Supertrees and the Tree of Life: generating a metaphylogeny for a diverse invertebrate family (Insecta : Diptera : Therevidae) using constraint trees and the parsimony ratchet to overcome low taxon overlap

Christine L. Lambkin[A,B,N], *John W. H. Trueman*[C], *David K. Yeates*[A], *Kevin C. Holston*[D,E],
Donald W. Webb[F], *Martin Hauser*[D,G], *Mark A. Metz*[D,H], *Hilary N. Hill*[I,J],
Jeffrey H. Skevington[K], *Longlong Yang*[I,L], *Michael E. Irwin*[D,M] *and Brian M. Wiegmann*[I]

[A]CSIRO Entomology, PO Box 1700, Canberra, ACT 2601, Australia.

[B]Queensland Museum, PO Box 3300, South Bank, Brisbane, Qld 4101, Australia.

[C]School of Botany and Zoology, Australian National University, Canberra, ACT 0200, Australia.

[D]Department of Natural Resources & Environmental Sciences, University of Illinois, 901 West Illinois Street,
 Urbana, IL 61801, USA.

[E]Research Division, Entomology, Swedish Museum of Natural History, PO Box 50007, SE-104 05 Stockholm,
 Sweden.

[F]Illinois Natural History Survey, Institute of Natural Resource Sustainability, University of Illinois,
 1816 South Oak Street, MC-652, Champaign, IL 61820, USA.

[G]Department of Food and Agriculture, Plant Pest Diagnostics Branch, 3294 Meadowview Road, Sacramento,
 CA 95832-1448, USA.

[H]United States National Museum of Natural History, 10th and Constitution Avenue NW, Washington,
 DC 20013, USA.

[I]Department of Entomology, North Carolina State University, Raleigh, NC 27695-7613, USA.

[J]16 N. Belle Grove Road, Catonsville, MD 21228, USA.

[K]Canadian National Collection of Insects, Agriculture and Agri-Food Canada, 960 Carling Avenue,
 K.W. Neatby Building Ottawa, ON K1A 0C6, Canada.

[L]The Hamner Institute for Health Sciences, 6 Davis Drive, PO Box 12137, Research Triangle Park,
 NC 27709-2137, USA.

[M]15634 E. Wandering Creek Place, Vail, AZ 85641, USA.

[N]Corresponding author. Email: christine.lambkin@qm.qld.gov.au

Abstract. The dipteran family Therevidae (stiletto flies) is cosmopolitan and has been the focus of many taxonomic and phylogenetic studies over the last 25 years. Despite this work, questions remain concerning the relationships between subfamilies, genera and generic groups and membership of those groups. We use the supertree method to produce an inclusive phylogeny for the family Therevidae from 24 phylogenetic studies using matrix representation with parsimony (MRP) analysis. The supertree method, one of the most common approaches to calculating globally inclusive phylogenies from smaller more exclusive analyses, produced the therevid metaphylogeny despite only 34% of the terminal taxa being found in more than one source tree. We describe a method for handling low taxon overlap in supertree analyses, in combination with the parsimony ratchet and constraint tree techniques. The supertree presented here is an overarching phylogenetic hypothesis of the Therevidae, incorporating extensive sampling of major lineages and summarising past phylogenetic work on the family. The inclusive metaphylogeny for 362 therevid taxa robustly retrieves the subfamilies Agapophytinae, Phycinae, Therevinae and Xestomyzinae, and the tribes Cyclotelini and Therevini. The Phycinae and Xestomyzinae form a clade, sister to the remaining Therevidae. The Australasian and South American *Taenogera* Kröber genus-group is monophyletic and sister to a clade of Therevinae and the Australian endemic Agapophytinae. The Therevinae consists of the *Anabarhynchus* Macquart genus-group of Australian, South American, New Caledonian and New Zealand taxa as sister to the non-Australasian 'higher Therevinae', which contains the tribes Cyclotelini and Therevini. The Therevini includes the *Hoplosathe* Lyneborg & Zaitzev, *Litolinga* Irwin & Lyneborg, *Baryphora* Loew, *Pandivirilia* Irwin & Lyneborg and *Thereva* Latreille generic-groups. MRP supertree methods can be used to produce inclusive metaphylogenies in situations where source trees have poor data overlap and low taxon overlap, and are therefore valuable in species-rich groups such as

arthropods. These methods may be necessary for constructing the 'Tree of Life', representing phylogenetic relationships among the millions of known species. However, our analyses show that in situations of source tree conflict, MRP supertree analyses present only the majority signal. We also show that conflict between source trees can be hidden in MRP supertrees, thus our results emphasise the need to evaluate the resulting clades with reference to the source trees.

Introduction

Stiletto flies belong to the Therevidae (Diptera : Brachycera), a medium-sized family that comprises ~1600 described species, which are found on every continent except Antarctica. Australia has the world's richest therevid fauna and the genera show an extraordinary degree of endemism, with 22 of the 24 described genera found only in that region. Therevids are flies of moderate to small size, often with patterned wings and silvery pruinescent markings on the body. In Australia many adult therevids are brilliantly coloured (Fig. 1*A*, *C–E*, *G*), with modifications in body shape and behaviour that aid in their mimicry of various wasps (Winterton *et al.* 2001). Adults frequent a wide variety of habitats, often in rather dry situations such as sand dunes or beaches. Little is known of the adult habits as they are generally secretive – some are found at flowers and many track drying creek beds. Some genera such as *Agapophytus* Guérin (Fig. 1*D*) are commonly found on tree trunks (Winterton and Irwin 2001) and *Ectinorhynchus* Macquart males (Fig. 1*E*) can often be seen hovering in small swarms above and among shrubs (Ferguson and Lambkin 2006). Adults are often collected in very large numbers in Malaise traps placed across flight paths in gullies (Lambkin *et al.* 2002). Alpha-diversity can be quite high; for example, the Warrumbungle Mountains area in New South Wales, Australia, is home to more than 100 species. The larvae are smooth, thin, white and vermiform, with a rather well developed head; the abdomen is secondarily divided into some 16 apparent segments and terminates in a pair of tiny pseudopods (Colless and McAlpine 1991). They are found mainly in sand or soil close to the surface, are voracious predators of other insect immatures, especially Coleoptera and Lepidoptera, and can give a painful bite (English 1950).

The Therevidae is currently divided into four subfamilies: Agapophytinae, Phycinae, Therevinae and Xestomyzinae. Lyneborg (1976) erected the first subfamily-level classification, splitting the Therevidae into two subfamilies, Phycinae and Therevinae, and naming the Xestomyzini and Phycini as tribes of the Phycinae. The Xestomyzini mainly occurs in southern Africa with one genus, *Henicomyia* Coquillett (Fig. 1*B*), found in the New World, and was elevated to subfamily level by Irwin and Webb (1992). The Phycinae (Fig. 1*F*) have their highest diversity in Africa and are well represented in the southern and central parts of the Palaearctic Region, spreading eastward into the Oriental Region and into the New World, but are not found in the Australian Region (Hauser 2005).

Winterton erected the endemic Australian subfamily Agapophytinae (Winterton *et al.* 2001) after tentatively establishing the *Taenogera*-group (Winterton *et al.* 1999*a*), stating that it is found in the Australasian region but may include some genera from Chile. The validity of the *Taenogera* genus-group was subsequently questioned (Winterton and Irwin 2001; Winterton *et al.* 2001), and later regarded as merely part of the Agapophytinae (Winterton 2006, 2007*a*, 2007*b*).

Half of the described genera and over 70% of the described species in Therevidae belong to the Therevinae, whose monophyly is supported by both molecular (Yang *et al.* 2000; Holston *et al.* 2007) and morphological (Lyneborg 1992, 2001; Winterton *et al.* 1999*a*, 2001) evidence. Phylogenetic relationships among the therevine genera have been poorly understood (Holston *et al.* 2007). Gaimari and Irwin (2000*a*, 2000*b*) described and characterised the therevine tribe Cyclotelini (Fig. 1*L*). Metz (2002) examined relationships amongst the genera of the higher Therevidae. Holston (2003) tested the monophyly of the genus *Thereva* Latreille (Fig. 1*H*). In a recent examination of the relationships among the therevine genera good support was found for the Cyclotelini, *Pandivirilia* Irwin & Lyneborg (Fig. 1*I*), *Litolinga* Irwin & Lyneborg and *Arenigena* Irwin & Lyneborg clades (Holston *et al.* 2007).

The number of therevid studies has increased particularly over the last 14 years, with the development of the NSF-PEET research program concentrating on monographic revisions and compilation of molecular datasets for phylogenetic analyses. Over the years of the Therevid PEET project both individuals and groups endeavoured to create a morphological matrix for the family, thwarted by difficulties in achieving consistency across subfamilies and generic groups with essentially species-level examinations. Attempts were then made to concatenate

Fig. 1. Therevidae. (*A*) *Taenogera* genus-group, *Johnmannia kosciuskoensis* Lambkin & Recsei from Tallaganda NP, NSW, Australia. Photograph by C. Lambkin. (*B*) Xestomyzinae, *Henicomyia* sp. from Guatemala (MEI 125377). Illustration (TAFKAM) digitally prepared by J. Marie Metz, funded by NSF (Therevid PEET project) and Schlinger Foundation. (*C*) Agapophytinae, *Acupalpa semirufa* Mann from Tallaganda NP, NSW, Australia. Photograph by C. Lambkin. (*D*) Agapophytinae, *Agapophytus pallidicornis* (Kröber) from Australia. Photograph by Shaun Winterton, QDPI, Brisbane, Australia. (*E*) *Taenogera* genus-group, *Ectinorhynchus latistria* (Walker) from Australia. Photograph by Cor Zonneveld, Netherlands. (*F*) Phycinae, *Orthactia pencillata* Lyneborg from Africa. Photograph by S. Winterton. (*G*) Agapophytinae, *Pipinnipons fascipennis* (Kröber) at mud, from Australia. Photograph by Dr C. Riley Nelson, Brigham Young University, Utah, USA. (*H*) Therevinae, *Thereva* group, *Thereva nobilitata* (Fabricius) from Volsted, Denmark. Photograph by Peter Krogh, Denmark. (*I*) Therevinae, *Pandivirilia* group, *Pandivirilia albifrons* (Say) from USA. Photograph by Tom Murray, Massachusetts, USA. (*J*) Therevinae, *Baryphora* group, *Baryphora speciosa* Loew from Hierapolis, Turkey. Photograph by Nicole Lartigau, France. (*K*) Therevinae, *Anabarhynchus* group, *Anabarhynchus kampmeierae* Irwin & Lyneborg. Photograph by S. Winterton. (*L*) Therevinae, Cyclotelini, *Ozodiceromyia argentata* (Bellardi) from USA. Photograph by Tom Murray.
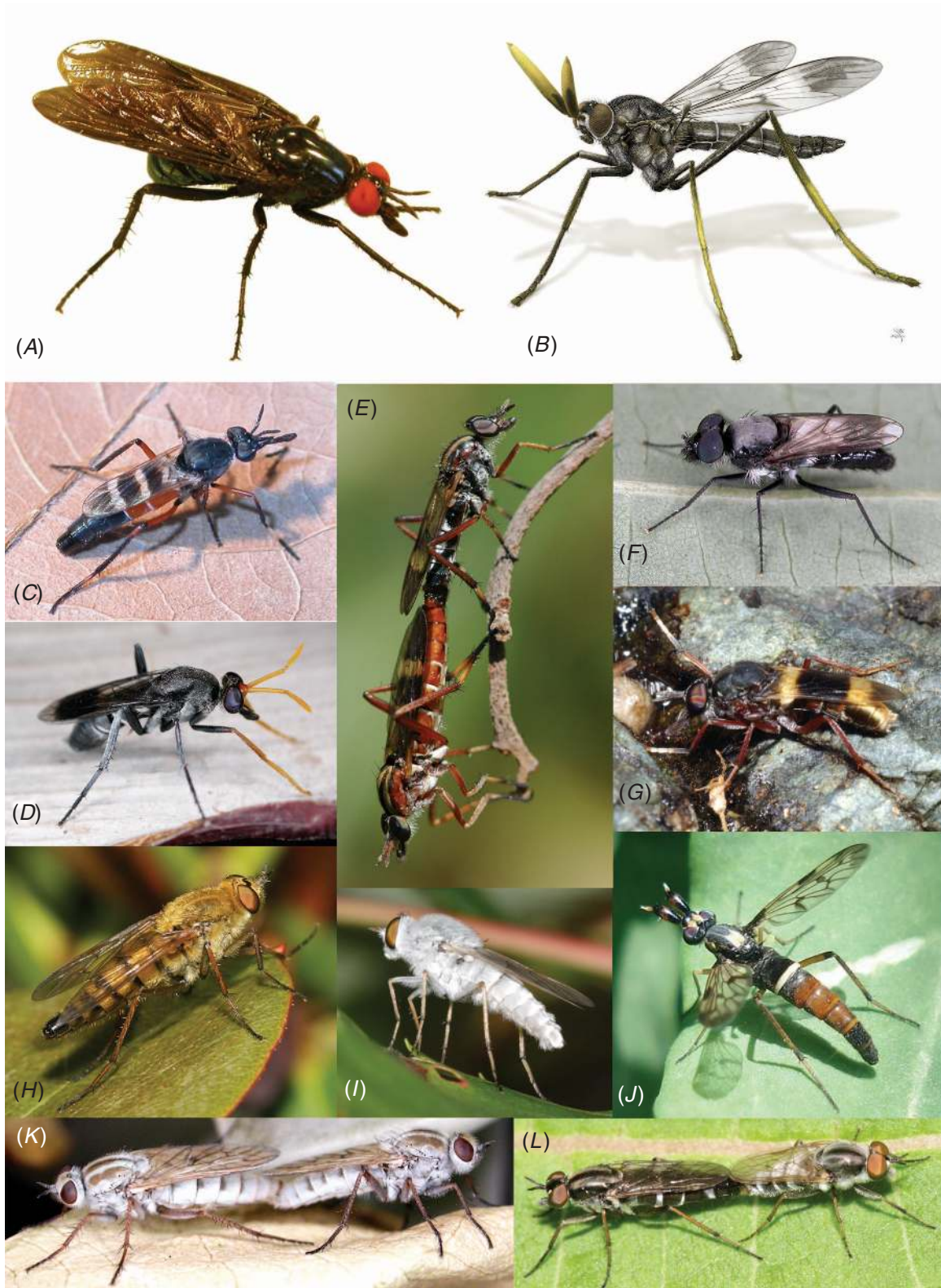
morphological matrices but this proved impossible because in too many cases the definition of the states for the same character differed significantly between studies, e.g.

- *Number of rows of postocular setae in female*: 1–3 poorly defined rows/one row/more than three rows (Winterton *et al.* 1999*a*),
- *Number of rows of postocular setae in female*: multiple/one well defined row/a single poorly defined row (Winterton and Irwin 2001),
- *Occipital setae other than dorsal row on female*: present/absent (Metz and Irwin 2000).

Of the 22 studies containing phylogenies that focus on therevids (Table 1), 15 focused on species-groups or intrageneric relationships (Webb and Irwin 1991, 1995; Winterton and Irwin 1999, 2001; Winterton *et al.* 1999*a*, 1999*b*, 2000; Gaimari and Irwin 2000*b*; Metz and Irwin 2000; Winterton 2000; Metz 2002; Hauser and Irwin 2003; Holston 2003; Metz and Webb 2003; Webb and Metz 2003). Six studies include molecular data (Yang *et al.* 2000; Winterton *et al.* 2001; Hill 2003; Holston 2003; Yeates *et al.* 2003; Hill and Winterton 2004), two are purely molecular and 20 from 12 different authors include morphological characters.

