JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH

Accepted on 9 March 2011
J Zool Syst Evol Res doi: 10.1111/j.1439-0469.2011.00620.x

[1]*Department of Integrative Biology, University of Guelph, Guelph, ON, Canada;* [2]*Jacobys alle 2, 1806 Frederiksberg C, Denmark;* [3]*Department of Entomology, Natural History Museum, London, UK;* [4]*Centre de Recherche d' Orléans, INRA, UR 633 Zoologie Forestière, Orléans, France;* [5]*University Museum of Zoology, Cambridge, UK;* [6]*57 rue Genot, B-4032 Chênée, Belgium;* [7]*Kerkhoflei 54, Hofstade (Zemst), Belgium*

Phylogenetic systematics of *Colotis* and associated genera (Lepidoptera: Pieridae): evolutionary and taxonomic implications

Vazrick Nazari[1]*, Torben B. Larsen[2], David C. Lees[3,4], Oskar Brattström[5], Thierry Bouyer[6], Guy Van de Poel[7] and Paul D. N. Hebert[1]

Abstract

We investigated the genetic diversity and phylogenetic placement of the butterflies in the genus *Colotis* and eight related pierid genera using sequence information from two mitochondrial and two nuclear genes. To establish the status of species, we initially barcoded 632 specimens representative of all genera and most species and subspecies in those genera. A subset was then selected for phylogenetic analysis where additional gene regions were sequenced: 16S rRNA (523 bp), EF-1α (1126 bp) and wg (404 bp). DNA barcode results were largely congruent with the traditional classification of species in the *Colotis* group, but deep splits or lack of genetic divergence in some cases supported either species-level differentiation or synonymy. Despite using information from four genes, the deeper nodes in our phylogeny were not strongly supported, and monophyly of the '*Colotis* group' and the genera *Colotis* and *Eronia* could not be established. To preserve the monophyly of *Colotis*, we revive the genus *Teracolus* for three outlying species previously in *Colotis* (i.e. *Colotis eris*, *Colotis subfasciatus* and *Colotis agoye*), as well as the genus *Afrodryas* for *Eronia leda*. The position of *Calopieris* is unresolved although it appears to be well outside the molecular variation in *Colotis* (s.l.). A dispersal/vicariance analysis suggested that major diversification in *Colotis* (s.str.) occurred in Africa with subsequent dispersal to India and Madagascar.

Key words: Pieridae – *Colotis* – phylogenetic – DNA barcoding – taxonomy

Introduction

The Pieridae include some of the most familiar butterflies, the cabbage whites and the grass yellows, yet the apparently long stable status of many of the species in this family are yet to be clarified by molecular means. Several recent papers have used information from gene sequences to shed light on the taxonomy or evolutionary history of the family or lower ranks therein (Morinaka et al. 2002; Braby 2005; Braby and Trueman 2006; Braby et al. 2006, 2007; Chew and Watt 2006; Braby and Pierce 2007; Wheat et al. 2007; Xu et al. 2007; Wheat and Watt 2008; Suárez et al. 2009). However, many groups remain largely unexamined. One such group consists of the genus *Colotis* and eight closely related genera (i.e. *Gideona*, *Eronia*, *Nepheronia*, *Pareronia*, *Ixias*, *Pinacopteryx*, *Hebomoia* and *Calopieris*), traditionally placed in different tribes within the subfamily Pierinae but brought together by weakly supported molecular phylogenies and tentatively named the '*Colotis* group' because of inferred paraphyly (Braby et al. 2006).

There is considerable historical confusion over the status of many taxa in this group, and various revisions present differing classifications and species numbers (Butler 1876; Sharpe 1898–1902; Aurivillius 1925; Talbot 1939; Peters 1952; Ackery et al. 1995). Aurivillius (1925) placed '*Herpaenia*' (= *Pinacopteryx*) in the Pierinae but identified three genera under the subfamily 'Teracolinae' (i.e. *Teracolus*, *Calopieris* and *Eronia*). He also identified three groups within the genus *Teracolus* (= *Colotis*), one of which he further divided into 12 subgroups based on wing colour patterns. Klots (1933) regarded *Colotis* and *Ixias* to be derived from Anthocharidini, considered *Eronia*, *Nepheronia* and *Pareronia* closer to Coliadinae, and expressed uncertainties about the placement of *Hebomoia* and *Pinacopteryx*. Talbot (1939), who used the form and colour of the apical spot as the basis for his classification, recognized *Gideona* as a separate genus and identified 16 species-groups within *Colotis*, one of which was *Colotis* (= *Calopieris*) *eulimene*. Peters (1952) also recognized *Eronia* and *Nepheronia* as separate genera within Teracolinae. Braby et al. (2006) stated, 'These (nine) genera may well comprise a separate lineage sister to the rest of Pierinae, but evidence for their monophyly is lacking'. In case the monophyly of the *Colotis* group is successfully established, the tribal name Teracolini Reuter 1896 is available (Hesselbarth et al. 1995; Braby et al. 2006).

The group is entirely Old World and, despite the powerful flight ability of some of its members, exhibits considerable endemism. *Hebomoia* (2 spp.), *Ixias* (10 spp.) and *Pareronia* (13 spp.) are distributed in India and the Oriental region, *Gideona* (1 sp.) is endemic to Madagascar, while *Calopieris* (1 sp.), *Eronia* (2 spp.), *Nepheronia* (4 spp.) and *Pinacopteryx* (1 sp.) are confined to Africa and the Arabian Peninsula. *Calopieris* is endemic to the Somali subregion and its only species, *Calopieris eulimene*, is extremely rare in collections, and its position within the *Colotis* group has not been substantiated through molecular data (Braby et al. 2006).

With about 46 species, *Colotis* is a relatively large genus with a centre of richness in the east African Savannah zone. Outside

Corresponding author: Vazrick Nazari (vazrick.nazari@agr.gc.ca)
Contributing authors: Torben B. Larsen (torbenlarsen@btinternet.com), David C. Lees (dclees@gmail.com), Oskar Brattström (ob269@cam.ac.uk), Thierry Bouyer (tbo@swing.be), Guy Van de Poel (ag.vdp@telenet.be), Paul D. N. Hebert (phebert@uoguelph.ca)

*Present address: Agriculture and Agri-Food Canada, 3058-C KW Neatby Bldg, 960 Carling Avenue, Ottawa, ON, K1A 0C6 Canada.

of Africa, five species are endemic to the Madagascar subregion (i.e. *Colotis evanthides*, *C. evanthe*, *C. guenei*, *C. mananhari* and *C. zoe*), one (*C. evagore*) occurs as far north as southern Spain, many penetrate the Arabian Peninsula and Iran, and seven species fly in India and Sri Lanka (i.e. *Colotis amata*, *C. danae*, *C. eucharis*, *C. fausta*, *C. liagore*, *C. phisadia* and *C. vestalis*), where *C. etrida* and *C. protractus* are endemic. Species of *Colotis* are one of the most prominent insect components of the savannah zone (*sensu* Larsen 1984); their larvae feed on various Capparaceae as well as *Salvadora persica* (Salvadoraceae), and clouds of them are often seen around the stands of their food plant or on other flowering plants. Most *Colotis* species are morphologically diverse and exhibit pronounced seasonal, sexual and individual variation. Wet- and dry-season forms of some species are phenotypically very different although transitions are common between the two extreme seasonal generations (Talbot 1939). Over the centuries, hundreds of infra-specific names have been applied to this wide spectrum of variation, and although many have been synonymized, some uncertainties remain. Some species within *Colotis* have been segregated at subgeneric levels, including *C.* (*Madais*) *fausta*, *C.* (*Cuneacolotis*) *agoye*, *C.* (*Teracolus*) *eris* and *C.* (*Teracolus*) *subfasciatus*, but these have not gained much support. Prior work on wing scale structure, pigmentation and iridescence in *Colotis* and related genera has also revealed a wide range of invisible (ultraviolet) sexual dimorphism and patterns invisible to the human eye (Stavenga and Arikawa 2006; Stavenga et al. 2006; Stavenga and Leertouwer 2007; Wijnen et al. 2007), but the lack of phylogenetic hypotheses for the origin and evolution of *Colotis* and related genera has been a stumbling block in understanding how these patterns may have evolved.

The occurrence of both widespread and narrowly restricted species of *Colotis* in three main zoogeographical regions makes the genus particularly attractive to investigate from a historical biogeography perspective. The Malagasy and Indian endemics present several intriguing questions: Was vicariance or dispersal the main process driving evolution of non-African endemic *Colotis*? Is it possible that, as in some other insects, *Colotis* originated in Madagascar and spread later to Africa and India (c.f. Zakharov et al. 2004; Monaghan et al. 2005)? Alternatively, if the ancestors of the non-African *Colotis* arrived through dispersal from Africa (c.f. Kodandaramaiah and Wahlberg 2007; Lohman et al. 2008; Aduse-Poku et al. 2009), how many dispersal events were involved? In this paper, we (i) use DNA barcoding to test the status of the available taxa in *Colotis* and related genera and (ii) attempt to reconstruct a phylogeny for the group with the available molecular data, and use it to test the higher level taxonomy and determine the origins of the genera within the *Colotis* group.

