Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies

Marko Mutanen1,*, Niklas Wahlberg2 and Lauri Kaila3

1Zoological Museum, Department of Biology, University of Oulu, PO Box 3000, Oulu 90014, Finland
2Laboratory of Genetics, Department of Biology, University of Turku, Turku 20014, Finland
3Finnish Museum of Natural History, University of Helsinki, PO Box 17, Helsinki 00014, Finland

Lepidoptera (moths and butterflies) represent one of the most diverse animals groups. Yet, the phylogeny of advanced ditrysian Lepidoptera, accounting for about 99 per cent of lepidopteran species, has remained largely unresolved. We report a rigorous and comprehensive analysis of lepidopteran affinities. We performed phylogenetic analyses of 350 taxa representing nearly 99 per cent of lepidopteran families. We found Ditrysia to be a monophyletic taxon with the clade Tischerioidea + Palaephatoidea being the sister group of it. No support for the monophyly of the proposed major internested ditrysian clades, Apodityria, Oblectomera and Macrolepidoptera, was found as currently defined, but each of these is supported with some modification. The monophyly or near-monophyly of most previously identified lepidopteran superfamilies is reinforced, but several species-rich superfamilies were found to be para- or polyphyletic. Butterflies were found to be more closely related to ‘microlepidopteran’ groups of moths rather than the clade Macrolepidoptera, where they have traditionally been placed. There is support for the monophyly of Macrolepidoptera when butterflies and Calliduloidea are excluded. The data suggest that the generally short diverging nodes between major groupings in basal non-tineoid Ditrysia are owing to their rapid radiation, presumably in correlation with the radiations of flowering plants.

Keywords: Lepidoptera; molecular systematics; phylogeny; rapid radiation

1. INTRODUCTION

Lepidoptera (moths and butterflies) with over 160 000 described and 500 000 estimated species are among the most diverse animal groups (Kristensen et al. 2007). Together with Coleoptera, Hymenoptera and Diptera, they cover well beyond one half of all described organism species (Hunt et al. 2007; Sharkey 2007; Yeates et al. 2007). Their fascinating appearance renders them the most adored group of insects. They have been bred for sericultural purposes for at least 5000 years, and the detailed life histories of countless numbers of species have been investigated since then. Lepidoptera include many serious pest species and they have been popular organisms in various model systems (Roe et al. 2010). The phylogenetics of some major lepidopteran groups have been intensively studied on the basis of their morphology, and more recently also based on DNA data. While the interrelationships of the earliest lineages of Lepidoptera have been extensively studied for decades (Kristensen & Skalski 1998; Kristensen et al. 2007), the largest radiation of Lepidoptera, the Ditrysia has eluded analytical studies until recently (Regier et al. 2009). However, even this analysis was based on a limited set of taxa and taxon sampling heavily biased towards ‘advanced’ Lepidoptera.

Hence, the order itself lacks a rigorous evolutionary framework. In addition to a lack of phylogenetic analyses with comprehensive taxon sampling, we can see two additional reasons for the existing situation. First, groups of ditrysian Lepidoptera are morphologically homogeneous, and their phylogenetic affinities are therefore especially difficult to unravel (Kristensen & Skalski 1998), and second, there has been a shortage of phylogenetically informative genetic markers suitable for routine phylogenomic analyses. Owing to recent pioneering work in designing appropriate nuclear genomic markers (Wahlberg & Wheat 2008), there is now a good set of suitable genetic markers for use in lepidopteran phylogenomics.

The aim of this research was to clarify broad patterns of lepidopteran affinities using a wide array of molecular markers and a comprehensive taxon sampling. We performed phylogenetic analyses of 350 taxa that represent 43 out of 45 recognized lepidopteran superfamilies and nearly 90 per cent of lepidopteran families. These taxa altogether represent over 99 per cent of described moth and butterfly species. The analyses are based on eight gene regions, of which seven are from the nuclear genome and one from the mitochondrial genome. Our routine phylogenomic analyses. Owing to recent pioneering work in designing appropriate nuclear genomic markers (Wahlberg & Wheat 2008), there is now a good set of suitable genetic markers for use in lepidopteran phylogenomics.

