Systematic Entomology

Systematic Entomology (2016), **41**, 112–132 DOI: 10.1111/syen.12143

Phylogeny and feeding trait evolution of the mega-diverse Gelechioidea (Lepidoptera: Obtectomera): new insight from 19 nuclear genes

JAE-CHEON SOHN[1,2,3], JEROME C. REGIER[1], CHARLES MITTER[1], DAVID ADAMSKI[2], JEAN-FRANÇOIS LANDRY[4], MARIA HEIKKILÄ[2,3], KYU-TEK PARK[5], TERRY HARRISON[6], KIM MITTER[1], ANDREAS ZWICK[7], AKITO Y. KAWAHARA[8], SOOWON CHO[9], MICHAEL P. CUMMINGS[10] and PATRIC SCHMITZ[11]

[1]Department of Entomology, University of Maryland, College Park, MD, U.S.A., [2]Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC, U.S.A., [3]Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington, DC, U.S.A., [4]Agriculture and Agri-Food Canada, C.E.F., Ottawa, Canada, [5]National Academy of Sciences, Seoul, Republic of Korea, [6]Department of Entomology, University of Illinois, Urbana, IL, U.S.A., [7]Australian National Insect Collection, CSIRO Ecosystem Science, Canberra, Australia, [8]Florida Museum of Natural History/McGuire Center for Lepidoptera and Biodiversity, University of Florida, Gainesville, FL, U.S.A., [9]Department of Plant Medicine, Chungbuk National University, Cheongju, Republic of Korea, [10]Laboratory of Molecular Evolution, Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD, U.S.A. and [11]PEPS, University of Hawaii, Honolulu, HI, U.S.A.

Abstract. The Gelechioidea (>18 000 species), one of the largest superfamilies of Lepidoptera, are a major element of terrestrial ecosystems and include important pests and biological model species. Despite much recent progress, our understanding of the classification, phylogeny and evolution of Gelechioidea remains limited. Building on recent molecular studies of this superfamily and a recently revised family/subfamily classification, we provide an independent estimate of among-family relationships, with little overlap in gene sample. We analysed up to five nuclear genes, totalling 6633 bp, for each of 77 gelechioids, plus up to 14 additional genes, for a total of 14 826 bp, in 45 of those taxa and all 19 outgroup taxa. Our maximum-likelihood (ML) analyses, like those of previous authors, strongly support monophyly for most multiply-sampled families and subfamilies, but very weakly support most relationships above the family level. Our tree looks superficially divergent from that of the most recent molecular study of gelechioids, but when the previous tree is re-rooted to accord maximally with ours, the two phylogenies agree entirely on the deepest-level divergences in Gelechioidea, and strongly though incompletely on among-family relationships within the major groups. This concordance between independent studies is evidence that the groupings (or at least the unrooted branching order) are probably accurate, despite the low bootstrap values. After re-rooting, both trees divide the families into three monophyletic groups: a 'Gelechiid Assemblage,' consisting of Gelechiidae and Cosmopterigidae; a 'Scythridid Assemblage,' consisting of Stathmopodidae, Scythrididae, Blastobasidae, Elachistidae, Momphidae, Coleophoridae and Batrachedridae; and a 'Depressariid Assemblage,' consisting of Autostichidae, Xyloryctidae, Lecithoceridae, Oecophoridae, Depressariidae and Lypusidae. Within the largest family, Gelechiidae, our results strongly support the pairing of Anomologinae with Gelechiinae, in accordance with a recent

Correspondence: Charles Mitter, Department of Entomology, University of Maryland, 4112 Plant Sciences Building, College Park, MD 20742, U.S.A. E-mail: cmitter@umd.edu

study of this family. Relationships among the other subfamilies, however, conflict moderately to strongly between studies, leaving the intrafamily phylogeny unsettled. Within the 'Scythridid Assemblage,' both trees support an 'SSB clade' consisting of Blastobasidae + (Scythrididae + Stathmopodidae), strongly resolved only in our results. Coleophoridae + Batrachedridae is supported, albeit weakly, in both trees, and only Momphidae differ in position between studies. Within the 'Depressariid Assemblage,' both trees support an 'AXLO' clade consisting of Autostichidae, Xyloryctidae, Lecithoceridae and Oecophoridae. The monophyly of this clade and relationships therein are supported weakly in previous results but strongly in ours. The recently re-defined family Depressariidae is paraphyletic in our tree, but the evidence against depressariid monophyly is very weak. There is moderate support for a core group of Depressariidae consisting, among the seven subfamilies we sampled, of Depressariinae, Aeolanthinae and Hypertrophinae. We show that gelechioids have a higher total number and percentage of species that are saprophagous as larvae than any other apoditrysian superfamily, that saprophagy is concentrated primarily in the 'AXLO clade,' and that the ancestral gelechioid condition was probably feeding on live plants. Among the living-plant feeders, concealed external feeding was probably the ancestral state. The multiple origins of internal feeding of various kinds, including leaf mining (otherwise almost unknown in Apoditrysia), are restricted mostly to the Scythridid and Gelechiid Assemblages. The traits that predispose or permit lineages to adopt these unusual life histories are worthy of study.

Introduction

The Gelechioidea are one of the most species-rich superfamilies of Lepidoptera and the most diverse of all among non-macroheterocerans, comprising 1478 genera and about 18 500 described species worldwide (van Nieukerken *et al.*, 2011). Given the difficulty of distinguishing species and the high proportion of undescribed species from nearly all faunistic regions (e.g. Hodges, 1998), gelechioid diversity may eventually prove to be much higher, possibly rivalling even the most species-rich macro-moth superfamilies (Powell *et al.*, 1998). Most gelechioids are very small (<10 mm in wingspan), although the largest reach nearly 70 mm in wingspan [e.g. Australian Xyloryctidae such as *Cryptophasa hyalinopa* (Lower) and *Thysiarcha ecclesiastis* (Meyrick) (I. McMillan, personal communication)]. Gelechioidea occur in nearly all eco-zones, including remote islands, polar regions and deserts, and can be a dominant herbivore group. On the one hand, gelechioids include many pest species. For example, the pink bollworm, *Pectinophora gossypiella* (Saunders) [Gelechiidae], has historically been one of the most destructive cotton pests in the world. The 15 gelechioids of economic importance in Europe listed by Carter (1984) collectively attack field crops and tree fruits as well as stored grains. On the other hand, some gelechioids are beneficial, serving as agents for weed biological control (Diatloff & Palmer, 1988; Shen & Xie, 1990; Boggs *et al.*, 1991; van Klinken *et al.*, 2003), and as model systems for the study of plant–insect interactions (van Dam & Bhairo-Marhé, 1992; Berenbaum & Passoa, 1999), sociality (Costa & Pierce, 1997) and mimicry (Hoare, 2005).

A reliable classification and phylogeny are indispensable for the organization, communication and prediction of observations about such an economically and scientifically important group of insects, and for understanding how the traits important to their pest management, such as their larval feeding habits, have evolved. Although much recent progress has been made (see next section), the state of systematics remains less advanced in Gelechioidea than in other large lepidopteran superfamilies. As detailed in the next section, highly divergent hypotheses have been proposed for both the delimitation of families and relationships among them. The goal of this paper is to contribute additional molecular evidence toward resolution of gelechioid phylogeny.

The taxa currently placed in Gelechioidea historically were scattered across Tineina, an early collective group name for microlepidopterans (Bruand, 1851; Stainton, 1854; Heinemann & Wocke, 1877; Meyrick, 1928), and Yponomeutidae (Stephens, 1829), until Fracker (1915) grouped them based on larval characters and first proposed superfamily status. Subsequently, the superfamily definition has been modified by multiple authors for their local faunas (e.g. Forbes, 1923; McDunnough, 1939; Common, 1970, 1990; Bradley, 1972; Kuznetsov & Stekol'nikov, 1978, 1984; also see reviews by Hodges, 1978; Kaila, 2004). Most of these studies proposed family-group names based on distinctive genera, resulting in a total of 71 such hypothesized groups to date (Appendix S1).

Gelechioidea belong to the clade Ditrysia, which makes up almost 98% of the extant Lepidoptera. They were formerly regarded as an early-diverging ditrysian lineage, together with more primitive superfamilies such as Tineoidea, Gracillarioidea and Yponomeutoidea (Kristensen & Skalski, 1998). Recent molecular phylogenetic studies (Mutanen *et al.*, 2010; Cho *et al.*, 2011; Bazinet *et al.*, 2013; Regier *et al.*, 2013; Kawahara & Breinholt, 2014; Timmermans *et al.*, 2014), however, have found convincing evidence that Gelechioidea are instead one of the

early-diverging groups within the advanced clade Obtectomera *sensu* van Nieukerken *et al.* (2011).

Morphological evidence for gelechioid monophyly is still limited. A long-recognized diagnostic feature of some gelechioid families is their characteristic ascending labial palpi. One apomorphy has long defined Gelechioidea: the haustellum covered with overlapping scales dorsobasally (Common, 1970; Hodges, 1986, 1998; Minet, 1990). This state is, however, parallelled in Pyraloidea, Choreutoidea and Millieriidae within Ditrysia. The mesothoracic leg of the gelechioid pupa has an invagination caused by the mesal meeting of the antennae, another possible synapomorphy (Minet, 1988; Passoa, 1995; Hodges, 1998; Kaila, 2004). The morphology-based phylogenetic analysis of Kaila (2004) found moderate support for the monophyly of Gelechioidea on the basis of three homoplastic abdominal structures as well as the abovementioned pupal character. In contrast, a molecular study by Kaila *et al.* (2011) and a combined molecular and morphological study by Heikkilä *et al.* (2014) showed weak support for the monophyly of Gelechioidea.

Relationships within gelechioids have been even more problematic. Minet (1990) critically revised the classification of the superfamily based on cladistic interpretation of morphological characters, although without presenting a formal analysis or cladogram. He recognized 17 families (Appendix S1). Fetz (1994) and Passoa (1995) conducted cladistic analyses using larval characters but obtained poorly-resolved phylogenies (Fig. 1A, B). Hodges (1998) presented a parsimony analysis with a larger character dataset, based on hypothesized groundplans, and provided a re-classification of Gelechioidea that postulated 15 families (Fig. 1E). This hypothesis was later challenged by Kaila (2004) using an even larger morphological dataset and an exemplar approach (Fig. 1F). Heppner (1998) also disagreed with Hodges (1998) and provided his own view on gelechioid phylogeny, but without explanation (Fig. 1D). Sinev (1992) and Lvovsky (2011) advanced very different views, and elevated Gelechioidea to an infraorder, Coleophoromorpha, with 3–6 superfamilies (Fig. 1C; Appendix S1). Their proposals, however, have not been widely accepted. All of these discrepancies in classification may stem in part from the very limited and inevitably biased taxon sampling characterizing nearly all systematic studies of Gelechioidea.

In the first application of molecular data to gelechioid phylogeny, Bucheli & Wenzel (2005) presented a parsimony analysis of a dataset combining morphology and two mitochondrial DNA markers. This study, however, yielded largely unresolved phylogenies (Fig. 2A), possibly due to very limited taxon and gene sampling. Mutanen *et al.* (2010) and Regier *et al.* (2013) included 30 and 54 gelechioids, respectively, in broad multi-gene studies across the Lepidoptera. The results of Mutanen *et al.* (2010), although mostly weakly supported, contributed to the reclassification by van Nieukerken *et al.* (2011), who recognized 21 families within Gelechioidea (Appendix S1). The first extensive molecular analysis of Gelechioidea (Kaila *et al.*, 2011) applied maximum-likelihood and Bayesian inference to a dataset comprising one mitochondrial and five nuclear genes sequenced for 109 ingroup taxa representing 32 of 37 known gelechioid subfamilies (Fig. 2B). Recently, Heikkilä *et al.* (2014) expanded the molecular dataset of Kaila *et al.* (2011) to 156 ingroup taxa and combined it with a morphological dataset containing 167 ingroup taxa. Based on their phylogeny, they proposed a revised classification of Gelechioidea that included 16 monophyletic families (their fig. 2). About a third of these, however, had very weak bootstrap support, and with one exception, relationships among the families were very weak. In this paper we seek to test and extend the conclusions of Heikkilä *et al.* (2014) by analysing an independent dataset of up to 19 genes sequenced in 70 gelechioids plus outgroups.

The ubiquity and mega-diversity of Gelechioidea has prompted multiple authors (e.g. Hodges, 1978; Powell *et al.*, 1998; Kaila *et al.*, 2011) to seek the reasons underlying their success. The evolution of diverse life-history traits in gelechioid larvae was proposed as a possible explanation (Kaila *et al.*, 2011). This adaptability may enable them to exploit resources that are unavailable to other animals, leading to explosive radiation as exemplified by *Hyposmocoma* in Hawaii (Rubinoff, 2008) that includes species adapted to an aquatic mode of life (Rubinoff & Schmitz, 2010) and *Eucalyptus*-associated Oecophoridae in Australia (Common, 1990). Gelechioids mainly feed on living plants, but they also include significant numbers of detritivores, fungivores and opportunistic feeders. Predation is rare among the lepidopterans, but several gelechioids prey on Sternorrhyncha such as aphids and scale insects (Pierce, 1995) or, in one remarkable example, on snails (Rubinoff & Haines, 2005; Schmitz & Rubinoff, 2011). Within phytophagous gelechioids there are also diverse feeding modes, such as mining leaves and boring in stems or buds. As a first step toward evaluating the possible role of ecological diversity in gelechioid diversification, Kaila *et al.* (2011) mapped larval feeding characters onto their phylogeny to look for evolutionary patterns. Using our phylogenetic results, we re-evaluate the hypotheses of larval feeding strategy evolution proposed by Kaila *et al.* (2011).