In an attempt to discern the relationships between subfamilies, genera and generic groups, all six studies that included molecular data and the morphologically based works of Winterton *et al.* (1999*a*), Gaimari and Irwin (2000*a*), and Metz (2002) examined a cross-section of the family or large groups previously recognised in the Therevidae. Despite this work, the relationships among the subfamilies and their status have not been generally accepted (Hauser 2005). For example, the analysis of Yang *et al.* (2000) supported a monophyletic Therevidae but could not resolve the relationships between the higher Therevidae, Phycinae and Xestomyzinae that was represented only by the genus *Henicomyia*. Later, Winterton *et al.* (2001) placed the Agapophytinae between Phycinae and the *Taenogera* group and Therevinae, without recognising the subfamily status of Xestomyzinae or the relationship between Xestomyzinae and Agapophytinae (Hauser 2005). Consequently, one of the least understood parts of the family are the basal clades and their relationship to one another (Hauser 2005). Questions also remain concerning the relationships between genera and generic groups and membership of those groups, mainly because of the inadequate taxon sampling included in analyses. An articulated gonocoxal process distinguishes the *Anabarhynchus* Macquart group (Fig. 1*K*), or 'lower' Therevinae, from 'higher' Therevinae but is also found in the *Taenogera*-group and Agapophytinae (Gaimari and Irwin 2000*a*; Winterton *et al.* 2001; Holston *et al.* 2007). Cyclotelini includes 10 therevine genera (Gaimari and Irwin 2000*a*) and five informal genus-groups have been proposed for all but 11 of the remaining

Table 1.   Source trees and taxon overlap used for matrix representation with parsimony (MRP) analysis of Therevidae

All studies based on morphological datasets unless otherwise stated

Study group	References	No. of terminal taxa	No. of matrix elements	No. of taxa overlap	No. of taxa with no overlap
Pallicephala Irwin & Lyneborg	Webb and Irwin (1991)	8	5	4	4
Chromolepida Cole	Webb and Irwin (1995)	5	3	2	3
Tabuda Walker	Webb and Irwin (1999)	10	8	6	4
Laxotela Winterton & Irwin	Winterton and Irwin (1999)	6	4	2	4
Nanexila Winterton & Irwin	Winterton *et al.* (1999*b*)	25	23	8	17
Taenogera Kröber genus-group	Winterton *et al.* (1999*a*)	22	19	9	13
Cyclotelini	Gaimari and Irwin (2000*a*)	21	19	12	9
Lindneria Kröber	Metz and Irwin (2000)	18	9	10	8
Ozodiceromyia Bigot	Gaimari and Irwin (2000*b*)	6	4	2	4
Bonjeania Irwin & Lyneborg	Winterton *et al.* (2000)	16	14	4	12
Acupalpa Kröber	Winterton (2000)	8	6	7	1
Therevidae – 28S	Yang *et al.* (2000)	33	27	33	0
Therevidae – EF1α	Yang *et al.* (2000)	38	28	37	1
Agapophytus Guérin	Winterton and Irwin (2001)	43	38	6	37
'higher Therevinae'	Metz (2002)	31	29	25	6
Therevinae	Metz (2002)	96	66	56	40
Ammonaios Irwin & Lyneborg	Hauser and Irwin (2003)	12	8	6	6
Therevidae – combined molecular	Hill (2003)	26	23	20	6
Therevinae – combined molecular	Holston (2003)	58	54	44	14
Thereva Latreille – combined morphological/molecular	Holston (2003)	43	40	24	19
Distostylus Metz & Webb	Metz and Webb (2003)	9	7	8	1
Pandivirilia Irwin & Lyneborg	Webb and Metz (2003)	28	23	11	17
Evocoidae Yeates, Irwin & Wiegmann – 28S	Yeates *et al.* (2003)	15	13	11	4
Acraspisoides Hill & Winterton – combined morphological/molecular	Hill and Winterton (2004)	27	23	19	8
Therevid supertree		362	493	124	238

genera, but group membership is difficult to establish for new or unplaced therevine taxa (Holston *et al.* 2007). We aim to produce a quantitative summary of this systematic work in an overarching metaphylogeny of relationships between subfamilies, genera and generic groups for over 350 therevid taxa to provide a basis for answering these questions and for future testing of higher-level evolutionary hypotheses.

Supertree methods: an overview

The most common approaches to calculating globally inclusive phylogenies from smaller, less inclusive (source) analyses are supertree methods (Sanderson *et al.* 1998) and supermatrix (Gatesy *et al.* 2002, 2004; Bininda-Emonds *et al.* 2003; Malia *et al.* 2003) or total evidence (Kluge 1989) methods. Supermatrices (Wiens and Reeder 1995; Sanderson *et al.* 1998; Springer and de Jong 2001) concatenate datasets into a single larger matrix, inserting missing values for incomplete datasets where the taxa from one analysis have not been scored in another (Fig. 2*B*). The supermatrix approach is to be preferred if the data are available (Nixon and Carpenter 1996; Sanderson *et al.* 1998; Gatesy and Arctander 2000; Gatesy *et al.* 2003, 2004; Hughes and Vogler 2004; Matthee *et al.* 2004; Gatesy and Baker 2005; Hill 2005) and if they are compatible in the sense of being arranged together and analysed optimally within a single matrix. This can be difficult with, for example, mixed character and distance data or morphological and molecular sequence data. The supermatrix approach presents a major challenge for combining data from different morphological studies including large numbers of taxa, such as for arthropods. The task of adequate taxon sampling and gathering sufficient information for phylogenetic analyses of large groups, which may contain thousands of species, is immense. Methods based on

biological character sampling are challenged at this scale unless data collection is carefully and appropriately coordinated to provide the data coverage necessary for analysis (Salamin and Davies 2004). While scattered molecular datasets for the same genes may pose problems in alignment across biodiverse groups, in a similar fashion, assessing homologies in morphological characters in datasets accumulated by several different authors across large and diverse groups poses considerable challenges (Cardillo *et al.* 2004). This is the case in the Therevidae, where 20 of the 22 phylogenetic studies from 12 different authors include morphological characters. For such groups, supertree analyses currently represent the best method of producing a summary of previous systematic work to provide a basis for testing higher-level evolutionary hypotheses.

Supertrees

Supertree methods were developed to overcome the challenges that arise when an inclusive, overall hypothesis of relationships is necessary for a group where previous phylogenetic studies used different data sources and significantly differing taxon samples. Supertree (Sanderson *et al.* 1998) meta-analyses combine a set of partially overlapping source trees (Pisani and Wilkinson 2002) to yield a single composite tree (Purvis 1995*b*) (Fig. 2*E*) or multiple most parsimonious composite trees (MPCT) that are summarised by strict consensus to generate a consensus composite tree (CCT) (Bininda-Emonds and Bryant 1998). Because supertrees are calculated from source trees (Fig. 2*C*), not source data, the support for supertree nodes and evidence for or against relationships should be sought from the source trees themselves (Bininda-Emonds *et al.* 1999). Supertree methods indicate where broad agreement exists and where it does not and therefore indicate areas requiring further study (Purvis 1995*a*).
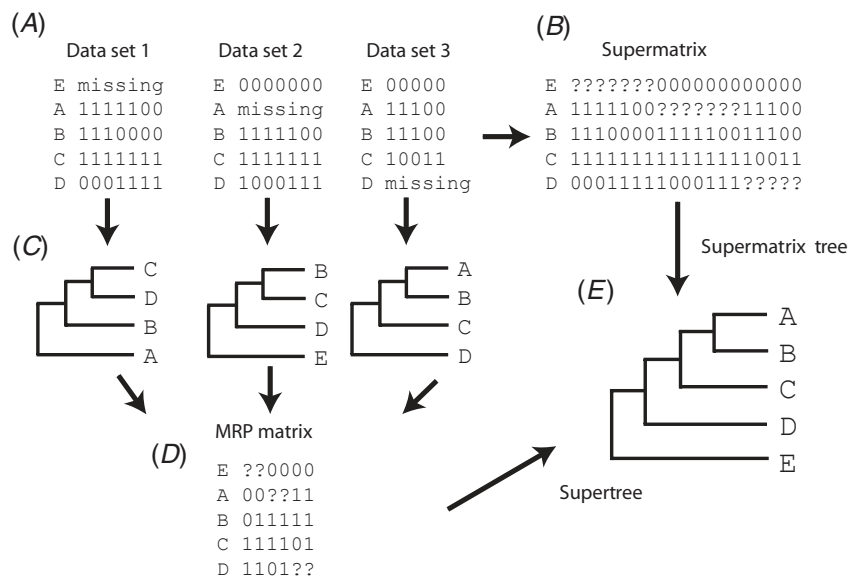

Fig. 2. Schematic of supertree and supermatrix methods. (*A*) Incomplete datasets with overlapping taxa. (*B*) Supermatrix: overlapping incomplete data matrices are concatenated. (*C*) Incomplete datasets produce different trees. (*D*) MRP matrix coded from trees produced from incomplete datasets. (*E*) Supertree produced from MRP matrix, Tree from supermatrix. Modified from de Queiroz and Gatesy (2007: 36, fig. 2).

MRP

Baum (1992) and Ragan (1992) proposed the most frequently used phylogenetic supertree method (Wilkinson *et al*. 2004; Beck *et al*. 2006), matrix representation with parsimony analysis (MRP), where each clade or node on a character-based source tree is scored as a single matrix element (Ragan 1992), coded '1'. Terminal taxa in the source tree not present in the clade are coded as '0'. A matrix is formed by the combination of the matrix elements to describe the clades from all source trees, with a missing-value code, e.g. '?', for terminal taxa not in that source tree (Fig. 2*D*). The matrix is then rooted by the addition of the MR-outgroup (Bininda-Emonds *et al*. 2005), the equivalent of a hypothetical 'ancestor' (except the matrix entries here represent its non-membership of any clade, not its possession of 'ancestral' character states), coded all '0', and analysed with parsimony. Phylogenetic supertree analyses require each source tree to minimally share two terminal taxa with other source trees (Bininda-Emonds *et al*. 2002). MRP analytical methods for constructing more-inclusive hypotheses have been developed (Baum 1992; Ragan 1992; Purvis 1995*b*; Bininda-Emonds and Bryant 1998), discussed (Ronquist 1996; Bininda-Emonds *et al*. 2003) and improved (Bininda-Emonds 2003). MRP appeals because of its ability to synthesise numerous small, disparate source trees (Fig. 2*C*) into a single more-inclusive (Beck *et al*. 2006) well resolved phylogeny (Bininda-Emonds and Sanderson 2001). This method effectively handles source trees with incompatible nodes without loss of resolution (Salamin and Davies 2004).

Previous supertree analyses

In the animal kingdom, phylogenetic supertree analyses have been completed for several groups of vertebrates: 226 tetrapods (Ruta *et al*. 2003), 98 species of advanced snakes (Kelly *et al*. 2003), 383 species of frogs (Summers *et al*. 2006), 47 tinamous (Aves) (Bertelli *et al*. 2002), 122 seabirds (Kennedy and Page 2002), 352 parrots (Munshi-South and Wilkinson 2006), 8 hammerhead sharks (Cavalcanti 2007), 146 Insectivora (Grenyer and Purvis 2003), 4501 of the 4554 extant species of mammals (Bininda-Emonds *et al*. 2007), 203 primates (Purvis 1995*a*), 271 Carnivora (Bininda-Emonds *et al*. 1999; Bininda-Emonds 2000), 79 caniform carnivores (Fulton and Strobeck 2006), all 925 Chiroptera (Jones *et al*. 2002), 80 lagomorphs (Stoner *et al*. 2003) and 171 Artiodactyla (Mahon 2004). Among the plants, supertrees have been prepared for angiosperm families with 379 terminal taxa (Davies *et al*. 2004), 403 genera of grasses (Salamin *et al*. 2002) and 108 species of *Symplocos* (Ericales) (Fritsch *et al*. 2006). Using published supermatrices that included at least 10 genes from each taxon and each gene with sequences from at least four taxa, Burleigh *et al*. (2006) created supertrees to examine support values for 70 metazoans, 69 green plants, 49 bilaterans and 8 yeast taxa. Within the invertebrates, until recently only relatively small supertree analyses had been completed, with 77 species of extinct stylophoran echinoderms (Ruta 2003), 39 egg parasitoid wasp taxa from the family Eulophidae (Hymenoptera) (Cuignet *et al*. 2007), 46 taxa of acorn weevils (Hughes and Vogler 2004) and 34 arthropod species including 21 beetles in a subfamily investigation of the Coleoptera (Hughes *et al*. 2006). Yeates *et al*. (2007) constructed the

dipteran supertree for 151 families, van der Linde and Houle (2008) presented a large arthropod supertree for 297 species of *Drosophila* and related genera and Beutel *et al*. (2008) produced the largest arthropod supertree analysis for 401 genera of the Adephaga, the second largest suborder of beetles.

Supertrees and low taxon overlap

As the degree of taxon overlap among source trees decreases, the ability of MRP methods to retrieve a model tree decreases significantly (Bininda-Emonds and Sanderson 2001). In previous MRP supertree analyses, taxon overlap has ranged from 100%, where there is one complete dataset containing all taxa (Eick *et al*. 2005; Robinson and Matthee 2005; Munshi-South and Wilkinson 2006; Cavalcanti 2007), to 6%, where 94% of the taxa are found in only one of the datasets (Sanderson *et al*. 1998). Low taxon overlap in phylogenetic supertree analyses increases the amount of missing data in the MRP matrix, which may increase the number of MPCT and result in lack of resolution in the CCT. Among the 24 source trees used in the MRP therevid supertree analysis, 238 of the 362 terminal taxa (not including outgroups) or 65.7% are found in only one source tree and, therefore, have no overlap (Table 1). Of the 493 total MRP elements in our matrix, individual taxa register from 230 to 490 missing elements (46.7–99.4%). Even those terminal taxa with the highest overlap, being found in seven or eight of the 24 source trees, are missing over 46% of matrix elements (Table 2).

The phylogenetic accuracy of analyses including incomplete datasets in simulated studies is positively correlated with the number of characters that can be scored across all taxa (Wiens and Reeder 1995; Wiens 2003, 2006). On the other hand, missing data may lead to a loss of resolution but not necessarily to misleading relationships (Kearney 2002; Malia *et al*. 2003; Fulton and Strobeck 2006). Essentially, the missing data problem reflects sampling of too few complete characters to accurately place some taxa on the tree (Wiens 2003, 2006; Fulton and Strobeck 2006), but the effects of missing data depend on the precise distribution of missing data across the taxa (Novacek 1992).

With only 34% of the taxa in more than one source tree, low taxon overlap made the development of a phylogenetic therevid supertree problematic. The problem of low taxon overlap can be largely ameliorated by 'seeding' supertree analyses with the inclusion of one reasonably complete study that included most terminal taxa, even if poorly resolved (Bininda-Emonds *et al*. 2002). Non-phylogenetic taxonomic treatments were used to create seed trees for supertrees of all extant shorebirds (Thomas *et al*. 2004), marsupials (Cardillo *et al*. 2004),

Table 2. Missing data for terminal taxa with highest taxonomic overlap in therevid source trees

Terminal taxa	Source trees	% of 493 matrix elements missing
Anabarhynchus tristis Bigot	8	46.7
Penniverpa festina (Coquillett)	8	49.1
Agapophytus albobasalis Mann	8	63.9
Cyclotelus pictipennis (Wiedemann)	7	50.1
Chromolepida pruinosa (Coquillett)	7	53.3

Artiodactyla (Mahon 2004), placental mammals (Beck *et al.* 2006) and adephagan beetle genera (Beutel *et al.* 2008). For the therevid supertree none of the source trees included a majority of the 362 terminal taxa for use as a seed tree.

