



Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogenetics of Coenonymphina (Nymphalidae: Satyrinae) and the problem of rooting rapid radiations

Ullasa Kodandaramaiah^{a,*}, Carlos Peña^b, Michael F. Braby^{c,d}, Roger Grund^e, Chris J. Müller^f, Sören Nylin^a, Niklas Wahlberg^{a,g}

^aDepartment of Zoology, Stockholm University, 106 91 Stockholm, Sweden

^bMuseo de Historia Natural, Universidad Nacional Mayor de San Marcos, Av. Arenales 1256 Apartado 14-0434, Lima-14, Peru

^cMuseum and Art Gallery Northern Territory, Darwin, NT 0801, Australia

^dSchool of Botany and Zoology, The Australian National University, Canberra, ACT 0200, Australia

^e9 Parkers Road, Torrens Park, Adelaide, SA 5062, Australia

^fMolecular Ecology Lab, Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

^gLaboratory of Genetics, Department of Biology, University of Turku, 20014 Turku, Finland

ARTICLE INFO

Article history:

Received 20 April 2009

Revised 10 August 2009

Accepted 12 August 2009

Available online 15 August 2009

Keywords:

Rapid radiation
Coenonymphina
Hypocystina
Outgroup rooting
Outgroup selection

ABSTRACT

We report a rapid radiation of a group of butterflies within the family Nymphalidae and examine some aspects of popular analytical methods in dealing with rapid radiations. We attempted to infer the phylogeny of butterflies belonging to the subtribe Coenonymphina *sensu lato* using five genes (4398 bp) with Maximum Parsimony, Maximum Likelihood and Bayesian analyses. Initial analyses suggested that the group has undergone rapid speciation within Australasia. We further analyzed the dataset with different outgroup combinations the choice of which had a profound effect on relationships within the ingroup. Modelling methods recovered Coenonymphina as a monophyletic group to the exclusion of *Zipaetis* and *Orsotriaena*, irrespective of outgroup combination. Maximum Parsimony occasionally returned a polyphyletic Coenonymphina, with *Argyronympha* grouping with outgroups, but this was strongly dependent on the outgroups used. We analyzed the ingroup without any outgroups and found that the relationships inferred among taxa were different from those inferred when either of the outgroup combinations was used, and this was true for all methods. We also tested whether a hard polytomy is a better hypothesis to explain our dataset, but could not find conclusive evidence. We therefore conclude that the major lineages within Coenonymphina form a near-hard polytomy with regard to each other. The study highlights the importance of testing different outgroups rather than using results from a single outgroup combination of a few taxa, particularly in difficult cases where basal nodes appear to receive low support. We provide a revised classification of Coenonymphina; *Zipaetis* and *Orsotriaena* are transferred to the tribe Eritina.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Rapid radiations have been documented in several groups of organisms – plants (Harris et al., 2000; Fishbein et al., 2001; Malcomber, 2002; Shaw et al., 2003), insects (Mardulyn and Whitfield, 1999; Jordal et al., 2000; Von Dohlen and Moran, 2000; Lockhart and Cameron, 2001; Braby and Pierce, 2007; Peña and Wahlberg, 2008), shrimps (Morrison et al., 2004), amphibians (Mahoney, 2001), snakes (Wiens et al., 2008), birds (Barker et al., 2004) and

* Corresponding author. Fax: +46 8 167 715.

E-mail addresses: ullasa.kodandaramaiah@zoologi.su.se (U. Kodandaramaiah), carlosp420@yahoo.com (C. Peña), michael.braby@nt.gov.au (M.F. Braby), rgrund@chariot.net.au (R. Grund), cmuller@bio.mq.edu.au (C.J. Müller), soren.nylin@zoologi.su.se (S. Nylin), niklas.wahlberg@utu.fi (N. Wahlberg).

mammals (Halanych and Robinson, 1999; Lin et al., 2002). Phylogenetic reconstructions of such groups have posed significant challenges (Whitfield and Lockhart, 2007). Resolution of nodes preceded by short branches (internal nodes) has proven to be difficult, even with large amounts of data (Tajima, 1983; Wiens et al., 2008; Hallström and Janke, 2008). Furthermore, rooting of trees with outgroups has been problematic, especially when the outgroup taxa are distant from the ingroup (Lin et al., 2002; Zanis et al., 2002; Slack et al., 2003; Blanga-Kanfi et al., 2009). Here, we report a rapid radiation of a group of butterflies within the subfamily Satyrinae (Nymphalidae) and examine in detail some aspects of popular analytical methods in dealing with rapid radiations.

Peña et al. (2006), in their phylogenetic study of the subfamily Satyrinae, identified several clades that did not conform to

the accepted morphological classification of the subfamily (Harvey, 1991), which was based on the monograph by Miller (1968). One major clade in Peña's study consisted of genera from the subtribe Hypocystina (represented by 13 genera), subtribe Coenonymphina (represented by *Coenonympha*), *Oressinoma* (Euptychiina) and *Orsotriaena* (then placed in Mycalesina). Although they did not include all genera in the two subtribes Coenonymphina and Hypocystina, it was clear that their Coenonymphina was paraphyletic and Hypocystina polyphyletic; hence, delineation of these two subtribes was not satisfactory. Peña et al. (2006) therefore informally proposed that Hypocystina should be synonymized with Coenonymphina and its studied genera included under Coenonymphina along with *Oressinoma* and *Orsotriaena*. Thus, Coenonymphina, as proposed by Peña et al. (2006), includes four distinct zoogeographic elements: Holarctic (*Coenonympha*), Australasian (many genera), Oriental (*Zipaetis*, *Orsotriaena*) and Neotropical (*Oressinoma*). More recently, Peña and Wahlberg (2008) and Peña et al. (in press) included some coenonymphine genera in their studies. In their analyses, *Zipaetis* and *Orsotriaena* formed a clade with members of Eritina and were together sister to a clade composed of the remaining members of Coenonymphina. A study of three Holarctic members of the group (*Lyela*, *Triphysa*, *Coenonympha*) showed that *Coenonympha* was paraphyletic with respect to *Lyela* and *Triphysa*; hence the latter two genera were synonymized under *Coenonympha* (Kodandaramaiah and Wahlberg, 2009). However, Kodandaramaiah and Wahlberg (2009) did not include data from the fourth Palaearctic genus, *Sinonympha*, and thus its affinity with other members of Coenonymphina remains unclear. Miller (1968) divided his Hypocystini into the *Xenica* and *Hypocysta* series. The *Xenica* series was composed of temperate species found mainly in the Australian mainland, Tasmania and New Zealand while the *Hypocysta* series consisted of genera distributed predominantly in New Guinea, the Solomon Islands and New Caledonia (Table 1).

Table 1

Current classification of Coenonymphina and revised classification scheme proposed in this study. An asterisk indicates that the genus was part of Miller's *Hypocysta* series and a double asterisk indicates that it belongs to his *Xenica* series.