The aim of this research was to clarify broad patterns of lepidopteran affinities using a wide array of molecular markers and a comprehensive taxon sampling. We performed phylogenetic analyses of 350 taxa that represent 43 out of 45 recognized lepidopteran superfamilies and nearly 90 per cent of lepidopteran families. These taxa altogether represent over 99 per cent of described moth and butterfly species. The analyses are based on eight gene regions, of which seven are from the nuclear genome and one from the mitochondrial genome. Our results show several unexpected patterns as well as corroborate many previous groupings.

2. MATERIAL AND METHODS

(a) Material acquisition and taxon sampling
The taxon sampling was planned largely based on the tentative lepidopteran phylogeny of Minet (1991) and Kristensen et al. (2007). Basically, all taxa at the subfamily level for which suitable material was available were included,
supplemented with many taxa for which phylogenetic affinities have remained ambiguous, a total of 350 taxa (see the electronic supplementary material, table S1). The study covers all the lepidopteran superfamilies except Simalithioidea and Whalleyaniaioidea, for which recent material is not known. Of the 124 lepidopteran families listed in Kristensen et al. (2007), 111 (89.5%) are included in this study (see the electronic supplementary material, table S1). They are divided into 252 named subfamilies, accounting for 75.2 per cent of the subfamilies listed in Kristensen et al. (2007). The classification and nomenclature follow Kristensen et al. (2007), except in Lypusidae and Cimelioidea, where subsequent classifications were followed (Yen & Minet 2007; Heikilä & Kaila 2010). The bulk of the DNA material was gathered from the DNA collections of the authors. Notable additions were received from the ATOlep DNA collection at the University of Maryland (http://www.leptree.net), the Australian National Collection (ANIC) at the Commonwealth Scientific and Industrial Research Organization, and several private collections.

(b) Molecular techniques

Usually, legs preserved in 100 per cent alcohol were used for DNA extraction, but sometimes also other body parts or larvae were used. In several cases, air-dried specimens up to 10 years old yielded DNA of good quality, indicating that also nuclear DNA of unrelaxed material kept in dry and cool conditions can successfully be amplified using standard extraction and sequencing protocols even after a relatively long time of preservation (Wahlberg & Wheat 2008). Notably, material of the ANIC and some private collections were found to be very useful, emphasizing the importance of proper preservation conditions in museum collections, and at the same time warning against water relaxation of specimens, as such an operation almost certainly fragments the nuclear DNA. Remaining parts of specimens were stored to serve as vouchers. Total genomic DNA was extracted and purified using Qiagen’s DNeasy extraction kit. DNA amplification and sequencing were carried out following protocol explained in detail elsewhere (Wahlberg & Wheat 2008). Sequencing was performed mainly with an ABI 3730 capillary sequencer (Oulu), and a smaller part with an ABI PRISM 3130×1 capillary sequencer (Turku). One mitochondrial (COI) and seven protein-coding nuclear gene regions (EF-1α, Wingless, RpS5, MDH, GAPDH, CAD and IDH) were sequenced, accounting for a total of 6303 bp with gaps. For details on the sequencing success and GenBank accession numbers of each species and gene, see the electronic supplementary material, table S2.

(c) Phylogenetic analyses

The sequence alignments were carried out manually using BOLD (Barcode of Life Data Systems, http://www.barcodinglife.org/views/login.php) (Ratnasingham & Hebert 2007), where reference specimens were available for many of the species used in this study (see the electronic supplementary material, table S1).

In a few cases where material of excellent quality was not available, sequence data were constructed by combining sequences from two individuals. This was done only if sequences successful in both individuals showed perfect or very close identity, suggesting no doubt about their conspecificity.

Since we were particularly interested in basal splitting events of ditrysian Lepidoptera, we evaluated the information value of third positions of codons in each gene separately. Changes in third positions represent mostly synonymous substitutions, and are hence unlikely to provide useful information at the deeper level of lepidopteran phylogeny, but may instead considerably increase the amount of uninformative ‘noise’ (homoplasy). After examining gene trees and a number of trials with various gene and taxon sets, we excluded third positions from the data in all genes except EF-1α, which evolves more slowly than any of the other genes (Wahlberg & Wheat 2008) and which, based on our evaluation, showed better resolution with third positions than without them. The trials with third positions included showed overall lower support values, and a few deviating groupings were regularly poorly supported and considered unlikely to be true. The effect of third positions was however not crucial as differences between the trials were not remarkable. With third positions excluded from all genes, except EF-1α, the data consisted of 4451 bp.