Taxonomic equivalents

MRP supertree matrices often include taxonomic equivalents (Wilkinson 1995), species for which there is so little information that they could be grouped equally parsimoniously with numerous other species. Cardillo *et al.* (2004) searched for taxonomic equivalents in their marsupial MRP matrix and removed six species from further analyses to allow the hidden resolution to be revealed. A search of the therevid MRP matrix (Additional File 1 [*therevidmatrix363t.nex*]; 'additional files' are available in an Accessary Publication on the *Invertebrate Systematics* website) revealed that 299 of the 362 terminal taxa were taxonomic equivalents.

Constraint tree

When generating a supertree for 113 species of pines from 20 source trees, Grotkopp *et al.* (2004) found that the position of one taxon was unstable, and attributed that problem to the single use of the taxon, with insufficient context to place it within the correct group. They used a constraint tree to hold that floating taxon in place, as had been recommended by Page (2002, 2004). For the therevid supertree, we developed a backbone constraint tree, restricting each genus to be monophyletic unless there was prior evidence of non-monophyly. With this modification of the MPR supertree method, we were able to overcome the limitations caused by low taxon overlap and place species in generic position even though they lacked information from source trees.

Supertrees and the consensus paradox

Consensus methods suffer from the 'consensus paradox' whereby the more a group is studied, the more alternative phylogenetic hypotheses are generated and the more unresolved the consensus tree becomes (Purvis 1995*b*). Supertree methods were developed, in part, to avoid the consensus paradox (Purvis 1995*b*; Ronquist 1996; Bininda-Emonds and Bryant 1998). However, the influence of individual source trees depends on their size and resolution, meaning that source trees are not combined equally during MRP. In situations of considerable conflict between source trees and hidden support for alternative topologies, MRP produces the same result as the strict consensus tree because the MRP matrix presents only the principal signal from each dataset (Pisani and Wilkinson 2002).

Low taxon overlap in phylogenetic supertree analyses may increase the number of MPCT and result in lack of resolution in the CCT. Loss of resolution in the supertree may, however, reflect either a lack of information or significant disagreement among the source trees (Jones *et al.* 2002). Problems arise in supertree analyses with very small numbers of source trees, and support decreases as taxon overlap decreases, especially when there is no largely complete source tree (Bininda-Emonds *et al.* 2003). The low levels of taxon overlap between the source trees available for the therevid supertree may affect the accuracy of supertree methods (Eulenstein *et al.* 2004). However as Eulenstein *et al.* (2004) found, increasing the number of input trees can greatly increase the accuracy of the supertree when there is low taxon overlap. By using 24 source trees from 22 studies we attempted to increase the accuracy of the MRP therevid supertree.

Adams and majority-rule consensus approaches

The phylogenetic position of terminal taxa unique to one source tree is a major issue for supertree construction because of their potential to introduce ambiguity into the supertree (Bininda-Emonds *et al.* 1999; Eulenstein *et al.* 2004). The rogue branch, where a single branch appears in different places in different trees, can be identified by using an Adams consensus (Adams 1972) as it preserves the internal structure while collapsing the rogue branch to the first common node (Ragan 1992). After finding at least 100 000 MPCT in the seabird MRP supertree analysis Kennedy and Page (2002) summarised the results using both Adams and strict consensus trees. The Adams consensus was presented as it retained more information by summarising the similarity of optimal trees and indicating which groups included difficult-to-place taxa. We also used the Adams consensus in the development of the therevid supertree to identify difficult-to-place taxa.

Majority-rule consensus has been used to summarise the set of supertrees (Beck *et al.* 2006) with the frequencies related to the support received from the source trees (Pisani and Wilkinson 2002; Grotkopp *et al.* 2004). We present a 50% majority-rule consensus for the therevid supertree to provide a direct reading of the number of MPCT that include a given clade, as did Ruta *et al.* (2003) for tetrapod data.

Methods

Source trees

MRP analysis for the Therevidae included a total of 24 source trees (Table 1) with two trees included from Holston (2003) [overall and generic], Metz (2002) [overall and more inclusive] and Yang *et al.* (2000) [28S and EF1α]. As suggested by Bininda-Emonds *et al.* (2003), taxonomic inclusiveness (the tree with the highest number of terminal taxa) guided the choice of the Hill and Winterton (2004) phylogeny as a source tree rather than the earlier study of the Agapophytinae completed by Winterton *et al.* (2001). The 24 source trees were coded manually into a MRP data matrix (Additional File 1) using MacClade 4 (Maddison and Maddison 2001): 363 terminal taxa (including MRP outgroup) and 493 matrix elements (Table 1). The 15 years of intense study of the Therevidae has resulted in synonymies and differences in species designation of terminal taxa included in source trees. Careful cross-referencing of publications such as recent revisions (Webb and Metz 2006, 2008; Webb 2007) and checking of published database numbers in the World database of Therevidae, *Mandala* (Kampmeier *et al.* 2004; Kampmeier and Irwin 2009), allowed names to be converted to the current nomenclature.

Analyses

Weighting

When source trees differ considerably in the amount of taxon overlap, reweighting methods have questionable significance (Ruta 2003). Therefore, in this study, matrix elements coding source trees have been given equal weights.

Parsimony ratchet

The parsimony ratchet (Nixon 1999) is a fast, broad, search method that is less likely to be trapped on suboptimal islands, because searches jump from one section of tree space to another (Sikes and Lewis 2001*b*) and has often been used for supertree construction (Jones *et al.* 2002; Cardillo *et al.* 2004; Mahon 2004; Thomas *et al.* 2004; Beck *et al.* 2006; Munshi-South and Wilkinson 2006; Beutel *et al.* 2008). PAUPRat (Sikes and Lewis 2001*a*) in conjunction with PAUP* (Swofford 2002) was used to perform the parsimony ratchet, with 25 batches of 250 iterations with a random sample of 20% of the characters reweighted at each iteration. This format follows Nixon's (1999) suggestion that multiple short searches cover more tree space. A backbone constraint tree was developed using MacClade 4 (Maddison and Maddison 2001) to constrain the monophyly of each genus and outgroup family in the therevid supertree. If there was phylogenetic evidence that a genus was not monophyletic, those species that defied the monophyly were not constrained to belong to the genus: the letters NM were appended to their name for rapid recognition. Genera represented by a single taxon in the MRP matrix were denoted *Gen*. The command 'enforce constraint' was added to each parsimony ratchet search. PAUP* was used for producing strict, semistrict, Adams and majority-rule CCT. The Adams CCT was examined and compared to the strict and majority-rule CCT to identify rogue taxa that were removed from some analyses, as suggested by Kennedy and Page (2002). Bootstrap values were used as a measure of nodal support, and were calculated in PAUP* with 500 bootstrap replicates using two random-addition sequence-starting trees per replicate and saving no more than one tree of length >1 each replicate. TNT (Goloboff *et al.* 2003) was used to verify that the shortest trees were found in the parsimony ratchets, with 500 random-addition traditional searches with tree bisection–reconnection (TBR) saving no more than 10 trees/replicate, followed by three rounds of parsimony ratchet and drift of 1500 iterations with a random sample of 45 of the characters reweighted by 4 at each iteration and 100 rounds of tree fusing.

Results

The parsimony ratchet of the matrix (Additional File 1) found 6275 MPCT compatible with the constraint tree, 5415 different and 1108 MPCT of length 593 (Additional File 2 [*therevid1108CT363t.tre*]). TNT found 6429 trees of the same length in a short time but, despite more severe perturbation, was unable to find shorter trees. The strict CCT of the 1108 MPCT is poorly resolved, with 69% of possible nodes. A search of the therevid MRP matrix (Additional File 1) revealed that 299 terminal taxa were taxonomic equivalents. Removing these would significantly affect the analyses. However, this does explain why there is so little resolution in the strict CCT. The Adams consensus of the therevid MPCT is extremely well resolved. The majority-rule CCT is reasonably resolved with 88% of possible nodes, with low percentage values for many major nodes. The more resolved majority-rule CCT suggests that although there is some conflict among source trees, the low levels of taxon overlap and different taxon combinations are leading to large numbers of MPCT. This is because terminal taxa that cannot be aligned to other taxa can be placed equally parsimoniously almost anywhere on the tree.

The Adams CCT was compared to the strict and majority-rule CCT of the 1108 MPCT of length 593 (Additional File 2) to identify rogue taxa. Ten such taxa (Table 3) were found; six occurred in only a single study. Eight of the 10 are the sole representatives of their respective genera (as indicated by *Gen*: Table 3), providing little grouping information, and their position would not have been influenced by the constraint tree. To further complicate their placement in the supertree, four of those eight taxa were present in only one source tree, in which their placement was unresolved. Both *Hermannula lanata* Kröber and *Thereva mirabilis* Lyneborg are cases of conflict between source trees. *Incoxoverpa borealis* (Cole) is the only representative of *Incoxoverpa* Webb & Irwin, whose position is unresolved in one source tree, is an outgroup in a second, and the third and only resolved source tree is small with 10 terminal taxa. *Thereva nebulosa* Kröber, while not the only representative of

Table 3. Rogue taxa identified through comparison of Adams, strict and majority-rule consensus composite tree (CCT)

'Gen' indicates a terminal taxon that is the only representative of that genus included in any source tree, and thus in the supertree

Rogue taxa	Source trees	Resolved	Outgroup?
Acraspisoides helviarta Hill & Winterton *Gen*	Hill and Winterton (2004)	No	No
Aristothereva eversmanni Zaitzev *Gen*	Metz (2002) [Therevinae]	No	No
Genus Chile	Metz (2002) [Therevinae]	No	No
Genus S	Hill (2003)	Yes	No
Apenniverpa venezuela Webb	Metz (2002) [Therevinae] (as New Genus Venezuela)	No	No
Hermannula lanata Kröber *Gen*	Metz (2002) [Therevinae]	Yes	No
	Holston (2003) [Therevinae]	Yes	No
Incoxoverpa borealis (Cole) *Gen*	Webb and Irwin (1999)	Yes	No
	Metz (2002) [Therevinae]	No	No
	Webb and Metz (2003)	Yes	Yes
Manestella Metz *Gen*	Metz (2002) [Therevinae]	Yes	Yes
	Hill (2003)	Yes	No
Thereva mirabilis Lyneborg	Hill (2003)	Yes	No
	Holston (2003) [Therevinae]	Yes	No
	Holston (2003) [*Thereva*]	Yes	No
Thereva nebulosa Kröber	Gaimari and Irwin (2000*a*)	Yes	No

Thereva, was present in only one source tree, in which there were no other *Thereva*, so there was no information on its placement within the genus. *Manestella* Metz is monotypic and found in only two source trees, as an outgroup taxon in one and as part of a resolved clade in monophyly with another rogue taxon, Genus S, in the second.

The 10 rogue taxa were removed from subsequent analyses, as suggested by Kennedy and Page (2002). A revised constraint tree was constructed. MRP analysis of the 353 taxa and 493 matrix elements (Additional File 3 [*therevidmatrix353t.nex*]) for 10 000 random-addition heuristic searches with TBR and Multrees, restricted to 1 million rearrangements/replicate, found 2399 MPCT compatible with the constraint tree, at L = 589. The parsimony ratchet on the revised data matrix found 6275 MPCT, compatible with the constraint tree, 4793 different, and 1189 MPCT were of length 588 (Additional File 4 [*therevid1189CT353t.tre*]). The strict CCT of the 1189 MPCT is reasonably resolved, with 74% of possible nodes. Bootstrap analysis of this dataset found support greater than 50% for 37% of the nodes, although 11 of those nodes were not present in the majority-rule CCT. The majority-rule CCT is well resolved, with 91% of possible nodes appearing on the tree, though with low percentage values for some major nodes.

The majority-rule CCT of the 1189 MPCT for 353 terminal taxa (Additional File 4) is shown in five parts (Fig. 3*A–E*) with the percentage of MPCT containing the node included above the branch to indicate support (100% indicates that the node is present in all MPCT, and therefore the strict and semistrict CCT). Bootstrap support values greater than 50% are shown below the branches. The 10 rogue taxa have been added to the majority-rule CCT with broad grey branches and are indicated in bold type on the therevid supertree. The position of the rogue taxa corresponds to their position on the majority-rule CCT of the 1108 MPCT for 363 terminal taxa (including the outgroup) (Additional File 2).

Discussion

Therevid supertree

The supertree presented here is the only overarching phylogenetic hypothesis of the Therevidae, and summarises all past phylogenetic work on the family. Although low taxon overlap resulted in over 1100 MPCT and caused significant lack of resolution in the CCT for the Therevidae, we were able to retrieve the subfamilies Agapophytinae, Phycinae, Therevinae and Xestomyzinae. The Phycinae and Xestomyzinae form a clade, sister to the remaining Therevidae. The Australasian and South American *Taenogera* Kröber genus-group is monophyletic, and sister to a clade of Therevinae and the Australian endemic Agapophytinae. The Therevinae consists of the *Anabarhynchus* Macquart genus-group of Australian, South American, New Caledonian, and New Zealand taxa as sister to the non-Australasian 'higher Therevinae', which contains the tribes Cyclotelini and Therevini. The Therevini includes the *Hoplosathe* Lyneborg & Zaitzev, *Litolinga* Irwin & Lyneborg, *Baryphora* Loew, *Pandivirilia* Irwin & Lyneborg, and *Thereva* Latreille genus-groups.

The therevid supertree is shown in five parts (Fig. 3*A–E*): Part 1, Xestomyzinae, Phycinae, and *Taenogera* genus-group (Fig. 3*A*); Part 2, Agapophytinae (Fig. 3*B*); Part 3, part of the Therevinae including the Cyclotelini (Fig. 3*C*); Part 4, part of the Therevini, including *Thereva* (Fig. 3*D*); and Part 5, *Pandivirilia* genus-group (Fig. 3*E*). The biogeographical regions occupied are indicated beside the taxonomic names in Fig. 3*A–E*. All estimates and numbers reported for taxonomic groups in the discussion below follow summaries from the comprehensive World database of Therevidae, *Mandala* (Kampmeier *et al.* 2004).

The **Xestomyzinae** (Figs 1*B*, 3*A*) includes an estimated 57 species in 12 genera (one undescribed) but has been very poorly sampled in phylogenetic analyses to date. Two source trees (28S and EF1α, Yang *et al.* 2000) included a single representative of the group, *Henicomyia hubbardi* Coquillett. *Hemigephyra* Lyneborg was added by Hill (2003). One of the least understood parts of the Therevidae are the basal clades and their relationship to one another (Hauser 2005). The Xestomyzinae was considered to be a tribe of the Phycinae (Lyneborg 1976, 1980; Irwin and Lyneborg 1981); however, evidence for subfamily status has accumulated through taxonomic study, and several publications now refer to the subfamily (Irwin and Webb 1992; Hauser and Irwin 2005*b*, 2005*c*). While the therevid supertree does not incorporate this new evidence concerning the status of the two subfamilies, the Xestomyzinae is present in all MPCT, with low bootstrap support (Fig. 3*A*). Studies including a greater taxon sampling of both subfamilies, especially of the Xestomyzinae, indicate that the Xestomyzinae and Phycinae do not form a basal clade, as is found in the supertree; rather, the Phycinae forms a clade sister to the remaining Therevidae, which includes a monophyletic Xestomyzinae (Hauser 2005).