Current classification (Peña et al., 2006)	Proposed classification
Coenonymphina	Coenonymphina
<i>Coenonympha</i>	<i>Argyronympha</i>
<i>Sinonympha</i> (tentative)	<i>Oressinoma</i>
<i>Altiapa</i>	<i>Sinonympha</i>
<i>Argyronympha</i> *	<i>Coenonympha</i>
<i>Dodonidia</i> **	<i>Erycinidia</i>
<i>Argyrohenga</i> **	<i>Oreixenica</i>
<i>Percnodaimon</i> **	<i>Paratisiphone</i>
<i>Erebiola</i> **	<i>Tisiphone</i>
<i>Erycinidia</i> *	<i>Nesoxenica</i>
<i>Harsiesis</i> *	<i>Dodonidia</i>
<i>Argynnina</i> **	<i>Erebiola</i>
<i>Geitoneura</i> **	<i>Argyrohenga</i>
<i>Heteronympha</i> **	<i>Percnodaimon</i>
<i>Nesoxenica</i> **	<i>Argynnina</i>
<i>Hyalodia</i> *	<i>Heteronympha</i>
<i>Hypocysta</i> *	<i>Geitoneura</i>
<i>Lamprolenis</i> *	<i>Platythima</i>
<i>Paratisiphone</i> **	<i>Harsiesis</i>
<i>Platythima</i> *	<i>Lamprolenis</i>
<i>Oreixenica</i> **	<i>Hypocysta</i>
<i>Tisiphone</i> **	<i>Altiapa</i>
<i>Zipaetis</i> *	<i>Hyalodia</i>
<i>Orsotriaena</i>	
<i>Oressinoma</i>	
	Moved to Eritina
	<i>Orsotriaena</i>
	<i>Zipaetis</i>

In summary, the tribe Coenonymphina now contains 24 genera, including members of the former Hypocystina, *Oressinoma* and *Coenonympha* (Parsons, 1998; Braby, 2000; Bozano, 2002; Peña et al., 2006, in press; Kodandaramaiah and Wahlberg, 2009). The inclusion of *Zipaetis*, *Orsotriaena* and *Sinonympha* is tentative, needing further corroboration. Our objectives at the outset were to define Coenonymphina and to clarify relationships of genera within the group. Our preliminary analyses indicated that the basal branches were very short and phylogenetic relationships of major lineages were unstable. We conducted more rigorous analyses with increased data to investigate the effects of outgroup choice and stability of clades recovered.

Hard or near-hard polytomies can theoretically result from rapid speciation where extremely short internal (basal) branches are followed by much longer branches (Braby et al., 2005; Shavit et al., 2007). If successive speciation events proceed rapidly, there is little time for synapomorphies to accumulate in the short period between two speciation events (Whitfield and Kjer, 2008). Even if synapomorphies are accumulated in the internal branch, phylogenetic reconstruction methods can be misled if at least two of the external (terminal) branches are much longer than the internal branch, since homoplastic changes on these longer branches can override the signal in the internal branch (long-branch attraction; Felsenstein, 1978; Bergsten, 2005). Long branches are thought to affect Maximum Parsimony (MP) methods more severely (Felsenstein, 1978; Philippe et al., 2005), but model-based methods such as Bayesian Inference (BI) and Maximum Likelihood (ML) are not immune to the problem (Siddall, 1998; Omilian and Taylor, 2001; Schwarz et al., 2004; Bergsten, 2005). Generally speaking, most problems of rapid radiations can be attributed to their resemblance to hard polytomies.

Although empirical studies can highlight many problems posed by rapid radiations, they cannot be used to pinpoint the sources of errors and estimate the accuracy of a given phylogenetic reconstruction method, because the 'true' phylogeny remains unknown. Simulation studies are useful in this respect. Studies by Holland et al. (2003) and Shavit et al. (2007) have suggested that the choice of outgroups can affect relationships within the ingroup significantly. In their studies, adding outgroups frequently disrupted ingroup relationships that were correctly inferred when analyzed without outgroups and the ingroup was most accurately recovered when no outgroup was used. However, the effect of different outgroup combinations on ingroup relationships has rarely been investigated in detail on 'real world' datasets. A typical phylogenetic analysis includes a single set of outgroup taxa that are presumed to be close enough to the ingroup taxa (Nixon and Carpenter, 1993). In this study, we explicitly test the effect of different outgroup combinations on ingroup relationships and compare them to the results from the analysis without outgroups.

It is quite likely that for a given rapid radiation there have been no substitutions in the genes used to infer the phylogeny. If this is true, the real phylogeny for that particular gene combination is a hard polytomy. All three methods MP, ML and BI have been shown to be biased towards a spurious resolved topology in such cases (Suzuki et al., 2002; Cummings et al., 2003; Shavit et al., 2007), with BI especially prone to returning high posterior probability values for non-existent nodes (Susko, 2008). Lewis et al. (2005) proposed a solution to this problem. They suggested a modification of the MCMC (Markov Chain Monte Carlo) algorithm currently used in BI so that less resolved tree topologies with polytomies could be allowed to compete with fully resolved topologies using the reversible-jump MCMC algorithm (Green, 1995). This method is potentially useful in analyses of rapid radiations to determine whether a hard polytomy is the best hypothesis for the genes comprising the dataset used in a study. We applied this method to our dataset to test for a hard polytomy.

2. Materials and methods

2.1. Data collection

We included 21 out of 22 genera of Coenonymphina (we were unable to procure samples of *Hyalodia*, which is endemic to western Papua), as well as *Zipaetis* and *Orsotriaena*. Sequences of relevant genera from Peña et al. (2006, in press) and Kodandaramaiah and Wahlberg (2009) were also included in the current study. Supplementary material 1 lists the samples used in this study along with their collection localities. Specimens were collected by the authors and collaborators and preserved either in ethanol or by desiccation. DNA was extracted from two legs using QIAGEN's DNeasy extraction kit (Hilden, Germany). We initially sequenced the same three genes used in Peña et al. (2006) and Kodandaramaiah and Wahlberg (2009) using protocols described therein. These are COI, EF1- and *wingless*, amounting to a total of 3090 bp. The combination of these three genes to reconstruct phylogenies at similar taxonomic levels has been proven in previous studies on butterflies (Wahlberg et al., 2005; Simonsen et al., 2006; Kodandaramaiah and Wahlberg, 2007; Wahlberg and Freitas, 2007; Silva Brandão et al., 2008). Initial analyses on this three gene dataset indicated poorly resolved basal nodes. To test whether additional data improves resolution, we sequenced two more genes, GAPDH and RPS5. These two genes have been found to be phylogenetically informative within various subfamilies of Nymphalidae (Peña and Wahlberg, 2008; Aduse-Poku et al., 2009; Wahlberg et al., 2009), and display variation on par with *wingless* (Wahlberg and Wheat, 2008). GAPDH was amplified with the primers Frigga and Burre, while the primers RpSfor and RpSrev amplified RPS5. Both primer pairs and protocols were adopted from Wahlberg and Wheat (2008). Outgroup sequence data were taken from Peña and Wahlberg (2008). Supplementary material 1 lists the Genbank accession numbers for samples used in the study.

2.2. Phylogenetic analyses

The combined dataset of five genes including those from eight outgroup taxa (henceforth the '8-outgroup' dataset) was analyzed using MP, ML and BI. MP analyses were conducted in TNT v 1.1 (Goloboff et al., 2008). New Technology searches (Goloboff, 1999; Nixon, 1999) consisting of Tree Fusion, Ratchet, Tree Drifting and Sectorial searches were performed on 1000 random additional replicates. Support for individual clades was calculated using Bootstrap proportions that were based on 1000 pseudo-replicates with 10 random replicates each.