The phylogenetic analyses were carried out with both the model-based (maximum likelihood and Bayesian inference) and parsimony methods. We rooted our trees with Micropterygidae, arguably the sister group to other Lepidoptera (Kristensen & Skalski 1998; Wiegmann et al. 2002; Kristensen et al. 2007). The maximum likelihood analyses were carried out under the GTR + G model, chosen by MODELLTEST v. 3.7, and the data were partitioned into eight gene regions. The maximum likelihood analysis was implemented using the online version of RAxML (http://phylobench.vital-it.ch/raxml-bb/index.php) (Stamatakis et al. 2008). Supports for nodes were evaluated with 1000 bootstrap replicates of the data. We considered groups supported by over 50 per cent and strongly supported by over 90 per cent bootstrap support values.

The Bayesian analyses were carried out on a subset of the data using BEAST v. 1.4.8. (Drummond & Rambaut 2007). This analysis specifically aimed at clarifying basal branching events in Ditrysia and was performed with a subset of 118 species, because independent runs of the full set of taxa failed to converge in 20 million generations. The data were partitioned into the mitochondrial gene region (COI), the full EF-1α gene region and the combined nuclear genes with third codon positions removed (CAD, GAPDH, IDH, MDH, RpS5 and Wingless). The tree prior was set to the birth–death process, the independent models for the three partitions were all set to the GTR + G model, while all other priors were left to defaults. The branch lengths were allowed to vary under a relaxed clock model with an uncorrelated lognormal distribution. The analyses were run three times independently for 20 million generations, with every 1000th generation sampled. Using TRACER software (part of the BEAST package), we confirmed that the three runs
Proc. R. Soc. B

Phylogeny of moths and butterflies M. Mutanen et al.

had converged to a stationary distribution after the burn-in
stage, which left a total of 36,000 samples describing the pos-
terior distribution. Note that the age of the root was
arbitrarily set to 100. Since Bayesian posterior probability
values have a tendency toward ‘overcredibility’ (Suzuki
et al. 2002; Kolaczkowski & Thornton 2009), we interpreted
the Bayesian posterior probabilities conservatively and con-
sidered groups supported only if over 0.9, or preferably full
1.0, posterior probability was achieved. However, we did
not consider posterior probability estimates entirely uninfor-
mative because groups strongly supported in posterior
probabilities were usually found in maximum likelihood
analysis as well.

The parsimony analysis was carried out using TNT
(Goloboff et al. 2008). We used the ‘modern technology
search’, including sectorial search, ratchet, drift and tree
fusing; data compressed; gap = 5th state; ‘set initial level’—
more than 60, random seed 100. The search was interrupted
at 169 h 8 min (2 392 708 427 536 rearrangements tried by
then). The best score (40 007 steps) was hit six times, and
with 19 trees retained. The trees were further swapped
using the tree bisection and reconnection in a ‘traditional
search’. A total of 11 808 equally parsimonious trees were
found. Bremer supports were calculated with a script for
TNT that uses anticonstraints (Peña et al. 2006).

3. RESULTS AND DISCUSSION

All the methods agreed on the broad patterns of relation-
ships recovered (figures 1 and 2; see the electronic
supplementary material, figures S1 and S2). The best
obtained maximum likelihood tree is presented in
figure 1 and in more detail in the electronic supplement-
ary material, figure S1. We used this tree as a basis for
most of our conclusions. A Bayesian tree of a ditrysian
subset of taxa is presented in figure 2. Results of this
tree were used to draw conclusions only when very high
posterior probability was achieved. The result of the
unconstrained parsimony analysis is in general agreement
with those obtained using model-based analyses (see the
electronic supplementary material, figure S2). The main
discrepancies are lower general resolution with many
different groups, and a general pattern of clustering of the
most divergent taxa together. This implies the vulner-
ability of the parsimony analysis to long-branch
attraction (Felsenstein 1978). This appears particularly
strong in the pattern of long-branched basal non-tineoid
ditrysians (buculacridic complex, Lithocolletiinae +
Phylloncistinae and most divergent putative zygaenoids,
Heterogniidae and Epiprypodidae), among which even
non-ditrysian Tischeridae and Nepticulidae, with very
long branches, grouped (see the electronic supplementary
material, figure S2). Another probable reason for the
underperformance of parsimony analysis is its vulner-
ability to large amounts of homoplasy (Whitfield & Kjer
2008); a feature almost inherently present in nucleotide
data. In such cases, likelihood or Bayesian methods apply-
ing complex evolutionary models have reported to
generally outperform parsimony methods (Gadagkar &
Kumar 2005; Gaucher & Miyamoto 2005).