The **Phycinae** contains an estimated 173 species in 15 genera (one undescribed). The therevid supertree includes 11 species of the Phycinae from nine genera (Figs 1*F*, 3*A*). Nine of these species were included in the molecular analyses of Yang *et al.* (2000), and the therevid supertree reflects these original relationships. The Phycinae was found in all MPCT, but bootstrap support is below 50%. Nodes defining phycine genera were constrained in analyses and, consequently, were found in 100% of the MPCT. Other nodes, both within and among genera, were not constrained and when found in 100% MPCT are supported. All seven of the unconstrained nodes in the Phycinae were found in all MPCT, and five of those relationships receive higher bootstrap support (Fig. 3*A*).

Australian therevids in the therevid supertree are found in the *Taenogera* genus-group (Fig. 3*A*), the subfamily Agapophytinae (Fig. 3*B*), and the *Anabarhynchus* genus-group in the Therevinae (Fig. 3*C*).

The ***Taenogera*-group** (Figs 1*A*, *E*, 3*A*), containing an estimated 150 species, is predominately Australian, but includes many recognised but undescribed New Caledonian species, species currently in *Ectinorhynchus* Macquart from New Zealand, and several genera from South America (*Entesia* Oldroyd, *Pachyrrhiza* Philippi, *Melanothereva* Malloch, and one undescribed genus). The only described genus not in any source tree is *Pachyrrhiza*, and four undescribed *Taenogera*-group genera are also not represented in source trees. There are indications from unpublished molecular studies that the *Taenogera*-group includes two clades that do not form a monophyletic group, and disparate
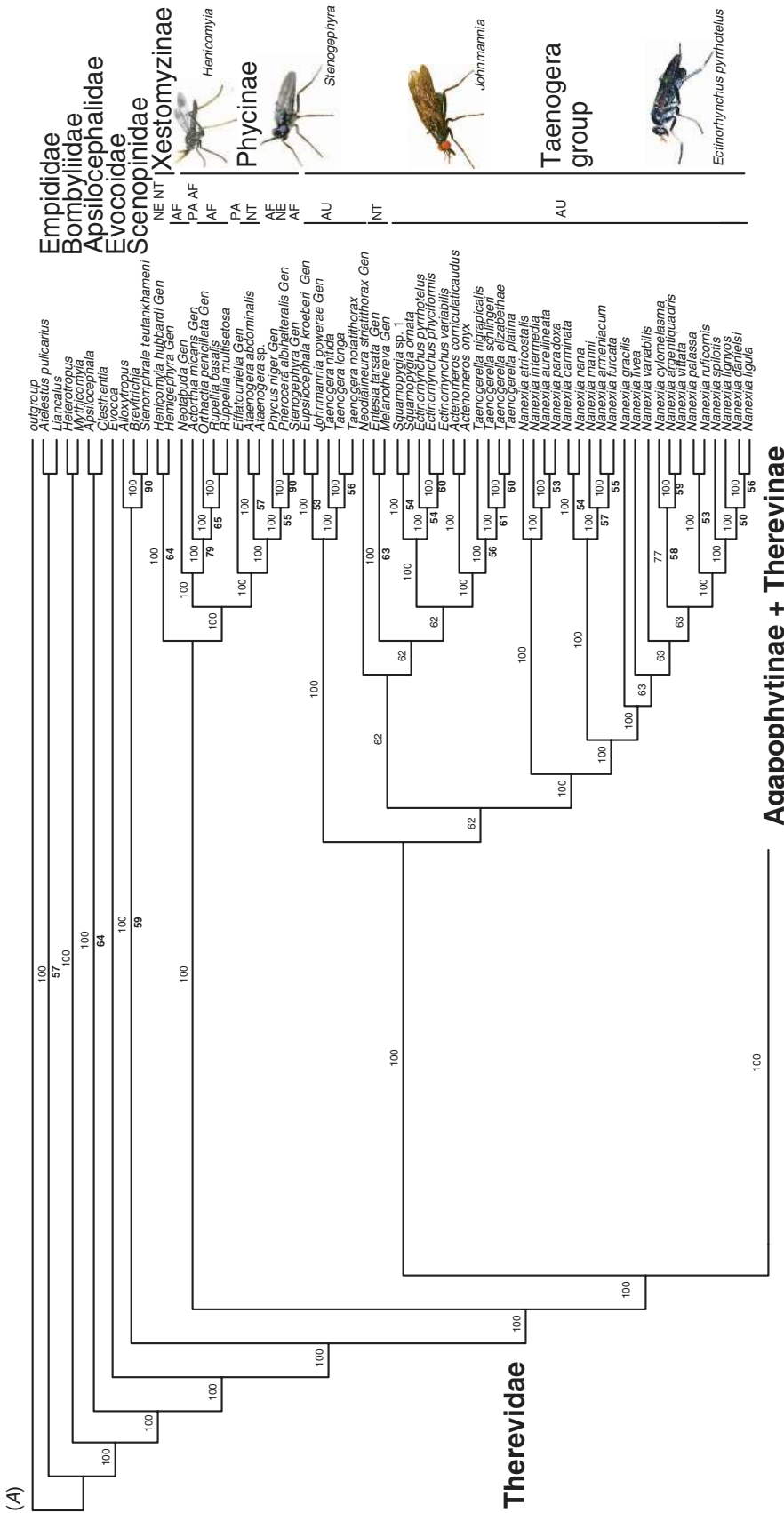
(A)

Empididae
Bombyliidae
Apsilocephalidae
Evocoidae
Scenopinidae
Xestomyzinae
Phycinae

Taenogera
group

Henicomyia
Stenogephyra
Johnmannia
Ectinorhynchus pyrrhotelus

outgroup
Atelestus pulicarius
Liancalus
Heterotropus
Mythicomyia
Apsilocephala
Clesthentia
Evocoa
Alloxytropus
Brevithrica
Stenomphrale teutankhameni
Henicomyia hubbardi Gen
Hemigephyra Gen
Neotabuda Gen
Actorthia micans Gen
Orthactia penicillata Gen
Ruppellia basalis
Ruppellia multisetosa
Effatouniella Gen
Ataenogera abdominalis
Ataenogera sp.
Phycus niger Gen
Pherocera labiflateralis Gen
Stenogephyra Gen
Eupsilocephala kroeberi Gen
Johnmannia powerae Gen
Taenogera nitida
Taenogera longa
Taenogera notatithorax
Neodialineura striatithorax Gen
Entesia tarsata Gen
Melanothereva Gen
Squamopygia sp. 1
Squamopygia ornata
Ectinorhynchus pyrrhotelus
Ectinorhynchus phyciformis
Ectinorhynchus variabilis
Actenomeros corniculaticaudus
Actenomeros onyx
Taenogerella nigrapicalis
Taenogerella schlingeri
Taenogerella elizabethae
Taenogerella platina
Nanexia atricostalis
Nanexia intermedia
Nanexia aurelineata
Nanexia paradoxa
Nanexia carminata
Nanexia nana
Nanexia manni
Nanexia armeniacum
Nanexia furcata
Nanexia gracilis
Nanexia livea
Nanexia variabilis
Nanexia cyclomelasma
Nanexia argentiquadris
Nanexia vitata
Nanexia palassa
Nanexia ruficornis
Nanexia spilotis
Nanexia lignyos
Nanexia danielsi
Nanexia ligula

NE NT
AF
PA AF
AF
PA
NT
AF
NE
AF

AU

NT

AU

Agapophytinae + Therevinae

Therevinae

Fig. 3. Therevid supertree. Relationships based on 50% majority-rule CCT with the percentage of the 1189 MPCT for 353 terminal taxa containing the node above the branch and bootstrap support values greater than 50% below the branch. 'Gen' indicates a terminal taxon that is the only representative of that genus included in any source tree, and thus in the supertree. Rogue taxa (indicated in bold type) have been added with broad grey branches in their position on the majority-rule CCT of the 1108 MPCT for 363 terminal taxa (including the outgroup). NM indicates that source trees provided phylogenetic evidence that terminal taxa were not part of a monophyletic group, and therefore were not constrained to be members of their respective genus in the MRP analysis. Biogeographical regions occupied by terminal taxa are indicated (AF, Africa; AU, Australasian; NC, New Caledonia; NE, Nearctic; NT, Neotropical; NZ, New Zealand; OR, Oriental; PA, Palaearctic). (*A*) **Part 1.** The Xestomyzinae, Phycinae, and largely Australasian *Taenogera* genus-group. Therevid images: Xestomyzinae, *Henicomyia* sp. by M. Metz; Phycinae, *Stenogephyra* sp. from South Africa by M. Hauser; *Taenogera* genus-group, *Johnmannia* sp. by C. Lambkin and *Ectinorhynchus pyrrhotelus* (Walker) by R. Nelson. (*B*) **Part 2.** The Australian endemic Agapophytinae. Therevid images: *Parapsilocephala* sp. and *Acupalpa semirufa* by C. Lambkin, *Agapophytus pallidicornis* by S. Winterton. (*C*) **Part 3.** Part of the Therevinae, including the *Anabarhynchus* genus-group and the tribe Cyclotelini. Therevid images: *Anabarhynchus* group, *Anabarhynchus kampmeierae* by S. Winterton. Cyclotelini, *Ozodiceromyia argentata* by T. Murray. (*D*) **Part 4.** Part of the Therevini containing the *Hoplosathe*, *Litolinga*, *Baryphora*, and *Thereva* generic-groups. Therevid images: *Baryphora* group *Baryphora speciosa* by N. Lartigau. *Thereva* group, *Thereva nobilitata* by P. Krogh. (*E*) **Part 5.** The *Pandivirilia* genus-group. Therevid images: *Penniverpa festina* (Coquillett) from USA by J. N. Dell, Georgia, USA; *Acrosathe novella* (Coquillett) from USA by G. Thompson, Queensland Museum, Australia; *Tabuda varia* (Walker) and *Pandivirilia albifrons* by T. Murray.
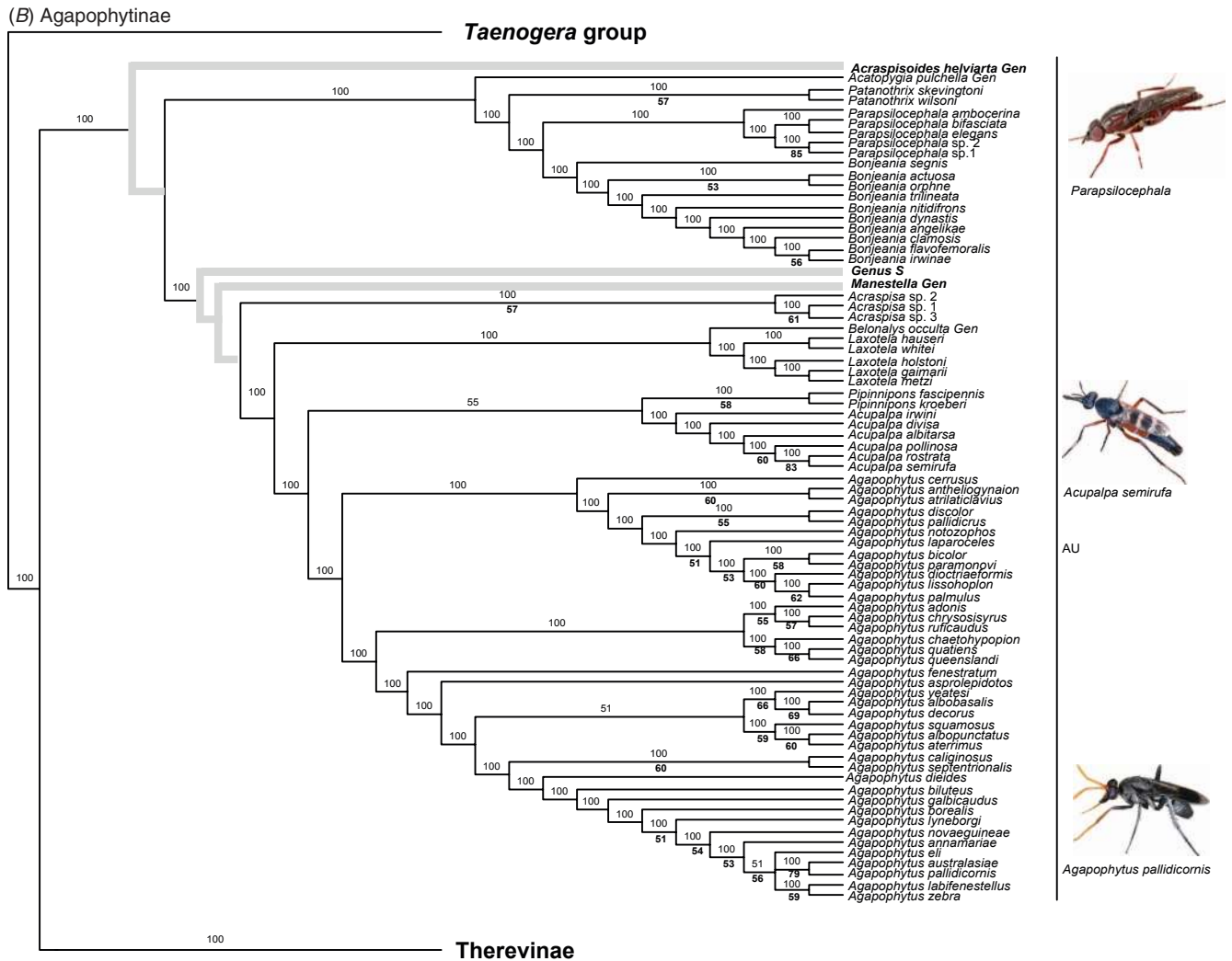
(*B*) Agapophytinae

Fig. 3.   (*continued*)

taxon sampling from those two groups in analyses may have provided the contradictory signals for *Taenogera*-group monophyly. However, a monophyletic *Taenogera*-group is supported by more than one source tree (Fig. 4*A–D*), and the therevid supertree indicates monophyly for the *Taenogera*-group in all MPCT, but bootstrap support for this node is below 50% (Fig. 3*A*). The MRP analysis (Fig. 3*A*) effectively summarises the current hypotheses for the *Taenogera*-group and implies that the group is monophyletic, sister to Agapophytinae plus Therevinae, and could be considered a separate subfamily. However, we do not propose a subfamily for the *Taenogera*-group because of hidden conflict in the supertree (see discussion below).