Materials and Methods

Taxon sampling

A total of 632 specimens representing all but one species in the genus *Colotis* as well as representatives from other genera in the '*Colotis* group' were examined (Table S1). In most cases, a dry leg accompanied by an image of the specimen was sent to the Biodiversity Institute of Ontario (Guelph, Canada). Vouchers were retained in the originating collections, and the voucher data are publicly available through the published project '*Colotis* of the World' (CLT) on the Barcode of Life Data Systems (BOLD; http://www.barcodinglife.org). Outgroups representing all subfamilies of Pieridae were selected from previous studies and sequences were obtained from GenBank (Table S2; Caterino and Sperling 1999; Caterino et al. 2001; Wahlberg et al. 2005; Narita et al. 2006; Braby et al. 2006, 2007; Nazari et al. 2007), and the phylogenetic trees were rooted using the Dismorphiinae + Pseudopontiinae clade, the well-established sister to Coliadinae + Pierinae (Braby et al. 2006). Because of the narrow focus of the study, no outgroups were selected from other lepidopteran families.

Among others, we could not sample the following taxa: *Colotis eunoma*, a rare dune specialist closely related to the *C. ione* group; *Nepheronia buquetii buchanani* from North Africa and the Arabian Peninsula, *C. evagore evagore* from Arabia; *C. evagore niveus*, endemic to the island of Socotra in Yemen and sometimes considered a separate species based on differences in morphology and genitalia; *C. agoye zephyrus* from Somalia; and *C. vesta amelia* from Western Africa.

Molecular techniques

The extraction of total genomic DNA, amplification and sequencing were performed in the Biodiversity Institute of Ontario using previously described protocols (Ivanova et al. 2006). Full-length mtDNA barcode sequences (i.e. 658 bp) were obtained for nearly all specimens, and based on results from sequence similarity (Neighbour-Joining) analyses and the quality of DNA, a subset of specimens was selected for additional gene sequencing (Table S2). Older or failed samples were targeted using six overlapping primer pairs designed for cytochrome oxidase I (COI) (Hausmann et al. 2009). Partial sequences from 16S ribosomal rRNA and two nuclear genes – Elongation Factor 1-alpha (EF-1α) and wingless (wg) – were also obtained using primers and protocols described previously (Brower and De Salle 1994; Aubert et al. 1999). Amplified DNA from all specimens was bidirectionally sequenced for each gene, and final sequencing products were run on an ABI 3730XL® DNA analyzer (Life Technologies, Foster City, CA, USA). Complementary strands were assembled into contigs and edited manually, and primers were removed using SEQUENCHER 4.5 (Gene Codes Corporation, Ann Arbor, MI, USA). Sequences were aligned using CLUSTALX 2.0 (Thompson et al. 1997), evaluated by eye and converted to Nexus using SE-AL 2.0a11 (Rambaut 2002). The 16S rRNA alignment was unambiguous and remarkably variable in the conserved sections. To eliminate noise caused by questionable homologies in loops in the secondary structure models of the 16S rRNA, we retested the data partition after excluding problematic bases without removing doublet characters using the program GBLOCKS 0.91b (Talavera and Castresana 2007); however, the results were similar. New sequences were deposited in GenBank, and accession numbers are provided in Tables S1 and S2. All data are also available publicly through the published project 'CLT' on the Barcode of Life Data System (BOLD; http://www.barcodinglife.org).

Phylogenetic analyses

Neighbour-Joining (NJ) trees for barcode data were constructed initially using the QUICKTREE algorithm (Howe et al. 2002) and under the Kimura two-parameter (K2P) model (Kimura 1980). Additional NJ and maximum parsimony (MP) analyses were conducted in PAUP* 4.0 β 10 (Swofford 2002). Heuristic searches for MP analysis were carried out with all characters equally weighted and under the tree bisection-reconnection (TBR) swapping algorithm with 100 random addition sequences. Bootstrapping of 100 replicates was conducted under the parsimony criterion with the default setting starting with a random seed and the TBR branch-swapping algorithm. Bremer support values were calculated using TREEROT v.3 (Sorenson and Franzosa 2007). The maximum likelihood (ML) tree was generated using PHYML 3.0 online (Guindon and Gascuel 2003), with the parameters of the best-fit model (GTR + Γ + I) selected previously using MULTIPHYL (Keane et al. 2007) and 100 bootstrap replicates. Haplotype diagrams (Figure S1) were constructed in TCS 1.21, with a 90% or 95% confidence limit for parsimony (Templeton et al. 1995). Shorter fragments of COI barcodes or those with ambiguous bases were excluded from haplotype analyses.

Bayesian posterior probabilities for the combined data set were first calculated using MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003) under the GTR + Γ + I model using one cold and three heated simultaneous Markov chain Monte Carlo (MCMC) chains, starting

with random initial trees and sampling every 100 generations. The analysis was allowed to continue for 5 500 000 generations until the average standard deviation of split frequencies fell below 0.01. Substitution rates were estimated as part of the analysis from default priors, and model parameters were allowed to vary. A total of 5000 trees corresponding to the burnin values estimated prior to initiation of each MCMC chain were discarded, and the majority rule consensus tree was generated using the remaining trees with posterior probabilities plotted on each node.

The same data set and model specifications were used to perform a parallel analysis in BEAST (Drummond and Rambaut 2007; tree not shown). Branch lengths were allowed to vary according to an uncorrelated lognormal distribution. The Yule process was selected as the tree prior. The MCMC analysis was run twice under default priors for 10 000 000 generations, sampling the chains every 1000 generations and yielding a total of 10 000 samples for each run, the first 1000 of which were later discarded as burnin. Node posterior probabilities and standard deviations were computed for each internal node using the Tree Annotator module implemented in BEAST. The estimated posterior probabilities by MrBayes and BEAST were comparable (Table S4). The decay values for each data partition were also calculated using TREEROT and are presented alongside the posterior probabilities in Table S4.

Dispersal/vicariance analysis

From the African diversity centre, an African origin for *Colotis* with later spread to India and Madagascar seems likely. This hypothesis was tested through a dispersal and vicariance analysis with DIVA (Ronquist 1997). Given a phylogeny, this method has frequently been used in the reconstruction of ancestral distributions in butterflies (e.g. Wahlberg et al. 2005; Kodandaramaiah and Wahlberg 2007; Nazari et al. 2007; Wiemers et al. 2009). Non-*Colotis* taxa were excluded from the DIVA analysis because of limited sampling. Unit areas were selected based on biogeographic zones for African butterflies (Carcasson 1964; Larsen 1984, 1991). Distributional data for each species were compiled in a Nexus file in MESQUITE 2.72 (Maddison and Maddison 2009) as presence/absence for each region, with the Bayesian phylogeny used for the analysis. Analyses were conducted with and without restriction of maximum number of areas for ancestral nodes. DIVA assigns a cost of zero to vicariance (i.e. allopatric speciation) and duplication (i.e. sympatric speciation) events and a cost of 1 per unit area to any dispersal and extinction events; thus, the best reconstructions are those that minimize the number of dispersals and extinctions under a parsimony criterion. To avoid accumulation of distribution areas towards the root of the phylogeny, constraints of 2, 3, 4 and 5 unit areas were imposed as the ancestral distribution, and the best construction (with the least number of dispersals = 25) was obtained when the maximum number of unit areas in ancestral distributions was set to 2.

Results

Attempts to sequence the selected genes for all species were not successful because of a lack of fresh material, and for some species only the COI barcodes were available for phylogenetic analysis (Table S2). The combined data set included 2812 positions, of which 1761 were constant, 187 were uninformative and 864 were parsimony-informative. The wg and the COI barcode partitions had a higher proportion of parsimony-informative characters (39.6% and 38.3% respectively) than 16S (25.2%) or EF-1α (26.1%) (Table S3).

Support for the deeper nodes in our phylogenetic trees of the combined data was generally weak or lacking, but many nodes were consistently recovered through various reconstruction methods (i.e. MP, ML, Bayesian) (Table S4). Our BEAST phylogeny (not shown) separated Pseudopontiinae from the rest of the taxa with very long branches near the base and made Pierinae paraphyletic by placing the Coliadinae within the subfamily. We did not observe a similar pattern in our ML or MrBayes trees, and these were more consistent with previous studies with better taxon sampling (Wheat et al. 2007).

The phylogenetic positions of several taxa were unstable throughout ML or Bayesian analyses, including *C. eulimene*, *Hebomoia glaucippe*, *Ixias pyrene* and *Eronia leda* (Figs 1 and 2). Across our analyses, *Calopieris* appeared as either sister to Pierinae, Anthocharidini or *Hebomoia* with no support. *Eronia leda* strayed from its supposed congener, *E. cleodora*, and appeared either close to *Pareronia* or as a sister to the *Colotis aurora* clade, in both cases with no support. The GenBank COI sequences for *Pareronia valeria* (AY954573, Braby et al. 2006) and *P. anais* (EF584868, Xu et al. 2007) are identical and possibly reflect misidentification.

None of the gene partitions or the combined analyses recovered the nine genera in the '*Colotis* group' as monophyletic. The genus *Colotis* was also not monophyletic, with three species (i.e. *C. agoye*, *C. subfasciatus* and *C. eris*) consistently forming a separate clade, which stayed outside the remaining *Colotis* and closer to *Gideona* and *Pinacopteryx*. The monophyly of a group consisting of all *Colotis* except [*C. eris* + (*C. subfasciatus* + *C. agoye*)] was supported by ML analysis, within which several distinct and relatively well-supported subclades were observed (Fig. 2):

Group I. *etrida*, *ephyia*.
Group II. *aurora*, *evarne*, *dissociatus*, *auxo*, *incretus*.
Group III. *antevippe*, *rogersi*, *euippe*, *pallene*, *lais*, *daira*, *evagore*, *evanthe*, *evanthides*.
Group IV. *liagore*, *evenina*.
Group V. *danae*, *annae* (*hildebrandti*), *guenei*.
Group VI. *protractus*, *fausta*, *amata*, *calais*, *vestalis*, *phisadia*.
Group VII. *ungemachi*, *doubledayi*, *chrysonome*, *vesta*, *aurigineus*.
Group VIII. *zoe*, *celimene*.
Group IX. *protomedia*, *halimede*, *pleione*, *venosa*, *mananhari*, *regina*, *hetaera*, *elgonensis*, *ione*, *erone*, (*eunoma*).