(a) Major phylogenetic patterns

Our results consistently support the sister group relation-
ship between Agathiphagidae and Heterobathmiidae. The
monophyly of Glossata, including all Lepidoptera except
Micropterigoidea, Agathiphagoidae and Heterobathmioi-
de, is supported. Of the other named clades below
Ditrysia, we found support for the monophy of Hetero-
neura, while our results do not support the monophy of
Coelolepida, Myoglossata and Neolepidoptera. Eulepi-
doptera become monophyletic after the inclusion of
Andesianidae, which always group with Incurvarioida,
not with Tischerioidea as suggested by Simonsen
(2009). We found support for the sister group relation-
ship between Tischerioidea and Palaeophatoidea, as
suggested in Davis (1986) and Nielsen (1989). They
together form the sister group to Ditrysia.

Morphological evidence of the monophy of Ditrysia
is convincing (Nielsen 1989; Kristensen & Skalski
1998), and this huge assemblage is supported by our mol-
ecular evidence, as well (but see note about parsimony
analysis above). Tineoidea may not be monophyletic, as
they are often divided into three separate and always
strongly supported lineages, Eriococcitae, Psychidae,
including Arrenophanidae, and Tineidae, including
Acrornphidae. The shift of Lypusidae from Tineoida to
Gelechioidea as in Heikkilä & Kaila (2010) is supported.
The monophy of non-tineoid Ditrysia is well supported,
although morphological evidence of this assemblage
has been scanty (Minet 1991; Kristensen & Skalski
1998).

While non-tineoid Ditrysia and many superfamilies
therein appear monophyletic, the relationships within
the clade remain largely questionable. Despite low sup-
port at many basal nodes, we do not consider the
observed patterns uninformative for two reasons. First,
the position of non-tineoid ditrysian superfamilies
remained constant in various analyses we carried out.
Second, there seems to be remarkable congruence
between our result and that of Regier et al. (2009),
even though the two studies are based largely on non-overlap-
ing data. Hence, we consider it justified to keep the
tentative branches uncollapsed. Regier et al. (2009)
concluded that the low ‘backbone’ supports are most likely
owing to the short internode branch lengths along the
‘backbone’, reflecting rapid radiation of major ditrysian
lineages in the past. This feature, common in many
other insect groups as well (Whitfield & Kjer 2008),
renders resolving the basal internode relationships
challenging.

The supposedly more advanced Ditrysia, i.e. super-
families other than Tineoida, Gracillarioidea, Yponomeutoidea and Gelechioidea, have been thought
to form three internested clades: Apoditrysia, Oblecto-
mecta and Macrolepidoptera (Minet 1986, 1991; Nielsen
1989). None of them appeared monophyletic. This
result is consistent with that of Regier et al. (2009) as
a broad pattern, but there are several differences in details.
Apoditrysia become monophyletic if Douglasidae and
Gelechioidea as in Kaila (2004) are included. The pattern
is largely the same as the one observed by Regier et al.
(2009), but with a significant conflict in the position of
Choreutoidea, Alucitoidea and Urodoidea, which Regier
et al. (2009) often found among the most basally branch-
ing ditrysian lineages after Tineoida. A reason behind
this discrepancy might be the rather unrepresentative
taxon sampling in lower Ditrysia in their study. Oblecto-
mecta are monophyletic if Copromorphoidea and
Immoidea are excluded. Immoidea were not included in
Regier et al. (2009) and Copromorphoidea was similarly found to often fall outside Obtectomera. Overall, we found Obtectomera to be a more coherent assemblage than did Regier et al. (2009). Macrolepidoptera sensu Minet (1991) get support only if butterflies (including Hedyloidea and Hesperioidea) and Calliduloidea are excluded; a result consistent with that of Regier et al. (2009).