The **Agapophytinae** (Figs 1*C*, *D*, *G*, 3*B*), containing an estimated 340 species, with 12 genera, is also monophyletic. This subfamily is represented in the therevid supertree by 78 species in 13 genera (one undescribed) and appears in all MPCT although bootstrap support is below 50%, and placement of most agapophytine genera is uncontroversial (Fig. 3*B*). However, the basal placement of the rogue taxon (Table 3) *Manestella* Metz, and the undescribed Genus S to Agapophytinae in the supertree

has interesting implications for the morphological diagnosis of the subfamily. Metz (2002) included *Manestella* as an outgroup in his analysis of Therevinae, thus providing no information on its placement within Therevidae (Fig. 4*C*). *Manestella* (as Genus M) and Genus S were included in the molecular analysis of Therevidae by Hill (2003), where they were placed as sister to the Agapophytinae, represented by *Acraspisa* Kröber and *Agapophytus* Guérin (Fig. 4*B*). The inclusion of these genera within the Agapophytinae is contentious because they lack one of the subfamily's diagnostic characters, the velutum on the fore and hind femora (Winterton *et al.* 2001). The placement of *Manestella* as sister to the remaining Agapophytinae and Genus S within the *Taenogera*-group, however, has been supported by recent, unpublished molecular results. The questionable placement of these terminal taxa, and their designation as rogue taxa in the therevid supertree, is a consequence of few informative characters. Although some source trees (Winterton *et al.* 1999*b*; Yang *et al.* 2000: 28S) depict Agapophytinae as paraphyletic (Fig. 4*D*), or that it would be rendered paraphyletic by separation of the *Taenogera* genus-
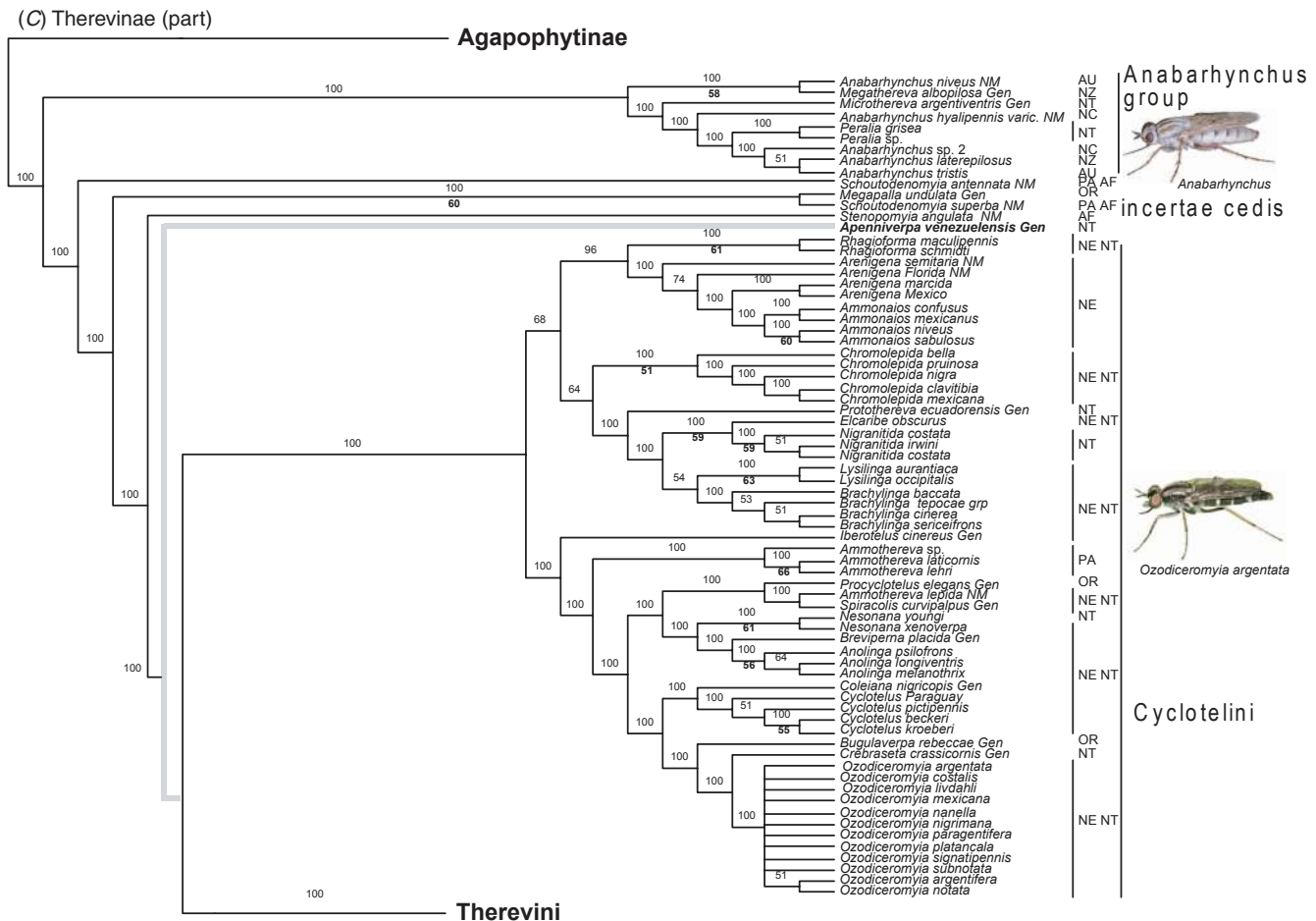
(*C*) Therevinae (part)

Agapophytinae

Anabarhynchus group

- *Anabarhynchus niveus* NM — AU
- *Megathereva albopilosa* Gen — NZ
- *Microthereva argentiventris* Gen — NT
- *Anabarhynchus hyalipennis varic.* NM — NC
- *Peralia grisea* — NT
- *Peralia* sp. —
- *Anabarhynchus* sp. 2 — NC
- *Anabarhynchus laterepilosus* — NZ
- *Anabarhynchus tristis* — AU

incertae cedis

- *Schoutedenomyia antennata* NM — PA AF
- *Megapalla undulata* Gen — OR
- *Schoutedenomyia superba* NM — PA AF
- *Stenopomyia angulata* NM — AF

Anabarhynchus

- ***Apenniverpa venezuelensis* Gen** — NT
- *Rhagioforma maculipennis* —
- *Rhagioforma schmidti* — NE NT
- *Arenigena semitaria* NM —
- *Arenigena Florida* NM —
- *Arenigena marcida* —
- *Arenigena Mexico* — NE
- *Ammonaios confusus* —
- *Ammonaios mexicanus* —
- *Ammonaios niveus* —
- *Ammonaios sabulosus* —
- *Chromolepida bella* —
- *Chromolepida pruinosa* — NE NT
- *Chromolepida nigra* —
- *Chromolepida clavitibia* —
- *Chromolepida mexicana* —
- *Prototereva ecuadorensis* Gen — NT
- *Elcaribe obscurus* — NE NT
- *Nigranitida costata* — NT
- *Nigranitida irwini* —
- *Nigranitida costata* —
- *Lysilinga aurantiaca* —
- *Lysilinga occipitalis* —
- *Brachylinga baccata* — NE NT
- *Brachylinga tepocae* grp —
- *Brachylinga cinerea* —
- *Brachylinga sericeifrons* —
- *Iberotelus cinereus* Gen —

Ozodiceromyia argentata

Cyclotelini

- *Ammothereva* sp. — PA
- *Ammothereva laticornis* —
- *Ammothereva lehri* —
- *Procyclotelus elegans* Gen — OR
- *Ammothereva lepida* NM — NE NT
- *Spiracolis curvipalpus* Gen —
- *Nesonana youngi* — NT
- *Nesonana xenoverpa* —
- *Breviperna placida* Gen —
- *Anolinga psilofrons* —
- *Anolinga longiventris* —
- *Anolinga melanothrix* — NE NT
- *Coleiana nigricopis* Gen —
- *Cyclotelus Paraguay* —
- *Cyclotelus pictipennis* —
- *Cyclotelus beckeri* —
- *Cyclotelus kroeberi* —
- *Bugulaverpa rebeccae* Gen — OR
- *Crebraseta crassicornis* Gen — NT
- *Ozodiceromyia argentata* —
- *Ozodiceromyia costalis* —
- *Ozodiceromyia livdahli* —
- *Ozodiceromyia mexicana* —
- *Ozodiceromyia nanella* —
- *Ozodiceromyia nigrimana* — NE NT
- *Ozodiceromyia paragentifera* —
- *Ozodiceromyia platancala* —
- *Ozodiceromyia signatipennis* —
- *Ozodiceromyia subnotata* —
- *Ozodiceromyia argentifera* —
- *Ozodiceromyia notata* —

Therevini

Fig. 3. (*continued*)

group (Fig. 4*E*), a monophyletic Agapophytinae is supported by a greater number of earlier analyses. This assessment is summarised in the therevid supertree (Fig. 3*B*).

The therevid supertree supports the monophyly of the **Therevinae** that is found in all MPCT, although bootstrap support is below 50% (Fig. 3*C–E*). Supertree analyses recovered two major clades of Therevinae, referred to in this paper as Cyclotelini and Therevini to preserve nomenclatural stability in this subfamily. 'Cyclotelini', as used in this study, extends the previous circumscription of this tribe but retains a monophyletic subclade status for the 10 original cycloteline genera (Gaimari and Irwin 2000*a*). 'Therevini' refers to a monophyletic group, found in all MPCT, that includes five generic-groups: *Baryphora*-, *Hoplosathe*-, *Litolinga*-, *Pandivirilia*-, and *Thereva*-groups. The Therevinae also contains the *Anabarhynchus* genus-group that has previously been referred to as 'lower' Therevinae to distinguish it from the diverse non-Australian clade of the remaining 'higher' Therevinae (Yang *et al*. 2000; Holston 2003; Holston *et al*. 2007). These qualifiers do not imply evolutionary significance for either therevine group.

The morphological analyses of Therevinae by Metz (2002) included several representatives of all genera, except four of the 11 monotypic therevine genera. These analyses, therefore,

supplied a backbone source tree for the Therevinae, despite being poorly resolved at many levels. The therevid supertree consequently lacks resolution within the Therevinae, and many of the nodes were found in a low proportion of the MPCT. In some cases, the lack of resolution within Therevinae results from contradictory source trees, but otherwise may be attributed to the lack of evidence for generic groupings in analyses by Metz (2002). For example, the rogue *Apenniverpa venezuela* Webb was included in only the source tree from Metz (2002), where it was sister to a clade formed by two taxa from the *Hoplosathe* genus-group, one from the *Litolinga* genus-group, and four from the *Baryphora* genus-group. The inclusion of other taxa in these generic-groups from other source trees resulted in an unresolved position of *A. venezuela* within the Therevinae in the supertree.

A group of five terminal taxa have *incertae cedis* status in Therevinae based on the supertree. Phylogenetic relationships in the therevid supertree for *Megapalla* Lyneborg and the non-monophyletic (NM) genera *Stenopomyia* Lyneborg and *Schoutedenomyia* Kröber reflect the topology of the source tree from the morphological analysis of the Therevinae by Metz (2002).

The ***Anabarhynchus* genus-group** (Figs 1*K*, 3*C*), including an estimated 190 species, has been poorly represented in phylogenetic analyses to date, but clearly includes the South
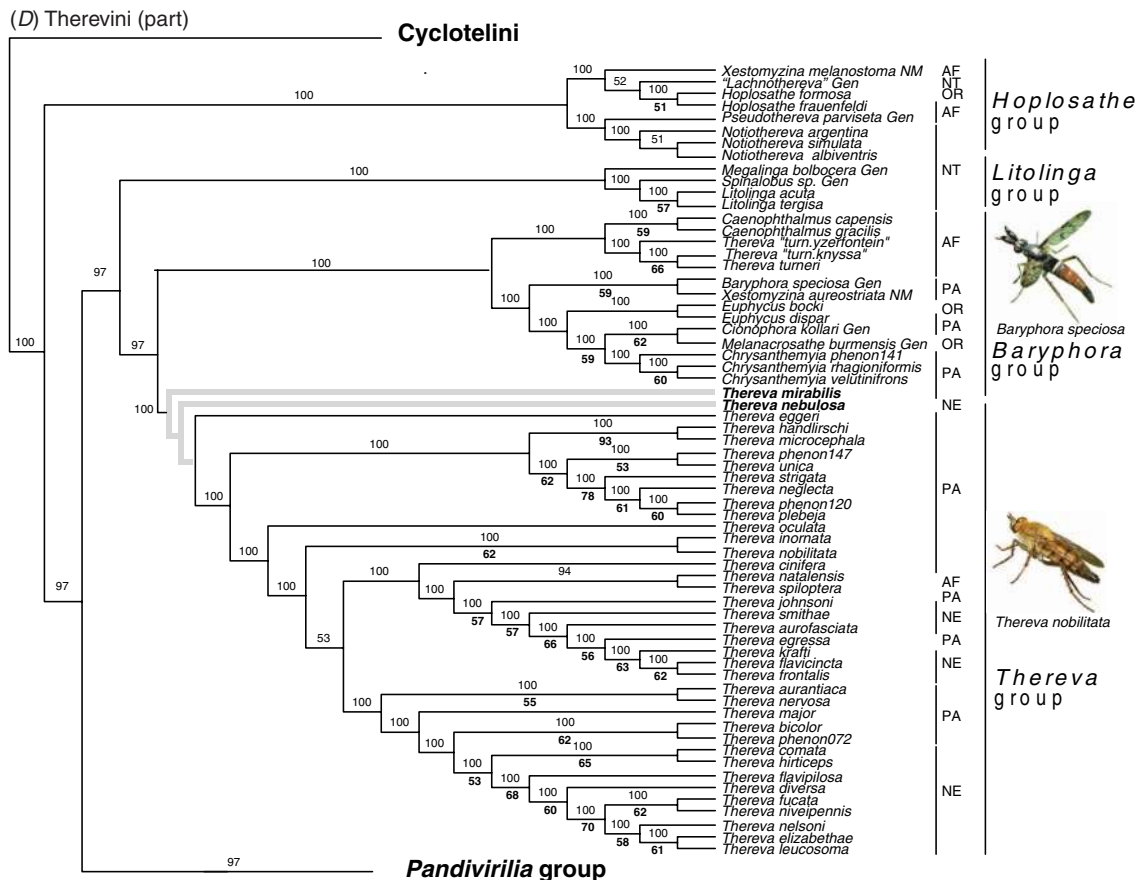
Fig. 3.   (*continued*)

American genera *Microthereva* Malloch and *Peralia* Malloch, the New Zealand genus *Megathereva* Lyneborg, and the Australasian genus *Anabarhynchus* Macquart. A single Australian species in this group, *Anabarhynchus tristis* Bigot, was included in several analyses (Yang *et al.* 2000; Winterton *et al.* 2001; Hill and Winterton 2004). Metz (2002) included seven taxa from South America, New Caledonia, and Australia in morphological analyses of Therevinae, and the relationships among these taxa in the therevid supertree essentially reflect his resulting topology. An additional New Caledonian *Anabarhynchus* was included by Hill (2003), and a second species of *Peralia* was added by Holston (2003). The *Anabarhynchus* genus-group was recovered as monophyletic in all MPCT. As in Metz (2002), the *Anabarhynchus* genus-group forms the sister clade to the remaining non-Australian Therevinae. The therevid supertree implies that the *Anabarhynchus* genus-group could be considered a separate tribe, the 'Anabarhynchini' in the Therevinae, or given separate subfamily status. Relationships in the therevid supertree, and taxonomic assessments of species-groups (Lyneborg 2001), suggest that *Anabarhynchus* is broadly paraphyletic, even with this limited sample of five taxa from Australia, New Caledonia, and New Zealand.

The tribe **Cyclotelini** (including *Penniverpa* Irwin & Lyneborg, and *Brachylinga* Irwin & Lyneborg and *Arenigena* Irwin & Lyneborg genus-groups) (Figs 1*L*, 3*C*) has been well

sampled in phylogenetic analyses, with 19 of the 22 genera and 58 of the estimated 280 species represented in the therevid supertree. Cyclotelini is monophyletic and found in all MPCT, and the relationships between the Cyclotelini and other therevine groups are clear in the therevid supertree, in contrast to results from previous studies (Gaimari and Irwin 2000*a*; Holston 2003). The therevid supertree reflects most of the internal relationships in the source tree for the Cyclotelini from Metz (2002), but includes an additional 24 taxa. Genera that were not monophyletic (NM) in either their source trees or the resulting supertree include *Brachylinga* (Metz 2002), *Ammothereva* Lyneborg (Metz 2002), and *Arenigena* (Metz 2002; Hauser and Irwin 2003). However, a recent revision (Webb and Metz 2006) placed *Brachylinga squamosa* (Hardy), the species that had rendered *Brachylinga* polyphyletic, into a new genus (*Elcaribe* Webb, composed of Caribbean endemics) and synonymysed it with *Elcaribe obscurus* (Coquillett). The monophyly of *Ammothereva* and *Arenigena* have not been evaluated in recent phylogenetic analyses of therevine genera.