The position of *C. ungemachi* as sister to two clades consisting of the more similar *C. pleione* group (VII) as well as the distinct *C. zoe* + *C. celimene* group (VIII) was supported only weakly by COI barcodes. Since *C. ungemachi* morphologically clearly belongs to the *pleione* (rather than *zoe*) species group, and considering the lack of nuclear sequences to support its current position, we included this species as part of (currently paraphyletic) *Colotis* group VII.

The MP analysis on EF-1α partition divided the *Colotis* group IX into two clades, consisting of (i) *C. protomedia*, *C. halimede* and *C. pleione*, and (ii) *C. mananhari*, *C. elgonensis*, *C. ione*, *C. hetaera* and *C. regina* (Fig. 3). This pattern also corresponds better with morphology, but it was not supported by the other genes or by the combined analysis.

Discussion

Systematics and biogeography

Despite exhaustive taxon sampling of the *Colotis* group and the potential combined support from two nuclear and two mitochondrial genes, the lack of resolution at the deeper nodes of our phylogeny is striking and highlights the possible

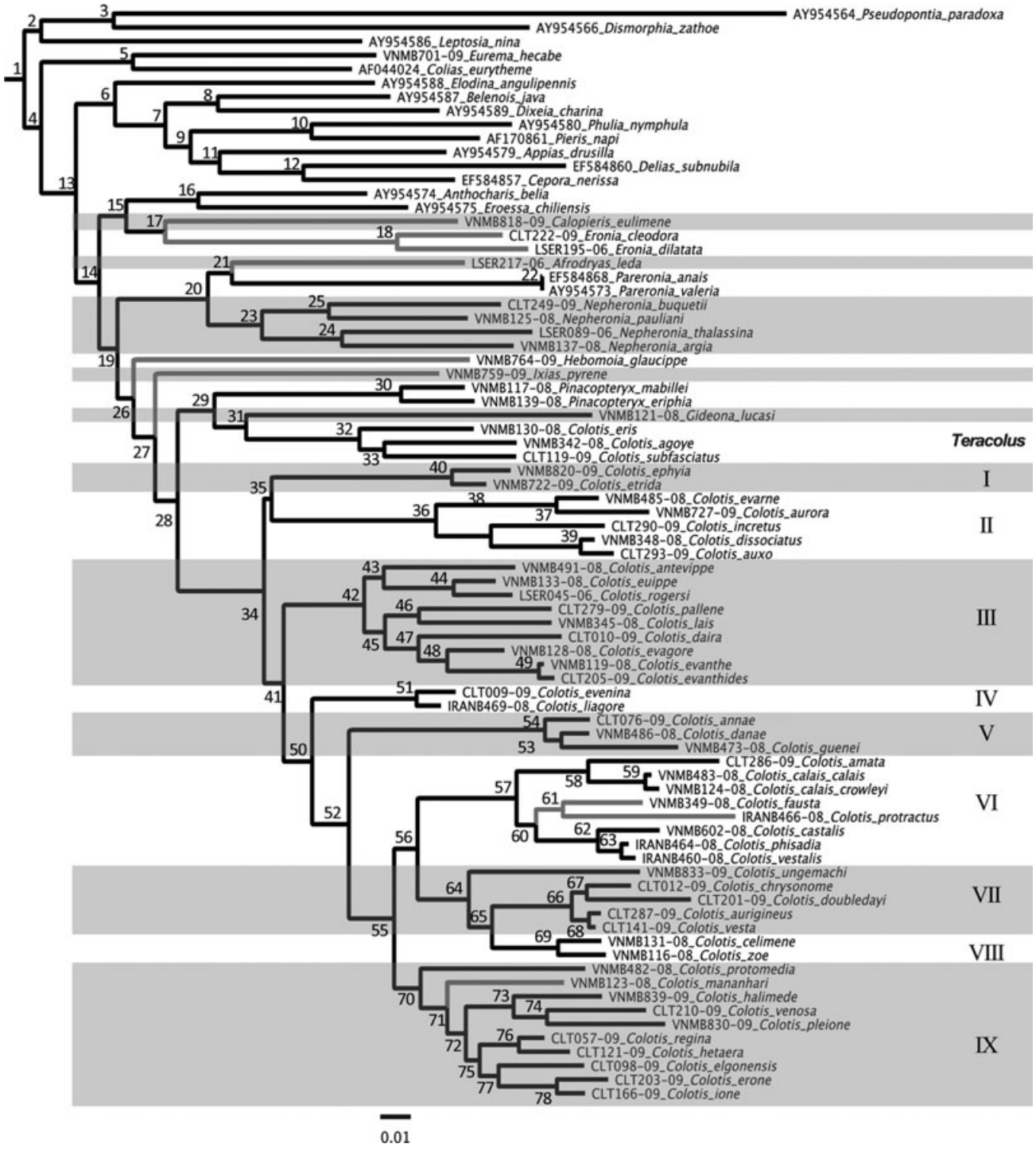

Fig. 1. Maximum likelihood (ML) phylogeny of combined data (Tree Length = 6818 steps, CI = 0.2232, RI = 0.4658, −ln *L* = −33610.72878). Nodes are numbered, and support values are presented in Table S4. Inferred groups of *Colotis* are identified with Roman numerals. Red branches show taxa with a varying position between ML and Bayesian phylogenies (Fig. 2)

inapplicability of this range of markers for inferring phylogenies at least in this group of butterflies. Some factors that may explain the low support values include the following: (i) poor outgroup sampling; (ii) extremely old age of radiation beyond the range of the markers; (iii) extinction of intermediate lineages or species and presence of floating taxa, like *Gideona* that reduce the overall support at the base or among the ingroup taxa; and (iv) presence of hard polytomies resulting from rapid speciation, where very short internal branches are followed by much longer branches (Kodandaramaiah and Wahlberg 2009; Kodandaramaiah et al. 2010). Some of these effects can be potentially alleviated by adding more basal taxa or removing rogue taxa from the analysis. However, our outgroup taxa included all major groups of Pieridae, and removal of the rogue taxa (e.g. *Hebomoia*, *Ixias* and *Calopieris*) did not affect the overall support values in a substantial way (data not shown). No hard or near-hard polytomies were evident in our trees, and previous studies employing molecular