Materials and methods

Taxon and gene sampling

The central goal of this study was to re-evaluate the relationships among the families and subfamilies of Gelechioidea postulated by Heikkilä *et al.* (2014), and the subfamily relationships within Gelechiidae reported by Karsholt *et al.* (2013). Our ingroup taxon sampling included 70 species, representing: 33 of the 71 suprageneric groups proposed by previous authors (Figs 1, 2; Appendix S1); 27 of the 39 subfamilies or families recognized by Hodges (1998); and all of the 16 families recognized by Heikkilä *et al.* (2014) except Pterolonchidae. We used the data of all 54 species of gelechioids sequenced by Regier *et al.* (2013), and added 16 more: *Batrachedra pinicolella* (Zeller) [Batrachedridae]; *Blastobasis* sp. [Blastobasidae]; *Coleophora artemisicolella* Bruand [Coleophoridae]; *Euclemensia bassettella* (Clemens) [Cosmopterigidae: Antequerinae]; Chrysopeleiinae, an undetermined genus and species [Cosmopterigidae]; *Anatrachyntis japonica* Kuroko [Cosmopterigidae: Cosmopteriginae];

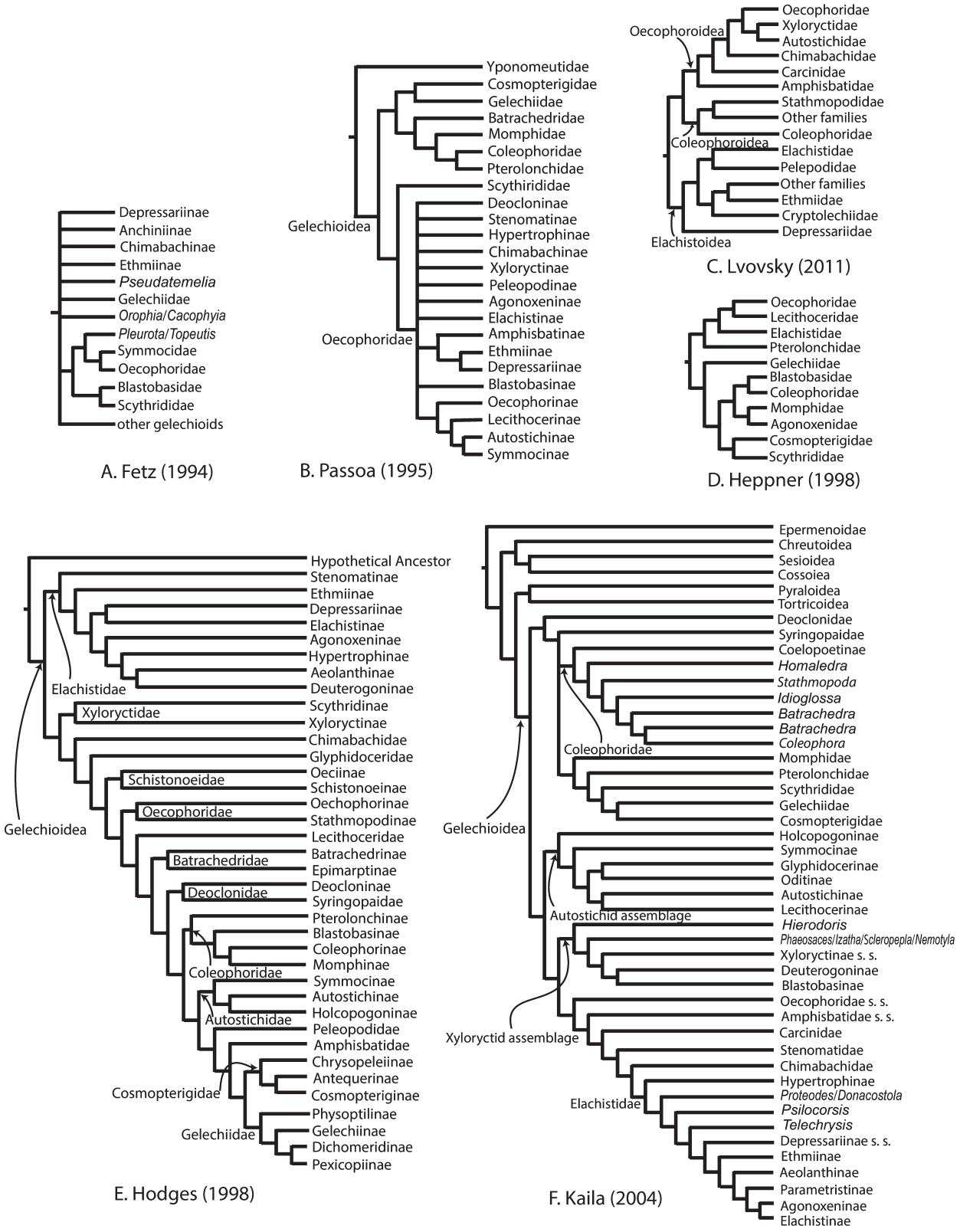

Fig. 1. Previous hypotheses of phylogenetic relationships in Gelechioidea, based on morphological data: (A) Fetz (1994), (B) Passoa (1995), (C) Lvovsky (2011), (D) Heppner (1998), (E) Hodges (1998), (F) Kaila (2004).

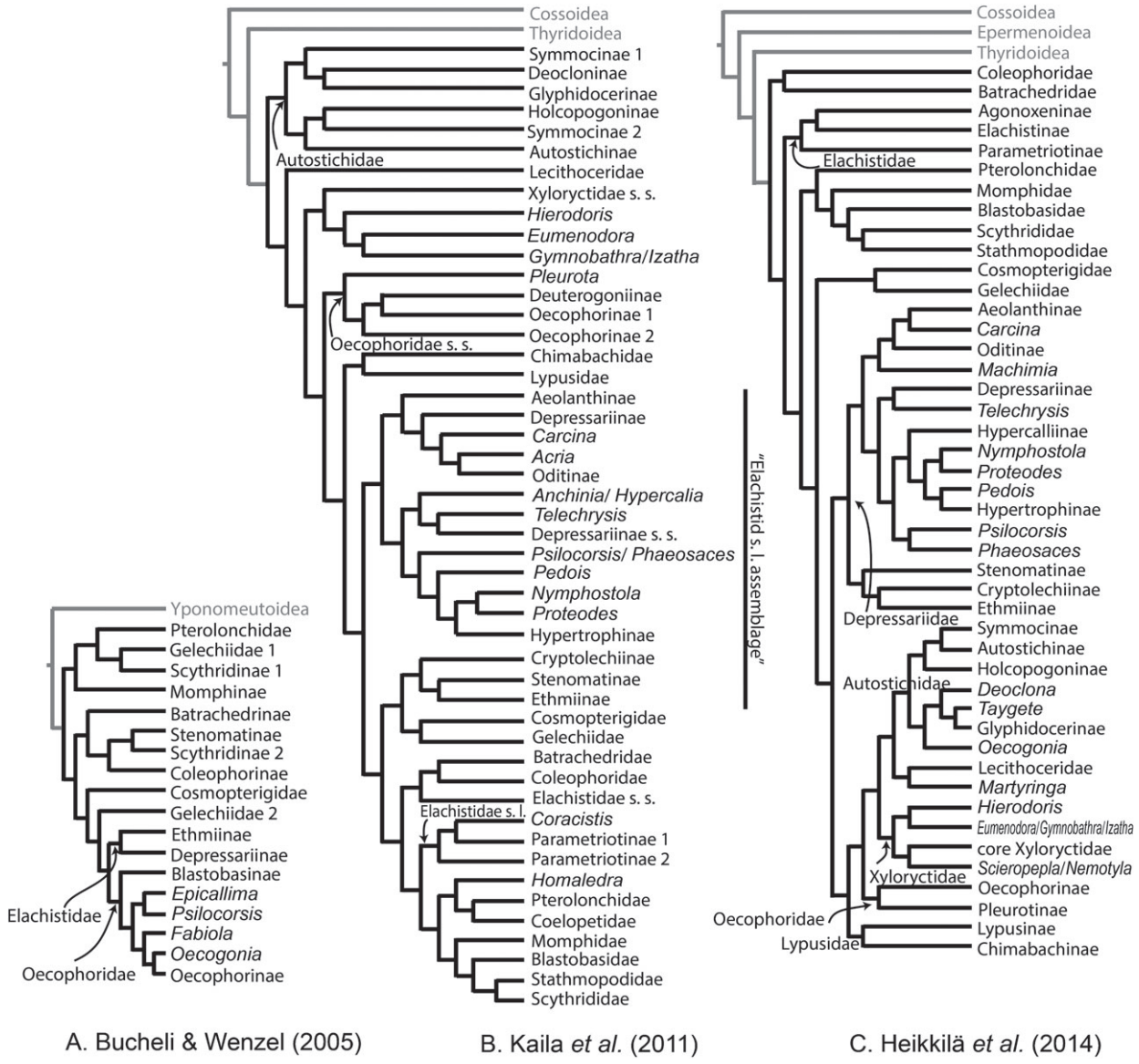

Fig. 2. Previous hypotheses of phylogenetic relationships in Gelechioidea, based on either molecular data alone (B) or combined molecular and morphological data (A, C): (A) Bucheli & Wenzel (2005), (B) Kaila *et al.* (2011), (C) Heikkilä *et al.* (2014). Grey branches and taxon names indicate outgroups.

Limnaecia sp. [Cosmopterigidae: Cosmopteriginae]; *Thiotricha biformis* (Omelko) [Gelechiidae: Thiotrichinae]; *Faristenia furtumella* Ponomarenko [Gelechiidae: Anacampsinae]; *Hypatima excellentella* Ponomarenko [Gelechiidae: Anacampsinae]; *Exoteleia pinifoliella* (Chambers) [Gelechiidae: Gelechiinae]; *Teleiodes pekunensis* Park [Gelechiidae: Gelechiinae]; *Torodora babeana* Park [Lecithoceridae]; *Tisis mesozosta* Meyrick [Lecithoceridae]; *Idioglossa miraculosa* (Frey) [incertae sedis]; and *Mimobrachyoma hilaropa* (Meyrick) [Oecophoridae]. As outgroups we included 19 species sequenced by Regier *et al.* (2013) representing 18 families and 14 superfamilies of Obtectomera (9) and non-obtectomeran Apoditrysia (5). The root of the entire tree, ingroups plus outgroups, was provisionally placed at Limacodidae + Zygaenidae + Cossidae, outside the Obtectomera.

Specimens for this study were obtained by our collecting and with the kind help of collectors around the world (see Acknowledgements). They were stored in 100% ethanol at −85°C, as a part of the ATOLep frozen tissue collection at the University of Maryland, College Park, U.S.A. The species sequenced and specimen accession numbers are listed in Appendix S2. DNA extraction used only the head and thorax for most specimens, leaving the abdomen and genitalia as a voucher, although the entire specimen was consumed for small species. DNA 'barcodes' were generated for all taxa, either by us using standard primer sequences with M13 tails

(Regier & Shi, 2005) or, more typically, by the All-Leps Barcode of Life project (http://www.lepbarcoding.org). *COI* DNA barcodes were checked against the Barcode of Life Data system reference library (Ratnasingham & Hebert, 2007) to confirm specimen identifications and also to facilitate future identification of specimens whose identity is still pending – namely species listed as 'sp.' or 'unidentified' in this report.

The gene sample for this study, consisting entirely of 19 protein-coding regions of nuclear genes, comprises two components. Firstly, the primary five gene loci, totalling 6633 bp, were sequenced for all taxa. These genes are: *CAD* (2928 bp; Moulton & Wiegmann, 2003), *DDC* (1281 bp; Fang *et al*., 1997), *enolase* (1134 bp; Farrell *et al*., 2001), *period* (888 bp; Regier *et al*., 1998) and *wingless* (402 bp; Brower & DeSalle, 1998). Secondly, an additional 14 gene regions – hence a total of 19 gene fragments (total 14 826 bp) – were sequenced for all of the outgroups and for 45 species of the ingroup taxa. These genes are *acetyl-coA carboxylase* (acc: 501 bp), *alanyl-tRNA synthetase* (*3070fin*: 705 bp), *AMP deaminase* (*268fin*: 768 bp), *gelsolin* (*109fin*: 552 bp), *glucose phosphate dehydrogenase* (*3007fin*: 621 bp), *glucose phosphate isomerase* (*8091fin*: 666 bp), *glutamyl- & prolyl-tRNA synthetase* (*192fin*: 402 bp), *histidyl-tRNA synthetase* (*265fin*: 447 bp), *nucleolar cysteine-rich protein* (*8028fin*: 324 bp), *phosphogluconate dehydrogenase* (*40fin*: 750 bp), *proteasome subunit* (*262fin*: 501 bp), *putative GTP-binding protein* (*42fin*: 840 bp), *tetrahydrofolate synthase* (*3017fin*: 594 bp) and *triosephosphate isomerase* (*197fin*: 444 bp). Details of these gene fragments and their PCR primer sequences can be found in Regier (2008). The 19 gene regions are a subset of 26 gene segments found to be phylogenetically informative across ditrysian Lepidoptera by Zwick *et al.* (2011) and Cho *et al.* (2011). One gelechioid species, *Caryocolum pullatella* (Tengström), was sequenced instead for only eight gene regions (*acc*, *CAD*, *DDC*, *enolase*, *109fin*, *265fin*, *268fin*, *3007fin*). GenBank numbers for these sequences are listed in Appendix S2.

Generation and analysis of DNA sequence data

A detailed protocol of all laboratory procedures is provided by Regier *et al.* (2008). Further descriptions, including gene amplification strategies, PCR primer sequences, and sequence assembly and alignment methods, can be found in Regier *et al.* (2008, 2009). Three distinct datasets that include all sequences were constructed. The first consists of unaltered nucleotides from all three nucleotide positions (nt123). The second (nt123_partition) contains the same nucleotides, but with these partitioned into two nonoverlapping character sets that separate nonsynonymous-only from mostly synonymous change. These two complementary character sets are called noLRall1nt2 and LRall1nt3 (see Table 1 in Regier & Zwick, 2011 for complete definitions; also see http://www.phylotools.com). Scripts to generate the two character sets are freely available (appendix 4 of Regier *et al.*, 2008; http://www.phylotools.com). The third dataset (nt123_degen1) is based on the degen1 approach of Regier *et al.* (2010; see also Zwick, 2010; Zwick *et al.*, 2012). The substitution model used in all analyses was GTR + gamma + I. This model was applied separately to each character subset in the partitioned analysis.

All phylogenetic analyses were based on the maximum-likelihood (ML) criterion as implemented in GARLI (Genetic Algorithm for Rapid Likelihood Inference; v2.0; Zwickl, 2011). We used the program default settings, including random stepwise addition starting trees, except that we halved the number of successive generations yielding no improvement in likelihood score that prompts termination (genthreshfortopoterm = 10 000), as suggested for bootstrapping in the GARLI manual. Each search for an optimal tree consisted of 970–1000 GARLI runs, whereas bootstrap analyses consisted of 708–750 pseudo-replicates, each based on 15 heuristic search replicates. Optimal-tree searches and bootstrap analyses were parallelized using Grid computing (Cummings & Huskamp, 2005) through The Lattice Project (Bazinet & Cummings, 2009). For consistency in the characterization of results, we will refer to bootstrap support of 70–79% as 'moderate,' 80–89% as 'strong' and ≥90% as 'very strong.'

The prepared tissue samples were processed for sequencing according to Regier *et al*. (2008). PCR amplicons were sequenced directly on a 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, U.S.A.). Sequences were edited and assembled using the TREV, PREGAP4 and GAP4 programs in the STADEN package (Staden, 1999) or the Geneious Pro v5.3.4 software package (Biomatters Ltd.). Multi-sequence alignments were made using the Translation Align option in Geneious Pro v5.3.4 (Protein alignment option: Geneious Alignment; Cost matrix: Blosum62; Gap open penalty: 12; Gap extension penalty: 3; Alignment type: Global alignment with free end gaps; + Build guide tree via alignment; Refinement iterations: 2). The final alignments were concatenated, separately for the 5 gene and 19 gene analyses, using Geneious Pro v5.3.4. Regions of uncertain alignment, totalling 1509 characters, were masked and excluded from subsequent analyses, using PAUP* 4.0b8 (Swofford, 2002).