The remaining Therevinae are monophyletic, found in all MPCT, and in this work are referred to as the '**Therevini**'. The 'Therevini' (Figs 1*H*, *J*, 3*D*, *E*) can be divided into five genus-groups: *Hoplosathe*-, *Litolinga*-, *Baryphora*-, *Thereva*-, and *Pandivirilia*-groups. Genera that are not monophyletic in either the Metz (2002) source trees or in the supertree are *Irwiniella*
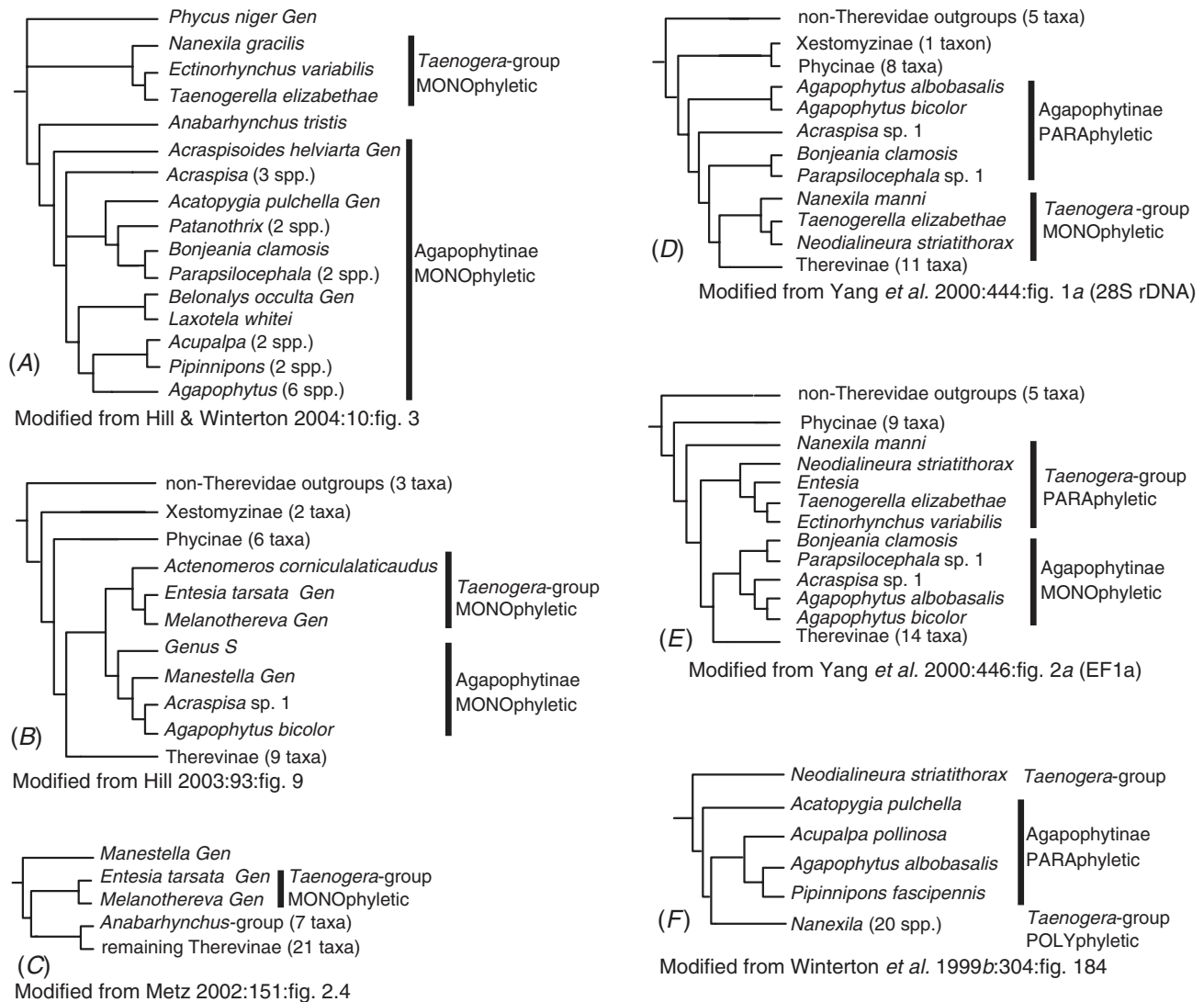
Fig. 3.   (*continued*)

Lyneborg, *Acrosathe* Irwin & Lyneborg, and *Stenopomyia* Lyneborg. While the 'Therevini' is found in all MPCT, the relationships between the genus-groups are found in only 97% of the MPCT, reflecting the lack of resolution in the Metz (2002) source tree. Holston *et al.* (2007) recently expanded the taxonomic sampling of the Holston (2003) source tree to 39 genera, with results from combined molecular analyses emphasising the lack of morphological evidence for relationships within 'Therevini' (Metz 2002). While the core circumscriptions of genus-groups indicated in the supertree were recovered, molecular data support different higher-level relationships, such as a sister group relationship between *Thereva* and a subclade of the *Pandivirilia*-group (i.e. *Nebritus* Coquillett, *Acrosathe*, and *Pandivirilia*) (Holston *et al.* 2007).

The therevid supertree provides a summary for the Therevini of the relationships among taxa from morphological analyses by Metz (2002) and the previous molecular analysis of genera in 'Therevini' by Holston (2003). The *Hoplosathe* genus-group in the therevid supertree, including eight of the estimated 18 species in three genera (one undescribed), is present in all MPCT. The *Litolinga* genus-group in the therevid supertree is monophyletic, and is found in all MPCT. The group in the analysis contains four of the estimated seven species, from three of the four genera (one undescribed). The *Baryphora* genus-group (Fig. 1*J*) includes 14 of an estimated 20 species, corresponds to the hypothesis of monophyly for the endemic Mediterranean genera suggested by Lyneborg (1983) and Metz (2002), and reflects the circumscription and resolution provided by the Holston (2003)

Fig. 4. Hidden conflict in supertree. The monophyly and non-monophyly of the *Taenogera*-group and Agapophytinae. Summary of six source trees. (*A*) Modified from Hill and Winterton (2004). (*B*) Modified from Hill (2003). (*C*) Modified from Metz (2002). (*D*) Modified from the 28S rDNA strict consensus of Yang *et al.* (2000). (*E*) Modified from the EF1α strict consensus of Yang *et al.* (2000). (*F*) Modified from Winterton *et al.* (1999*b*). 'Gen' indicates a terminal taxon that is the only representative of that genus included in the supertree.

analysis. In the therevid supertree, the *Baryphora* genus-group is monophyletic and fully resolved in all MPCT. *Xestomyzina* Kröber is not monophyletic and, as in the source tree from Metz (2002), members of the genus are found in the *Baryphora*- and *Hoplosathe* genus-groups. The *Thereva* genus-group (Fig. 1*H*) includes 38 of an estimated 180 species considered monophyletic in previous taxonomic and phylogenetic studies (Holston 2003; Holston and Irwin 2005; Holston *et al.* 2007). The monophyly of the group in the supertree analysis was predetermined by the constraint on the genus *Thereva* excluding the *turneri*-group and *Thereva analis* Kröber. The *Pandivirilia* genus-group (Figs 1*I*, 3*E*), which includes 87 of the estimated 190 species and 29 genera (one undescribed) in the therevid supertree, was found in 97% of MPCT. The limited resolution of the *Pandivirilia* genus-group, largely provided by the Holston (2003) molecular source tree, is due to topological differences with the phylogenies

from the other morphological analyses in which monophyly was assumed for restricted samples of *Pandivirilia*-group genera (Webb and Irwin 1999; Webb and Metz 2003). The conflict among the four source trees is also reflected in the lower percentage of MPCT that support the intergeneric relationships (Fig. 3*E*), and the difficulty in placing several rogue taxa in the *Pandivirilia* genus-group.

Biogeography

Endemism for recognised therevid genera includes 17 of 31 in the Palaearctic Region, 28 of 29 in the Australasian Region, 11 of 31 in the Nearctic Region, 25 of 35 in the Neotropical Region, and 24 of 31 in the Afrotropical Region. In the Afrotropical Region, there have been major radiations of Xestomyzinae and Phycinae, and Holarctic temperate zones

encompass major radiations of 'Therevini.' Most of the Australian fauna results from radiations in the endemic subfamily Agapophytinae, the *Taenogera* genus-group, and the therevine *Anabarhynchus* (Fig. 1*K*), the most species-rich therevid genus. The geographic distribution of terminal taxa in the therevid supertree is indicated in Fig. 3*A–E*.

The current distribution of 'lower' therevines can be explained by continental vicariance events, especially the breakup of Gondwanaland during the early Cretaceous (95 million years ago). The fossil therevid species *Rhagiophyrne bianalis* Rohdendorf, known from deposits in southern Kazakhastan (Mostovski 1998), is the oldest (152–158 million years ago) putative fossil therevid. This species, fossil remains of other lower brachycerous flies (Grimaldi and Engel 2005), and divergence estimates for lower brachyceran lineages (Wiegmann *et al.* 2003) suggest a late Jurassic origin for Therevidae. The high proportion of southern and south-west African endemics in Xestomyzinae and the occurrence of the xestomyzine genus *Henicomyia* Coquillett (Fig. 1*B*) in the New World, however, has been considered evidence for an early Cretaceous Gondwanan origin, after which xestomyzines dispersed from South America to North America over the Mesoamerican land bridge during the past four million years. A recently described fossil indicates that Xestomyzinae occurred in North America more than 20 million years ago (Hauser and Irwin 2005*b*), suggesting an earlier evolutionary link between the Neotropical and Nearctic faunas (Hauser 2005). A pre-Miocene barrier to latitudinal movement has previously been disputed (Marshall *et al.* 1979; Rowell and Flook 2004; Wüster *et al.* 2005). The phylogenetic hypothesis presented here also questions a barrier to latitudinal movement four million years ago in light of the presence of the Caribbean Cyclotelini genus *Elcaribe* Webb, and the *Pandivirilia* genus-group, and the monotypic taxon *Ambrodolon* Metz from Dominican Amber (mid-Miocene 15–20 million years ago) (Iturralde-Vinent and MacPhee 1996; Iturralde-Vinent 2001), which would have dispersed from either the northern or, more likely, the southern continent. *Ambrodolon* is the only higher therevid fossil, and also the most recent fossil. All other described purported therevid fossils were discovered to be Phycinae, Xestomycinae, or were not Therevidae (Hauser and Irwin 2005*a*, 2005*b*; Hauser 2007; Hauser and Fisher 2007). This might indicate a relatively young age for the higher Therevidae, but the absence of fossil records older than 20 million years only provides a minimal age for the group, not evidence that the group is not older. The classification of a Baltic amber (45 million years ago) fossil (*Arctogephyra* Hauser) as a xestomyzine (Hauser 2007) documented the presence of this subfamily in the Palaearctic region, even though this group is now restricted to southern Africa and the New World. This implies that the Xestomyzinae may have been much more widespread and not restricted to Gondwana in the past. Phycinae are found throughout the Afrotropical Region, where they have their highest generic diversity, the Palaearctic Region through southern Asia, the Nearctic region (mainly the south-west), and the Neotropical Region from north-eastern Brazil to the eastern edge of the Andes, with one endemic genus *Ataenogera* Kröber (Hauser and Webb 2007) from southern Mexico to Argentina. The distribution of the genus *Phycus* Walker indicates dispersal from Africa through Asia into the New World. The same scenario could be assumed for the other three North American genera, which are not found in the Neotropical part of Mexico. The presence of *Ataenogera*, which is endemic to the neotropics, could be explained by a Cretaceous trans-Atlantic dispersal (Pitman *et al.* 1993; Smith *et al.* 1994). Distribution patterns in the Australasian clades reflect a southern Gondwanan distribution, with radiations predating the separation of South America from southern temperate Cretaceous Australia and Antarctica (Smith *et al.* 1994). The *Taenogera* genus-group is the most diverse Gondwanan clade, encompassing a mix of endemic New Caledonian, Chilean, Australian, and one largely Australian genus with New Zealand species. The *Anabarhynchus* genus-group has its centre of diversity in Australia but also occurs in New Caledonia, New Guinea, Chile, and New Zealand.

Distribution patterns for Therevinae suggest a more complex series of dispersal and vicariance events for resolved areas of the phylogeny. The absence of endemic Australasian 'higher' therevine genera could have resulted from a Laurasian–Gondwanan vicariance, after which 'higher' therevine lineages dispersed into the Southern Hemisphere (Gaimari and Irwin 2000*a*); however, the presence of *Taenogera* genus-group genera and basal Cyclotelini in the Southern Hemisphere suggests that it is equally likely to have been northern dispersal (Metz 2002). The only quantitative cladistic biogeographical analysis (Rosen 1978; Nelson and Platnick 1981; Humphries and Parenti 1999) of a therevine clade was conducted for Cyclotelini (Gaimari and Irwin 2000*a*), in which vicariant events in the Western Hemisphere were linked to generic-level divergences. Considering the widely allopatric distributions of closely related cycloteline genera, and the similar biogeographical patterns in three cycloteline subclades, Gaimari and Irwin (2000*a*) also identified likely routes of transcontinental dispersal from Asia and within the Americas. Holston *et al.* (2007) focused on the high level of regional endemism in therevine groups, suggesting that climate restrictions may represent an important, uninvestigated issue in therevine biogeography. 'Lower' and 'higher' Therevinae are associated with southern temperate and northern tropical South America, respectively, and an ecological vicariant scenario accommodates their divergence in early Tertiary South America as well as extensive intercontinental dispersals of the 'higher' therevine lineages. Tropical lineages of 'higher' Therevinae may have dispersed into the boreotropical forest in the Northern Hemisphere (65–35 million years ago) with trans-American interchange across the Protocaribbean Archipelago (100–49 million years ago). Climate associations have been implicated in the distinct biogeographical separation between northern and southern South America (Sanmartín and Ronquist 2004) and may explain why Phycinae and Xestomyzinae, which are associated with tropical and subtropical climates, are absent from Chile and Australia despite their divergence before Gondwanan vicariance (Holston *et al.* 2007).

Hidden conflict in the therevid supertree

Although the monophyly of the *Taenogera*-group has been questioned (Winterton *et al.* 1999*a*; Winterton 2006), the therevid supertree indicates that the *Taenogera*-group is a monophyletic clade found in all MPCT (Fig. 3*A*). Independent

phylogenetic analyses including *Taenogera*-group genera have provided sparse and contradictory evidence concerning group monophyly, mainly due to insufficient taxon sampling of this group in larger phylogenetic analyses, as summarised in Fig. 4. Four studies, each including at most three *Taenogera*-group taxa, presented evidence for monophyly, shown in Fig. 4*A* (Hill and Winterton 2004), Fig. 4*B* (Hill 2003), Fig. 4*C* (Metz 2002), and Fig. 4*D* (Yang *et al*. 2000: 28S). Two analyses suggested that the *Taenogera*-group was not monophyletic, but paraphyletic as in Fig. 4*E* (Yang *et al*. 2000: EF1α), and polyphyletic as in Fig. 4*F* (Winterton *et al*. 1999*b*). However, this conflict is not apparent in the supertree. The *Taenogera*-group is found in 100% of the MPCT (Fig. 3*A*), and is present in the strict supertree because the MRP analysis searches for the shortest trees. With four source trees supporting monophyly, and two indicating non-monophyly, the more frequent result is a monophyletic *Taenogera*-group, found in all MPCT and thus in the strict supertree. The conflict is overridden by the method of analysis, and hidden completely. The conflict between source trees is reflected in the poor support for groupings, both within and between genera in the *Taenogera*-group. Of the 26 unconstrained nodes, nine are not found in all MPCT (Fig. 3*A*). The conflict is hidden completely by MRP analysis, which instead reflects the principal phylogenetic signal from source trees in building supertrees.