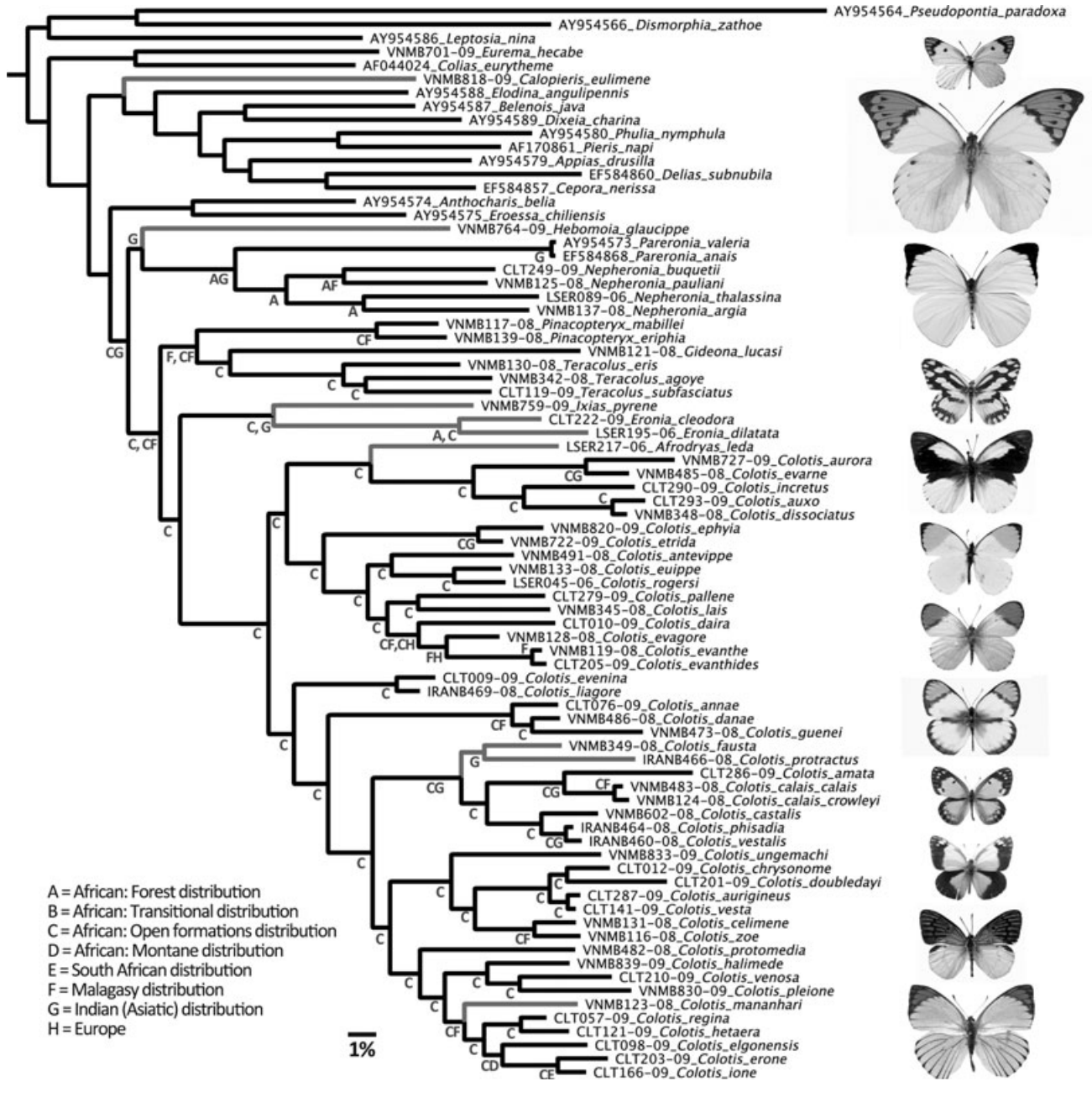

Fig. 2. Bayesian tree for Pieridae inferred through MrBayes analysis (Tree Length = 6806 steps, CI = 0.2470, RI = 0.4670, −ln L = −39738.22794). The results of the DIVA analysis are given for *Colotis* and related taxa, with maxtrees set to 2 ancestral areas. Coloured clades indicate *Colotis* groups identified through this study, and red branches highlight taxa with a varying position between Bayesian and maximum likelihood phylogenies (Fig. 1). Figured species are, from top to bottom, *Calopieris eulimene*, *Hebomoia glaucippe*, *Nepheronia argia*, *Pinacopteryx eriphia*, *Teracolus eris*, *Colotis incretus*, *Colotis evanthe evanthides*, *Colotis evenina*, *Colotis amata*, *Colotis ungemachi*, *Colotis zoe* and *Colotis ione*

clock methods do not report an extremely old age for the group (Braby et al. 2007; Wheat et al. 2007).

The position of *Calopieris* as sister to *Eronia* (Fig. 1), Pierinae (Fig. 2) or *Hebomoia* + Anthocharidini (BEAST, not shown) is also not well supported, but it strongly suggests that the morphologically distinct *Calopieris* belongs well outside *Colotis* sensu lato. Similarly, *Ixias* and *Hebomoia* have unstable positions across our molecular reconstructions. The deep divergence between *E. leda* and *E. cleodora* barcodes (15.9 ± 0.21%), their completely different colour patterns and their paraphyly strongly suggest that these two species do not belong in the same genus.

The position of the Malagasy genus *Gideona* as sister to *Pinacopteryx* + *Teracolus*, although weakly supported, is consistent with previous findings (Braby et al. 2006). The adult of *Gideona lucasi* is superficially reminiscent of *Hebomoia*, and in the past this species has been included in the genus *Colotis*. Our DIVA analysis suggests that the last common ancestor of this group with *Ixias* and *Eronia* flew in Africa, and subsequently dispersed to Madagascar (*Gideona*) and India (*Ixias*, *Eronia*). Independent dispersals to Madagascar are also noted for *Nepheronia pauliani* and *C. mananhari*, possibly much later than *Gideona*. Most other Malagasy *Colotis* show shallow divergence from their sister taxa and hence seem to

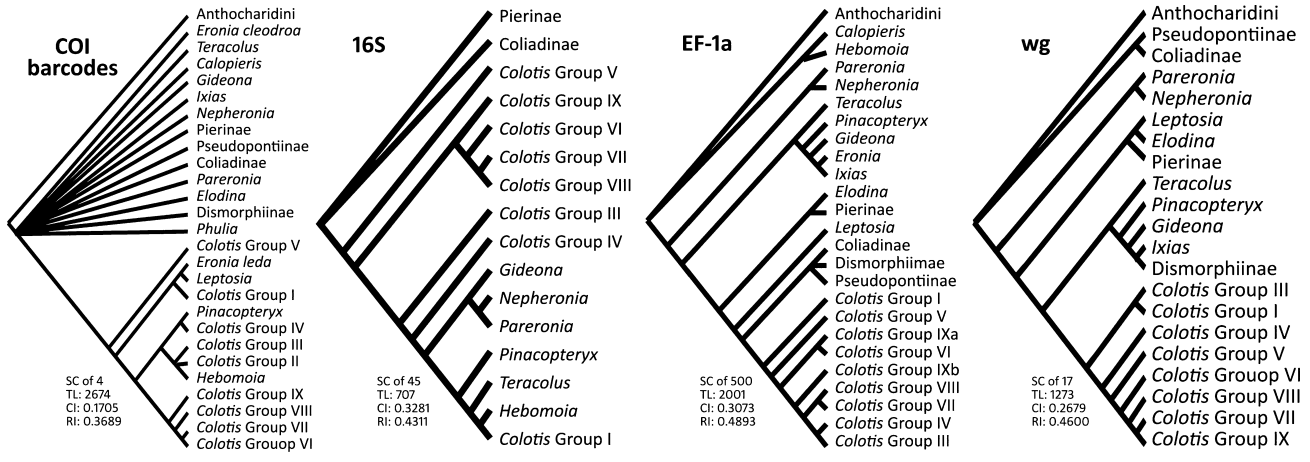

Fig. 3. Consensus trees resulting from Maximum Parsimony analysis for each gene partition. SC = strict consensus of (*n*) trees, TL = tree length, CI = consistency index, RI = retention index

have arrived on the island more recently, including *C. calais crowleyi* and *C. evanthides*, derived from common ancestors in Africa (*C. calais*) or Madagascar (*C. evanthe*). Multiple independent dispersals from Africa to Madagascar have been documented in other butterflies, including those in the genus *Charaxes* (Aduse-Poku et al. 2009).

In all of our phylogenies, the taxa *eris*, *subfasciatus* and *agoye* always formed a well-supported clade outside the *Colotis*, usually close to *Pinacopteryx* and *Gideona*. Shared differences between these species and other *Colotis* have been noted, including an acute forewing apex, a shorter aedeagus and a produced valval apex (Henning et al. 1997), as well as differences in larval morphology (Larsen 1992). A separate generic status for these three species is required.

Both Indian endemic species, *C. aurora* (group I) and *C. etrida* (group II), seem to have common ancestors with African species (*C. evarne* and *C. ephyia*, respectively) which range as far north as Sudan, Chad, Saudi Arabia and Yemen (Fig. 2). The shallow divergence between these species and their presence in the Arabian Peninsula rejects the possibility of an old vicariance speciation resulting from the severance of the Indian plate from Africa or Madagascar and instead suggests a more recent vicariance or dispersal event in the Near East that resulted in isolation of *C. etrida* and *C. aurora* in the Indian subcontinent.

With *Teracolus* removed, and ignoring the unsupported position of *Afrodryas leda* in our Bayesian phylogeny, *Colotis* forms a monophyletic clade that is primarily supported by the EF-1α gene (Figs 1 and 2; Table S4). Despite unique distributions and several autapomorphies, the status of *Madais* Swinhoe (1909) or *Cuneacolotis* Henning et al. (1997) is not supported as subgenera under *Colotis*. Instead, we prefer to recognize 'species groups' that formed well-supported clusters.

DNA barcoding and taxonomy

DNA barcoding (Hebert et al. 2003, 2004; Hajibabaei et al. 2006) is rapidly becoming the most important tool for species identification and discovery. Application of DNA barcoding to assess mitochondrial sequence diversity among material examined in our study highlighted the inconsistencies between current taxonomy and genetic variation in many of the species. Divergent populations showing deep splits (>4%) observed within some species supported their separate species status, including the Malagasy *N. buquetii pauliani* (9.5 ± 0.3%) and *Pinacopteryx eriphia mabillei* (5.8 ± 0.3%), as well as the Tanzanian *Eronia cleodora* (4.6 ± 0.9%) restricted to the coastal forests in eastern Kenya and Tanzania and with a much wider black border on its wings.

The divergence between the two subspecies of *Teracolus agoye* (i.e. ssp. *agoye* and spp. *bowkeri*, 2.0 ± 0.3%) supports the traditional status of these taxa (Larsen 1992). No subspecies are currently recognized under *Teracolus eris*; however, the presence of several distinct mitochondrial lineages within *eris* suggests an overlooked geographic structure that requires further investigation (Fig. 4).

DNA barcode divergence between *C. auxo* (South Africa) and *C. dissociatus* (Botswana) is very shallow (1.0 ± 0.1%). These two live in very different habitats in South Africa with no overlap: while *auxo* occupies the coastal forests and the bushveld, *dissociatus* mainly thrives in savannah habitats (TBL observation). Both have a white to pale yellow ground colour, with *dissociatus* being smaller, generally paler and often with no trace of black margin on the inner edge of the orange tip of the forewings. Our *C. auxo* from Kenya/Tanzania (i.e. ssp. *incretus*) were significantly divergent from both of these (11.2 ± 0.1%). These are usually larger with much deeper yellow ground colour and more pointed forewings.