Our deliberately incomplete gene sampling included 45 out of 70 ingroup taxa sequenced for the full set of 19 genes. The effectiveness of such a sampling scheme, in theory, might be undercut by phylogenetic artifacts resulting from the large blocks of missing data (Wiens, 1998, 2003; Lemmon *et al*., 2009). To ensure that our results are not subject to such artifacts, we carried out parallel analyses on the full, deliberately incomplete 19 gene dataset and on a reduced gene sample, the 'five-gene complete matrix,' comprising only the five gene regions sequenced in all ingroup and outgroup taxa.

Evolution of feeding habits

Kaila *et al*. (2011) reconstructed the evolution of larval feeding substrates and feeding modes in Gelechioidea. Our primary goal was to test if our results agree or disagree with the evolutionary patterns drawn by Kaila *et al*. (2011). To do so, we mapped the larval feeding substrates and feeding modes of our taxon sample on the molecular phylogeny, using mostly the

procedures and a simplified version of the categories described by Kaila *et al*. (2011). Three categories of larval feeding substrate were recognized: (i) live plant tissues (herbivory); (ii) dead plant or fungi (saprophagy or fungivory); (iii) predation, parasitism or feeding on animal remains. Feeding modes refer to how larvae attack their host plants. The categories include: (i) larvae boring in shoots, stems, or flowering parts of live or dead plants (internal feeding); (ii) larvae consuming the exterior and interior tissues of host plants while concealed in a shelter or exposed (external feeding); and (iii) larvae mining foliage tissues (leaf-mining). We regarded a lineage as leaf-mining if the larva mines in leaves for most of its development. Kaila *et al.* (2011) coded feeding substrates and feeding modes at the species level in their taxon sample. In contrast, we generalized those codes to the generic level. For species in the same genus as one studied by Kaila *et al.* (2011), we applied the state assigned by those authors. Our taxon sampling, however, included at least 31 genera which Kaila *et al*. (2011) did not include. For these, we searched relevant literature, compiled life-history data, and coded the predominant feeding substrate and feeding mode for each genus. Generalization of larval feeding habits at the generic level masks variation, incompleteness, and bias in such data, introducing errors. For this reason, we did not attempt any formal statistical approach, although we did compute parsimony optimizations on our molecular phylogeny.

Results

Figure 3 shows the best-score ML tree found from 1000 GARLI searches on the 19-gene, 89-taxon unpartitioned nt123 dataset. Bootstrap values for all five analyses, including character codings nt123 and degen1 for the 5-gene and 19-gene datasets, and the partitioned 19-gene nt123 dataset, are superimposed on each node of this tree. Monophyly for Gelechioidea was very strongly supported (node 1, BP = 95%, 19-gene nt123). Within Gelechioidea, 43 of the 69 nodes (62.3%) had strong bootstrap support (≥80%) from at least one analysis (Fig. 3). The most robust phylogenies came from the nt123 analysis of the 19-gene deliberately incomplete dataset: the fraction of nodes with bootstrap support of ≥70, ≥80 and ≥90% were 69.6, 60.9 and 55.1% respectively. The partitioned nt123 analysis with 19 genes gave a topology and node supports similar to those of the unpartitioned nt123 analysis. These two analyses showed less than 5% bootstrap difference at all nodes supported by bootstraps ≥70% in at least one analysis except for node 26 in Fig. 3. The 19-gene, nonsynonymous only (=degen1) dataset recovered 56.5, 52.2 and 46.4% of nodes with bootstrap values of ≥70, ≥80 and ≥90%, respectively. Most of the nodes strongly supported in the degen1 analysis were also strongly supported by the partitioned and unpartitioned nt123 analyses, but for three nodes (28, 53 and 55), only the degen1 analysis gave strong support. There was no apparent sign of signal conflict between synonymous and nonsynonymous changes. There were also no suggestions of phylogenetic artifacts arising from the large blocks of missing data in the 19-gene dataset, as the five-gene dataset with fewer missing data gave essentially identical topologies except for nodes which are very weakly supported in all analyses.

We sequenced two or more exemplars from 19 of the 33 previously postulated suprageneric groups sampled from Gelechioidea (Appendix S1). Fifteen of these 19 groups were monophyletic as sampled here with ≥75% bootstrap support. Monophyly of the *Autosticha* group was weakly supported (BP ≤ 53%). The *Gelechia*, *Lecithocera* and *Torodora* groups were nonmonophyletic. Monophyly was recovered by at least one of our analyses for 14 of the 16 families proposed by Heikkilä *et al*. (2014), usually with strong support. Pterolonchidae were not sampled, and Depressariidae were found to be paraphyletic. We also obtained mostly strong resolution of relationships within and among subfamilies in the largest family, Gelechiidae. Most nodes defining relationships between families, however, were very poorly supported.

Discussion

It has been challenging to establish a widely-accepted classification for the mega-diverse superfamily Gelechioidea. Two recent studies conducted rigorous phylogenetic analyses using the largest morphological and molecular sets to date. From the results, Heikkilä *et al*. (2014) and Karsholt *et al*. (2013) updated the classifications, respectively, of Gelechioidea and their largest included family, Gelechiidae (>4000 species). In this section, we first review the agreements and disagreements of our results with those two studies and other previous hypotheses. We then re-examine the evolutionary trends in gelechioid larval feeding habits postulated by Kaila *et al*. (2011), based on the new phylogeny.

Our study represents almost completely independent evidence from that used in Heikkilä *et al*. (2014) and Karsholt *et al*. (2013). The gene samples overlap minimally, the only two genes in common being *CAD* and *wingless*, and we mostly used different exemplar species and genera. We included fewer taxa than those studies, a point to which we return below, but doubled the number of genes sampled. We performed only ML analyses, and restrict our comparisons to the ML results of Heikkilä *et al*. (2014) and Karsholt *et al*. (2013), excluding their Bayesian analyses, because ML bootstraps and Bayesian posterior probabilities appear not to be directly comparable (Suzuki *et al*., 2002; Douady *et al*., 2003; Lewis *et al*., 2005). In the treatment below, the phrase 'tree of Heikkilä *et al*. (2014)' refers to those authors' combined molecular and morphological ML analysis unless otherwise specified. We compare bootstrap values for shared nodes between our study and those of Heikkilä *et al.* (2014) and Karsholt *et al.* (2013) in Table S1.

Phylogenetic position of Gelechioidea and basal divergences within the superfamily

Minet (1986, 1991) proposed that Gelechioidea was the sister group to either Apoditrysia or Yponomeutoidea, based on shared characters of the male genitalia and the labial

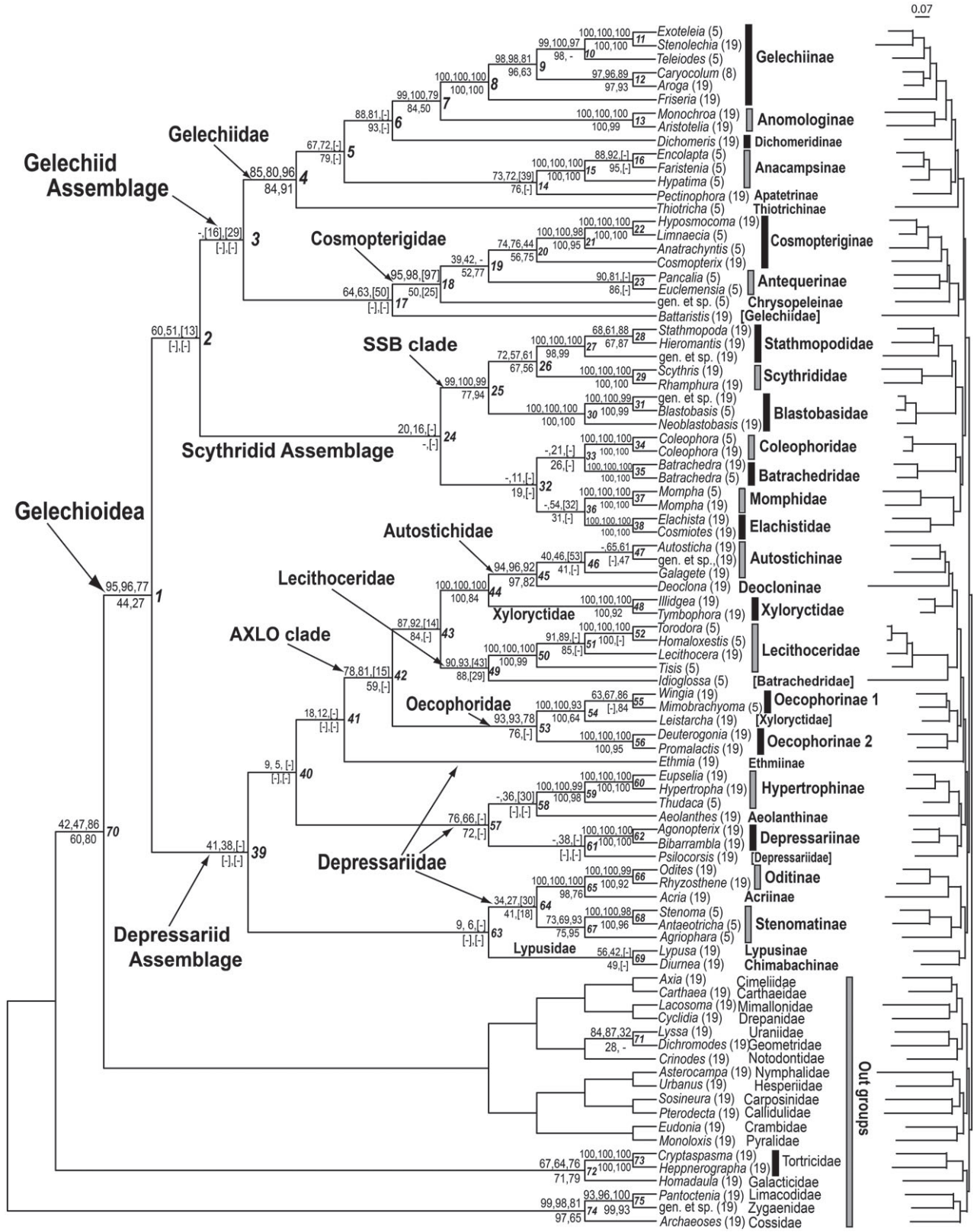

Fig. 3. Legend on next page.

palpi. Kaila (2004) found instead that Gelechioidea fall within Apoditrysia, and argued that they possess the synapomorphies thereof. Kaila's hypothesis was corroborated by several molecular studies (Regier *et al*., 2009; Mutanen *et al*., 2010; Regier *et al*., 2013). Using RNA-Seq data, Bazinet *et al*. (2013) found strong evidence that, within Apoditrysia, Gelechioidea were nested within the Obtectomera s.l., as sister group to Pterophoroidea + Thyridoidea. Heikkilä *et al*. (2014) found that choice of outgroups critically affected the backbone relationships of major gelechioid groups. Therefore, they regarded their rooting as tentative. However, they included Thyridoidea among their outgroups and recovered it next to Gelechioidea, consistent with Bazinet *et al*. (2013), Timmermans *et al*. (2014) and Kawahara & Breinholt (2014). The obtectomeran association of Gelechioidea was also supported by the present study (Fig. 3: node 70, 86% BP in the 19-gene, degen1 analysis).

In attempting to summarize the huge diversity of Gelechioidea, multiple authors have divided the superfamily into a small number of major groups. For example, Kuznetzov & Stekolnikov (1984) recognized four major groups within Gelechioidea (their infraorder Coleophoromorpha), including some taxa now placed in Copromorphoidea. Minet (1990) recognized an 'XS group' of nine families based on two larval apomorphies. Sinev (1992) proposed splitting Gelechioidea into six major groups. Passoa (1995) expanded Oecophoridae to encompass more than half of the other families, based on his examination of immature stages (Fig. 1B). Hodges (1998) postulated a basal divergence within Gelechioidea, one clade bearing eight subfamilies within Elachistidae and the other including the remaining 14 families (Fig. 1E). Kaila (2004) also postulated a basal divergence into two groups, which he informally referred to as the 'gelechiid' and 'oecophorid' lineages (Fig. 1F). Following Kuznetzov & Stekolnikov (1984) and Sinev (1992), Lvovsky (2011) raised Gelechioidea to infraorder status and recognized three superfamilies: Oecophoroidea, Coleophoroidea, and Elachistoidea (Fig. 1C).

The foregoing groupings, however, strongly contradict each other, and none can be said to be strongly supported. Homoplasy of morphological character states has hindered resolution of the basal divergences in Gelechioidea (Kaila *et al*., 2011). Molecular data, with their potentially huge character sets, probably offer the best hope of solving this problem. However, all molecular studies so far (Kaila *et al*., 2011; Heikkilä *et al*., 2014), including this one, have yielded mostly very weak support for deeper divergences within Gelechioidea, despite often finding strong support at the family level and below. This pattern may reflect rapid radiation of the families.

For this reason, we applied an additional criterion for reliably identifying deeper divergences, namely, concordance between studies (Regier *et al*., 2009; Mutanen *et al*., 2010). A grouping should gain credence if it is recovered by two or more independent studies, even if it is not strongly supported in any one study. Following this logic, we sought to determine what major groups, if any, are in common between our results and the maximum likelihood tree of Heikkilä *et al*. (2014). On first inspection, the two trees appear substantially different (Figs 2C, 3). However, this discord could be due largely to differences in the placement of the root, on which the evidence is especially weak in both studies. To remove this potential source of conflict, we deleted the outgroups from the family-level tree of Heikkilä *et al*. (2014) and re-rooted it in such a way as to minimize its conflict with our rooted tree. We placed more credence in our rooting because it is based on more genes and more outgroups, and has higher bootstrap values at the base.

Following this re-rooting, the main features of the two trees become remarkably similar (Fig. 4). Each can be divided into three major lineages. (i) The first, which we term the 'Gelechiid Assemblage,' consists of Gelechiidae plus Cosmopterigidae. It has 100% bootstrap in the analysis of Heikkilä *et al*. (2014), lower in the present study. (ii) In both trees, the sister group to the Gelechiid Assemblage, which we term the 'Scythridid Assemblage,' contains the families Stathmopodidae, Scythrididae, Blastobasidae, Momphidae, Elachistidae, Coleophoridae and Batrachedridae. Pterolonchidae, not sampled in the present study, also fall into this group in the Heikkilä *et al*. (2014) tree. The 'Scythridid Assemblage' has vanishingly small bootstrap support in both studies (4% in Heikkilä *et al*., 2014; ≤20% in the present study) but its identical composition in both strongly suggests that it is real. (iii) The first two assemblages together are sister group to what we term the 'Depressariid Assemblage,' which consists of the families Autostichidae, Xyloryctidae, Lecithoceridae, Oecophoridae, Depressariidae and Lypusidae. This assemblage too has low bootstrap support, but identical composition, in both studies.