This problem is not limited to the *Taenogera*-group. The monophyly of the Agapophytinae has never been questioned, as all genera have elongate velutum patches on the ventral surfaces of the femora of the fore- and hind-legs, and nearly all have a velutum patch on the posteroventral surface of the male gonocoxites, characters that are not found elsewhere in the Therevidae (Winterton *et al*. 2001). Three source trees presented evidence for monophyly, shown in Fig. 4*A* (Hill and Winterton 2004), Fig. 4*B* (Hill 2003), and Fig. 4*E* (Yang *et al*. 2000: EF1α). Two source trees suggested that the Agapophytinae was paraphyletic if the *Taenogera*-group was recognised and separated, as in Fig. 4*D* (Yang *et al*. 2000: 28S) and Fig. 4*F* (Winterton *et al*. 1999*b*). With three source trees supporting monophyly, and two indicating paraphyly, the most parsimonious result is a monophyletic Agapophytinae, found in 100% MPCT (Fig. 3*B*), and thus in the strict supertree. Again, the conflict is hidden completely.

In situations where there is considerable conflict between source trees, MRP removes weak suboptimal signal, and represents only the principal signal from a dataset (Pisani and Wilkinson 2002) by resolving conflict in favour of the most frequent topology (Moore *et al*. 2006). Parsimonious analytical methods used in supertree construction may insinuate that there is no conflict between the source trees when it does exist. A node on a strict supertree does not mean that the node is present in all source trees. The node is only present in all *Most Parsimonious Component Trees*. As a summary, the supertree indicates only where the majority of opinion lies. Meta-analyses such as supertrees are supposed to point out where broad agreement exists, and where it does not (Purvis 1995*a*); however, it is clear that the supertree may hide conflict.

Conclusions

The supertree presented here is the only overarching phylogenetic hypothesis of the family Therevidae, and was produced using MRP methods, despite extremely low taxon overlap between source trees. While low taxon overlap increased the amount of missing data in the MRP matrix, the consequent increase in the number of MPCT was reduced by using a topological constraint. The parsimony ratchet was used to find all (putative) MPCT. Lack of resolution in the CCT was avoided by using Adams and majority-rule consensus techniques. The therevid supertree summarises all past phylogenetic work on the family. The subfamilies Agapophytinae, Phycinae, Therevinae, and Xestomyzinae were retrieved. The Phycinae and Xestomyzinae form a clade, sister to the remaining Therevidae. The Australasian and South American *Taenogera* genus-group is monophyletic, and sister to a clade of the Australian Agapophytinae and Therevinae. The Therevinae consists of the *Anabarhynchus* genus-group of Australian, South American, New Caledonian, and New Zealand taxa as sister to the non-Australasian 'higher Therevinae', which is further divided into the tribes Cyclotelini and Therevini. The Therevini includes the *Hoplosathe*, *Litolinga*, *Baryphora*, *Thereva*, and *Pandivirilia* genus-groups.

MRP supertree construction can produce supported nodes that hide conflict in the source trees. A node on a strict supertree does not mean that the node is present in all source trees although a node on a strict supertree must be present in all MPCTs. The supertree indicates the majority opinion for phylogenetic relationships from source trees, where broad agreement exists, but does not necessarily indicate where there is disagreement. These analyses of Therevidae include examples of hidden source tree conflict in MRP supertrees. Counts of the numbers of MPCT with and without the node will not reveal this conflict, and neither will bootstrapping by MRP element.

While supertrees combine phylogenies, supermatrices concatenate the data. Such supermatrices may be more directly, though equivalently compromised by data duplication than supertree matrices unless gene realignment or other editing removes redundant characters. The supermatrix approach requires the accumulation of primary biological data. For small or well studied groups, this remains a possibility. However, for huge groups potentially containing thousands of species (our example contains only 362, but ~1600 Therevidae have been described), scoring and accumulation of non-molecular homologous characters is not a viable option. Supertree analyses, by combining the phylogenies of smaller groups, even with minimal taxon overlap, can provide a realistic alternative for the construction of the 'Tree of Life'.

Author contributions

CLL designed the study, carried out the phylogenetic analyses, and drafted the manuscript; JWHT assisted with the design of the analytical methods. MAM, HNH, KCH and MH provided unpublished phylogenetic results for analysis. DWW, MH and KCH checked taxonomic names. KCH, MH, MAM and CLL wrote the section on biogeography. MAM, HNH, KCH, MH and the remaining authors assisted in the design of the study, collection of source trees, discussion of the results, and editing and revision of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

References

Adams, E. N. (1972). Consensus techniques and the comparison of taxonomic trees. *Systematic Zoology* **21**, 390–397. doi: 10.2307/2412432

Baum, B. R. (1992). Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**, 3–10. doi: 10.2307/1222480

Beck, R. M. D., Bininda-Emonds, O. R. P., Cardillo, M., Liu, F.-G. R., and Purvis, A. (2006). A higher-level MRP supertree of placental mammals. *BMC Evolutionary Biology* **6**, 1–14. doi: 10.1186/1471-2148-6-93

Bertelli, S., Giannini, N. P., and Goloboff, P. A. (2002). A phylogeny of the tinamous (Aves: Palaeognathiformes) based on integumentary characters. *Systematic Biology* **51**, 959–979. doi: 10.1080/10635150290102492

Beutel, R. G., Ribera, I., and Bininda-Emonds, O. R. P. (2008). A genus-level supertree of Adephaga (Coleoptera). *Organisms, Diversity & Evolution* **7**, 255–269. doi: 10.1016/j.ode.2006.05.003

Bininda-Emonds, O. R. P. (2000). Factors influencing phylogenetic inference: a case study using the mammalian carnivores. *Molecular Phylogenetics and Evolution* **16**, 113–126. doi: 10.1006/mpev.2000.0772

Bininda-Emonds, O. R. P. (2003). Novel versus unsupported clades: assessing the qualitative support for clades in MRP supertrees. *Systematic Biology* **52**, 839–848.

Bininda-Emonds, O. R. P., and Bryant, H. N. (1998). Properties of matrix representation with parsimony analyses. *Systematic Biology* **47**, 497–508.

Bininda-Emonds, O. R. P., and Sanderson, M. J. (2001). Assessment of the accuracy of matrix representation with parsimony analysis supertree construction. *Systematic Biology* **50**, 565–579. doi: 10.1080/10635150 1750435112

Bininda-Emonds, O. R. P., Gittleman, J. L., and Purvis, A. (1999). Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews of the Cambridge Philosophical Society* **74**, 143–175. doi: 10.1017/S0006323199005307

Bininda-Emonds, O. R. P., Gittleman, J. L., and Steel, M. A. (2002). The (super)tree of life: procedures, problems, and prospects. *Annual Review of Ecology and Systematics* **33**, 265–289. doi: 10.1146/annurev.ecolsys.33.010802.150511

Bininda-Emonds, O. R. P., Jones, K. E., Price, S. A., Grenyer, R., Cardillo, M., Habib, M., Purvis, A., and Gittleman, J. L. (2003). Supertrees are a necessary not-so-evil: a comment on Gatesy *et al.* *Systematic Biology* **52**, 724–729. doi: 10.1080/10635150390235647

Bininda-Emonds, O. R. P., Beck, R. M. D., and Purvis, A. (2005). Getting to the roots of matrix representation. *Systematic Biology* **54**, 668–672. doi: 10.1080/10635150590947113

Bininda-Emonds, O. R. P., Cardillo, M., Jones, K. E., MacPhee, R. D. E., Beck, R. M. D., Grenyer, R., Price, S. A., Vos, R. A., Gittleman, J. L., and Purvis, A. (2007). The delayed rise of present-day mammals. *Nature* **446**, 507–512. doi: 10.1038/nature05634

Burleigh, J. G., Driskell, A. C., and Sanderson, M. J. (2006). Supertree bootstrapping methods for assessing phylogenetic variation among genes in genome-scale data sets. *Systematic Biology* **55**, 426–440. doi: 10.1080/10635150500541722

Cardillo, M., Bininda-Emonds, O. R. P., Boakes, E., and Purvis, A. (2004). A species-level phylogenetic supertree of marsupials. *Journal of Zoology* **264**, 11–31. doi: 10.1017/S0952836904005539

Cavalcanti, M. J. (2007). A phylogenetic supertree of the hammerhead sharks (Carcharhiniformes: Sphyrnidae). *Zoological Studies* **46**, 6–11.

Colless, D. H., and McAlpine, D. K. (1991). Diptera. In 'The Insects of Australia: A Textbook for Students and Research Workers'. (Eds I. D. Naumann, P. B. Carne, J. F. Lawrence, E. S. Nielsen, J. P. Spradbery, R. W. Taylor, M. J. Whitten and M. J. Littlejohn.) pp. 717–786. (Melbourne University Press: Melbourne.)

Cuignet, M., Hance, T., and Windsor, D. M. (2007). Phylogenetic relationships of egg parasitoids (Hymenoptera: Eulophidae) and correlated life history characteristics of their neotropical Cassidinae hosts (Coleoptera, Chrysomelidae). *Molecular Phylogenetics and Evolution* **42**, 573–584. doi: 10.1016/j.ympev.2006.09.005

Davies, T. J., Barraclough, T. G., Chase, M. W., Soltis, P. S., Soltis, D. E., and Savolainen, V. (2004). Darwin's abominable mystery: insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 1904–1909. doi: 10.1073/pnas.0308127100

de Queiroz, A., and Gatesy, J. (2007). The supermatrix approach to systematics. *Trends in Ecology & Evolution* **22**, 34–41. doi: 10.1016/j.tree.2006.10.002

Eick, G. N., Jacobs, D. S., and Matthee, C. A. (2005). A nuclear DNA phylogenetic perspective on the evolution of echolocation and historical biogeography of extant bats (Chiroptera). *Molecular Biology and Evolution* **22**, 1869–1886. doi: 10.1093/molbev/msi180

English, K. (1950). Notes on the morphology and biology of *Anabarrhynchus fasciatus* Macq. and other Australian Therevidae. *Proceedings of the Linnean Society of New South Wales* **75**, 345–359.

Eulenstein, O., Chen, D., Burleigh, J. G., Fernandez-Baca, D., and Sanderson, M. J. (2004). Performance of flip supertree construction with a heuristic algorithm. *Systematic Biology* **53**, 299–308. doi: 10.1080/10635150490423719

Ferguson, D. J., and Lambkin, C. L. (2006). Behavioral observations of Australian stiletto flies from south-eastern New South Wales (Diptera: Therevidae). *Australian Entomologist* **33**, 123–132.

Fritsch, P. W., Cruz, B. C., Almeda, F., Wang, Y. G., and Shi, S. H. (2006). Phylogeny of *Symplocos* based on DNA sequences of the chloroplast trnC–trnD intergenic region. *Systematic Botany* **31**, 181–192. doi: 10.1600/036364406775971877

Fulton, T. L., and Strobeck, C. (2006). Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. *Molecular Phylogenetics and Evolution* **41**, 165–181. doi: 10.1016/j.ympev.2006.05.025

Gaimari, S. D., and Irwin, M. E. (2000*a*). Phylogeny, classification and biogeography of the cycloteline Therevinae (Insecta: Diptera: Therevidae). *Zoological Journal of the Linnean Society* **129**, 129–240. doi: 10.1111/j.1096-3642.2000.tb00012.x

Gaimari, S. D., and Irwin, M. E. (2000*b*). Revision of the *mexicana*-group of the cycloteline genus *Ozodiceromyia* (Diptera: Therevidae). *Proceedings of the Entomological Society of Washington* **102**, 561–600.

Gatesy, J., and Arctander, P. (2000). Hidden morphological support for the phylogenetic placement of *Pseudoryx nghetinhensis* with bovine bovids: a combined analysis of gross anatomical evidence and DNA sequences from five genes. *Systematic Biology* **49**, 515–538. doi: 10.1080/10635159950127376

Gatesy, J., and Baker, R. H. (2005). Hidden likelihood support in genomic data: can forty-five wrongs make a right? *Systematic Biology* **54**, 483–492. doi: 10.1080/10635150590945368

Gatesy, J., Mathee, C., DeSalle, R., and Hayashi, C. (2002). Resolution of a supertree/supermatrix paradox. *Systematic Biology* **51**, 652–664. doi: 10.1080/10635150290102311

Gatesy, J., Amato, G., Norell, M., DeSalle, R., and Hayashi, C. (2003). Combined support for wholesale taxic atavism in gavialine crocodylians. *Systematic Biology* **52**, 403–422. doi: 10.1080/10635150390197037

Gatesy, J., Baker, R. H., and Hayashi, C. (2004). Inconsistencies in arguments for the supertree approach: supermatrices versus supertrees of Crocodylia. *Systematic Biology* **53**, 342–355. doi: 10.1080/10635150490423971

Goloboff, P. A., Farris, J. S., and Nixon, K. C. (2003). TNT: Tree Analysis Using New Technology. Program and documentation, available from the authors and at www.zmuc.dk/public/phylogeny [Verified May 2009].

Grenyer, R., and Purvis, A. (2003). A composite species-level phylogeny of the 'Insectivora' (Mammalia: Order Lipotyphla Haeckel, 1866). *Journal of the Zoological Society of London* **260**, 245–257. doi: 10.1017/S0952836903003716

Grimaldi, D., and Engel, M. S. (2005). 'Evolution of the Insects.' (Cambridge University Press: Cambridge.)

Grotkopp, E., Nek, M. R., Sanderson, M. J., and Rost, T. L. (2004). Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses. *Evolution* **58**, 1705–1729.

Hauser, M. (2005). Systematics and evolution of the basal lineages of Therevidae (Insecta: Diptera). Ph.D. Thesis. University of Illinois, Urbana–Champaign, IL.

Hauser, M. (2007). Baltic amber Therevidae and Apsilocephalidae (Diptera). *Studia Dipterologica* **14**, 37–59.

Hauser, M., and Fisher, E. M. (2007). The identity of the fossil *Psilocephala tarsalis* (Statz, 1940) (Diptera: Asiloidea). *Zootaxa* **1465**, 65–68.

Hauser, M., and Irwin, M. E. (2003). The Nearctic genus *Ammonaios* Irwin and Lyneborg 1981 (Diptera: Therevidae). *Annals of the Entomological Society of America* **96**, 738–765. doi: 10.1603/0013-8746(2003)096[0738:TNGAIA]2.0.CO;2

Hauser, M., and Irwin, M. E. (2005*a*). Fossil Therevidae (Insecta: Diptera) from Florissant, Colorado (upper Eocene). *Journal of Systematic Palaeontology* **3**, 393–401.

Hauser, M., and Irwin, M. E. (2005*b*). A new remarkable Xestomyzinae (Insecta, Diptera, Therevidae) genus from Mexican Amber. *Zootaxa* **1008**, 39–45.

Hauser, M., and Irwin, M. E. (2005*c*). The subfamily Xestomyzinae (Diptera: Therevidae) new to Madagascar, with description of four new species. *African Invertebrates* **46**, 181–202.

Hauser, M., and Webb, D. W. (2007). A revision of the New World stiletto fly genus *Ataenogera* Kröber (Diptera: Therevidae: Phycinae) with description of two new species. *Zootaxa* **1530**, 41–67.

Hill, H. N. (2003). Phylogenetic utility of two new nuclear genes, Opsin and CAD, within the stiletto flies (Diptera: Therevidae). M.Sc. Thesis. North Carolina State University, Raleigh, NC.

Hill, H. N., and Winterton, S. L. (2004). *Acraspisoides* gen. nov. (Diptera: Therevidae: Agapophytinae): a new genus of stiletto flies from Australia. *Zootaxa* **414**, 1–15.