Colotis phisadia and *C. vestalis* are two closely related species often distinguishable by their forewing ground colour (i.e. pink-salmon in *C. phisadia*, white in *C. vestalis*). They share the same larval host (*Salvadora persica*, Nazari 2003) and occur sympatrically from Iran to Pakistan and Gujarat in India. *Colotis phisadia* is a migratory species (Gardner and Howarth 2007) and extends into Arabia and Africa, while *C. vestalis* spreads deeper into Pakistan and wet habitats as far as Delhi; it is conspicuously absent from Arabia but re-appears in the dry parts of East Africa with a radically different morphology (i.e. ssp. *castalis*). With the exception of the latter population, all other *phisadia* and *vestalis* examined in our study were nearly identical in their DNA barcodes and shared haplotypes (Figure S1). Intermediate phenotypes, often with a yellow ground colour, are not uncommon when the two occur in sympatry (e.g. in Southern Iran; see Nazari 2003). *Colotis protractus*, traditionally considered a subspecies of *C. phisadia*, occurs sympatrically with *vestalis* as well as *phisadia* in

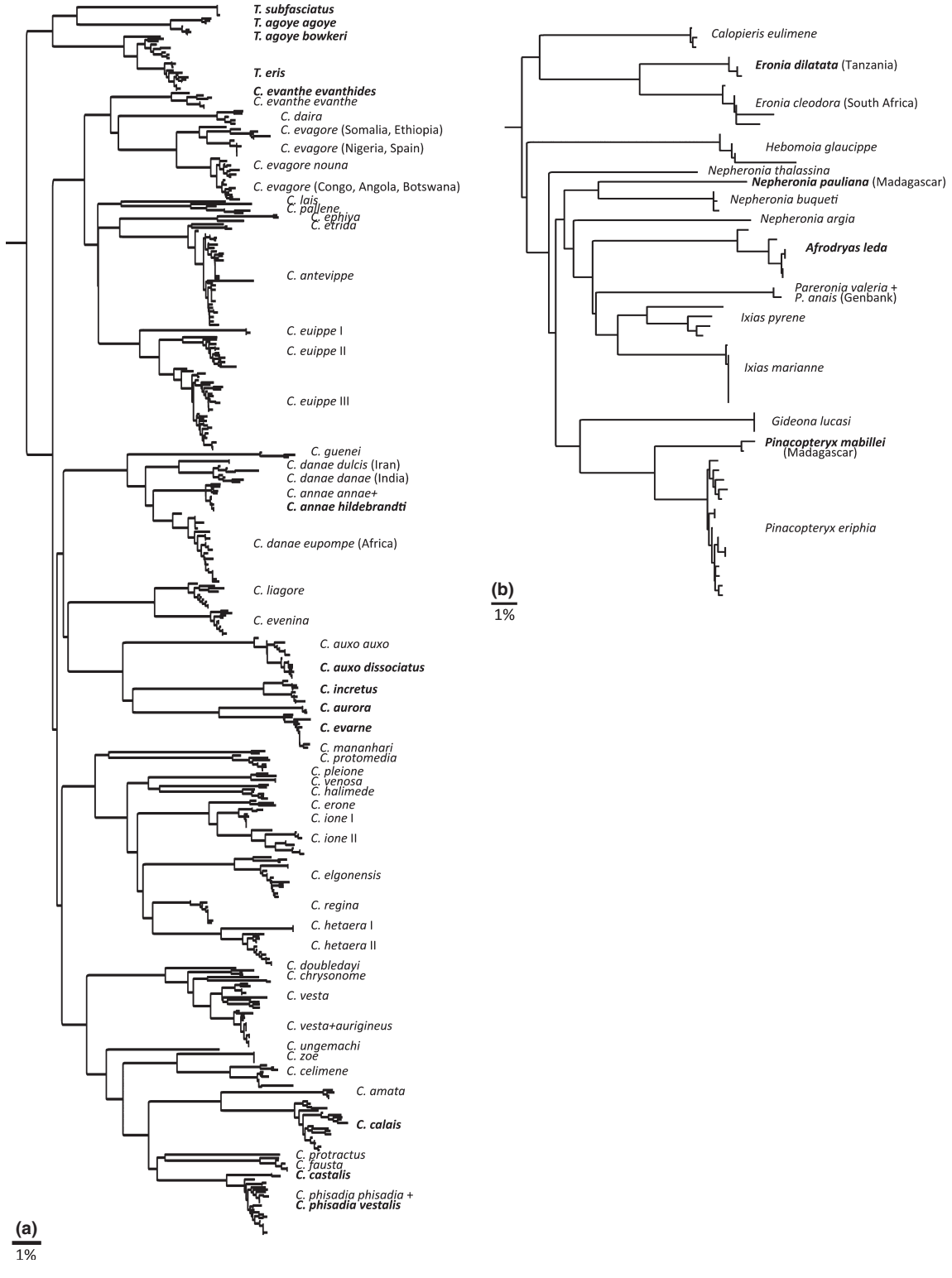

Fig. 4. Neighbour-joining trees of the cytochrome oxidase I barcodes for *Colotis* (a) and associated genera (b). Taxa with changed ranks or combinations are in bold

Southern Iran to Gujarat and is immediately distinguishable by the pink-salmon ground colour on both wings and blue forewing apical spots. Its barcode also clearly separates it as a distinct species closer to *C. fausta* (Fig. 4).

Divergent clusters within *C. evagore* demonstrate a geographic structure, with populations in (i) Spain, Morocco, Tunisia and Nigeria, corresponding to ssp. *nouna*, (ii) Congo, Angola and Botswana, and (iii) Somalia and Ethiopia. A similar situation in *C. euippe* is taxonomically more difficult to explain: while individuals from Gambia, Nigeria, Cameroon and Somalia form a cluster (II), those from Congo, Angola, Namibia, Tanzania and Uganda form two separate clusters (I, III) with geographic overlap. Minor but consistent differences in coloration of the forewing tip as well as the underside

of the wings in these two clusters suggests the possibility of an overlooked species, but further investigation is needed (TB observation). Our specimens of *C. rogersi* did not show any barcode differences from *C. euippe*; however, misidentifications could not be ruled out.

The Congolese *C. pallene*, with uniquely wide submarginal bands on the upperside of the hindwings and a notable sequence divergence from the Botswana population (3.2 ± 0.8%), presents another case of a potentially overlooked species flagged by DNA barcoding. Examination of additional material from South Africa and Tanzania is required before a separate status for this population can be invoked.

Despite a wide range geographic sampling of *C. antevippe*, small barcode variation observed among these populations (1.06 ± 0.04%) does not support the recognition of subspecies within *C. antevippe*. The stability of mitochondrial DNA over a large area with a high climatic variation can be explained by a rapid recent range expansion, a phenomenon that deserves further investigation.

Colotis evanthides is a rare species confined to the Comoros, Aldabra, Assumption, Cosmoledo and Astove Islands in the Indian Ocean. On the basis of similarities in genitalia, Bernardi (1954) hypothesized a close affinity between *C. evanthides* and the Malagasy *C. evanthe*, the African *C. aurora* and the Indian *C. etrida*. Thus, *C. evanthides* was later termed 'the Lemurian link' (Cogan and Hutson 1971). However, our phylogenetic analysis does not support these four species as a natural group.

Within *C. celimene*, individuals from Nigeria (ssp. *sudanicus*), Niger, Congo and Kenya show only minor variation in their barcodes (0.3 ± 0.1%). Several subspecies currently recognized under *C. celimene* (e.g. *sudanicus*, *amina*, *angusi*, etc.) may be redundant; however, the Namibian ssp. *pholoe* is divergent (2.5 ± 0.2%). The Somali ssp. *praeclarus* has been recently proposed as a distinct species (Bouyer 2010).

Colotis chrysonome, *C. aurigineus*, and *C. vesta* are closely related species with overlapping ranges in eastern Africa. *Colotis doubledayi*, a rare species confined to the Namib dry zone in southern Africa, is another member of this group with a distinct barcode haplotype (Fig. 4). Within the widespread *C. chrysonome*, we observed gaps (3.3 ± 0.5%) between populations from Sudan (type locality), Kenya and Ethiopia, suggesting undetected differentiation (Fig. 4). The often larger and variable *C. vesta* formed two separate clusters with specimens from Congo (Katanga) appearing in both. One of these clusters (i.e. Kenya, Tanzania and Congo) also included the sympatric *C. aurigineus* (i.e. Kenya, Tanzania and Uganda) (Fig. 4). Presence of shared haplotypes between these otherwise readily recognizable sister species when in sympatry suggests gene flow between the two (Figure S1).

Within the montane species *C. elgonensis*, the disjunct ssp. *glauningi* from Nigeria seems to be distinct. The Congo (Katanga) ssp. *nobilis* also shows some differentiation; however, it falls within a cluster with all other populations from Uganda, Burundi, Kenya and Congo (Kivu Nord), including ssp. *elgonensis* and ssp. *basilewskyi* from Uganda and ssp. *kenia* from Kenya, which show minimal variation (0.5 ± 0.3%). The three montane distribution areas are well separated from each other, with the Nigerian population isolated by over 2000 km from Kivu.

The distance between *C. ione* (I) from Somalia, Tanzania and Kenya (Eastern, Nyaza) and those (II) from Congo and the Kenya (Coast) (3.3 ± 0.2%) is comparable to the distance from their closest sister species, *C. erone* (i.e. 3.7% and 3.6%, respectively), which supports a distinction at species level. No subspecies are currently recognized under *C. ione* (Ackery et al. 1995). Further sampling of populations (e.g. from South Africa) is needed before a third species is recognized in this complex.

The type locality of *C. hetaera* is 'Endara' (i.e. Mt. Ndara in Coast, Kenya, not Zanzibar as indicated by Ackery et al. 1995). Several individuals of *C. hetaera*, including a yellow female (GVDP087) from the Golini Forest in Kenyan Coast (I), were divergent from other subspecies (i.e. *aspasia*, *ankolensis* and *lorti*) (II). However, a white female from the same forest (GVDP082) fell within the second group. The males in the two clusters are similar. This indicates the possible coexistence of a cryptic species sympatric with *C. hetaera* in the Kenyan Coast.