ML analyses were performed in RAxML III (Stamatakis et al., 2008) assuming the GTR + G model of substitution, which was chosen by the software jModelTest (Posada, 2008). The dataset was partitioned into five categories corresponding to the genes, with model parameters estimated individually for each partition. Support for the nodes recovered was estimated from 1000 bootstrap replicates. BI analyses were performed in MrBayes 3.2 (Ronquist and Huelsenbeck, 2003). Partitioning of the dataset was as in the ML analysis and model parameters were unlinked between partitions. The MCMC chains were run either for 5,000,000 generations or until the average standard deviation of split frequencies dropped below 0.01.

We performed a series of analyses to test whether outgroup choice affected relationships inferred within the ingroup. We first conducted the MP, ML and BI analyses described above on a dataset with 23 outgroups by adding 15 additional outgroups. This dataset will henceforth be referred to as the '23-outgroup' dataset. The same analyses were also repeated without outgroup taxa (henceforth the 'ingroup' dataset). Furthermore, we performed

tests to distinguish between effects of outgroup number (outgroup sampling density) and varying outgroup combination. We put together eight datasets, each having eight randomly chosen outgroups from the 23-outgroup dataset in addition to *Zipaetis* and *Orsotriaena*. These datasets were analyzed in TNT and compared to results from the MP analysis on the 8-outgroup dataset. We also performed MP, ML and BI analyses with only COI, EF1- and *wingless* to make them comparable with the many published studies that have used these three genes.

Partitioned Bremer support (PBS) values were used to estimate conflict between partitions (Baker and DeSalle, 1997), and were calculated in TNT using a script written by one of the authors (Carlos Peña; available from www.zmuc.dk/public/phylogeny/tnt/scripts/pbsup.run). Due to anomalous results, the PBS analyses were conducted on the 23-outgroup, 8-outgroup and the ingroup only datasets. Partition Congruence Index (PCI) values were used to summarize these PBS values (Brower, 2006). If all partitions support a particular node, the PCI value equals the total Bremer support (BS) value from all partitions for that node. The PCI value decreases with increasing conflict between partitions and becomes negative with high amounts of conflict (Brower, 2006).

The 8-outgroup dataset was analyzed with the software Phycas v1.1.2 (www.phycas.org) to test whether a tree with one or more polytomous nodes was a better hypothesis than trees with a fully resolved topology. The MCMC algorithms in the most widely used software such as MrBayes and BEAST (Drummond and Rambaut, 2007) only consider and evaluate the likelihoods of fully resolved topologies. Thus, less resolved topologies with one or more polytomies do not feature in the posterior distribution of trees. This leads to spurious resolution in groups whose true phylogenies include hard polytomies. To overcome this shortcoming, Lewis et al., 2005 proposed that less resolved topologies should be evaluated in addition to completely resolved topologies. However, less resolved topologies are of different dimension compared to a completely bifurcate pattern because they have fewer branch length parameters. Thus their evaluation necessitates a jump between models of differing dimensions, which cannot be done by the MCMC in MrBayes and BEAST. Lewis and colleagues offered a solution that involves a modification to MCMC called the reversible-jump MCMC (Green, 1995). This modification allows the MCMC run to explore a tree space that includes all topologies: fully bifurcate, less resolved and completely star-like. This algorithm has been implemented in Phycas. Currently, Phycas does not allow partitioning of the dataset. The GTR + G model was imposed on the combined dataset. A gamma distribution with mean 0.5 was used as the prior for the gamma shape parameter. An exponential distribution with mean 1.0 was set as the prior for both the ratio of the rate of transitions to the rate of transversions and the mean of the branch lengths. Polytomies and resolved topologies were assigned equal priors. The MCMC chain was run for 60,000 cycles after discarding the first 1000 trees as burnin (note that one cycle involves more calculations in Phycas and is not the equivalent of one generation in MrBayes), with trees sampled every 10 cycles.

3. Results

The combined dataset including seven outgroups consisted of 4447 bp, of which 1490 bp were parsimony informative. We will refer to the Bayesian analysis in MrBayes as MB to distinguish it from the Phycas analysis. The trees from the three analyses (MP, MB and ML) on the 8-outgroup dataset were in conflict at several nodes. Fig. 1 depicts the ML tree and Supplementary material 2a and 3a show the MP and MB trees. Basal branches in ML and MB were shorter than terminal branches and the nodes that they preceded generally had poor support. Three nodes on the 50% majority