This tripartite scheme could certainly be overturned by further evidence, particularly as regards the root of the tree, which differs dramatically among studies; it is the unrooted branching structure that is robust across studies. [For example, the relationships among the three gelechioids in the RNA-Seq study of Bazinet *et al.* (2013) suggest that the root lies within the Depressariid Assemblage.] Nonetheless, we believe that our scheme, including the rooting inferred in our analyses, is the best-founded working hypothesis of high-level gelechioid classification to date. Therefore, we use it to structure the more detailed discussion below. As we will see, there is much inter-study congruence within these high-level assemblages as well. Figure 4C shows the strict consensus of the two phylogenies at the family level (with Pterolonchidae omitted). Nine

Fig. 3. Maximum-likelihood (ML) estimate of phylogenetic relationships in Gelechioidea obtained from 500 GARLI searches under a GTR + gamma + I model. Topology shown is best tree from the unpartitioned 19-gene nt123 analysis. Bootstrap values (1000 pseudo-replicates) above branches for the 19-gene unpartitioned nt123 (left), partitioned nt123 (middle) and degen1 analyses (right); and below branches for the five-gene unpartitioned nt123 (left) and degen1 (right) analyses. Hyphen (-) denotes bootstrap value <50%. Square brackets, node not present in the best ML tree for that analysis. Node numbers, in bold italics, are used to organize text presentation of phylogeny. Numbers in parentheses next to genus names indicate numbers of genes attempted. Corresponding phylogram of best ML tree is shown in the right-hand column.

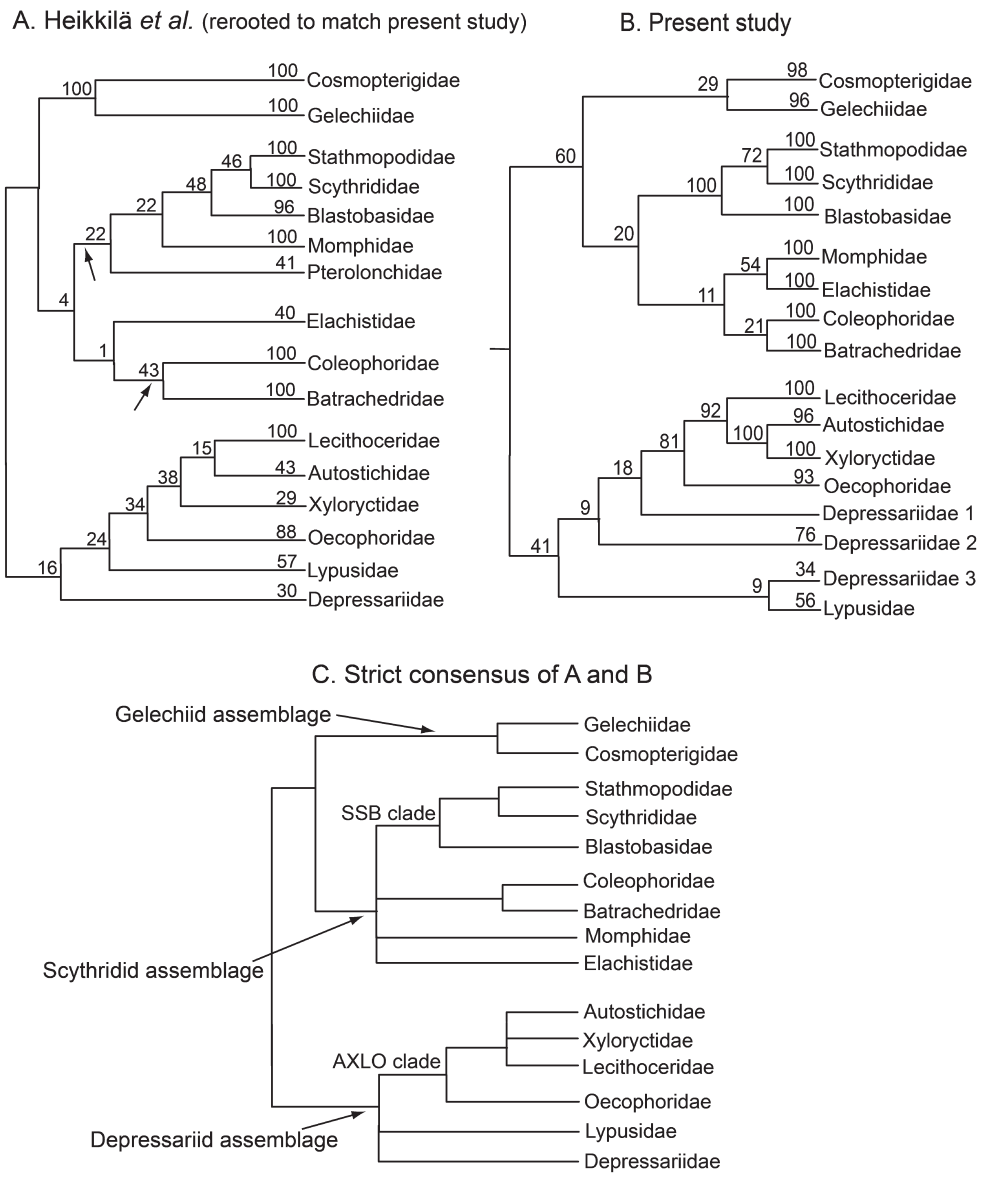

Fig. 4. Comparison and strict consensus of gelechioid phylogenies simplified to the family level. (A) Heikkilä *et al.* (2014), ML tree with bootstraps. Arrows denote rootings found by Heikkilä *et al.* (B) This study, ML tree with bootstraps. (C) Strict consensus of (A) and (B), omitting Pterolonchidae, not sampled in the present study.

of the possible 13 nodes (69%) are resolved, with the remaining irresolution split evenly between conflicting arrangements within the Scythridid and Depressariid Assemblages.

The 'Gelechiid Assemblage' (node 3)

Heikkilä *et al*. (2014) found 100% BP support for the grouping of Gelechiidae with Cosmopterigidae. The same relationship was recovered by multiple previous studies using both exclusively molecular (Kaila *et al*., 2011) and exclusively morphological data (Hodges, 1998; Kaila, 2004). This clade, which contains about 6400 species (numbers here and throughout from Heikkilä *et al*., 2014) is supported by three apomorphies which are to some extent homoplastic (Heikkilä *et al*., 2014): (i) the male sternum VIII as a lobe covering the bases of valvae; (ii) the presence of two lamellae on the anteromedian process of the metafurca; and (iii) a unique structure of female tergum + sternum VIII, where the laterally fused tergum + sternum VIII are medially incised. As further evidence, Hodges (1998) also cited four homoplastic characters of the male genitalia and wing venation. In the present study, the gelechiid/ cosmopterigid clade (Fig. 3: node 3) was also recovered, albeit only in the 19-gene nt123 and degen1 analyses, and with bootstrap support ≤29%. In our results, *Battaristis emissurella* (Walker), currently placed in Gelechiidae, was identified as sister group to Cosmopterigidae but with weak

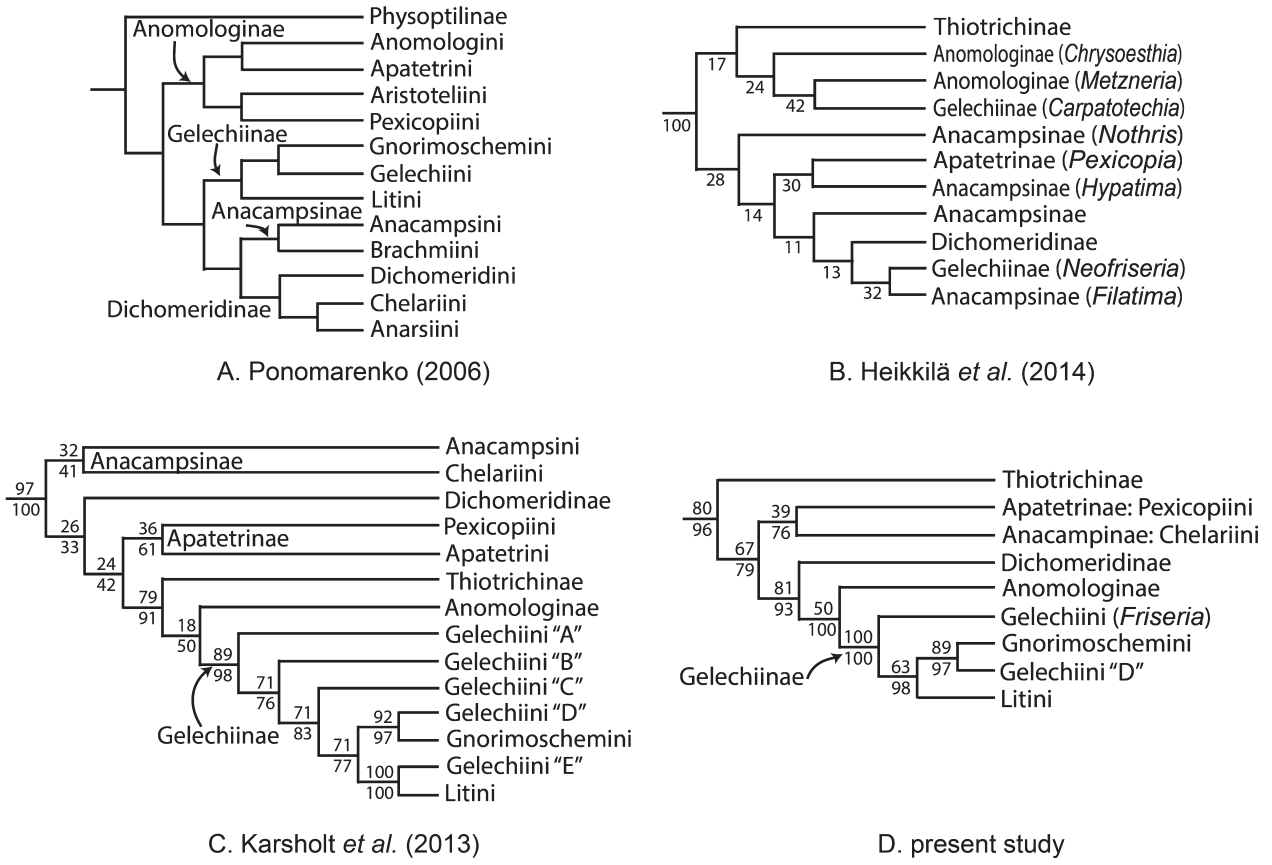

Fig. 5. Previous hypotheses of phylogenetic relationships in Gelechiidae. (A) Ponomarenko (2006), (B) Heikkilä *et al.* (2014) with bootstrap values below branches, (C) Karsholt *et al.* (2013) with smallest and largest bootstrap vales above and below branches respectively. (D) Our study with smallest and largest bootstrap values above and below branches respectively, Gelechiini 'D' corresponding to Karsholt *et al.* (2013).

support (Fig. 3: node 17, BP = 63–64). This relationship needs further attention, because no morphological synapomorphies are evident.

Gelechiidae (4700 spp.) are one of the few gelechioid families clearly defined by synapomorphies: (i) gnathos comprising a pair of lateral, articulated, symmetric sclerites with an articulated, mesial hook (Hodges, 1998); and (ii) the presence of a row of narrow, anteriorly directed scales on the forewing R vein in females (Kaila, 2004). Gelechiid monophyly has been strongly supported by previous molecular (Kaila *et al*., 2011) and combined molecular/ morphological studies (Karsholt *et al*., 2013; Heikkilä *et al*., 2014), as well as by our results (Fig. 3: node 4, ≥80% BP in all analyses).

Ponomarenko (2006) proposed a subfamilial and tribal classification of Gelechiidae, based on her phylogenetic studies of morphological characters, primarily genital musculature (Fig. 5A). Karsholt *et al*. (2013) critically reviewed this classification and proposed seven subfamilies (Fig. 5C), six of which were included in their analyses. The three subfamilies that were multiply sampled in our tree were recovered with very strong support (Fig. 5D), whereas this was not true in the tree of Heikkilä *et al*. (2014; Fig. 5B). Our tree agreed with that of Karsholt *et al*. (2013) in grouping Anomologinae + Gelechiinae, but there was conflict between studies, sometimes strongly supported, regarding the remaining relationships. Our tree places Dichomeridinae as sister group to Anomologinae + Gelechiinae, with BP = 88, whereas Karsholt *et al*. (2013) instead place Thiotrichinae as nearest relative to that pair with average bootstrap support of 85%. In the Karsholt *et al*. (2013) tree, relationships among the other subfamilies are very weakly supported (average BP ≤ 33%); Anacampsinae branches off first, followed by Dichomeridinae and then Apatetrinae. In contrast, our tree places Thiotrichinae at the base, with 19-gene bootstraps ≤79, and joins Apatetrinae and Anacampsinae as sister groups, with 19-gene bootstraps ≤73. Given this unusual level of conflict, we regard subfamily relationships within Gelechiidae as unsettled, apart from the pairing of Anomologinae + Gelechiinae (19-gene BP up to 100% in our analyses). In our analyses, there are three examples within Gelechiidae of higher bootstraps for five genes than for 19 genes, suggesting conflicts in phylogenetic signal among genes that may need to be taken into account in resolving subfamily relationships.

Cosmopterigidae (1730 spp.) comprise three subfamilies –Cosmopteriginae, Antequerinae, and Chrysopeleinae – according to Hodges (1998), who defended

monophyly for the family based on two parallelisms, one polymorphic apomorphy and one reversal. Sinev (1992) recognized a fourth subfamily, Scaeosophinae (not sampled here). The three sampled subfamilies form a monophyletic group in molecular studies (Kaila *et al*., 2011; our study: node 18, 95–98% BP from all 19-gene analyses). The interrelationships of these subfamilies, however, remain uncertain. Kaila (2004) observed one apomorphy shared by Cosmopteriginae and Antequerinae: projection of the tuba analis characteristically dorsad of the often asymmetrical uncus in the male genitalia. Most of our analyses recovered Cosmopteriginae and Antequerinae together, but support for this relationship was only weak to moderate (Fig. 3: node 19, 77% in the five gene degen1 analysis).

The 'Scythridid Assemblage' (node 24)

Within this assemblage, which totals about 3400 species, Kaila *et al*. (2011) and Heikkilä *et al*. (2014) recovered, albeit with weak support, a somewhat unconventional grouping consisting of the families Stathmopodidae, Scythrididae and Blastobasidae. Our results further suggest that this node, which we term the SSB Clade, is real, as it was very strongly supported by all of our 19-gene analyses (Fig. 3: node 25; BP $\geq$ 99). Possible morphological synapomorphies include a submental pit in the larva, a sclerotized ring around SD1 and SD2 (where SD2 is minute), and a small opening proximoposterior to SD1. The submental pit is also found in some Oecophoridae, Lecithoceridae, Batrachedridae and Pterolonchidae (Hodges, 1978; Kaila, 2004). It has been asserted (Hodges, 1978) that the submental pit is present in Xyloryctidae. However, in Xyloryctidae, it is actually a pair of sclerotized submental grooves, now interpreted to be an independent character (Heikkilä *et al.*, 2014). The submental pit and the pair of sclerotized grooves can occur separately or together as in some Lecithoceridae.