Gracillarioidea never come out as a monophyletic entity, with Douglasiidae consistently coming out as an apoditrysian taxon. Lithocolletiinae and the remaining Gracillarioidea are not always associated with each

Figure 1. Overview of the 350-taxon RAxML maximum likelihood analysis. The tree was rooted on Micropteryx, a taxon likely to be a sister group to all other Lepidoptera. Non-ditrysian clades are all shown in black. Major ditrysian branches are coloured and their content indicated at the superfamily level. Putative ditrysian clades are shown by arrows in the middle of the circle. A solid line indicates complete inclusion and a dashed line partial inclusion in the named clade.
Figure 2. Cladogram of the Bayesian tree of the 118-taxon subset of ditrysian taxa. Posterior probabilities estimated under the GTR + G model (three BEAST runs of 20 million generations each) are shown above the branches. Ditrysian superfamilies are shown on the right.
other. The status of the bucculatricid complex as a gracillarioid taxon is ambiguous. Bucculatricidae are linked with Ogmograptis and Tritymba, both Australian taxa, the first tentatively associated with Bucculacidae, and the second formerly placed in Plutellidae (Yponomeutoidea). This complex may alternatively be linked with Zygaenoidea. The remaining Gracillarioidea are found to be either a sister group to, or embedded in, Yponomeutoidea. While almost consistently occurring, this connection is strongly supported only in a Bayesian analysis. Yponomeutoidea (with Tritymba excluded) are usually found to be monophyletic (although not in parsimony analysis or Bayesian analysis of a subset of taxa), but similarly consist of only loosely connected taxa. Lyonetioidea appear polyphyletic and Yponomeutoidea paraphyletic, with Plutellidae and Lyonetiinae embedded therein. Gypsipheritigidae are paraphyletic, with Acrolepiidae and the New Zealand, Tasmanian ‘megaplutellids’ included.

While being found within the core Macrolepidoptera, the other members of the ‘butterfly assemblage’ of Minet (1991), including Cimelioidea, Geometroidea and Drepanoidea, do not form a clade. Regier et al. (2009) came to the same conclusion. Within Macrolepidoptera, the status of Mimallonidae as a sister taxon to all the others is often found, which fits their peculiarity of having both microlepidopteran and macrolepidopteran morphological features (Minet 1991). Like Regier et al. (2009), we found Mimallonidae to be an unstable taxon, which in various trials associated with ‘microlepidoptera’, Bombycoidea or other groups, or even formed its own lineage, leaving its affinities unclear.

Of the most species-rich ditrysian superfamilies, the monophyly of Tortricoidea, Pyraloidea and Noctuoidea are supported, with the exception that Doidae associate with Drepanoidea instead of Noctuoidea. This finding is identical with that observed in Regier et al. (2009). Gelechioidea also appear monophyletic, although with low statistical support. A particularly problematic assemblage of ditrysian Lepidoptera concentrates around Cossoidea, Sesioidea and Zygaenoidea. None of these morphologically heterogeneous superfamilies appear monophyletic, but together they form a near-monophyletic assemblage, with the enigmatic and unstable zygaenoid families Epipyrionidae and Cyclotornidae and, curiously, Tinthiinae of Sesiidae falling outside. This pattern is similar to that first suggested by Scott (1986), and observed also by Regier et al. (2009), albeit with less comprehensive taxon sampling. Even though Tinthiinae cannot be firmly associated with any other taxa, they never group with Sesiaidea, and hence the wasp-like appearance of Sesiaidea appears to have evolved twice independently. The unassigned Australian Heliocosma group is found within this loose assemblage of taxa.