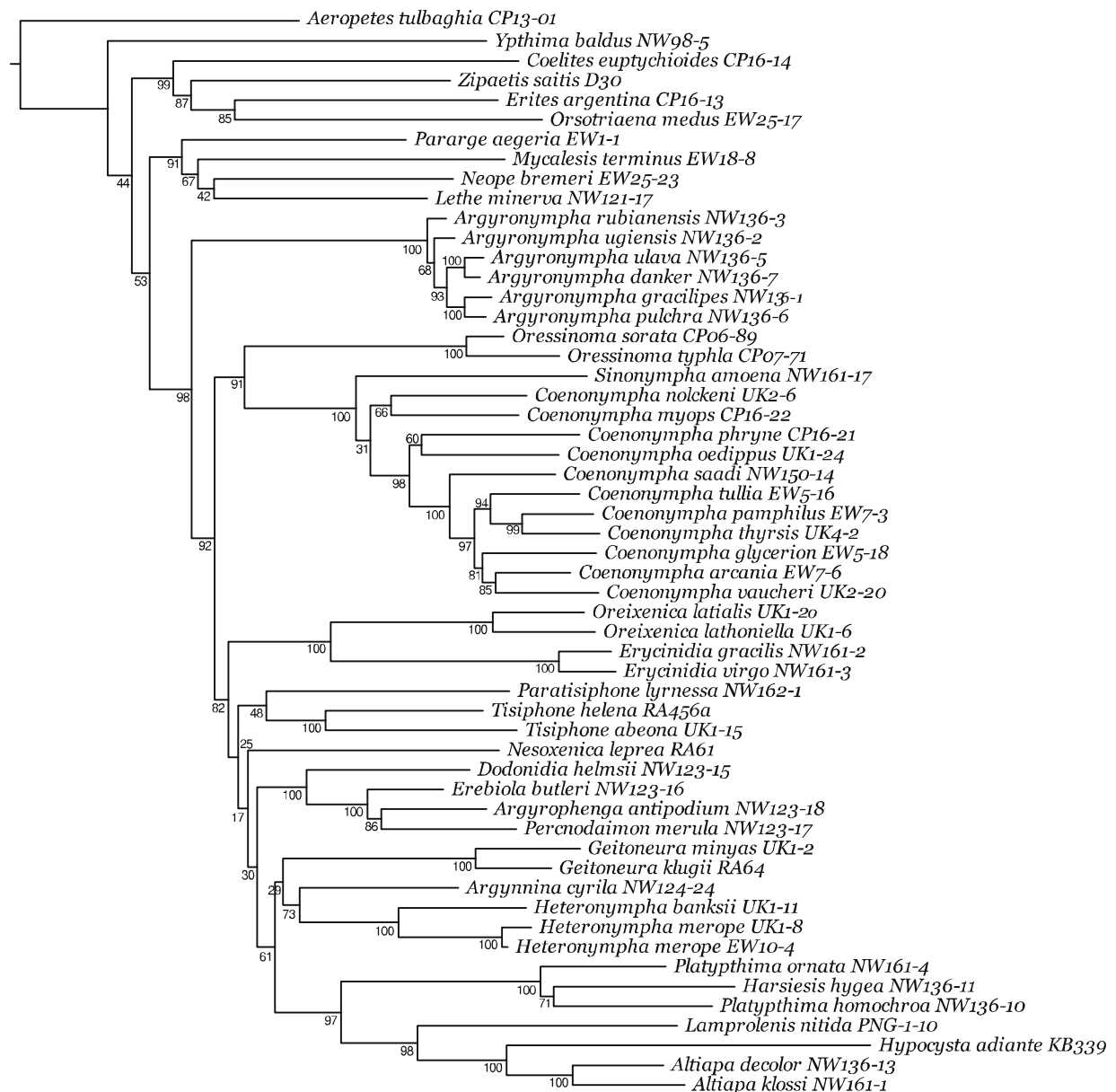


Fig. 1. Maximum Likelihood tree from the RAXML analysis on the 8-outgroup dataset. Numbers below branches are bootstrap values.

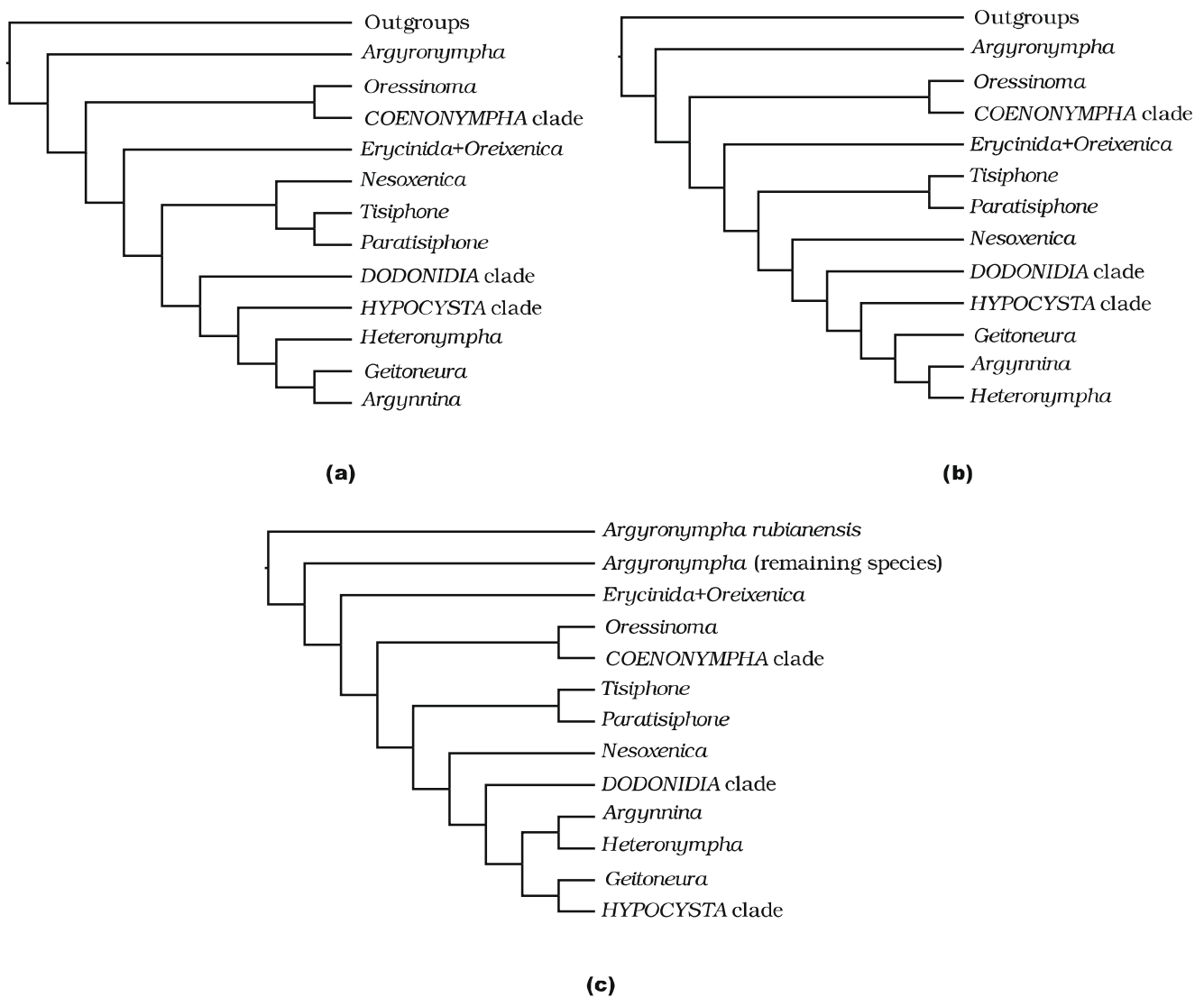
rule consensus tree from MB were unresolved (Supplementary material 2a). Only one MP tree was recovered (Supplementary material 3a).

Sinonympha was sister to *Coenonympha* with strong support in the model-based analyses (MB: 1.0 posterior probability; ML: 100% bootstrap). *Sinonympha* was nested within *Coenonympha* and they were together monophyletic with strong support in the MP analysis (100% bootstrap). In all three analyses the genera comprising Coenonymphina excluding *Zipaetis* and *Orsotriaena* formed a monophyletic group, with good support in MB and ML (MP: 53, MB: 1.0, ML: 98), although relationships within this clade differed between the three. This clade will henceforth be referred to as the Coenonymphina clade.

Zipaetis and *Orsotriaena* formed a clade with *Erites* and *Coelites* (both Satyrini: Eritina) with good support in MB and ML (MB: 1.0, ML: 99) but this clade was not sister to the Coenonymphina clade. This clade was also recovered in the MP tree, but with poor support (<50), and as in the other two analyses not sister to the Coenonymphina clade.

All analyses on the 23-outgroup dataset indicated the monophyly of the Coenonymphina clade (MP: 66, MB: 1.0; ML: 100). The relationships within the clade differed significantly among the three trees. They were also different from respective analyses (i.e., analyzed with the same method) on the 8-outgroup dataset. The MB tree was less resolved than the 8-outgroup tree at two nodes. The *Zipaetis*–*Orsotriaena*–*Erites*–*Coelites* clade was recovered in all analysis, with strong support in MB and ML (MB: 1.0; ML: 93; MP: <50).

The Coenonymphina clade was used in the ingroup analyses. As with the previous sets of analyses, the different analyses recovered discordant relationships. For all three methods, recovered relationships differed from those analyzed using the same method but with outgroups. Fig. 2(a,b,c) depict the three different topologies recovered in the three ML analyses. Fig. 3(a,b,c) and Supplementary material 3b show the same for the MP and MB analyses, respectively. Analyses with only COI, EF1- α and *wingless* showed similar patterns of discordance among the three methods of analysis and the effect of outgroups on ingroup relationships (results not shown).