Within the SSB clade there is moderate support for the grouping of Stathmopodidae and Scythrididae to the exclusion of Blastobasidae, seen in all our analyses (Fig. 3, node 24; BP $\leq$ 72), and likewise by Kaila *et al*. (2011) and Heikkilä *et al*. (2014). Heikkilä *et al*. (2014) proposed a possible synapomorphy for Stathmopodidae + Scythrididae, the similarly expanded ductus seminalis. Earlier authors, however, reached different conclusions about relationships in the SSB clade. MacKay (1972) suggested that Scythrididae are a highly specialized group of Blastobasidae because the larvae possess a submental pit. This similarity, however, appeared to be homoplasious among gelechioids (Hodges, 1978) and thus its phylogenetic value was doubted by Kaila (2004). A close relationship between Blastobasidae and Stathmopodidae was proposed by Minet (1986), based on three putative synapomorphies, but this was later disputed by Common (1994). Heikkilä *et al*. (2014) noted that Blastobasidae and Stathmopodidae share a sclerotized ridge extending postero-medially from the lateral rod forming a window to the lateroposterior corners of the first tergum (their fig. 9). All of these proposals need further testing with increased and unbiased taxon sampling.

Stathmopodidae, Scythrididae and Blastobasidae are strongly monophyletic families in Heikkilä *et al*. (2014) and in our phylogeny, although the taxon samples are small. The monophyly of Stathmopodidae (100 spp.) is substantiated by an apomorphous abdominal tergum with spiniform setae on the posterior margins (Hodges, 1998), as well as by their characteristic resting posture and the presence of stiff setae on the hind legs (Heikkilä *et al*., 2014). Hodges (1998) recognized two additional autapomorphies: a ventrodistal sclerotized projection on the wall of the phallus and paired signa in the female genitalia with an inwardly-directed flange. The proposed apomorphies for Stathmopodidae should be re-examined, however, as the distribution of these traits within the family has not been thoroughly documented.

The monophyly of Scythrididae (650 spp.) is supported by the following apomorphies: the male genitalia with a sclerotized manica (Landry, 1991; Hodges, 1998); an ankylosed phallus [Landry, 1991; Scoble, 1992; but see Landry (1991: 183) and Kaila (2004) for discussion on the interpretation of this character]; the female ductus seminalis arising broadly from the posterior-most part of the corpus bursae (Landry, 1991; Hodges, 1998); the often very narrow ductus bursae (Landry, 1991); the larva with secondary setae, especially near the prolegs (Stehr, 1987; Scoble, 1992); the long and thin larval stipular setae, rarely found elsewhere in Gelechioidea (Kaila, 2004); and, the presence of a smaller spiracle on the larval abdomen VII than on other segments (Heikkilä *et al.*, 2014).

The monophyly of Blastobasidae (300 spp.) is supported by a unique combination of homoplasious and nonhomoplasious apomorphies (Adamski & Brown, 1989; Hodges, 1998). These include: the presence of a pterostigma between Sc and R_1 of the forewing; the base of CuA_2 of the forewing being perpendicular or sub-perpendicular to the cubitus; a subcubital retinaculum on the female forewing; a divided valva with a proximal flange near the base; an internal sclerite of the phallus; a setose anellus surrounding the phallus; spiniform setae on at least the posterior half of abdominal terga I–VIII in males and I–VI or I–VII in females; and larval stage with a submental pit and a small opening proximoposterior to SD1 (parallelled in Autostichidae: Glyphidocerinae and Xyloryctidae).

In our tree, the SSB clade is sister group to a lineage consisting of all the other members of the Scythridid Assemblage, specifically, (Coleophoridae + Batrachedridae) + (Momphidae + Elachistidae) (Fig. 3: node 24). Apart from the addition of Pterolonchidae, the trees of Kaila *et al*. (2011) and Heikkilä *et al*. (2014) differ from ours (Fig. 3) only in placing Momphidae as sister group to the SSB clade. These relationships are very weakly supported, but their similarity across studies suggests that they are real. Especially notable is the pairing of Coleophoridae + Batrachedridae, which has now recurred in multiple studies, both morphological and molecular (Kaila, 2004; Kaila *et al*., 2011; Heikkilä *et al*., 2014; present study). A possible synapomorphy is that both Coleophoridae and Batrachedridae have tergal spines distributed in two patches on each segment. Tergal spines in two patches are found in other groups as well (e.g. Momphidae), but when otherwise present in gelechioids, tergal spines tend to be arranged differently.

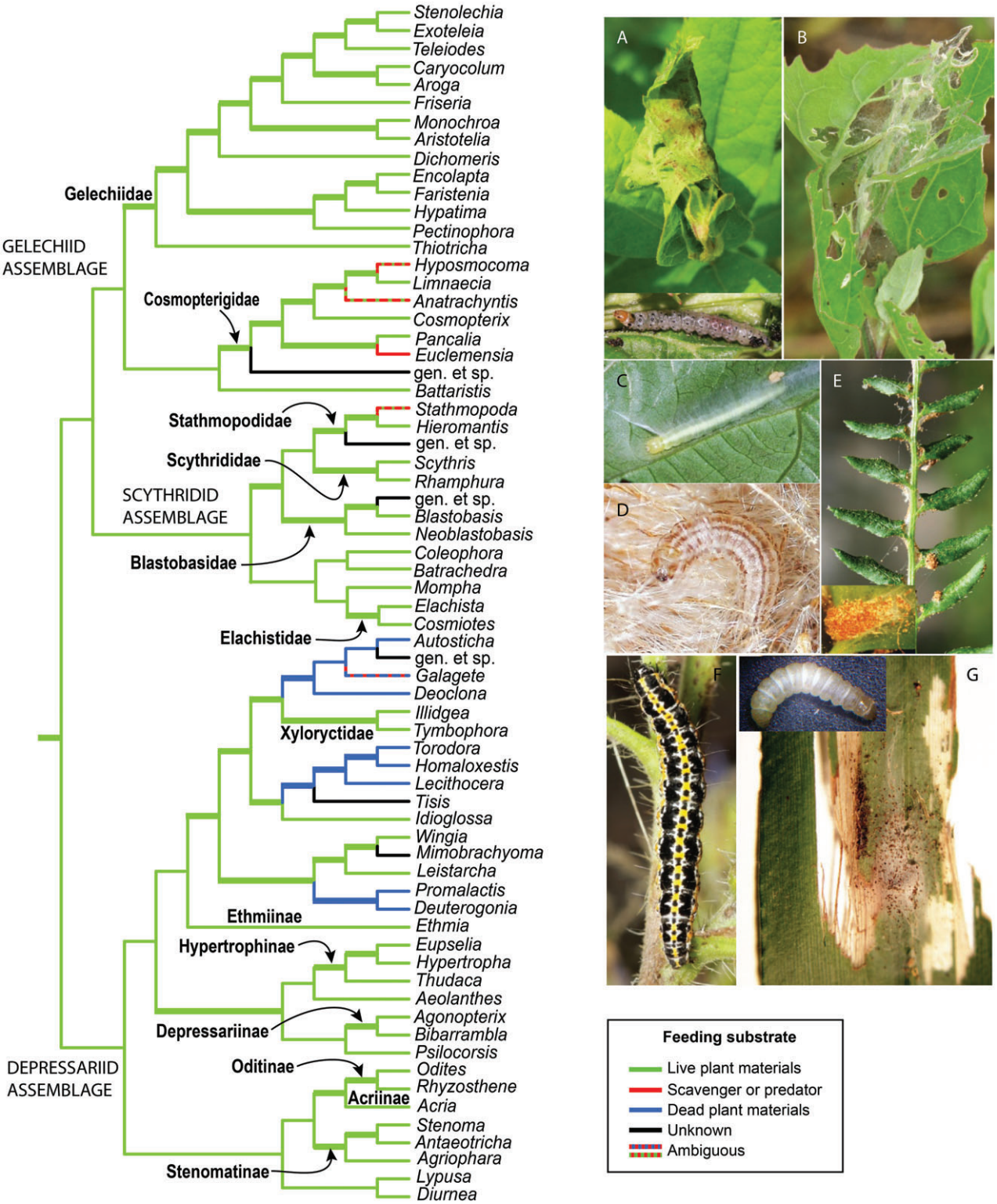

Fig. 6. Legend on next page.

The 'Depressariid Assemblage' (node 39)

Our tree is also very similar to that of Heikkilä *et al*. (2014) within the 'Depressariid Assemblage.' Most notably, the two trees share an 'AXLO' clade, consisting of Autostichidae, Xyloryctidae, Lecithoceridae and Oecophoridae, that excludes Depressariidae and Lypusidae. A version of this group also occurs, but in paraphyletic form, at the base of the tree of Kaila *et al*. (2011). The AXLO clade has weak support in Heikkilä *et al*. (2014), but moderate to strong support in our results (Fig. 3: node 42, 78 and 81% BP, 19-gene unpartitioned and partitioned nt123, respectively). Within the AXLO clade, both trees also group Autostichidae, Xyloryctidae and Lecithoceridae to the exclusion of Oecophoridae. Among the first three families, the Heikkilä *et al*. (2014) tree pairs Autostichidae with Lecithoceridae, then these two with Xyloryctidae, but bootstrap support is very low (BP = 15–38). In contrast, the present study groups Autostichidae and Xyloryctidae, with 100% bootstrap support, and these with Lecithoceridae with 92% support. The O(L(A + X)) arrangement thus seems clearly better founded than the alternative of O(X(L + A)).

Morphological synapomorphies for the 'AXLO' clade have not been found, but Kaila *et al*. (2011) noted that this set of families can be characterized as having the majority of species detritophagous as larvae. At least some morphological evidence may support the grouping of Autostichidae + Xyloryctidae (BP = 100). Hodges (1978) noted morphological similarities between the two families, although he later placed them in different clades (Hodges, 1998), and Lvovsky (2011) cited two synapomorphies: hindwing veins Rs and M_1 connate, a character also present in Lecithoceridae (and many other gelechioids), but absent in Oecophoridae (Kaila, 2004), and larval abdominal segments I–VIII with a pore near seta SD1. No morphological synapomorphies are known for the strongly supported clade Autostichidae + Xyloryctidae + Lecithoceridae (BP = 92).

Heikkilä *et al*. (2014) redefined Autostichidae (650 spp.) and included six subfamilies. The monophyly of Autostichidae was, however, weakly supported by their data. In our study, only Autostichinae and Deocloninae were included, but they were strongly grouped (Fig. 3: node 45, >82% in all analyses). These subfamilies span the basal divergence within Autostichidae as inferred by Heikkilä *et al*. (2014), thus our result lends credence to the monophyly of the entire family. No clear morphological synapomorphies are known for Autostichidae (Heikkilä *et al*., 2014).

Our sample of supposed Xyloryctidae consisted of the Australian genera *Illidgea*, *Tymbophora* and *Leistarcha*. *Leistarcha* was originally associated with Oecophoridae but later transferred to Xyloryctidae by Common (1996). The xyloryctid association of *Leistarcha* is disputed by genital features (McMillan, 2013), and our results strongly confirm its placement in Oecophoridae. Although *Xylorycta*, the type genus of Xyloryctidae, was not included in our study, Heikkilä *et al*. (2014) showed that this genus is closely related to *Tymbophora*. Therefore, we are confident that the strongly supported clade consisting of *Illidgea* and *Tymbophora* in our phylogeny corresponds to Xyloryctidae (Fig. 3: node 48, 92–100% BP in all analyses). The family is monophyletic but weakly supported in Heikkilä *et al*. (2014), who sampled ten diverse genera. No definitive morphological synapomorphies are known for Xyloryctidae s.l. (500 spp.), although a core group has a unique fusion of the uncus to the tegumen in the male genitalia (Heikkilä *et al*., 2014).

Gozmány (1978) supported monophyly for Lecithoceridae (1200 spp.), citing one autapomorphy and at least five homoplastic morphological characters, and recognized three subfamilies: Torodorinae, Lecithocerinae and Ceuthomadarinae. Lecithocerid monophyly was also supported by Hodges (1998) and Heikkilä *et al*. (2014). Our analyses included only Torodorinae and Lecithocerinae, and these were strongly grouped (Fig. 3: node 50, 99–100% BP in all analyses). Possible synapomorphies for Lecithoceridae (Heikkilä *et al*., 2014) include long antennae, a characteristic shape of the male gnathos, and groups of secondary setae in the larva. In our results the enigmatic genus *Idioglossa*, previously assigned to Oecophoridae (Hodges, 1983; Common, 1996), Batrachedridae (Sugisima & Arita, 2000), and Coleophoridae (Kaila, 2004), and deleted as a rogue taxon by Heikkilä *et al*. (2014), is strongly grouped with Lecithoceridae (Fig. 3: node 49, 88–93% BP in all nt123 analyses). *Idioglossa* shares larval characters with Lecithoceridae that support their close relationship: paired submental grooves (also present in some Xyloryctidae) (Heikkilä *et al*., 2014) and secondary setae on the D and SD pinaculae, a unique character in Gelechioidea (L. Kaila, personal communication). The association of *Idioglossa* with Lecithoceridae may necessitate re-evaluation of the synapomorphies for Lecithoceridae.

Hodges (1978, 1998) also included *Odites* within Lecithoceridae but did not assign it to a subfamily. Lvovsky (1996) designated a separate subfamily, Oditinae, for the *Odites*-group. The lecithocerid association of Oditinae was, however, challenged by Kaila (2004). Recent studies using molecular data alone (Kaila *et al*., 2011) or combined molecular and

Fig. 6. Maximum parsimony reconstruction of the evolution of feeding-substrate use among gelechioid genera included in our analyses. Feeding substrate categories include live plant tissues (green), dead plant material or fungi (blue), scavenger or predator (red) and uncertain (black). Branches with two alternating colours indicate an ambiguous state. Thickened branches are supported by ≥70% bootstrap in at least one analysis. (A–G) Representative phytophagous species of Gelechioidea. (A) Larval nest and larva (inset) of *Anacampsis solemnella* (Christoph) [Gelechiidae] on *Deutzia parviflora* Bunge; (B) larva of *Scythris sinensis* (Felder et Rogenhofer) [Scythrididae] and its nest on *Chenopodium album* var. *centrorubrum* Makino; (C) larva of *Acria ceramitis* Meyrick [Peleopodidae] on *Acalypha australis* L.; (D) larva of *Limnaecia phragmitella* Stainton [Cosmopterigidae] in head of *Typha* sp.; (E) larval damage (inset: a silken nest) of *Stathmopoda aenea* (Braun) [Stathmopodidae] on Christmas fern (*Polystichum acrostichoides* (Michx.) Schott); (F) larva of *Ethmia bipunctella* (Fabricius) [Ethmiidae] on *Echium vulgare* L.; (G) silken nest and leaf skeletonized by larva (inset) of *Idioglossa miraculosa* (Frey) [incertae sedis] on deer-tongue grass (*Dichanthelium clandestinum* (L.) Gould). Photo credits: Tristan Bantock (F), Ian Kimber (D); Terry Harrison (E, G); Jae-Cheon Sohn (A–C).

morphological data (Heikkilä *et al*., 2014) also favoured the exclusion of Oditinae from Lecithoceridae. Our results clearly support this hypothesis (Fig. 3: nodes 43 and 65; BP 87–100), placing Oditinae instead in Depressariidae, as do Heikkilä *et al*. (2014). This placement is in better accord with life history, as most Depressariidae, like Oditinae, are live plant feeders, whereas most Lecithoceridae are saprophages.