Several small apoditrysian superfamilies form loose coalitions in the basal region of ditrysian Lepidoptera. Even though each of these superfamilies is found to be monophyletic, their relationships to other groups remain without good support. An assemblage formed of Alucitoidea, Pterophoroidea, Epermenioidea, Schreckensteinioidea, Copromorphoidea and Urodoidae is often found, supplemented with Douglasiidae (Graclarioidae) and Millieriinae (Choreutoidea). Near that lies another loose concentration of small superfamilies, including Immidoidea, Galacticoidea and the core Choreutoidea. These three superfamilies alternatively form a monophylum with Tortricoidea. Milleriinae never associate with Choreutoidea, and Choreutoidea as currently delimited are thus polyphyletic.

True butterflies (Papilionoidea) appear paraphyletic with good support, both Hesperioidea and Hedyloidea nesting within them. A similar result was reported earlier based on molecular data, but this result was obscured by morphological evidence (Wahlberg et al. 2005), and later again by Regier et al. (2009). Also supporting the finding of Regier et al. (2009), Papilionidae seems to be the sister lineage of the other butterflies and skippers. Thysrioidea and Callidulioidea are found to be closely related to butterflies, with the association of Thysrioidea and butterflies being strongly supported in the Bayesian analysis.

The position of Hyblaeoidea, with only one species sampled, is unstable, and we cannot make firm statements about their affinities. They associate sometimes with Pyraloidea, but more often with Thysrioidea and butterflies. With Hyblaeoidea excluded, the clade including Pyraloidea and core Macrolepidoptera becomes well supported, placing most lepidopterans with tympanal organs together. In Macrolepidoptera, Bombycoidea form a monophyletic group, with Anthelidae included, as recently suggested (Regier et al. 2008; Zwick 2008). Lasiocampoidea are not found as a sister group to it, contradicting the findings of Regier et al. (2008). The core groups of Drepanoidea are monophyletic, with Doidae and Cimelioidea associating with them and Epicopeiidae falling outside, forming a sister group to Lasiocampoidea. This unexpected finding may also be supported by morphology (J. Holloway 2009, personal communication). The association of Lasiocampoidea with Epicopeiidae or the placement of Lasiocampoidea as distinct from Bombycoidea are, however, not well supported findings, and firm statements of their affinities cannot be drawn. Drepanoidea were not found to be a sister group to Geometroidea in any trials.

The monophyly of Geometroidea remains uncertain. Uroniidae, often with Sematuridae, usually form a monophylum with Geometridae, but this is not supported in the Bayesian analysis of the limited subset of taxa. The position of Sematuridae is similarly unstable. It may be the sister group to Uraniiidae or Geometridae, may be nested within Geometridae, or alternatively, may form its own lineage within Macrolepidoptera. With the exception of the affinities of Epicopeiidae, these observations are in general agreement with those reported by Regier et al. (2009). Noctuoidea without Doidae are a well-supported monophyletic entity. They are divided into six well-supported clades, of which the isolation of the Euteliinae + Stictopterinae clade from ‘quadriﬁne’ Noctuidae is a novel finding.

(b) Lower-level interrelationships

An overview of lower-level lepidopteran taxa that were not found monophyletic compared with the classiﬁcation of Kristensen et al. (2007) is presented in the electronic supplementary material, table S3. The putative monophyly of Tineoidea cannot be ruled out with certainty. The superfamily is divided into three well-supported clades, which follows the traditional division (Robinson 1988;
Davies & Robinson 1998) otherwise, but with Arrhenophanidae and Acrolithidae embedded in Psychidae and Tineidae, respectively. Within Psychidae, many interrelationships are well-resolved. Typhonidae get support as being a sister to the other taxa. The next splitting event is found between Placodolinae and the remaining Psychidae. Narycinae are found paraphyletic with respect to Taleporiinae. In Tineidae, interrelationships remain mostly weakly supported. Harmacoidea often form a sister group to the remaining Tineidae. The two representatives of Myrmecozelinae do not form a monophyletic group, as postulated by Robinson 2009.

In Yponomeutoidea, our results strongly support the monophyly of Ypsolophinae as consisting of Ypsolophinae and Ochsenheimerinae 1990). In Yponomeutidae, Prydinae and Attevinae are often found as sisters, but in the presented maximum likelihood tree this is obscured by the inclusion of the unstable Cemiostominae (Lyontiidae) within it. The two subfamilies of Lyontiidae never formed a monophylum, but both groups are somewhat unstable and branch leading to Cemiostominae is long. Bedelliidae are also unstable, though always found within Yponomeutoidea. The New Zealand and Tasmanian ‘megaplutellids’, here represented by Prodirix and Doxophyris, are found close to Orthotelinae of Glyphipterigidae, as suggested by Heppner 2003. Overall, although we tentatively consider Yponomeutoidea monophyletic, it is heterogeneous, and many interrelationships remain to be clarified.