COENONYMPHA clade - *Coenonympha*, *Sinonympha*
 DODONIDIA clade - *Dodonidia*, *Erebiola*, *Percnodaimon*, *Argyrophenga*
 HYPOCYSTA clade - *Platyptilima*, *Lamprolenis*, *Erycinidia*, *Hypocysta*, *Altiapa*

Fig. 2. Topologies recovered in the three Maximum Likelihood analyses in RaxML. (a) 23-outgroup dataset. (b) 8-outgroup dataset. (c) Ingroup analyzed without outgroups, rooted with *Argyronympha rubianensis* for comparison with a and b.

The PBS analyses showed high to extremely high conflict at many nodes in the 23-outgroup and the 8-outgroup analyses (especially among outgroup relationships). The conflict did not have similar patterns in the two analyses, rather in the 23-outgroup analysis the conflict was found randomly in the different gene regions, whereas in the 8-outgroup dataset COI and EF1- α were systematically in conflict with the other three gene regions, even for the same nodes found in the 23-outgroup analysis (Supplementary material 3 and 4). Interestingly, in the ingroup only PBS analysis, most conflict disappeared (Supplementary material 4). We believe this has something to do with outgroups being long branch taxa and these are interacting with the long branches of the ingroup taxa in unpredictable ways. Investigating this in more detail is beyond the scope of this paper and will be addressed elsewhere.

Where tested, all genera except *Platyptilima* and *Coenonympha* were monophyletic with good support. *Harsiesis hygea* was nested within the two *Platyptilima* species used in this study. Similarly, *Coenonympha* was paraphyletic with respect to *Sinonympha* in ingroup and all MP analyses. Basal nodes received poor support in all analyses. Clades that received strong support were recovered in all trees and showed little or no conflict in the PBS analysis. *Coenonympha-Sinonympha* was one such clade. Among the Australasian genera, three clades - '*Dodonidia-Erebiola-Percnodaimon-Argyrophenga*', '*Erycinidia-Oreixenica*' and '*Hypocysta-Altiapa-Lamprolenis-Platyptilima-Harsiesis*' were well-supported and recovered in all analyses. We refer to these clades that were recovered in all above mentioned analyses as 'stable' clades. Additionally, members of the Coenonymphina clade to the exclusion of *Argyronympha* formed a clade that was recovered in all analyses as the sister

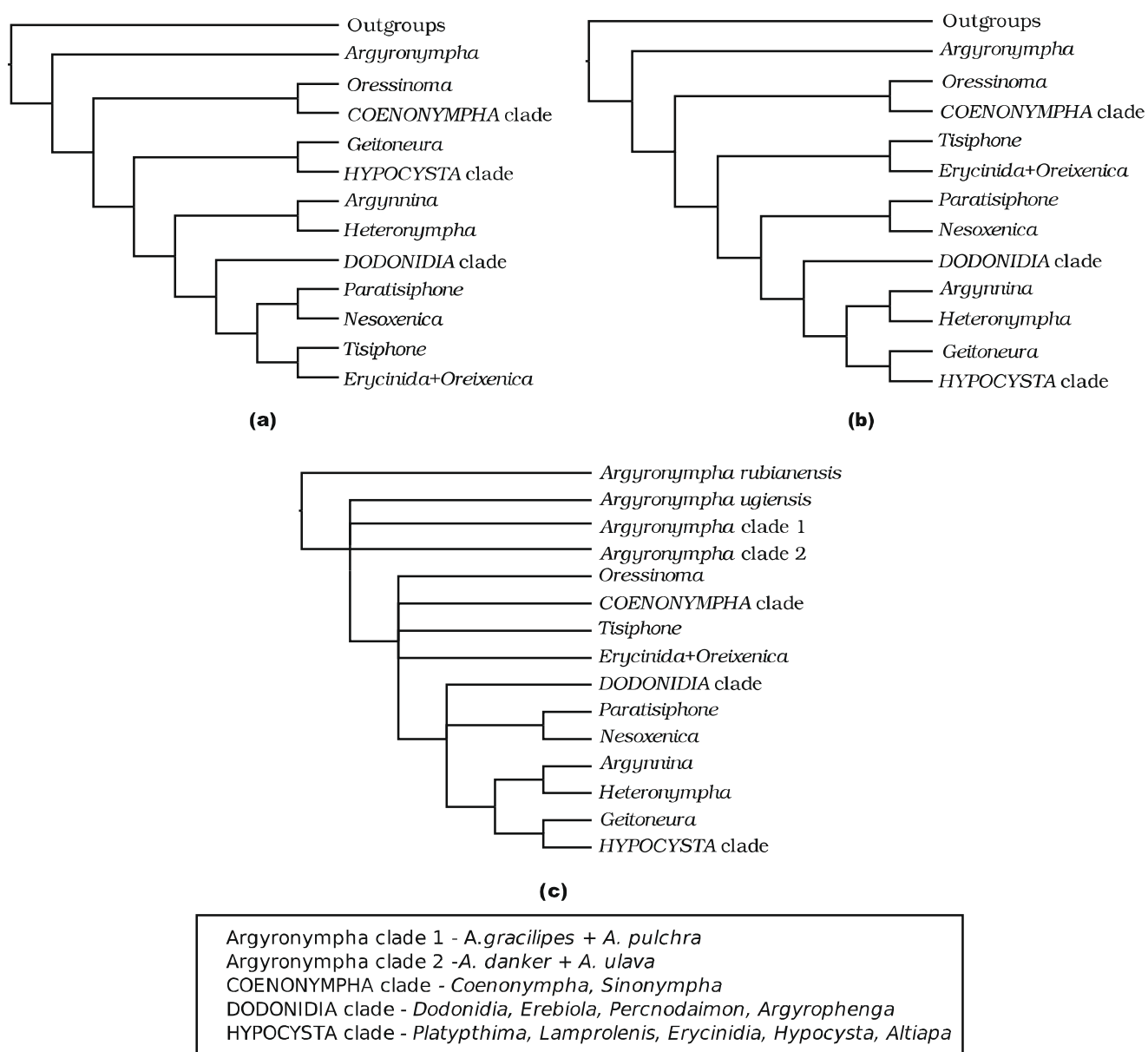


Fig. 3. Strict consensus topologies recovered in the Maximum Parsimony analyses in TNT. (a) 23-outgroup dataset. (b) 8-outgroup dataset. (c) Ingroup analyzed without outgroups, rooted with *Argyronympha rubianensis* for comparison with a and b. Note that *Argyronympha* is monophyletic in the unrooted tree.

group to *Argyronympha*. Furthermore, *Oressinoma* was sister to *Coenonympha* + *Sinonympha* with strong support in all 6 model-based analyses and this relationship was also recovered with weak support (<50%) in the MP analysis with 8 and 23 outgroups.

The eight MP analyses with randomly chosen combinations of 8 outgroups resulted in eight unique topologies (Supplementary material 5). The stable clades mentioned above were recovered in all eight analyses. However, the Coenonymphina clade was not monophyletic in five instances (Supplementary material 5b, c, d, f, and g). *Paratisiphone* was sister to *Nesoxenica* in all eight trees, and this clade was also recovered in the first three MP analyses (Fig. 3).

The Bayesian analysis in Phycas recovered the Coenonymphina clade with strong support (Posterior Probability 1.0; Supplementary material 6). The topology was slightly more resolved than the MB topology of the 8-outgroup dataset and all the stable clades appeared with strong support. Relationships were congruent with those from the MrBayes analysis, with the exception of the position

of *Paratisiphone* as either sister to *Tisiphone* in the MB analysis or sister to *Nesoxenica* in the Phycas analysis (Supplementary material 2a, 6).