Classification within Lecithoceridae is incompletely understood. For example, both our study and that of Heikkilä *et al*. (2014) strongly show Torodorinae to be nested within Lecithocerinae (Fig. 3, node 52; BP = 100). The lecithocerine genus *Homaloxestis*, sister group to *Torodora* in our analyses, had been regarded as the closest relative to *Lecithocera*, due to shared features of abdominal morphology (Gozmány, 1978). This strong discrepancy between molecular and morphological data suggests the need for reevaluation of the autapomorphies for Torodorinae and Lecithocerinae.

Oecophoridae (3400 species) once included the largest number of subfamilies within Gelechioidea (Passoa, 1995). Heikkilä *et al*. (2014) found no support for Oecophoridae in this broad sense and included only two subfamilies, Oecophorinae and Pleurotinae in this family, whose monophyly was strongly supported (88% BP). We sampled only Oecophorinae, recovering monophyly with very strong support (Fig. 3: node 53, 93% BP, 19-gene nt123 analysis). Within Oecophoridae, our results support a close relationship between the *Deuterogonia*- and *Oecophora*-groups, the latter here represented by *Promalactis* (Fig. 3, node 56, BP 95–100%, all analyses), in accord with previous morphological (Saito, 2005) and molecular evidence (Kaila *et al*., 2011). Three Australian genera, *Wingia*, *Mimobrachyoma* and *Leistarcha*, constitute another strongly monophyletic clade in our results (Fig. 3: node 54, >90% BP, 19-gene analyses). Common (1994, 1997) stated that Australian oecophorids including *Wingia* and *Mimobrachyoma* share a male gnathos with arms united to form a short, dorsally concave lobe. Of the three genera, *Wingia* was included in Heikkilä *et al*. (2014) and grouped with putative oecophorid genera such as *Hofmannophila*, *Philobota* and *Borkhausenia*.

Depressariidae (2300 species) as redefined by Heikkilä *et al*. (2014) represents a major change in gelechioid classification. Its history is closely linked to that of Elachistidae, another family that has undergone multiple changes in definition since the 1990s. Hodges (1998) expanded Elachistidae to include the *Stenoma*-, *Depressaria*-, *Hypertropha*- and *Deuterogonia*-groups, previously associated with Oecophoridae, as well as the *Ethmia*- and *Agonoxena*-groups, each of which has often constituted its own family. This concept was partly followed by Kaila (2004), who however excluded the *Stenoma*- and *Deuterogonia*-groups from Elachistidae. The 'Elachistoidea' *sensu* Lvovsky (2011) roughly correspond to this broad delineation of Elachistidae. Lvovsky (2011) proposed five synapomorphies for this clade, with the emphasis on pupal characters whose phylogenetic value was doubted by Heikkilä *et al*. (2014). Kaila *et al*. (2011) and Heikkilä *et al*. (2014) disputed the monophyly of the 'Elachistidae s.l.', and transferred many groups therein to Depressariidae as redefined by Heikkilä *et al*. (2014). They restricted Elachistidae to Elachistinae, Agonoxeninae and Parametriotinae. Our results, like those of Heikkilä *et al.* (2014), favoured the strict concept of Elachistidae, which unlike Depressariidae belongs to our 'Scythridid Assemblage' (Fig. 3).

Heikkilä *et al*. (2014) included ten subfamilies and several unplaced genera in their re-defined Depressariidae, which is monophyletic in their tree but with weak support (30% BP). Depressariidae in their sense is paraphyletic in our tree: Ethmiinae are grouped with the AXLO clade, whereas a very weakly supported clade consisting of Oditinae, Acriinae and Stenomatinae is sister group to Lypusidae. These conflicting groupings are, however, very weakly supported (nodes 41, 63; BP = 9 and 18 < 20) and are possibly a consequence of the absence of Cryptolechiinae in our study, which formed a clade with Stenomatinae and Ethmiinae in Heikkilä *et al.* (2014) and Kaila *et al.* (2011). Thus, our results do not rule out monophyly for Depressariidae, and in that sense are compatible with the revised definition of the family. Heikkilä *et al*. (2014) cited characters in the male genitalia and the pupal abdomen as possible synapomorphies, but also recognized the presence of frequent reversals and parallelisms.

Relationships among the seven depressariid subfamilies that we sampled differ substantially between our study and that of Heikkilä *et al*. (2014) but are very weakly supported in both, with two exceptions. First, the subfamilies Hypertrophinae, Aeolanthinae and Depressariinae, together with the problematic genus *Psilocorsis*, are united with moderate support in our study (Fig. 3, node 57, BP = 76, 19-gene nt123 analysis). Hypercalliinae, not sampled in this study, grouped with Depressariinae in Heikkilä *et al*. (2014), and may also belong here. This group is a plausible candidate for 'core Depressariidae.' A similar grouping was found by Heikkilä *et al.* (2014), except that Aeolanthinae was grouped instead, albeit very weakly, with Oditiinae, Acriinae, *Carcina* and *Machimia* (the last two not sampled here). The 'core Depressariidae' are morphologically diverse and no synapomorphies are yet known. *Psilocorsis* had long been associated with Depressariinae (e.g. Nye & Fletcher, 1991). Hodges (1998) transferred it to Amphisbatidae. Minet (1990) and Kaila (2004) placed it within the broadly defined Elachistidae. Two recent studies using molecular data (Kaila *et al.*, 2011; Heikkilä *et al.*, 2014) found very strong support for *Psilocorsis* as the sister group to a previously oecophorine genus, *Phaeosaces*, not included in our analyses. Kaila *et al.* (2011) recovered *Psilocorsis* + *Phaeosaces* nested within a clade containing Depressariinae, Hypertrophinae and Hypercalliinae with strong support (BP = 82). The relationships of *Psilocorsis* + *Phaeosaces* to other subfamilies of the possible 'core Depressariidae' remain ambiguous.

The only other notably supported relationship among depressariid subfamilies in our results is the grouping of Oditinae + Acriinae (node 65, BP = 100) to the exclusion of Stenomatinae, in the putative clade that is sister group to Lypusidae. The same pairing occurs in the tree of Heikkilä *et al.* (2014), with 79% BP. This robust grouping seems real, and a search for morphological synapomorphies merits further effort. A third notable subfamily grouping is the pairing of Ethmiinae and Cryptolechiinae, BP = 79, (only) in the tree of Heikkilä *et al.* (2014). The latter subfamily is not sampled here; had it been, it might

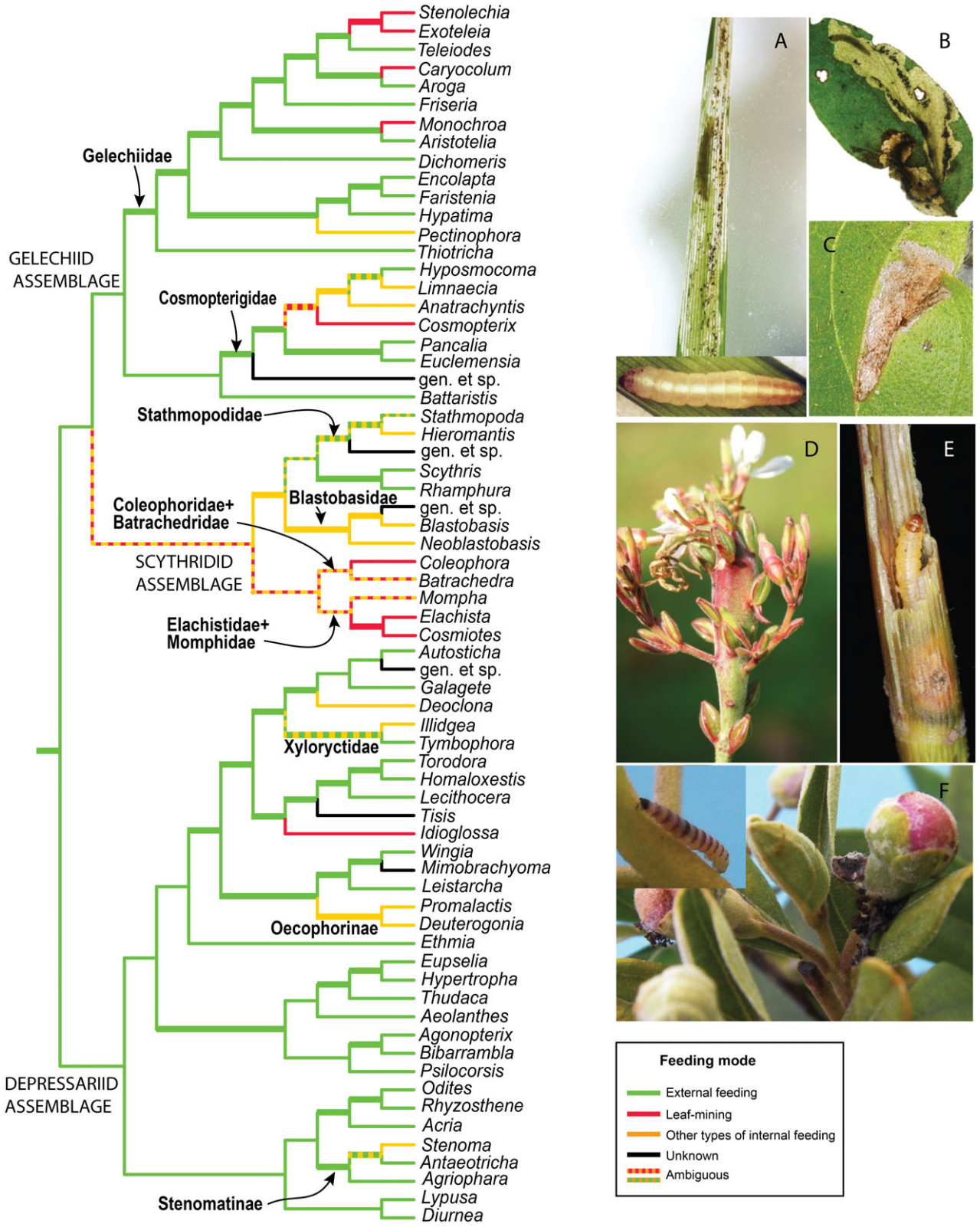

Fig. 7. Legend on next page.

have helped connect *Ethmia* with Stenomatinae as postulated in Heikkilä *et al.* (2014). Clearly there is much further work to be done on the relationships among depressariid subfamilies.

Finally, our results support, albeit weakly, the re-definition of Lypusidae by Heikkilä *et al.* (2014), as our representatives of their Lypusinae and Chimabachinae are grouped together (Fig. 3: node 69, 56% BP, 19-gene nt123). The same pairing was recovered by Kaila *et al.* (2011) and Heikkilä *et al.* (2014), the former with 83% bootstrap. Thus, the grouping seems to be real. Heikkilä *et al.* (2014) proposed a tongue-shaped, setose lobe on the male transtilla as a synapomorphy for Lypusinae and Chimabachinae, a character state otherwise found only in Cryptolechiinae.

Molecular phylogeny: summary and conclusions

1 In our analyses, as in Heikkilä *et al.* (2014), support for monophyly of most multiply-sampled families and subfamilies is strong, but most relationships above the family level are very weakly supported.
2 Nonetheless, when the tree of Heikkilä *et al*. (2014) is re-rooted to agree maximally with our tree, the two trees agree entirely on the deepest-level divergences in Gelechioidea. This concordance between independent studies is evidence that the groupings (or at least the unrooted branching structure) are real, despite the low bootstraps.
3 After re-rooting, both trees divide the gelechioid families into three monophyletic groups: a 'Gelechiid Assemblage' consisting of Gelechiidae and Cosmopterigidae; a 'Scythridid Assemblage' consisting of Stathmopodidae, Scythrididae, Blastobasidae, Elachistidae, Momphidae, Coleophoridae and Batrachedridae; and a 'Depressariid Assemblage' consisting of Autostichidae, Xyloryctidae, Lecithoceridae, Oecophoridae, Depressariidae and Lypusidae.
4 The Gelechiid Assemblage is weakly supported in our results, but has 100% bootstrap in Heikkilä *et al.* (2014).
5 Within the largest family, Gelechiidae, our results strongly support the pairing of Anomologinae with Gelechiinae, also seen in the gelechiid study by Karsholt *et al*. (2013), albeit with weak support. Relationships among the other subfamilies, however, conflict moderately to strongly between studies, leaving intrafamily phylogeny unsettled.
6 Within the Scythridid Assemblage, both trees support an 'SSB clade' consisting of Blastobasidae + (Scythrididae + Stathmopodidae). These relationships are supported weakly by Heikkilä *et al.* (2014) but strongly by our results. Coleophoridae + Blastobasidae is supported, albeit weakly, in both trees, and only Momphidae differs in position between them.
7 Within the Depressariid Assemblage, both trees support an 'AXLO' clade consisting of Autostichidae, Xyloryctidae, Lecithoceridae and Oecophoridae. Monophyly of this clade and relationships therein are supported weakly in Heikkilä *et al.* (2014) but strongly in our results. Depressariidae, monophyletic in Heikkilä *et al*. (2014), are paraphyletic with respect to both the AXLO clade and Lypusidae in our tree, but the evidence against depressariid monophyly is very weak. There is moderate support for a core group of Depressariidae consisting, among the seven subfamilies we sampled, of Depressariinae, Aeolanthinae and Hypertrophinae.

Evolutionary trends in feeding habits

Our phylogeny has mostly low bootstraps at deep levels, yet its concordance with the independent study of Heikkilä *et al.* (2014) is strong evidence for the approximate accuracy of its groupings. We therefore consider it a sufficiently reliable estimate for use in exploring evolutionary patterns in larval feeding habits. Figure 6 shows the distribution of larval feeding substrates across our tree. Our results overall are very similar to those of Kaila *et al.* (2011).