Our findings provide support for the position of Chlidanotinae as a sister group to the other two subfamilies of Tortricidae. Olethreutinae and Tortricinae are both found to be monophyletic and sister groups to each other. Isonomemis, considered an unusual member of Coprophorinae (Dugdale et al. 1998), shows affinities with Alucitinae. This association is supported by morphology (L. Kaila 2009, personal observation). Tineidae are paraphyletic, with Alucitinae nested within it. Neither Cossinae nor Sesioidea form monophyletic assemblages, but are embedded within each other; a result generally consistent to that reported in Regier et al. 2009). Zygaenoidea are found subordinate to this assemblage. Sesioidea are found to be polyphyletic, with Tinhininae falling outside the Sesioideae–Cossioidea assemblage and Castniidae, Brachodidae and Sesidae each associating with various taxa of Cossioidea or the unassigned Helicocosa group of species. Regier et al. 2009 also found Castniidae always associating with Cossioidea rather than Sesioidea. Affinities and composition of the Australian Helicocosa group remain to be examined in more detail. Pestocera is a species group under Psychidae in collections and lists (Nielsen et al. 1996), although this group has never really been studied. We find it related to Helicocosa. Cossinae are found to be polyphyletic, with Dudgeonidae, Metarbelinae and Cossulinae associated with groups other than ‘Core Cossidae’ i.e. Cossinae + Zeuzerinae. Alternatively, Cossidae can be considered paraphyletic as Sesioidea and Zygaenoidea fall within it in this study.

Zygaenoidea are a heterogeneous group of moths defined by hardly any shared characters (Epstein et al. 1998). Zygaenoidea are consistently found subordinate to the Cossioidea–Sesioida complex, as reported in Regier et al. 2009). With the exception of the parasitic

Epipyropidae and Cyclotornidae, which have very long branches and cannot safely be placed anywhere, the superfamily is usually found to be monophyletic, and most interrelationships within the superfamily remain without strong support. The ‘core Zygaenoidae’ are divided into two lineages, roughly following the patterns found by Regier et al. 2009), but with some notable exceptions. Dalceridae are not found as a part of the ‘limacodid group’, but closer to Zygaenidae and a sister group to Phaudinae, which in turn are supported as falling outside of Zygaenidae, as recently suggested (Fänger et al. 1998; Nichius et al. 2006). Lacturidae, Heterogynidae, Dalceridae and Phaudinae together form a sister group to ‘core Zygaenidae’. The ‘limacodid group’ is consistently found, but with Dalceridae excluded and with inclusion of Himantopteridae and Anomoecotidae, which were not included in Regier et al’s 2009 study. These two closely related families group with Aidiidae, together forming a strongly supported clade. It is therefore probable that Himantopteridae, Anomoecotidae, Somabrachyidae and Aidiidae form a monophyletic clade within the Limacodidae, rendering this family paraphyletic. Even though there is no molecular evidence supporting the placement of Epipyropidae and Cyclotornidae within Zygaenoidea, the morphology of their immature stages is in favour of this position (Epstein et al. 1998).

While all maximum likelihood analyses supported the groupings of Hodges 1998) and Kaila 2004), but most affinities remain without support. Xyloryctinae get support as being a sister to Coleophorinae rather than Scythridinae. Autostichidae become monophyletic only with Glyphidoceridae and Deoclonidae included. Statilomopoda is not closely related to Oecophora, and it is therefore unlikely that Oecophoridae (Hodges 1998) are monophyletic, a result also observed by Kaila 2004). Elachistidae (Hodges 1998 or Kaila 2004) never form a monophyletic entity. The division of Pyraloidea into Pyralidae and Crambidae 1992) is supported with high confidence. In Pyralidae, the result is consistent with that presented in Regier et al. 2009). In Crambidae, our results are not consistent with those reported earlier 2002), with their main division between ‘spilomeline’ and ‘pyraustine’ groups not supported. Our results broadly agree with the results of Regier et al. 2009).