4. Discussion

We first consider results from the MP, ML and MB analyses. Concordance between different methods of analysis, especially between MP and model-based methods is considered to be an indication of strong phylogenetic signal and stability of the relationships inferred (Kim 1993; Brooks et al., 2007; Wahlberg and Freitas, 2007; Wahlberg and Wheat, 2008). The monophyly of the Coenonymphina clade is very strongly supported in Maximum Likelihood and Bayesian analyses. This clade was also recovered in most MP analyses, albeit with low support. The lower support in MP is possibly due to long-branch attraction artefacts, which are difficult to identify because of the presence of multiple long branches within the clade. In instances where the monophyly of the clade was dis-

rupted, it was only because *Argyronympha* grouped with outgroup taxa. The MB and ML analyses indicate that the branch leading to this genus is longer than in other genera. In light of these results and considering the stability of the clade excluding *Argyronympha*, we regard the Coenonymphina clade to be reasonably stable. This is also in agreement with Peña et al. (in press), who found *Argyronympha* to be sister to the rest of the Coenonymphina clade in their broad study of the tribe Satyrini. We thus propose that taxa forming the Coenonymphina clade, including *Argyronympha*, should henceforth be classified in the subtribe Coenonymphina.

On a similar note, the *Zipaetis*–*Orsotriaena*–*Erites*–*Coelites* group was monophyletic in all analyses. *Erites* and *Coelites* are currently placed in Eritina, although Peña et al. (in press) did not find them to group together. *Zipaetis* was tentatively placed within Hypocystina by Miller (1968), who noted that the morphology of *Zipaetis* was 'exceptionally aberrant'. It is clear that *Zipaetis* does not belong in Coenonymphina and should be moved to Eritina, along with *Orsotriaena* (see also Peña et al., in press).

Relationships within the Coenonymphina clade changed with change in outgroup composition, irrespective of the analytical method used. The 23-outgroup, 8-outgroup and ingroup datasets had the same composition of terminals within Coenonymphina, while the outgroup sampling density differed. Analyses on datasets with eight randomly chosen outgroup taxa all had the same taxon sampling density, yet the results differed in inferred relationships within Coenonymphina. This shows that ingroup relationships can be affected by changes both in outgroup sampling density and combination of taxa used as outgroups.

Other studies have shown that distant outgroups can attach themselves randomly within the ingroup due to long-branch attraction, thereby disturbing ingroup relationships (Lin et al., 2002; Slack et al., 2003; Fric et al., 2007; Gatesy et al., 2007). We believe the disruption of monophyly of Coenonymphina in some analyses (e.g., Supplementary material 4a and d) is due to such long branch attraction. In all other instances, however, the monophyly of Coenonymphina was intact and outgroups always attached themselves to the branch between *Argyronympha* and its sister group. Despite being attached to the same branch, outgroups had an impact on relationships within Coenonymphina. This phenomenon has been shown previously in simulation studies (Shavit et al., 2007), but to our knowledge this is the first study to demonstrate it in an empirical dataset. The effect of outgroups is most pronounced in the basal parts of the Coenonymphina clade.

The finding that outgroup choice can affect ingroup relationships has implications for the use of phylogenies in biology. Although it is acknowledged that weakly supported relationships can only result in poorly supported inferences based on the phylogeny, in many cases a single topology is chosen to be used for further analyses of divergence time estimates, biogeography, character evolution, etc. The most common approach is to use a small set of outgroups in the analysis with the intention of rooting trees. In the analyses used here, rooting with outgroups is a *post hoc* action with no effect on the tree search. Yet, the current study indicates that the taxa included as outgroups in the analysis can affect relationships inferred among the taxa designated as the ingroup. Our results show that it can be very useful and important to evaluate the effects of varying outgroup combinations, both in terms of recovering a monophyletic ingroup and in obtaining a better indication of the robustness of relationships within it.

Within the Coenonymphina clade, the branches in the MB and ML trees are characterized by the presence of short basal branches followed by much longer branches. There is a clear trend for basal nodes to be weakly supported and well-supported groups to be preceded by longer branches. The combination of the three genes, COI, EF1- and *wingless*, that has resolved relationships at similar

taxonomic levels time and again fails to do so for this particular dataset. Furthermore, the tree topology with respect to basal nodes was sensitive to the choice of outgroups, a phenomenon that has been only been reported in simulation studies of near-hard polytomies (Shavit et al., 2007). In summary, these findings strongly suggest that the members of the clade are the result of a rapid radiation and conform to properties of near-hard polytomies.

One reason for poorly supported basal nodes is that there might not have been any mutational changes in the five genes during the short periods between early speciation events in the group. In this case, the true phylogeny for the five genes used is a hard polytomy. Although various software have had the option of collapsing nodes if branch lengths are zero, they are clearly biased towards resolved trees (Lewis et al., 2005; Shavit et al., 2007). The method of Lewis and colleagues implemented in Phycas can theoretically detect true hard polytomies. In our analysis, the consensus tree from Phycas was in fact more resolved than the MB tree with the same outgroups. There was one polytomy within Coenonymphina; however, inspection of sampled trees indicated that this polytomy was not solely because it appeared in more than 50% of the samples, but also because of conflict between trees. Overall, this analysis suggests that hard polytomous nodes within Coenonymphina do not have a higher posterior probability than resolved topologies.

Since all genera except one are included in the study, it is unlikely that better sampling of extant taxa will increase resolution within this group. Increased data from more genes may result in better resolution and stronger support for the basal nodes. The amount of data analyzed here is on the higher side compared with the average molecular systematic study on animal taxa. How much more data needs to be added before the basal branches become stable? The advent of the phylogenomic era in phylogenetics has allowed several authors to utilize genomic-scale data in an attempt to resolve difficult phylogenies. Hallström and Janke (2008) used data from 3012 genes (amounting to 2,844,615 bp) in their phylogenomic analysis of placental mammals but were unable to resolve divergences that occurred in less than a 4 my time-frame. They surmised that historic processes such as hybrid speciation or introgression may obscure speciation patterns. Other studies (Hackett et al., 2008; Wiens et al., 2008) have found similar patterns. We side with Rokas and Carroll (2006) in their view that there will be several bushes in the tree of life and that these are not necessarily failures of phylogenetic methodology, but a portrayal of the kind of evolution that has taken place. Coenonymphina is perhaps one of these.

4.1. Systematic relationships within Coenonymphina

Miller placed *Erycinidia* in his *Hypocysta* series on the basis of midleg tibial spurs being absent, but in fact spurs are present in *Erycinidia*, but are very reduced (Grund 2006). In this study, *Erycinidia* is sister to *Oreixenica* (*Xenica* series) with strong support in all analyses. The two genera share the morphological character of a strongly sclerotised annular juxtal-ring around the aedeagus of the male genitalia (Grund, 2006). Even with *Erycinidia* moved to the *Xenica* series, Miller's classification into two series is not supported by our results. Genera in the *Hypocysta* series form a clade, but the *Xenica* series is paraphyletic with respect to this clade.

