nägele, Stuttgart