Conclusion

Our study highlights the paraphyly of the nine genera currently known as the *Colotis* group and underscores the need for additional gene sampling to resolve the phylogenetic relationships within the family Pieridae. Major diversification in *Nepheronia*, *Pinacopteryx* and *Colotis* seems to have occurred in Africa during the Eocene, while *Hebomoia*, *Ixias* and *Pareronia* diversified in the Oriental region. The Malagasy species seem to represent a combination of old and new arrivals to the Island, while the Indian endemics appear to be derived from common ancestors that lived in the area between northeast Africa and Arabia to India.

DNA barcoding largely supported the traditional species taxa in *Colotis* and related genera, but we found evidence for exclusion of three species (i.e. *C. eris*, *C. subfasciatus* and *C. agoye*) from the genus *Colotis* and the taxon *leda* from *Eronia*, for which we assign new generic status (i.e. *Teracolus* and *Afrodryas*). Several cases of undetected species-level variation, or lack thereof, were also flagged by DNA barcoding, where relevant taxonomic changes are proposed. To detect such taxonomic discrepancies, future studies should also aim for even more comprehensive population sampling, particularly in the genera *Ixias* and *Pareronia* that were sparsely sampled in our study.

Acknowledgements

We thank Jean-Francois Landry, Don Lafontaine, and Jeremy DeWaard (Canada), Robert Robbins and Claire Kremen (USA), John Tennent and Keith Stiff (UK), Rodolphe Rougerie and Isabelle Meusnier (France), Roger Vila (Spain), Christopher Wheat (Finland), Robin van Velzen and Hein Leertouwer (the Netherlands), Alison Cameron (Ireland), Wolfgang ten Hagen (Germany), Alireza Naderi (Iran), Steve Collins (Kenya), and Gerardo Lamas (Peru) for providing specimens or support. Doekele G. Stavenga (Groningen, the Netherlands) generously provided his data and commentary on UV reflection patterns in Pierinae, which will be published separately. We also thank the staff at the NHM (UK), RMCA (Belgium) and Naturalis (the Netherlands) Museums and the Royal Antwerp Entomological Society (Belgium). We also thank Michael Braby (Australia), Niklas Wahlberg (Finland) and anonymous JZSER reviewers who provided helpful comments on the earlier drafts of this manuscript. DCL was supported by a STUDIUM fellowship and INRA. Some Malagasy collections were collected under NSF grant DEB-0072713 to B. L. Fisher and C. E. Griswold. This research was supported through funding to the Canadian Barcode of Life Network from Genome Canada (through

the Ontario Genomics Institute), NSERC and other sponsors listed at http://www.bolnet.ca.

Résumé

Systématique et Phylogenie de Colotis *et des genres apparentés (Lépidoptères: Pieridae): implications évolutives et taxonomiques*

Nous avons étudié la diversité génétique et la position phylogénétique des papillons du genre *Colotis* et de huit autres genres de Piéridés apparentés, à l'aide de quatre gènes, deux mitochondriaux et deux nucléaires. Afin d'établir le statut taxonomique des espèces, nous avons dans un premier temps établi les codes-barres génétiques de 523 spécimens représentant tous les genres et la plupart des espèces et sous-espèces constituantes. Nous avons ensuite choisi un sous-ensemble de ces taxons que nous avons soumis à une analyse phylogénétique après avoir établi les séquences de portions des gènes suivants : 16S ARNr (523 bp), EF-1α (1126 bp) et wg (404 bp). Les résultats des codes-barres génétiques concordent bien dans l'ensemble avec la classification traditionnelle des espèces du groupe *Colotis*, mais quelques cas de divergence génétique prononcée, ou à l'inverse des cas de similitudes génétiques, indiquent soit des différenciations spécifiques insoupçonnées, soit des cas de synonymie. Malgré l'analyse de données provenant de quatre gènes, les branchements de base de notre arbre phylogénétique sont peu fiables et nous n'avons pas réussi à établir la monophylie du 'groupe *Colotis*', ni des genres *Colotis* et *Eronia*. Afin de maintenir la monophylie de *Colotis*, nous avons rétabli le genre *Teracolus* pour y transférer trois espèces de *Colotis* (*Colotis eris*, *Colotis subasciatus* et *Colotis agoye*), et le genre *Afrodryas* pour inclure *E. leda*. La position de *Calopieris* n'est pas résolue mais ce genre présente un profil bien au-dehors de la variation observée dans *Colotis* (s.l.). Une analyse de vicariance/dispersion indique qu'une diversification importante des *Colotis* (s. str.) aurait eu lieu en Afrique, suivie d'une dispersion vers l'Inde et Madagascar.

References

Ackery PR, Smith CR, Vane-Wright RI (1995) Carcasson's African Butterflies: An Annotated Catalogue of the Papilionoidea and Hesperioidea of the Afrotropical Region. British Museum (Natural History), London, UK.

Aduse-Poku K, Vingerhoedt E, Wahlberg N (2009) Out-of-Africa again: a phylogenetic hypothesis of the genus *Charaxes* (Lepidoptera: Nymphalidae) based on five gene regions. Mol Phylogenet Evol **53:**463–478.

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

Aurivillius C (1925) Pieridae. In: Seitz A (ed.), Die Gross-Schmetterlinge der Erde. 13. Band: Die Afrikanischen Tagfalter. Alfred Kernen, Stuttgart, pp 29–69.

Bernardi G (1954) Notes sur la taxonomie de deux *Colotis* malgaches. Le Naturaliste malgache **6:**57–60.

Boisduval JBA (1847) Catalogue des Lépidoptères recuellis par M. Delegorgue pendant les années 1838–1844 à Port-Natal, au pays de Amazoulous et dans la contrée de Massilicatzi. In: Delegorgue A (ed.), Voyage dans l'Afrique australe, **2:**585–602.

Bouyer T (2010) *Colotis praeclarus* (Butler, 1886) (Lepidoptera, Pieridae). Entomologia Africana **15:**22–24.

Braby MF (2005) Afrotropical mistletoe butterflies: larval food plant relationships of *Mylothris* Hübner (Lepidoptera: Pieridae). J Nat Hist **39:**499–513.

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

Braby MF, Trueman JWH (2006) Evolution of larval host plant associations and adaptive radiation in pierid butterflies. J Evol Biol **19:**1677–1690.

Braby MF, Vila R, Pierce NE (2006) Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. Zool J Linn Soc **147:**239–275.

Braby MF, Pierce NE, Vila R (2007) Phylogeny and historical biogeography of the subtribe Aporiina (Lepidoptera: Pieridae): implications for the origin of Australian butterflies. Biol J Linn Soc **90:**413–440.

Brower AVZ, De Salle R (1994) Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. Ann Entomol Soc Am **87:**702–716.

Butler AG (1869) Descriptions of new Rhopalocera from the collection of Herbert Druce Esq. Cistula Entomol **1:**1–16.

Butler AG (1876) Revision of the lepidopterous genus *Teracolus* with descriptions of the new species. Proc Zool Soc Lond **1876:**126–165.

Carcasson RH (1964) A preliminary survey of the zoogeography of African butterflies. East Afr Wildl J **2:**122–157. [reprinted in Ackery et al. 1995].

Caterino MS, Sperling FAH (1999) *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. Mol Phylogenet Evol **11:**122–137.

Caterino MS, Reed RD, Kuo MM, Sperling FAH (2001) A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). Syst Biol **50:**106–127.

Chew FS, Watt WB (2006) The Green-veined White (*Pieris napi* L.), its pierine relatives, and the systematics dilemmas of divergent character sets (Lepidoptera, Pieridae). Biol J Linn Soc **88:**413–435.

Cogan BH, Hutson AM (1971) Preliminary observations on the affinities and composition of the insect fauna of Aldabra. Phil Trans R Soc Lond B Biol Sci **260:**315–325.

Cramer P (1780) De Uitlandsche Kapellen voorkomende in de drie waereld-deelen Asia, Africa en America. Amsteldam and Utrecht. **4** [part]**:**1–28.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol **7:**214.

Drury D (1773) Illustrations of Natural History. White, London.

Fabricius JC (1775) Systema Entomologiae. Flensburgi and Lipsiae.

Gardner D, Howarth B (2007) A butterfly migration: blue-spotted Arab at Jabal Barakah, western Abu Dhabi. Tribulus **17:**98.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52:**696–704.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. Proc Natl Acad Sci USA **103:**968–971.

Hausmann A, Hebert PDN, Mitchell A, Rougerie R, Sommerer M, Edwards T, Young CJ (2009) Revision of the Australian *Oenochroma vinaria* Guenée, 1858 species-complex (Lepidoptera: Geometridae, Oenochrominae): DNA barcoding reveals cryptic diversity and assesses status of type specimen without dissection. Zootaxa **2239:**1–21.

Hebert PDN, Cywinska A, Ball SL, Dewaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci **270:**313–321.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci USA **101:**14812–14817.

Hecq J (1975) Le complexe *Colotis eucharis* F. (Lepidopt. Pieridae). Lambillionea **75:**4–10.

Henning GA, Henning SF, Joannou JG, Woodhall SE (1997) Living Butterflies of the Southern Africa. Umdaus Press, Hatfield, South Africa.

Hesselbarth G, Van Oorschot H, Wagener S (1995) Die Tagfalter der Türkei, I. Selbstverlag Sigbert Wagener, Bocholt.