Platyphthima appears to be paraphyletic with respect to *Harsiesis*, but with our sampling we are unable to say whether *Platyphthima* really is paraphyletic, or whether the species *P. homochroa* should be transferred to *Harsiesis*. Parsons (1998) noted that the two genera are superficially similar, but that *P. homochroa* is quite different from other species of *Platyphthima*. *Harsiesis* and *Platyphthima* are the only genera of the *Hypocysta* series with absent or very reduced brachia in the male genitalia (Parsons, 1998) that are not replaced by a significant basal gnathos sclerotisation as in *Oressinoma*,

which also has no brachia. The four genera endemic to New Zealand – *Dodonidia*, *Erebiola*, *Percnodaimon* and *Argyrophenga* – have all descended from a common ancestor. *Paratisiphone* seems to be sister to *Tisiphone*, although this relationship is not recovered in all analyses. The relationship of *Sinonympha* to *Coenonympha* is unclear. Although some analyses suggest that they are sister genera, the former is nested within the latter in others, especially in the MP analyses. We refrain from proposing any taxonomic changes to these two genera.

5. Summary and conclusions

The subtribe Coenonymphina to the exclusion of *Zipactis* and *Orosotriaena* was a stable monophyletic clade. We revise the classification of the subtribe by removing the latter genera. These two genera comprise a clade with *Erites* and *Coelites* and we propose that they should henceforth be classified under Eritina. Outgroup choice had a significant effect on the topology within the in-group. The results indicate the Coenonymphina is a near-hard polytomy caused by rapid radiation. We tested whether a hard polytomy is a better hypothesis to explain the patterns found in the group, but were unable to find conclusive evidence for this. The current study includes 21 of 22 genera in the subtribe and increased taxon sampling is unlikely to improve the resolution. Addition of more data may do so, but we believe the group will remain difficult to resolve even with large amounts of data.

Acknowledgments

We thank the Swedish Research Council (Vetenskapsrådet) and the Academy of Finland (Grant No. 118369) for funding this project. We are grateful to Tania Deodati and Valerio Sbordoni for the *Sinonympha ameona* DNA extract. We also thank Julien Leneveu for providing some of the sequences used in this study. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University (put in URL:<http://cbsuapps.tc.cornell.edu/index.aspx>).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.08.012.

References

- Aduse-Poku, K., Vingerhoedt, E., Wahlberg, N., 2009. Out-of-Africa again: a phylogenetic hypothesis of the genus *Charaxes* (Lepidoptera: Nymphalidae) based on 5 gene regions. *Mol. Phylogenet. Evol.* 53, 463–478.
- Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Syst. Biol.* 46, 654–673.
- Barker, F.K., Cibois, A., Schikler, P., Feinstein, J., Cracraft, J., 2004. Phylogeny and diversification of the largest avian radiation. *Proc. Natl. Acad. Sci. USA* 101, 11040–11045.
- Bergsten, J., 2005. A review of long-branch attraction. *Cladistics* 21, 163–193.
- Blanga-Kanfi, S., Miranda, H., Penn, O., Pupko, T., DeBry, R.W., Huchon, D., 2009. Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evol. Biol.* 9, 71.
- Bozano, G.C., 2002. Satyrinae part III. In: Bozano, G.C. (Ed.), *Guide to the Butterflies of the Palearctic Region*. Omnes Artes, Milano, pp. 1–71.
- Braby, M.F., 2000. *Butterflies of Australia: Their Identification, Biology and Distribution*. CSIRO Publishing.
- Braby, M.F., Pierce, N.E., 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, M.F., Trueman, J.W.H., Eastwood, R., 2005. When and where did troidine butterflies (Lepidoptera: Papilionidae) evolve? Phylogenetic and biogeographic evidence suggests an origin in remnant Gondwana in the Late Cretaceous. *Invertebr. Syst.* 19, 113–143.
- Brooks, D.R., Bilewicz, J., Condy, C., Evans, D.C., Folinsbee, K.E., Fröbisch, J., Halas, D., Hill, S., McLennan, D.A., Mattern, M., Tsuji, L.A., Ward, J.L., Wahlberg, N., Zamparo, D., Zanatta, D., 2007. Quantitative phylogenetic analysis in the 21st century. *Revista Mexicana de Biodiversidad* 78, 225–252.
- Brower, A.V.Z., 2006. The how and why of branch support and partitioned branch support, with a new index to assess partition incongruence. *Cladistics* 22, 378–386.
- Cummings, M.P., Handley, S.A., Myers, D.S., Reed, D.L., Rokas, A., Winka, K., 2003. Comparing bootstrap and posterior probability values in the four-taxon case. *Syst. Biol.* 52, 477–487.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.
- Felsenstein, J., 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.* 27, 401–410.
- Fishbein, M., Hibscher-Jetter, C., Oltis, D.E.S., Hufford, L., 2001. Phylogeny of Saxifragales (Angiosperms, Eudicots): analysis of a rapid, ancient radiation. *Syst. Biol.* 50, 817–847.
- Fric, Z., Wahlberg, N., Pech, P., Zrzavý, J., 2007. Phylogeny and classification of the *Phengaris-Maculinea* clade (Lepidoptera: Lycaenidae): total evidence and phylogenetic species concepts. *Syst. Entomol.* 32, 558–567.
- Gatesy, J., deSalle, R., Wahlberg, N., 2007. How many genes should a systematist sample? Conflicting insights from a phylogenomic matrix characterized by replicated incongruence. *Syst. Biol.* 56, 355–363.
- Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics* 15, 415–428.
- Goloboff, P.A., Farris, J.S., Nixon, K., 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24, 774–786.
- Green, P.J., 1995. Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82, 711–732.
- Grund, R., 2006. Some new morphological observations in Australian region satyrs (Lepidoptera: Nymphalidae: Satyrinae). *Victorian Entomol.* 36, 28–29.
- Hackett, S.J., Kimbal, R.T., Reddy, S., Bowie, R.C.K., Braun, E.L., Braun, M.J., Chojnowski, J.L., Cox, W.A., Han, K.-L., Harshman, J., Huddlestone, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., Yuri, T., 2008. A phylogenomic study of birds reveals their evolutionary history. *Science* 320, 1763–1768.
- Halanych, K.M., Robinson, T.J., 1999. Multiple substitutions affect the phylogenetic utility of cytochrome b and 12S rDNA data: examining a rapid radiation in leporid (Lagomorpha) evolution. *J. Mol. Evol.* 48, 369–379.
- Hallström, B.M., Janke, A., 2008. Resolution among major placental mammal interordinal relationships with genome data imply that speciation influenced their earliest radiations. *BMC Evol. Biol.* 8, 162.
- Harris, D.J., Poulsen, A.D., Frimodt-Møller, C., Preston, J., Cronk, Q.C.B., 2000. Rapid radiation in *Aframomum* (Zingiberaceae): evidence from nuclear ribosomal DNA internal transcribed spacer (ITS) sequences. *Edinb. J. Bot.* 57, 377–395.
- Harvey, D.J., 1991. Higher classification of the Nymphalidae, Appendix B. In: Nijhout, H.F. (Ed.), *Higher Classification of the Nymphalidae*, Appendix B. Smithsonian Institution Press, Washington, DC, pp. 255–273.
- Holland, B.R., Penny, D., Hendy, M.D., 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock—a simulation study. *Syst. Biol.* 52, 229–238.
- Jordal, B., Normark, B., Farrell, B., 2000. Evolutionary radiation of an inbreeding haplodiploid beetle lineage (Curculionidae, Scolytinae). *Biol. J. Linn. Soc.* 71, 483–499.
- Kim, J., 1993. Improving the accuracy of phylogenetic estimation by combining different methods. *Syst. Biol.* 42, 331–340.
- 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.
- Lewis, P.O., Holder, M.T., Holsinger, K.E., 2005. Polytomies and Bayesian phylogenetic inference. *Syst. Biol.* 54, 241–253.
- Lin, Y., McLennan, P., Gore, A., Phillips, M., Ota, R., Hendy, M., Penny, D., 2002. Four new mitochondrial genomes and the increased stability of evolutionary trees of mammals from improved taxon sampling. *Mol. Biol. Evol.* 19, 2060–2070.
- Lockhart, P., Cameron, S., 2001. Trees for bees. *Trends Ecol. Evol.* 16, 84–88.
- Mahoney, M.J., 2001. Molecular systematics of *Plethodontidae* and *Aneides* (Caudata: Plethodontidae: Plethodontini): phylogenetic analysis of an old and rapid radiation. *Mol. Phylogenet. Evol.* 18, 174–188.
- Malcomber, S.T., 2002. Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution* 56, 42–57.
- Mardulyn, P., Whitfield, J.B., 1999. Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of Microgastrinae (Hymenoptera; Braconidae): evidence of a high diversification rate in this group of parasitoids. *Mol. Phylogenet. Evol.* 12, 282–294.
- Miller, L.D., 1968. The higher classification, phylogeny and zoogeography of the Satyridae (Lepidoptera). *Memoirs Am. Entomol. Soc.* 24, 1–174.
- Morrison, C.L., Rios, R., Duffy, J.E., 2004. Phylogenetic evidence for an ancient rapid radiation of Caribbean sponge-dwelling snapping shrimps (Synalpheus). *Mol. Phylogenet. Evol.* 30, 563–581.
- Nixon, K.C., 1999. The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics* 15, 407–414.
- Nixon, K.C., Carpenter, J.M., 1993. On outgroups. *Cladistics* 9, 413–426.
- Omilian, A.R., Taylor, D.J., 2001. Rate acceleration and long-branch attraction in a conserved gene of cryptic daphniid (Crustacea) species. *Mol. Biol. Evol.* 18, 2201–2212.