A salient fact to be explained about gelechioid larval habits is their unusually extensive departure from typical lepidopteran phytophagy. As Fig. 6 illustrates, the majority of the gelechioid genera we sampled feed on living plant tissues, but the alternative habit of feeding on dead organic matter (usually plants) or fungi (saprophagy/fungivory) is also widespread. We attempted to quantify the prevalence of saprophagy in Gelechioidea and compare it to that in other ditrysian superfamilies. To approximate the total number of gelechioid saprophages, we summed the diversities of the families (from Heikkilä *et al.*, 2014) in which saprophagy is identified, with varying degrees of confidence, to be the dominant habit. These are: Autostichidae (650 spp.), Lecithoceridae (1200 spp.) and Oecophoridae (3300 spp.). The sum of their diversities is 5150 species. This total omits the minority of saprophages scattered through several other families (e.g. Xyloryctidae, Stathmopodidae), but overestimates saprophage numbers in some of the foregoing families, because they also contain some true phytophages. We estimate that saprophages constitute about 28.6% of gelechioid species (5150 of 18 000). Appreciable numbers of saprophages/fungivores are otherwise found in only a few ditrysian superfamilies/families: Tineoidea: Tineidae

Fig. 7. Maximum-parsimony reconstruction of feeding mode evolution among gelechioid genera included in our analyses. Feeding mode categories include leaf mining (red), other forms of internal feeding (yellow), external feeding (green) and uncertain (black). Branches with two alternating colours indicate an ambiguous state. Thickened branches are supported by ≥70% bootstrap in at least one analysis. (A–F) Representative species of Gelechioidea whose larvae are leaf miners (A–C) or other forms of internal feeder (D–F). (A) Leaf mine and larva (inset) of *Elachista leucofrons* Braun [Elachistidae] on Poaceae; (B) leaf mine of *Gnorimoschema shepherdiae* (Priest) [Gelechiidae] on *Shepherdia canadensis* (L.) Nutt.; (C) leaf mine of *Cosmopterix pulchrimella* (Chambers) [Cosmopterigidae] on pellutory (*Parietaria* sp.); (D) gall induced by larva of *Mompha rufocristatella* (Chambers) [Momphidae] on flower stem of *Gaura biennis* L.; (E) larva of *Blastobasis repartella* (Dietz) [Blastobasidae] in stem of *Panicum virgatum* L.; (F) larva (inset) and larval damage of *Metharmostis multilineata* Adamski [Cosmopterigidae] on flower buds of *Rhodomyrtus tomentosa* (Aiton) Hassk. Photo credits: David Adamski (B, E, F); Terry Harrison (A, D); Mark Lawlor (C).

(3000 spp.; Regier *et al*., 2015); Tortricidae: Tortricinae: Epitymbiini (100 spp.); Pyraloidea: Pyralidae: Pyralinae (900 spp.) and Galleriinae (300 spp.); and Noctuoidea: Erebidae: Herminiinae (1000 spp.) plus several much smaller groups. Thus, no other ditrysian superfamily contains anywhere near as many saprophages as Gelechioidea, and except in Tineoidea, none has a higher percentage of saprophages.

A first step toward understanding the unusual prevalence of saprophagy in gelechioids is to ascertain the distribution of that trait across the phylogeny. Just as saprophagy does not appear to be randomly distributed across superfamilies, it appears nonrandomly distributed within Gelechioidea. Most of the saprophages fall in the AXLO clade of the Depressariid Assemblage. Autostichidae (650 spp.) are saprophagous with a single exception (Kaila *et al.*, 2011), as are Lecithoceridae (1200 spp.) with the exception of *Idioglossa.* Oecophoridae (3400 spp.) are mainly saprophagous but contain a significant minority of live plant feeders. Xyloryctidae (500 spp.) is the only family in this clade that consists mainly of live plant feeders, though it also contains saprophages. Although Blastobasidae (about 484 spp.) in the SSB clade of the Scythridid Assemblage are thought to be typically saprophagous (Powell, 1980; Hodges, 1998; Powell *et al.*, 1998), there are over twice as many published and unpublished records of phytophagy as saprophagy among species in this family worldwide (D. Adamski, unpublished data). Finally, there are scattered saprophagous species or groups thereof in multiple otherwise phytophagous groups (Kaila *et al.*, 2011). The phytophagy/saprophagy distinction thus seems to be largely conserved at the level of family and often of larger clades.

Our rooting, in contrast to that of Kaila *et al.* (2011), strongly implies that the ancestor of extant gelechioids, like those of all of their possible sister groups among Obtectomera, exhibited host-specific feeding on live plants, as argued by Hodges (1978). Saprophagy would then have arisen secondarily, multiple times. It is relatively easy to count scattered origins of saprophagy in otherwise phytophagous lineages. Even with greater taxon sampling and phylogeny resolution, however, it will be difficult to determine the number of transitions between phytophagy and saprophagy, and their directionality, in the AXLO clade, where the two habits are more extensively intermingled. The most we can say, perhaps, is that the probability of evolving and retaining saprophagy is much higher in the AXLO clade than elsewhere in the phylogeny. The compelling question then is why. The genomic basis of the origin of phytophagy from saprophagy/fungivory has begun to be studied in earnest (Goldman-Huertas *et al.*, 2015). Gelechioidea offer the opportunity to study the converse process. What are the traits (and their genomic bases), and/or the ecological circumstances, that permit or pre-dispose a lineage to adopt saprophagy? For example, is evolution of a broad host range a precursor to saprophagy? These issues have been little explored.

Among phytophagous gelechioids there is also great variation in mode of feeding (Kaila *et al.*, 2011). As shown in Fig. 7, the majority of the genera we sampled are external live-plant feeders, like most other Obtectomera and, quite probably, the ancestral gelechioid. External-feeding gelechioids almost always construct a shelter, of greatly variable kinds (Kaila *et al.*, 2011). Internal feeding of various kinds, including leaf mining, stem, seed and fruit boring, galling (Hanson *et al*., 2014) and others, appears to have arisen multiple times. These origins are not randomly distributed across the phylogeny. Rather, the great majority of internal feeders fall in the Gelechiid and Scythridid Assemblages. As is still clearer in Kaila *et al.* (2011) than in our smaller samples, internal feeding occurs to some extent in every family therein, whereas it is very rare or absent from the Depressariid Assemblage. Internal feeding appears to be dominant in several families, including Elachistidae and Cosmopterigidae. Internal versus external feeding does not appear to be tightly conserved, however, as external feeders can also be found in virtually all families of the Gelechiid and Scythridid Assemblages. To quantify the degree of conservation would require much additional sampling. What we can say at present is that the propensity to adopt internal feeding is substantially higher in these two assemblages than in the Depressariid Assemblage, even among the phytophages in the latter. Leaf mining is a particularly noteworthy form of internal feeding in Gelechioidea because it is otherwise almost unknown in Apoditrysia and especially Obtectomera. Leaf mining is mainly restricted to the Gelechiid and Scythridid Assemblages, where it is especially prevalent in Elachistidae, Coleophoridae, Cosmopterigidae, and some subclades of Gelechiidae, probably with many separate origins (Kaila *et al.*, 2011; Karsholt *et al*., 2013). As in the case of saprophagy, the cause of these phylogenetically clumped distributions of internal feeding is worth pursuing.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/syen.12143

Table S1. Comparison of bootstrap support for phylogenetic relationships found in two previous studies (Karsholt *et al*., 2013, for Gelechiidae; Heikkilä *et al*., 2014, for Gelechioidea) and the present study. ‘Singleton’, group was represented by a single exemplar, thus its monophyly cannot be tested.

Appendix S1. Position in previous classifications for 71 suprageneric groups of Gelechioidea proposed since Common (1970). Names have been updated where necessary so as to homogenize nomenclature across classifications. ‘N/A’, suprageneric group not considered or not clearly mentioned in that classification. Yellow blocks, suprageneric groups included in our analyses.

Appendix S2. A spreadsheet showing specimens sequenced, their classification, specimen accession number, number of genes attempted, total sequence length obtained and GenBank accession numbers.

Acknowledgements

We are indebted to many generous colleagues who provided specimens for this study, including Leif Aarvik, David Agassiz,

Joaquin Baixeras, John W. Brown, Timothy P. Friedlander, Jaga Giebultowicz, Peter Hattenschwiler, Ronald Hodges, Bernard Landry, David C. Lees, Marcus J. Matthews, Wolfram Mey, Andrew Mitchell, Ebbe S. Nielsen, Kenji Nishida, Olle Pellmyr and Jadranka Rota. We thank three contributors to the 'UK moths' website (www.ukmoths.org.uk), Tristan Bantock, Ian Kimber, and Mark Lawlor, for providing photos of gelechioid larval feeding, and Ian McMillan for information about Australian Xyloryctidae. The grid-based phylogenetic analyses were made possible by expert support from Adam Bazinet. Financial support was provided by the U.S. National Science Foundation's Assembling the Tree of Life program, award numbers 0531626 and 0531769; U.S. National Science Foundation awards DBI-0755048, DEB1355028 and DEB-1355023; and the Hatch funds of the Maryland Agricultural Experiment Station. Finally, we thank the two anonymous reviewers for their very helpful comments.

References

Adamski, D. & Brown, R.L. (1989) Morphology and systematics of North American Blastobasidae (Lepidoptera: Gelechioidea). *Mississippi Agricultural and Forestry Experiment Station Technical Bulletin*, **165**, i–iii and 1–70.

Bazinet, A.L. & Cummings, M.P. (2009) The Lattice Project: a grid research and production environment combining multiple grid computing models. *Distributed & Grid Computing. Science Made Transparent for Everyone. Principles, Applications and Supporting Communities* (ed. by M.H.W. Weber), pp. 2–13. Tectum Publishing House, Marburg.

Bazinet, A.L., Cummings, M.P., Mitter, K.T. & Mitter, C. (2013) Can RNA-Seq resolve the rapid radiation of advanced moths and butterflies (Hexapoda: Lepidoptera: Apoditrysia)? An exploratory study. *PLoS ONE*, **8**, e82615. DOI: 10.1371/journal.pone.0082615.

Berenbaum, M.R. & Passoa, S. (1999) Generic phylogeny of North American Depressariinae (Lepidoptera: Elachistidae) and hypotheses about coevolution. *Annals of the Entomological Society of America*, **92**, 971–986.

Boggs, K.W., Good, W.R., Harris, P. & Nowierski, R.M. (1991) *Metzneria paucipunctella* Zeller (Lepidoptera, Gelechiidae), a moth introduced against spotted knapweed: its feeding strategy and impact on two introduced *Urophora* spp. (Diptera; Tephrididae). *The Canadian Entomologist*, **123**, 1001–1007.

Bradley, J.D. (1972) A check list of the British insects. *Handbooks for the Identification of British Insects*, 2nd edn (revised), Vol. **11**, Part 2 (ed. by J.D. Bradley, D.S. Fletcher and P.E.S. Whalley), pp. 1–153. Royal Entomological Society of London, London.

Brower, A.V.Z. & DeSalle, R. (1998) Mitochondrial vs. nuclear DNA sequence evolution among nymphalid butterflies: the utility of Wingless as a source of characters for phylogenetic inference. *Insect Molecular Biology*, **7**, 1–10.

Bruand, C.T. (1851) Catalogue systématique et synonymique des Lépidoptères du département du Doubs, Tineides. *Mémoires de la Société d'émulation du Doubs*, **3**, 23–68.

Bucheli, S.R. & Wenzel, J. (2005) Gelechioidea (Insecta: Lepidoptera) systematics: a reexamination using combined morphology and mitochondrial DNA data. *Molecular Phylogenetics and Evolution*, **35**, 380–394.

Carter, D.J. (1984) *Pest Lepidoptera of Europe with Special Reference to the British Isles*. Dr W. Junk Publishers, Dordrecht.

Cho, S., Zwick, A., Regier, J.C. *et al.* (2011) Can deliberately incomplete gene sample augmentation improve a phylogeny estimate for the advanced moths and butterflies (Hexapoda: Lepidoptera)? *Systematic Biology*, **60**, 782–796.

Common, I.F.B. (1970) Lepidoptera (moths and butterflies). *The Insects of Australia*, Chapter 36, pp. 765–866. CSIRO, Melbourne.

Common, I.F.B. (1990) *Moths of Australia*. Melbourne University Press, Carlton.

Common, I.F.B. (1994) *Oecophorine Genera of Australia I. The* Wingia *Group*, *Monographs on Australian Lepidoptera*, Vol. **3**. CSIRO Division of Entomology, Canberra.

Common, I.F.B. (1996) Oecophoridae. *Checklist of the Lepidoptera of Australia*, *Monographs on Australian Lepidoptera*, Vol. **4** (ed. by E.S. Nielsen, E.D. Edwards and T.V. Rangsi), pp. 59–89. CSIRO Division of Entomology, Canberra.

Common, I.F.B. (1997) *Oecophorine Genera of Australia. II. The* Chezala*,* Philobota *and* Eulechria *Groups (Lepidoptera: Oecophoridae)*, *Monographs on Australian Lepidoptera*, Vol. **5**. CSIRO Division of Entomology, Canberra.

Costa, J.T. & Pierce, N.E. (1997) Social evolution in the Lepidoptera: ecological context and communication in larval societies. *The Evolution of Social Behavior in Insects and Arachnids* (ed. by J.C. Choe and B.J. Crespi), pp. 407–442. Cambridge University Press, Cambridge.

Cummings, M.P. & Huskamp, J.C. (2005) Grid computing. *EDUCAUSE Review*, **40**, 116–117.

van Dam, N.M. & Bhairo-Marhé, S.K. (1992) Induced chemical defence in *Cynoglossum officinale*. *Proceedings of the 8th International Symposium on Insect-Plant Relationships* (ed. by S.B.J. Menken, J.H. Visser and P. Harrewijn), pp. 79–80. Kluwer Academic, Dordrecht.

Diatloff, G. & Palmer, W.A. (1988) The host specificity of *Aristotelia ivae* Busck (Gelechiidae) and *Lorita baccharivora* Pogue (Tortricidae), two Microlepidoptera selected as biological control agents for *Baccharis halimifolia* (Asteraceae) in Australia. *Proceedings of the Entomological Society of Washington*, **90**, 458–461.

Douady, C.J., Delsuc, F., Boucher, Y., Doolittle, W.F. & Douzery, E.J.P. (2003) Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution*, **20**, 248–254.

Fang, Q., Cho, S., Regier, J.C. *et al.* (1997) A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative of relationships within Heliothinae (Lepidoptera: Noctuidae). *Systematic Biology*, **46**, 269–283.

Farrell, B.D., Sequeira, A.S., O'Meara, B. *et al.* (2001) The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution*, **55**, 2011–2027.

Fetz, R. (1994) Larvalmorphologische beiträge zum phylogenetischen system der ehemaligen Oecophoridae (Lepidoptera, Gelechioidea). *Neue Entomologische Nachrichten*, **33**, 1–273.

Forbes, W.T.M. (1923) *The Lepidoptera of New York and Neighboring States. Primitive Forms, Microlepidoptera, Pyraloids, Bombyces*, *Cornell University Agricultural Experiment Station Memoir*, Vol. **68**. Cornell University Press, Ithaca, New York.

Fracker, S.B. (1915) The classification of lepidopterous larvae. *Illinois Biological Monographs*, **2**, 1–161.

Goldman-Huertas, B., Mitchell, R.F., Lapoint, R.T., Faucher, C.P., Hildebrand, J.G. & Whiteman, N.K. (2015) Evolution of herbivory in Drosophilidae linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 3026–3031.

Gozmány, L.A. (1978) *Microlepidoptera Palearctica*, Band 5 Lecithoceridae. Goecke & Evers, Keltern.