Howe K, Bateman A, Durbin R (2002) QuickTree: building huge neighbour-joining trees of protein sequences. Bioinformatics **18:**1546–1547.

Hübner J (1823) Sammlung exotischer Schmetterlinge, 2. In: Hübner J (1819–1827). Jacob Hübner Verlag, Augsburg.

Ivanova NV, Dewaard JR, Hebert PDN (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. Mol Ecol Notes **6:**998–1002.

Keane TM, Naughton TJ, McInerney JO (2007) MultiPhyl: a high-throughput phylogenomics webserver using distributed computing. Nucleic Acids Res **35:**W33–W37.
Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol **16:**111–120.
Klots AB (1933) A generic revision of the Pieridae (Lepidoptera) together with a study of the male genitalia. Entomol Am (NS) **12:**139–242.
Klug JCF (1829) . Symbolae physicae seu icones et descriptiones insectorum quae ex itinere per Africam borealem et Asiam occidentalem Fridrichi Guilelmi Hemprich et Christiani Godofredi Ehrenberg Medicinae et Chirurgiae Doctorum studio novae aut illustratae redierunt. Decas Prima (1929), Berolini: ex Officina Academia, Impensis G. Reimer.
Kodandaramaiah U, Wahlberg N (2007) Out-of-Africa origin and dispersal mediated diversification of the butterfly genus *Junonia* (Nymphalidae: Nymphalinae). J Evol Biol **20:**2181–2191.
Kodandaramaiah U, Wahlberg N (2009) Phylogeny and biogeography of *Coenonympha* butterflies (Nymphalidae: Satyrinae) – patterns of colonization in the Holarctic. Syst Entomol **34:**315–323.
Kodandaramaiah U, Peña C, Braby MF, Grund R, Müller CJ, Nylin S, Wahlberg N (2010) Phylogenetics of Coenonymphina (Nymphalidae: Satyrinae) and the problem of rooting rapid radiations. Mol Phylogenet Evol **54:**386–394.
Larsen TB (1984) Butterflies of Saudi Arabia and its Neighbours. Stacey International, London.
Larsen TB (1991) The Butterflies of Kenya and their Natural History. Oxford University Press, Oxford.
Larsen TB (1992) On the status of *Colotis agoye* (Wallengren, 1857) and its subspecies. Metamorphosis **3:**5–10.
Larsen TB (2005) Butterflies of West Africa. Apollo Books, Stenstrup, Denmark.
Lohman DJ, Peggie D, Pierce NE, Meier R (2008) Phylogeography and genetic diversity of a widespread Old World butterfly, *Lampides boeticus* (Lepidoptera: Lycaenidae). BMC Evol Biol **8:**301.
Maddison WP, Maddison DR (2009) Mesquite: a modular system for evolutionary analysis. Version 2.72. Available from: http://mesquiteproject.org (accessed December 2009).
Monaghan MT, Gattolliat JL, Sartori M, Elouard JM, James H, Derleth P, Glaizot O, de Moor F, Vogler AP (2005) Trans-oceanic and endemic origins of the small minnow mayflies (Ephemeroptera, Baetidae) of Madagascar. Proc R Soc Lond B Biol Sci **272:**1829–1836.
Morinaka S, Miyata T, Tanaka K (2002) Molecular phylogeny of the Eichhorni group of *Delias* Hübner, 1819 (Lepidoptera: Pieridae). Mol Phylogenet Evol **23:**276–287.
Narita S, Nomura M, Kato Y, Fukatsu T (2006) Genetic structure of sibling butterfly species affected by *Wolbachia* infection sweep: evolutionary and biogeographical implications. Mol Ecol **15:**1095–1108.
Nazari V (2003) Butterflies of Iran. Dayereye Sabz publications, Tehran.
Nazari V, Zakharov EV, Sperling FAH (2007) Phylogeny, historical biogeography, and taxonomic ranking of Parnassiinae (Lepidoptera, Papilionidae) based on morphology. Mol Phylogenet Evol **42:**131–156.
Peters W (1952) A Provisional Checklist of the Butterflies of the Ethiopian Region. E.W. Classey, Feltham, Middlesex.
Rambaut A (2002) Sequence alignment editor, version 2.0. Available from: http://evolve.zoo.ox.ac.uk/software/Se-Al/Main.html (accessed December 2009).
Ronquist F (1997) DIVA (dispersal-vicariance analysis) ver. 1.2. Computer program and manual. Available from http://www.ebc.uu.se/systzoo/research/diva/diva.html (accessed January 2010).
Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:**1572–1574.
Sharpe EM (1898–1902) Monographiae Entomologicae. 1. A Monograph of the Genus Teracolus. Lovell Reeve, London.
Sorenson MD, Franzosa EA (2007) TREEROT, Version 3. Boston University, Boston, MA. Available from http://people.bu.edu/msoren/TreeRot.html (accessed December 2009).
Staudinger O (1884) Exotische Tagfalter in systematischer Reinfolge mit Berücksichtigung neuer Arten. In: Staudinger O, Schatz E (eds), Exotische Schmetterlinge. 4, Theil I, VII, pp 43–44.
Stavenga DG, Arikawa K (2006) Evolution of color and vision of butterflies. Arthropod Struct Dev **35:**307–318.
Stavenga DG, Leertouwer HL (2007) Colourful butterfly wings: scale stacks, iridescence and sexual dichromatism of Pieridae. Entomol Ber (Amsterdam) **67:**158–164.
Stavenga DG, Giraldo MA, Hoenders BJ (2006) Reflectance and transmittance of light scattering scales stacked on the wings of pierid butterflies. Optics Express **14:**4880–4890.
Stoneham HF (1957) Two new generic names, and new subspecies and forms of Pieridae. Bull. Stoneham Mus. **70:**1–3.
Suárez NM, Betancor E, Pestano J (2009) Intraspecific evolution of Canarian *Euchloe* (Lepidoptera: Pieridae) butterflies, based on mtDNA sequences. Mol Phylogenet Evol **51:**601–605.
Swainson W (1833) (ed.) Zoological illustrations, or original figures and descriptions of new, rare or interesting animals selected chiefly from the classes of ornithology, entomology and conchology, (2) 3, [ix]. James Moyes, London.
Swinhoe C (1909) New species of Indo-Malayan and African Lepidoptera. Ann Mag Nat Hist **3:**89–98.
Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, MA.
Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol **56:**564–577.
Talbot G (1939) Revisional notes on the genus *Colotis* Hubn. (Lepid. Pieridae) with a systematic list. Trans R Entomol Soc Lond **88:**173–233.
Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics **140:**767–782.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTALX Windows inference: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acid Res **25:**4876–4882.
Wahlberg N, Braby MF, Brower AVZ, de Jong R, Lee MM, Nylin S, Pierce N, Sperling FAH, Vila R, Warren AD, Zakharov E (2005) Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. Proc R Soc Lond B Biol Sci **272:**1577–1586.
Wallengren HDJ (1857) Kafferlandets Dag-fjärilar, insamlade åren 1838–1845 af J. A. Wahlberg/Lepidoptera Rhopalocera in Terra Caffrorum annis 1838–1845 collecta a J. A. Wahlberg. K Svenska Vetensk Akad Handl **2:**1–55.
Wheat CW, Watt WB (2008) A mitochondrial-DNA-based phylogeny for some evolutionary-genetic model species of *Colias* butterflies (Lepidoptera, Pieridae). Mol Phylogenet Evol **47:**893–902.
Wheat CW, Vogel H, Wittstock U, Braby MF, Underwood D, Mitchell-Olds T (2007) The genetic basis of a plant–insect coevolutionary key innovation. Proc Natl Acad Sci USA **104:**20427–20431.
Wiemers M, Keller A, Wolf M (2009) ITS2 secondary structure improves phylogeny estimation in a radiation of blue butterflies of the subgenus *Agrodiaetus* (Lepidoptera: Lycaenidae: *Polyommatus*). BMC Evol Biol **9:**300.
Wijnen B, Leertouwer HL, Stavenga DG (2007) Colors and pterin pigmentation of pierid butterfly wings. J Insect Physiol **53:**1206–1217.
Xu L, Hao JS, Zhu GP, Yin XB, Pan HC, Huang DY, Zhang XP (2007) Molecular phylogenetic relationships of some species and genera in Pierinae and Coliadnae [Coliadinae] (Pieridae) based on partial sequence of mitochondrial CO I and Cyt b genes. Acta Zootaxon Sinica **32:**842–850.
Zakharov EV, Smith CR, Lees DC, Cameron A, Vane-Wright RI, Sperling FAH (2004) Independent gene phylogenies and morphology demonstrate a Malagasy origin for a wide-ranging group of swallowtail butterflies. Evolution **58:**2763–2782.

Appendix: Proposed taxonomic changes

(1) The taxon *leda* was originally described in the genus *Dryas* (Boisduval 1847), a homonym of *Dryas* Hübner 1823; currently the valid genus for *Dryas iulia* (Heliconiinae). Butler (1869) subsequently placed *leda* under *Eronia* Hübner, 1823. *Afrodryas* Stoneham (1957) was described as a replacement name for *Dryas* Boisduval (1847); but later treated as a synonym of *Eronia*. Our results support the arrangement for *leda* by Boisduval (1847) and Stoneham (1957). Accordingly, we revive the genus *Afrodryas* and the combination *Afrodryas leda* (**comb. rev.**).

(2) We found that *Colotis agoye*, *C. subfasciatus* and *C. eris* consistently formed a well-supported clade separate and outside all other *Colotis*. Here, we revive the oldest available generic name for these species, *Teracolus* Swainson, (1833) (**stat. rev.**) (Type species: *Papilio subfasciatus*) to represent this clade.