- Parsons, M., 1998. The Butterflies of Papua New Guinea: Their Systematics and Biology. Academic Press, London.
- Peña, C., Nylin, S., Wahlberg, N., in press. The radiation of Satyrini butterflies (Nymphalidae: Satyrinae): a challenge for phylogenetic methods. *Zool. J. Linn. Soc.*
- Peña, C., Wahlberg, N., 2008. Prehistorical climate change increased diversification of a group of butterflies. *Biol. Lett.* 4, 274–278.
- Peña, C., Wahlberg, N., Weingartner, E., Kodandaramaiah, U., Nylin, S., Freitas, A.V.L., Brower, A.V.Z., 2006. Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data. *Mol. Phylogenet. Evol.* 40, 29–49.
- Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N., Delsuc, F., 2005. Heterotachy and long-branch attraction in phylogenetics. *BMC Evol. Biol.* 5, 50.
- Posada, D., 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256.
- Rokas, A., Carroll, S.B., 2006. Bushes in the tree of life. *PLoS Biol.* 4, e352.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Schwarz, M.P., Tierney, S.M., Cooper, S.J.B., Bull, N.J., 2004. Molecular phylogenetics of the allodapine bee genus *Braunsapis*: A-T bias and heterogeneous substitution parameters. *Mol. Phylogenet. Evol.* 32, 110–122.
- Shavit, L., Penny, D., Hendy, M.D., Holland, B.R., 2007. The problem of rooting rapid radiations. *Mol. Biol. Evol.* 24, 2400–2411.
- Shaw, A.J., Cox, C.J., Goffinet, B., Buck, W.R., Boles, S.B., 2003. Phylogenetic evidence of a rapid radiation of pleurocarpous mosses (Bryophyta). *Evolution* 57, 2226–2241.
- Siddall, M.E., 1998. Success of parsimony in the four taxon case: long-branch repulsion by likelihood in the Farris zone. *Cladistics* 14, 209–220.
- Silva Brandão, K.L., Wahlberg, N., Francini, R.B., Azeredo-Espin, A.M.L., Brown Jr., K.S., Paluch, M., Lees, D.C., Freitas, A.V.L., 2008. Phylogenetic relationships of butterflies of the tribe Acraeini (Lepidoptera, Nymphalidae, Heliconiinae) and the evolution of host plant use. *Mol. Phylogenet. Evol.*, 46.
- Simonsen, T.J., Wahlberg, N., Brower, A.V.Z., de Jong, R., 2006. Morphology, molecules and Fritillaries: approaching a stable phylogeny for Argynnini (Lepidoptera: Nymphalidae). *Insect Syst. Evol.* 37, 405–418.
- Slack, K., Janke, A., Penny, D., Arnason, U., 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. *Gene* 302, 43–52.
- Stamatakis, A., Hoover, P., Rougemont, J., 2008. A fast bootstrapping algorithm for the RAxML web-servers. *Syst. Biol.* 57, 758–771.
- Susko, E., 2008. On the distributions of bootstrap support and posterior distributions for a star tree. *Syst. Biol.* 57, 602–612.
- Suzuki, Y., Glazko, G.V., Nei, M., 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proc. Natl. Acad. Sci. USA* 99, 16138–16143.
- Tajima, F., 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics* 105, 437–460.
- Von Dohlen, C.D., Moran, N.A., 2000. Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. *Biol. J. Linn. Soc.* 71, 689–717.
- Wahlberg, N., Brower, A.V.Z., Nylin, S., 2005. Phylogenetic relationships and historical biogeography of tribes and genera in the subfamily Nymphalinae (Lepidoptera: Nymphalidae). *Biol. J. Linn. Soc.* 86, 227–251.
- Wahlberg, N., Freitas, A.V.L., 2007. Colonization of and radiation in South America by butterflies in the subtribe Phyciodina (Lepidoptera: Nymphalidae). *Mol. Phylogenet. Evol.* 44, 1257–1272.
- Wahlberg, N., Weingartner, E., Warren, A.D., Nylin, S., 2009. Timing major conflict between mitochondrial and nuclear genes in species relationships of *Polygonia* butterflies (Nymphalidae: Nymphalini). *BMC Evol. Biol.* 9, 92.
- Wahlberg, N., Wheat, C.W., 2008. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Syst. Biol.* 57, 231–242.
- Whitfield, J.B., Kjer, K.M., 2008. Ancient rapid radiations of insects: challenges for phylogenetic analysis. *Annu. Rev. Entomol.* 53, 449–472.
- Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. *Trends Ecol. Evol.* 22, 258–265.
- Wiens, J.J., Kuczynski, C.A., Smith, S.A., Mulcahy, D., Sites Jr., J.W., Townsend, T.M., Reeder, T.W., 2008. Branch length, support, and congruence. testing the phylogenomic approach with 20 nuclear loci in snakes. *Syst. Biol.* 57, 420–431.
- Zanis, M.J., Soltis, D.E., Soltis, P.S., Mathews, S., Donoghue, M.J., 2002. The root of the angiosperms revisited. *Proc. Natl. Acad. Sci. USA* 99, 6848–6853.