In Lasiocampoidae, Chionopsychinae were found to be the first diverging lineage and Lasiocampinae were a sister to Peocicocampinae, the latter finding being contradictory to that observed by Regier et al. 2008). Within Geometridae, our findings are in moderate agreement with those of Young 2006) and agree almost perfectly with those observed by other recent studies 2007; Regier et al. 2009; Wahlberg et al. in press). Larentiinae + Sterrhinae are found to form the sister group to all other Geometridae. Archieaeinae, formerly considered the first diverging lineage, are supported as subordinate to Larentiinae + Sterrhinae and sister to the lineage comprising Geometrinae, Oenochrominae and Ennominae.

While all maximum likelihood analyses supported Oenosandridae as being the sister group to all other Nocuoidea, this result was not unambiguously achieved in the Bayesian analysis. The monophyly of Notodontidae, Nolidae, Erebidae and Noctuidae is supported.
Euteliinae + Stictopterinae form a separate, strongly supported lineage, which is found as a sister group to Noctuinae. Formerly, the group has been associated with ‘quadrefine Noctuidae’ or Erebidae (Lafontaine & Fibiger 2006; Mitchell et al. 2006). As first observed by Weller et al. (1994) and later further supported by Mitchell et al. (2006), the arctiids and lymantriids, formerly considered valid families within Noctuoidea, are deeply embedded in Noctuaidae. We found them both within Erebidae, which is consistent with the findings of Mitchell et al. (2006).

4. CONCLUSIONS
This study, to our knowledge the most comprehensive analytical study on ditrysian lepidopteran phylogeny so far, suggests that the ditrysian Lepidoptera are monophyletic, with Tischerioidea + Palaephatoidea being their sister group. The superfamilies and most families are connected to each other with short nodes, but the findings are, in spite of the weak statistical support for many groupings, often in close agreement with another recent study (Regier et al. 2009). The non-tineoid Ditrysia is well supported. Even though none of the proposed major clades of more advanced Ditrysia, i.e. Apoditrysia, Obiectocera and Macrolepidoptera was supported as currently delineated, they get support from the present analysis after some adjustment. Many recognized superfamilies and families were found to be either para- or polyphyletic, though generally with weak support. The generally short nodes supporting the monophyly of most superfamilies imply a rapid radiation in the past, presumably in concert with the diversification of flowering plants. We anticipate that a more comprehensive taxon and gene sampling, supported by a rigorous analysis of comprehensive morphological data, will, in the future, provide a more robust backbone for the phylogeny of Lepidoptera.

A notable amount of material for this study was provided by the Lepture project, headed by C. Mitter (US NSF award no. DEB0515699), J. Junnilainen, U. Jürvete, J. Kaihit, A. Kallies, A. Karhu, O. Karsholt, T. Klemetti, C. Kokx, H. Kronholm, J. Kulberg, J.-F. Landry, J. Lehto, P. Malinen, V. S. Montey, T. Mutanen, E. Nieuwerken, K. Nupponen, O. Pellmyr, T. Roslin, P. Välimäki, B. Wåström and J. Zwier. We are grateful to M. Heikkilä for commenting and editing the text, K. Kosola for improving language, and L. Kvist, P. Kärkkäinen and R. Zahiri for their technical help. We also thank R. Hoare, M. Horak and C. Young for hosting our visits to Australia and New Zealand, C. Peña for providing efficient facilities for sequence handling, two anonymous referees for useful comments, as well as M. Horak, J. Holloway, C. Mitter and many other colleagues for fruitful discussions. This work was funded by the Academy of Finland through grants awarded to L.K. (project 1110906) and N.W. (projects 118369 and 129811).

REFERENCES
Heppner, J. B. 2003 Primitive sedge moths from New Zealand and Tasmania: transfer of Proditrix and relatives...


Young, C. J. 2006 Molecular relationships of the Australian Ennominae (Lepidoptera: Geometridae) and implications for the phylogeny of the Geometridae from molecular and morphological data. *Zootaxa* **1264**, 1–147.