(3) We also recognize the divergent populations *Nepheronia pauliani* (**stat. nov.**), *Pinacopteryx mabillei* (**stat. nov.**) and *Eronia dilatata* (**stat. rev.**) as good species.

(4) The taxon *incretus* was originally described as a separate species based on a single female from Mamboia (Tanzania) and later treated as a subspecies of *Colotis auxo* (Hecq 1975). Examination of the holotype in the NHM confirmed the identity of our divergent Kenya/Tanzania specimens. We therefore reinstate *C. incretus* as a valid species and consider *dissociatus* an ecological (savannah) subspecies of *C. auxo* (**stat. nov.**).

(5) The correct name for the Indian species (*Papilio*) *eucharis* (Fabricius 1775) (junior primary homonym of *Delias eucharis* Drury 1773) is *Colotis aurora* (Cramer 1780; Larsen 2005). Considering the large gap between the Indian and the African *C. aurora* (5.34 ± 0.03%), we recognize the African population as a separate species, *C. evarne* (Klug 1829) (**stat. rev.**).

(6) The Indian populations of *C. amata* are also deeply divergent from their African counterparts (6.6 ± 0.4%), and morphologically distinct (i.e. wider wings and lighter salmon coloration). We did not examine the *amata* populations from Arabia to Senegal, which may prove to have some intermediate status. However, on the basis of the available data we reinstate all non-Indian *amata* to *C. calais* (**stat. rev.**) and recognize the subspecies *C. calais williami* (Namibia and Angola) and *C. calais crowleyi* (Madagascar).

(7) *Colotis danae* from India (ssp. *danae*) and Iran (ssp. *dulcis*) are divergent from African populations (3.7 ± 0.9% and 3.9 ± 0.3% respectively) and from one another (3.4 ± 0.2%), while only minor variation exists within the African *C. danae* (0.5 ± 0.2%). We recognize the subspecies *C. danae danae*, *C. danae dulcis* and *C. danae eupompe* (=*pseudocaste* **syn. nov.**). The larger taxon *annae* Wallengren, 1857 flies in dry parts of South Africa, from Natal to Namibia and north to Zambia and Botswana. *Colotis hildebrandti* Staudinger (1884) is similarly large and ranges from southern Kenya to central Zambia. Their barcodes were similar (0.38 ± 0.05%). We reinstate *C. annae* (**stat. rev.**) as a good species, with ssp. *hildebrandti* as its northern subspecies (**stat. nov.**).

(8) We treat *C. phisadia* and *C. vestalis* as a single species, with an Arabian/African subspecies (i.e. *C. phisadia phisadia*) and an Asian subspecies (i.e. *C. phisadia vestalis* **stat. nov.**). We also reinstate *C. castalis* (**stat. rev.**) as a separate species, described from Tanzania and found in dry habitats through Kenya to Somalia.

(9) *Colotis evanthe* and *C. evanthides* are so similar in morphology as well as DNA barcodes (1.11 ± 0.04%) that we suggest the name *evanthides* be used as a subspecies of *C. evanthe* (**stat. nov.**).

(10) *Idmais vesta* Reiche, (1850) is a junior primary homonym of *Idmais vesta* Boisduval (1847). Under Art. 23.9.2 of the International Code for Zoological Nomenclature, to maintain usage of *vesta* Reiche, the name *Idmais vesta* Boisduval (1847) needs to be designated as a 'nomen oblitum', and *Idmais vesta* Reiche, (1850) as a 'nomen protectum'. This taxonomic act shall be properly completed in a separate publication.

Revised list of species and subspecies

Unexamined taxa are marked by an asterisk (*). The *Colotis* groups I–IX are listed according to the results of this study. For additional synonymy and references to original descriptions, see Ackery et al. (1995).

(1) Genus *Hebomoia* Hübner, (1819)
- *H. glaucippe* (Linnaeus, 1758)
- *H. leucippe* (Cramer, 1775)*

(2) Genus *Gideona* Klots, 1933
- *G. lucasi* (Grandidier, 1867)

(3) Genus *Ixias* Hübner, (1819)
- *I. flavipennis* Grose-Smith, 1885*
- *I. kuehni* Röber, 1891*
- *I. malumsinicum* Thieme, 1897*
- *I. marianne* (Cramer, 1779)
- *I. paluensis* Martin, 1914*
- *I. piepersii* (Snellen, 1877)*
- *I. pyrene* (Linnaeus, 1764)
- *I. reinwardtii* (Vollenhoven, 1860)*
- *I. venilia* (Godart, 1819)*
- *I. vollenhovii* (Wallace, 1867)*

(4) Genus *Calopieris* Aurivillius, 1899
- *C. eulimene* (Klug 1829)

(5) Genus *Nepheronia* Butler, 1870
- *N. argia* (Fabricius 1775)
- *N. buquetii* (Boisduval, 1836)
- *N. pauliani* Bernardi, 1959 (stat. nov.)
- *N. pharis* (Boisduval, 1836)
- *N. thalassina* (Boisduval, 1836)

(6) Genus *Pareronia* Bingham, 1907
- *P. argolis* (C. Felder and R. Felder, 1860)*
- *P. avatar* (Moore, 1858)*
- *P. aviena* Fruhstorfer, 1910*
- *P. boebera* (Eschscholtz, 1821)*
- *P. ceylanica* (C. Felder and R. Felder, 1865)*
- *P. chinki* Joicey and Noakes, 1915*
- *P. anais* Lesson, 1837 (=*P. hippia* (Fabricius, 1787))
- *P. iobaea* (Boisduval, 1832)*
- *P. nishiyamai* Yata, 1981*
- *P. paravatar* Bingham, 1907*
- *P. phocaea* (C. Felder and R. Felder, 1861)*
- *P. tritaea* (C. Felder and R. Felder, 1859)*
- *P. valeria* (Cramer, 1776)

(7) Genus *Eronia* Hübner, (1823)
- *E. cleodora* Hübner, (1823)
- *E. dilatata* Butler, 1888 (stat. rev.)

(8) Genus *Afrodryas* Stoneham (1957) (stat. rev.)
- *A. leda* (Boisduval 1847) (comb. rev.)

(9) Genus *Pinacopteryx* Wallengren (1857)
- *P. eriphia* (Godart, 1819)

P. mabillei (Aurivillius, 1899) (stat. nov.)
(10) Genus *Teracolus* Swainson, (1833) (stat. rev.)
T. agoye (Wallengren 1857) (comb. rev.)
T. eris (Klug 1829) (comb. rev.)
T. subfasciatus Swainson (1833) (comb. rev.)
(11) Genus *Colotis* Hübner, (1819)
Group I.
C. etrida (Boisduval, 1836)
C. ephyia (Klug 1829)
Group II.
C. aurora (Cramer, 1780)
C. evarne (Klug 1829) (stat. rev.)
C. incretus (Butler, 1881) (stat. rev.)
C. auxo (Lucas, 1852)
ssp. *dissociatus* (Butler, 1897) (stat. nov.)
Group III.
C. antevippe (Boisduval, 1836)
C. rogersi (Dixey, 1915)
C. euippe (Linnaeus, 1758)
C. pallene (Hopffer, 1855)
C. lais (Butler 1876)
C. daira (Klug 1829)
C. evagore (Klug 1829)
C. evanthe (Boisduval, 1836)
ssp. *evanthides* (Holland, 1896) (stat. nov.)
Group IV.
C. liagore (Klug 1829)
C. evenina (Wallengren 1857)
Group V.
C. annae (Wallengren 1857) (stat. rev.)
ssp. *hildebrandti* (Staudinger 1884) (stat. nov.)
C. guenei (Mabille, 1877)
C. danae (Fabricius 1775)
ssp. *eupompe* (Klug 1829) (=*pseudacaste* (Butler 1876) (syn. nov.)
ssp. *dulcis* (Butler 1876)
Group VI.
C. protractus (Butler 1876)
C. fausta (Olivier, 1804)
C. amata (Fabricius 1775)
C. calais (Cramer, 1775) (stat. rev.)
C. castalis (Staudinger, 1885) (stat. rev.)
C. phisadia (Godart, 1819)
ssp. *vestalis* (Butler 1876) (stat. nov.)
Group VII.
C. zoe (Grandidier, 1867)
C. celimene (Lucas, 1852)
C. praeclarus (Butler, 1886)
Group VIII.
C. ungemachi (Le Cerf, 1922)
C. doubledayi (Hopffer, 1862)
C. chrysonome (Klug 1829)
C. vesta (Reiche, 1850)
C. aurigineus (Butler, 1883)
Group IX.
C. protomedia (Klug 1829)
C. halimede (Klug 1829)
C. pleione (Klug 1829)
C. venosa (Staudinger 1884)
C. mananhari (Ward, 1870)
C. regina (Trimen, 1863)
C. hetaera (Gerstaecker, 1871)
C. elgonensis (Sharpe, 1891)
C. ione (Godart, 1819)
C. erone (Angas, 1849)
C. eunoma (Hopffer, 1855)*

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. TCS haplotype networks for cytochrome oxidase I barcodes in the *Colotis vesta* (a) and *Colotis phisadia* (b) groups of species.

Table S1. Material examined and GenBank Accession numbers.

Table S2. Taxa and sequences used in phylogenetic reconstructions.

Table S3. Summary of character partitions.

Table S4. Support values for clades supported through both maximum likelihood and Bayesian analyses.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.