Hanson, P., Nishida, K. & Gómez-Laurito, J. (2014) Insect galls of Costa Rica and their parasitoids. *Neotropical Insect Galls* (ed.

by G.W. Fernandes and J.C. Santos), pp. 497–517. Springer Science + Business Media, Dordrecht.

Heikkilä, M., Mutanen, M., Kekkonen, M. & Kaila, L. (2014) Morphology reinforces proposed molecular phylogenetic affinities: a revised classification for Gelechioidea (Lepidoptera). *Cladistics*, **30**, 563–589.

Heinemann, H.v. & Wocke, M.F. (1877) *Die schmetterlinge Deutschlands und der Schweiz, systematisch Bearbeitet*, *Zweite Abtheilung Kleinschmetterlinge, Band II. Die Motten und Federmotten, Heft II*. Vieweg, Braunschweig.

Heppner, J.B. (1998) Classification of Lepidoptera, Part 1. Introduction. *Holarctic Lepidoptera*, **5** (Suppl. 1), 1–148.

Hoare, R.J.B. (2005) Hierodoris *(Insecta: Lepidoptera: Gelechioidea: Oecophoridae), and Overview of Oecophoridae*, *Fauna of New Zealand 54*. Mannaaki Whenua Press, Canterbury.

Hodges, R.W. (1978) *The Moths of America North of Mexico, Fascicle 6.1 Gelechioidea (in Part): Cosmopterigidae*. E. W. Classey Ltd & The Wedge Entomology Research Foundation, London.

Hodges, R.W. (1983) Gelechioidea. *Check List of the Lepidoptera of America North of Mexico, including Greenland* (ed. by R.W. Hodges, T. Dominick, D.R. Davis, D.C. Ferguson, J.G. Franclemont, E.G. Munroe and J.A. Powell), pp. 11–25. E. W. Classey Ltd & The Wedge Entomology Research Foundation, London.

Hodges, R.W. (1986) *The Moths of America North of Mexico, Fascicle 7.1 Gelechioidea: Gelechiidae (in part) Dichomeridinae*. The Wedge Entomology Research Foundation, Washington, District of Columbia.

Hodges, R.W. (1998) The Gelechioidea. *Handbook of Zoology*, *Vol. IV: Arthropoda: Insecta, Part 35, Lepidoptera, Moths and Butterflies* (ed. by N.P. Kristensen), pp. 131–158. Springer, Berlin.

Kaila, L. (2004) Phylogeny of the superfamily Gelechioidea (Lepidoptera: Ditrysia): an exemplar approach. *Cladistics*, **20**, 303–340.

Kaila, L., Mutanen, M. & Nyman, T. (2011) Phylogeny of the megadiverse Gelechioidea (Lepidoptera): adaptations and determinants of success. *Molecular Phylogenetics and Evolution*, **61**, 801–809.

Karsholt, O., Mutanen, M., Lee, S. & Kaila, L. (2013) A molecular analysis of the Gelechiidae (Lepidoptera, Gelechioidea) with an interpretative grouping of its taxa. *Systematic Entomology*, **38**, 334–348.

Kawahara, A.Y. & Breinholt, J.W. (2014) Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20140970.

van Klinken, R.D., Fichera, G. & Gordo, H. (2003) Targeting biological control across diverse landscapes: the release, establishment, and early success of two insects on mesquite (*Prosopis* spp.) insects in Australian rangelands. *Biological Control*, **26**, 8–20.

Kristensen, N.P. & Skalski, A.W. (1998) Phylogeny and palaeontology. *Handbook of Zoology*, *Vol. IV: Arthropoda: Insecta, Part 35, Lepidoptera, Moths and Butterflies* (ed. by N.P. Kristensen), pp. 7–25. Walter de Gruyter, Berlin.

Kuznetsov, V.I. & Stekol'nikov, A.A. (1978) The taxonomic position and phylogenetic relations of the superfamily Coleophoroidea (Lepidoptera: Oecophoridae, Coleophoridae, Ethmiidae) inferred from the functional morphology of male genitalia. *Entomologicheskoe Obozrenie*, **57**, 131–149.

Kuznetzov, V.I. & Stekol'nikov, A.A. (1984) System and phylogenetic associations of families and superfamilies of gelechioid Lepidoptera of the infraorder Papilionomorpha (Lepidoptera: Copromorphoidea, Elachistoidea, Coleophoroidea, and Gelechioidea) with use of functional morphology of genitalia of males. *Trudy Instituta Zoologii, Akademiya Nauk SSSR*, **112**, 3–68.

Landry, J.-F. (1991) Systematics of Nearctic Scythrididae (Lepidoptera: Gelechioidea): phylogeny and classification of supraspecific taxa, with a review of described species. *Memoirs of the Entomological Society of Canada*, **160**, 1–341.

Lemmon, A.R., Brown, J.M., Stanger-Hall, K. & Lemmon, E.M. (2009) The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. *Systematic Biology*, **58**, 130–145.

Lewis, P., Holder, M.T. & Holsinger, K.E. (2005) Polytomies and Bayesian phylogenetic inference. *Systematic Biology*, **54**, 241–253.

Lvovsky, A.L. (1996) Composition of the genus *Odites* Wlsm. and its position in the classification of the Gelechioidea (Lepidoptera). *Entomologicheskoe Obozrenie*, **75**, 650–659.

Lvovsky, A.L. (2011) Comments on the classification and phylogeny of broad-winged moths (Lepidoptera, Oecophoridae *sensu lato*). *Entomologicheskoe Obozrenie*, **90**, 892–913.

MacKay, M.R. (1972) Larval sketches of some Microlepidoptera, chiefly North American. *Memoirs of the Entomological Society of Canada*, **88**, 1–83.

McDunnough, J.H. (1939) Check list of the Lepidoptera of Canada and the United States of America, Part 2. Microlepidoptera. *Memoirs of the Southern California Academy of Sciences*, **2**, 1–171.

McMillan, I. (2013) *Xyloryctine Moths of Australia. Genera and Species* [WWW document]. URL http://xyloryctinemothsofaustralia.blogspot.com/p/xyloryctinae.html [accessed on 15 March 2014].

Meyrick, E. (1928) *A Revised Handbook of British Lepidoptera*. Watkins & Doncaster, London.

Minet, J. (1986) Ebauche d'une classification moderne de l'ordre des Lépidoptères. *Alexanor*, **14**, 291–313.

Minet, J. (1988) Quelques caractères pré-imaginaux permettant de maintenir les Pterolonchidae au sein des Gelechioidea (Lepidoptera: Ditrysia). *Annales de le Société Entomologique de France (New Series)*, **24**, 375–376.

Minet, J. (1990) Remaniement partiel de la classification des Gelechioidea, essentiellement en fonction de caractères préimaginaux (Lepidoptera Ditrysia). *Alexanor*, **16**, 239–255.

Minet, J. (1991) Tentative reconstruction of the ditrysian phylogeny (Lepidoptera: Glossata). *Entomologica Scandinavica*, **22**, 69–95.

Moulton, J.K. & Wiegmann, B.M. (2003) Evolution and phylogenetic utility of CAD (rudimentary) among Mesozoic-aged eremoneuran Diptera (Insecta). *Molecular Phylogenetics and Evolution*, **31**, 363–378.

Mutanen, M., Wahlberg, N. & Kaila, L. (2010) Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. *Proceedings of the Royal Society B*, **277**, 2839–2848.

van Nieukerken, E.J., Kaila, L., Kitching, I.J. *et al.* (2011) Order Lepidoptera. *Animal Biodiversity: An Outline of Higher-Level Classification and Survey of Taxonomic Richness* (ed. by Z.-Q. Zhang). *Zootaxa*, **3148**, 212–221.

Nye, W.B. & Fletcher, D.S. (1991) *The Generic Names of Moths of the World*, *Vol. 6: Microlepidoptera*. Natural History Museum, London.

Passoa, S.C. (1995) *Larval and pupal systematics of Nearctic Amphisbatinae and Depressariinae (Lepidoptera: Oecophoridae)*. PhD Thesis, University of Illinois, Urbana, Illinois.

Pierce, N. (1995) Predatory and parasitic Lepidoptera: carnivores living on plants. *Journal of the Lepidopterists' Society*, **49**, 412–453.

Ponomarenko, M.G. (2006) Subfamily Dichomeridinae (Lepidoptera, Gelechiidae): phylogeny, classification, and position in the system of gelechiid moths. *Entomologicheskoe Obozrenie*, **85**, 375–384.

Powell, J.A. (1980) Evolution of larval food preferences in Microlepidoptera. *Annual Review of Entomology*, **25**, 133–159.

Powell, J.A., Mitter, C. & Farrell, B. (1998) Evolution of larval food preferences in Lepidoptera. *Handbook of Zoology*, *Vol. IV: Arthropoda: Insecta, Part 35, Lepidoptera, Moths and Butterflies* (ed. by N.P. Kristensen), pp. 403–422. Walter de Gruyter, Berlin.

Ratnasingham, S. & Hebert, P.D.N. (2007) Barcoding. BOLD: the barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364. DOI: 10.1111/j.1471- 8286.2007.01678.x.

Regier, J.C. (2008) *Protocols, Concepts, and Reagents for Preparing DNA Sequencing Templates*. V. 12/4/08 [WWW document]. URL http://www.umbi.umd.edu/users/jcrlab/PCR_primers.pdf [accessed on 25 January 2010].

Regier, J.C. & Shi, D. (2005) Increased yield of PCR product from degenerate primers with nondegenerate, nonhomologous 5' tails. *BioTechniques*, **38**, 34–38.

Regier, J.C. & Zwick, A. (2011) Sources of signal in 62 protein-coding nuclear genes for higher-level phylogenetics of arthropods. *PLoS ONE*, **6**, e23408.

Regier, J.C., Fang, Q.Q., Mitter, C., Peigler, R.S., Friedlander, T.P. & Solis, M.A. (1998) Evolution and phylogenetic utility of the *period* gene in Lepidoptera. *Molecular Biology and Evolution*, **15**, 1172–1182.

Regier, J.C., Shultz, J.W., Ganley, A.R.D. *et al.* (2008) Resolving arthropod phylogeny: exploring phylogenetic signal within 41 kb of protein-coding nuclear gene sequence. *Systematic Biology*, **57**, 920–938.

Regier, J., Zwick, A., Cummings, M. *et al.* (2009) Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. *BioMed Central Evolutionary Biology*, **9**, 280. DOI: 10.1186/1471-2148-9-280.

Regier, J.C., Shultz, J.W., Zwick, A. *et al.* (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature*, **463**, 1079–1083.

Regier, J.C., Mitter, C., Zwick, A. *et al.* (2013) A large-scale, higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and butterflies). *PLoS ONE*, **8**, e58568.

Regier, J.C., Mitter, C., Davis, D.R. *et al.* (2015) A molecular phylogeny and revised classification for the oldest ditrysian moth lineages (Lepidoptera: Tienoidea), with implications for ancestral habits of the mega-diverse Ditrysia. *Systematic Entomology*, **40**, 409–432.

Rubinoff, D. (2008) Phylogeography and ecology of an endemic radiation of Hawaiian aquatic case-bearing moths (*Hyposmocoma*: Cosmopterigidae). *Philosophical Transactions of the Royal Society B*, **363**, 3459–3465.

Rubinoff, D. & Haines, W.P. (2005) Web-spinning caterpillar stalks snails. *Science*, **309**, 575.

Rubinoff, D. & Schmitz, P. (2010) Multiple aquatic invasions by an endemic, terrestrial Hawaiian moth radiation. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 5903–5906.

Saito, T. (2005) Immature stages of two species of the genus *Deuterogonia* (Lepidoptera, Oecophoridae) in Japan, with remarks on the systematic position of the genus. *Tinea*, **18** (Suppl. 3), 45–54.

Schmitz, P. & Rubinoff, D. (2011) Ecologically and morphologically remarkable new Cosmet moth species of the genus *Hyposmocoma* (Lepidoptera: Cosmopterigidae) endemic to the Hawaiian Islands, with reference to the spectacular diversity of larval cases. *Annals of the Entomological Society of America*, **104**, 1–15.

Scoble, M.J. (1992) *The Lepidoptera. Form, Function and Diversity*. Oxford University Press, Oxford.

Shen, S. & Xie, X. (1990) Study on the bionomics of *Beijinga utila* Yang, an important natural enemy of scale insects belonging to *Eulecanium*. *Scientia Silvae Sinicae*, **26**, 30–38.

Sinev, S.Y. (1992) On the system and phylogeny of the Gelechioidea *s.l.* (Lepidoptera). *Entomologicheskoe Obozrenie*, **71**, 143–159.

Staden, R. (1999) *Staden Package*. MRC Laboratory of Molecular Biology, Cambridge. [WWW document]. URL http://www.mrclmb.cam.ac.uk/pubseq/ [accessed on 19 September 2007].

Stainton, H.T. (1854) *Insecta Britannica, Lepidoptera: Tineina*, Vol. **3**. Reeve & Benham, London.

Stehr, F.W. (1987) Scythrididae (Gelechioidea). *Immature Insects*, Vol. **1** (ed. by F.W. Stehr), pp. 393–394. Kendall/Hunt, Dubuque, Iowa.

Stephens, J.F. (1829) *A Systematic Catalogue of British Insects*. Baldwin & Cradock, London.

Sugisima, K. & Arita, Y. (2000) A new species of a gelechioid genus, *Idioglossa* Walsingham (Lepidoptera, Batrachedridae, Batrachedrinae), from Japan. *Transactions of the Lepidopterological Society of Japan*, **51**, 319–336.

Suzuki, Y., Glazko, G.V. & Nei, M. (2002) Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 16 138–16 143.

Swofford, D.L. (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer Associates, Sunderland, Massachusetts.

Timmermans, M.J.T.N., Lees, D.C. & Simonsen, T.J. (2014) Towards a mitogenomic phylogeny of Lepidoptera. *Molecular Phylogenetics and Evolution*, **79**, 169–178.

Wiens, J.J. (1998) Does adding characters with missing data increase or decrease phylogenetic accuracy? *Systematic Biology*, **47**, 625–640.

Wiens, J.J. (2003) Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology*, **52**, 528–538.

Zwick, A. (2010) *Degeneracy Coding Web Site*. *PhyloTools*, 10 February 2010 [WWW document]. URL http://www.phylotools.com/ptdegen1webservice.htm [accessed on 1 September 2011].

Zwick, A., Regier, J.C., Mitter, C. & Cummings, M.P. (2011) Increased gene sampling yields robust support for higher-level clades within Bombycoidea (Lepidoptera). *Systematic Entomology*, **36**, 31–43.

Zwick, A., Regier, J.C. & Zwickl, D.J. (2012) Resolving discrepancy between nucleotides and amino acids in deep-level arthropod phylogenomics: differentiating serine codons in 21-amino-acid models. *PLoS ONE*, **7**, e47450. DOI: 10.1371/journal.pone.0047450.

Zwickl, D.J. (2011) *GARLI 2.0* [WWW document]. URL http://code.google.com/p/garli [accessed on 3 May 2011].

Accepted 4 August 2015
First published online 4 September 2015