Hill, R. (2005). Integration of morphological data sets for phylogenetic analysis of amniota: the importance of integumentary characters and increased taxonomic sampling. *Systematic Biology* **54**, 530–547. doi: 10.1080/10635150590950326

Holston, K. C. (2003). A phylogenetic assessment of Thereva Latreille (Insecta: Diptera: Therevidae). Ph.D. Thesis. University of Illinois, Urbana–Champaign, IL.

Holston, K. C., and Irwin, M. E. (2005). Revision of the Nearctic *Thereva* (Diptera: Asiloidea: Therevidae). *Studia Dipterologica* **13**(Supplement), 1–219.

Holston, K. C., Irwin, M. E., and Wiegmann, M. E. (2007). Monophyly and phylogenetic relationships of *Thereva* and therevine genus groups (Insecta: Diptera: Therevidae) based on EF-1a, 28S rDNA, and mitochondrial 16S rDNA sequences. *Invertebrate Systematics* **21**, 279–296. doi: 10.1071/IS06005

Hughes, J., and Vogler, A. P. (2004). The phylogeny of acorn weevils (genus *Curculio*) from mitochondrial and nuclear DNA sequences: the problem of incomplete data. *Molecular Phylogenetics and Evolution* **32**, 601–615. doi: 10.1016/j.ympev.2004.02.007

Hughes, J., Longhorn, S. J., Papadopoulou, A., Theodorides, K., de Riva, A., Mejia-Chang, M., Foster, P. G., and Vogler, A. P. (2006). Dense taxonomic EST sampling and its applications for molecular systematics of the Coleoptera (beetles). *Molecular Biology and Evolution* **23**, 268–278. doi: 10.1093/molbev/msj041

Humphries, C. J., and Parenti, L. R. (1999). 'Cladistic Biogeography.' (Oxford University Press: Oxford.)

Irwin, M. E., and Lyneborg, L. (1981). Therevidae. In 'Manual of Nearctic Diptera'. (Eds J. F. McAlpine, B. V. Peterson, G. E. Shewell, H. J. Teskey, J. R. Vockeroth and D. M. Wood.) pp. 513–523. (Research Branch, Agriculture Canada: Ottawa.)

Irwin, M. E., and Webb, D. W. (1992). Brasilian Therevidae (Diptera): a checklist and description of species. *Acta Amazonica* **21**, 85–121.

Iturralde-Vinent, M. A. (2001). Geology of the amber-bearing deposits of the Greater Antilles. *Caribbean Journal of Science* **37**, 141–167.

Iturralde-Vinent, M. A., and MacPhee, R. D. E. (1996). Age and paleogeographical origin of Dominican Amber. *Science* **273**, 1850–1852. doi: 10.1126/science.273.5283.1850

Jones, K. E., Purvis, A., MacLarnon, A., Bininda-Emonds, O. R. P., and Simmons, N. B. (2002). A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biological Reviews of the Cambridge Philosophical Society* **77**, 223–259. doi: 10.1017/S1464793101005899

Kampmeier, G. E., and Irwin, M. E. (2009). Meeting the interrelated challenges of tracking specimen, nomenclature, and literature data in Mandala. In 'Diptera Diversity: Status, Challenges and Tools'. (Eds T. Pape, R. Meier and D. Bickel.) pp. 406–435. (Brill Academic Publishers: Leiden.)

Kampmeier, G. E., Irwin, M. E., and Algmin, K. (2004). Mandala 5.3: a database system for systematics & biodiversity studies. Diptera Data Dissemination Disk II. (CD-ROM). (North American Dipterists Society.) http://www.inhs.illinois.edu/research/mandala/

Kearney, M. (2002). Fragmentary taxa, missing data, and ambiguity: mistaken assumptions and conclusions. *Systematic Biology* **51**, 369–381. doi: 10.1080/10635150252899824

Kelly, C. M. R., Barker, N. P., and Villet, M. H. (2003). Phylogenetics of advanced snakes (Caenophidia) based on four mitochondrial genes. *Systematic Biology* **52**, 439–459. doi: 10.1080/10635150390218132

Kennedy, M., and Page, R. D. M. (2002). Seabird supertrees: combining partial estimates of Procellariiform phylogeny. *The Auk* **119**, 88–108. doi: 10.1642/0004-8038(2002)119[0088:SSCPEO]2.0.CO;2

Kluge, A. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* **38**, 7–25. doi: 10.2307/2992432

Lambkin, C. L., Power, N., and Starick, N. T. (2002). Mt Kosciuszko: biodiversity of flies (Diptera). In 'Biodiversity in the Snowy Mountains'. (Ed. K. Green.) pp. 76–87. (Australian Institute of Alpine Studies: Jindabyne, NSW.)

Lyneborg, L. (1976). A revision of the therevine stiletto-flies (Diptera: Therevidae) of the Ethiopian region. *Bulletin of the British Museum of Natural History (Entomology)* **33**, 191–346.

Lyneborg, L. (1980). Family Therevidae. In 'Catalogue of the Diptera of the Afrotropical Region'. (Eds R. W. Crosskey, B. H. Cogan, P. Freeman, A. C. Pont, K. J. V. Smith and H. Oldroyd.) pp. 314–320. (British Museum (Natural History): London.)

Lyneborg, L. (1983). The west Mediterranean genus *Chrysanthemyia* Becker (Insecta, Diptera, Therevidae). *Steenstrupia (Copenhagen)* **9**, 77–82.

Lyneborg, L. (1992). Therevidae (Insecta: Diptera). In 'Fauna of New Zealand/Ko te Aitanga Pepeke o Aotearoa 24'. pp. 1–139. (DSIR Plant Protection/Te Wahanga Manaaki Tupu: Auckland.)

Lyneborg, L. (2001). 'The Australian stiletto-flies of the *Anabarhynchus* genus-group (Diptera: Therevidae).' (Apollo Books: Stenstrup, Denmark.)

Maddison, D. R., and Maddison, W. P. (2001). 'MacClade 4. Analysis of Phylogeny and Character Evolution. Version 3.' (Sinauer Associates, Inc.: Sunderland, MA.)

Mahon, A. (2004). A molecular supertree of the Artiodactyla. In 'Phylogenetic Supertrees: Combining Information to Reveal the TRENDS in Ecology and Evolution of Life'. (Ed. O. R. P. Bininda-Emonds.) pp. 411–437. (Kluwer Academic.)

Malia, M. J., Lipscomb, D. L., and Allard, M. W. (2003). The misleading effects of composite taxa in supermatrices. *Molecular Phylogenetics and Evolution* **27**, 522–527. doi: 10.1016/S1055-7903(03)00020-4

Marshall, L. G., Butler, R. F., Drake, R. E., Curtis, G. H., and Tedford, R. H. (1979). Calibration of the Great American Interchange. *Science* **204**, 272–279. doi: 10.1126/science.204.4390.272

Matthee, C. A., Jansen van Vuuren, B., Bell, D., and Robinson, T. J. (2004). A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. *Systematic Biology* **53**, 433–447. doi: 10.1080/10635150490445715

Metz, M. A. (2002). Systematics within the Therevinae (Insecta: Diptera: Therevidae). Ph.D. Thesis. University of Illinois, Urbana–Champaign, IL.

Metz, M. A., and Irwin, M. E. (2000). Revision of *Lindneria* Kröber, with the description of two new genera *Insulatitan* and *Ambradolon*, a fossil from Dominican amber (Diptera: Therevidae: Therevinae). *Annals of the Entomological Society of America* **93**, 977–1018. doi: 10.1603/0013-8746(2000)093[0977:ROLKBW]2.0.CO;2

Metz, M. A., and Webb, D. W. (2003). *Distostylus* gen. nov., a monotypic therevine genus (Diptera: Asiloidea: Therevidae) from the Caribbean island of Dominica. *Zootaxa* **222**, 1–12.

Moore, B. R., Smith, S. A., and Donoghue, M. J. (2006). Increasing data transparency and estimating phylogenetic uncertainty in supertrees: approaches using nonparametric bootstrapping. *Systematic Biology* **55**, 662–676. doi: 10.1080/10635150600920693

Mostovski, M. B. (1998). A brief review of brachycerous flies (Diptera, Brachycera) in the Mesozoic, with descriptions of some curious taxa. In 'Proceedings of the First International Palaeoentomological Conference, Moscow'. pp. 103–110.

Munshi-South, J., and Wilkinson, G. S. (2006). Diet influences life span in parrots (Psittaciformes). *The Auk* **123**, 108–118. doi: 10.1642/0004-8038(2006)123[0108:DILSIP]2.0.CO;2

Nelson, G., and Platnick, N. (1981). 'Systematics and Biogeography: Cladistics and Vicariance.' (Columbia University Press: New York.)

Nixon, K. C. (1999). The Parsimony Ratchet, a new method for rapid parsimony analysis. *Cladistics* **15**, 407–414. doi: 10.1111/j.1096-0031.1999.tb00277.x

Nixon, K. C., and Carpenter, J. M. (1996). On simultaneous analysis. *Cladistics* **12**, 221–241. doi: 10.1111/j.1096-0031.1996.tb00010.x

Novacek, M. J. (1992). Fossils, topologies, missing data, and the higher level phylogeny of Eutherian mammals. *Systematic Biology* **41**, 58–73.

Page, R. D. M. (2002). Modified mincut supertrees. In 'Proceedings of the Second Workshop on Algorithms in Bioinformatics (WABI2002) (Lecture Notes in Computer Science 2452)'. pp. 1–14. (Rome, Italy.)

Page, R. D. M. (2004). Taxonomy, supertrees, and the tree of life. In 'Phylogenetic Supertrees: Combining Information to Reveal the TRENDS in Ecology and Evolution of Life'. (Ed. O. R. P. Bininda-Emonds.) pp. 247–265. (Kluwer Academic: Dordrecht.)

Pisani, D., and Wilkinson, M. (2002). Matrix representation with parsimony, taxonomic congruence, and total evidence. *Systematic Biology* **51**, 151–155. doi: 10.1080/106351502753475925

Pitman, W. C., Cande, S., LaBrecque, J., and Pindell, J. (1993). Fragmentation of Gondwana: The Separation of Africa from South America. In 'Biological Relationships between Africa and South America'. (Ed. P. Goldblatt.) pp. 15–34. (Yale University Press: New Haven, CT.)

Purvis, A. (1995*a*). A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **348**, 405–421. doi: 10.1098/rstb.1995.0078

Purvis, A. (1995*b*). A modification to Baum and Ragan's method for combining phylogenetic trees. *Systematic Biology* **44**, 251–255.

Ragan, M. A. (1992). Phylogenetic inference based on matrix representation of trees. *Molecular Phylogenetics and Evolution* **1**, 53–58. doi: 10.1016/1055-7903(92)90035-F

Robinson, T. J., and Matthee, C. A. (2005). Phylogeny and evolutionary origins of the Leporidae: a review of cytogenetics, molecular analyses and a supermatrix analysis. *Mammal Review* **35**, 231–247. doi: 10.1111/j.1365-2907.2005.00073.x

Ronquist, F. (1996). Matrix representation of trees, redundancy, and weighting. *Systematic Biology* **45**, 247–253. doi: 10.2307/2413617

Rosen, D. E. (1978). Vicariant patterns and historical explanation in biogeography. *Systematic Zoology* **27**, 159–188. doi: 10.2307/2412970

Rowell, C. H. F., and Flook, P. K. (2004). A dated molecular phylogeny of the Proctolabinae (Orthoptera, Acrididae), especially the Lithoscirtae, and the evolution of their adaptive traits and present biogeography. *Journal of Orthoptera Research* **13**, 35–56. doi: 10.1665/1082-6467(2004)013[0035:ADMPOT]2.0.CO;2

Ruta, M. (2003). A species-level supertree for stylophoran echinoderms. *Acta Palaeontologica Polonica* **48**, 559–568.

Ruta, M., Jeffery, E. J., and Coates, M. I. (2003). A supertree of early tetrapods. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **270**, 2507–2516. doi: 10.1098/rspb.2003.2524

Salamin, N., and Davies, T. J. (2004). Using supertrees to investigate species richness in grasses and flowering plants. In 'Phylogenetic Supertrees: Combining Information to Reveal the TRENDS in Ecology and Evolution of Life'. (Ed. O. R. P. Bininda-Emonds.) pp. 461–486. (Kluwer Academic: Dordrecht.)

Salamin, N., Hodkinson, T. R., and Savolainen, V. (2002). Building supertrees: an empirical assessment using the grass family (Poaceae). *Systematic Biology* **51**, 136–150. doi: 10.1080/106351502753475916

Sanderson, M., Purvis, A., and Henze, C. (1998). Phylogenetic supertrees: assembling the trees of life. *Trends in Ecology & Evolution* **13**, 105–109. doi: 10.1016/S0169-5347(97)01242-1

Sanmartín, I., and Ronquist, F. (2004). Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* **53**, 216–243. doi: 10.1080/10635150490423430

Sikes, D. S., and Lewis, P. O. (2001*a*). PAUPRat: PAUP* implementation of the parsimony ratchet. Beta software. Version 1. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT.)

Sikes, D. S., and Lewis, P. O. (2001*b*). Software manual for PAUPRat: A tool to implement parsimony ratchet searches using PAUP*. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT. Available at http://www.ucalgary.ca/~dsikes/software2.htm [Verified May 2009].

Smith, A. G., Smith, D., and Funnell, B. M. (1994). 'Atlas of Mesozoic and Cenozoic Coastlines.' (Cambridge University Press: Cambridge.)

Springer, M. S., and de Jong, W. W. (2001). Which mammalian supertree to bark up? *Science* **291**, 1709–1711. doi: 10.1126/science.1059434

Stoner, C. J., Bininda-Emonds, O. R. P., and Caro, T. (2003). The adaptive significance of coloration in lagomorphs. *Biological Journal of the Linnean Society. Linnean Society of London* **79**, 309–328. doi: 10.1046/j.1095-8312.2003.00190.x

Summers, K., McKeon, C. S., and Heying, H. (2006). The evolution of parental care and egg size: a comparative analysis in frogs. *Proceedings of the Royal Society B: Biological Sciences* **273**, 687–692.

Swofford, D. (2002). 'PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b.10.' (Sinauer Associates: Sunderland, MA.)

Thomas, G. H., Wills, M. A., and Székely, T. (2004). A supertree approach to shorebird phylogeny. *BMC Evolutionary Biology* **4**, 28. doi: 10.1186/1471-2148-4-28

van der Linde, K., and Houle, D. (2008). A supertree analysis and literature review of the genus *Drosophila* and closely related genera (Diptera: Drosophilidae). *Insect Systematics & Evolution* **39**, 241–267.

Webb, D. W. (2007). A new genus and new species of Nearctic Therevidae (Insecta: Diptera) from southern New Mexico. *Zootaxa* **1495**, 41–46.

Webb, D. W., and Irwin, M. E. (1991). A revision of the Nearctic species of *Dialineura* Rondani and *Pallicephala* Irwin and Lyneborg (Diptera: Therevidae: Therevinae). *Proceedings of the Entomological Society of Washington* **93**, 868–898.

Webb, D. W., and Irwin, M. E. (1995). The new world genus *Chromolepida* Cole (Diptera, Therevidae, Therevinae). *Proceedings of the Entomological Society of Washington* **97**, 197–224.

Webb, D. W., and Irwin, M. E. (1999). Revision of *Tabuda* Walker and *Tabudamima* Irwin and Lyneborg, with the description of a new genus *Incoxoverpa* Webb and Irwin (Diptera: Therevidae: Therevinae). *Annals of the Entomological Society of America* **92**, 644–674.

Webb, D. W., and Metz, M. A. (2003). The Nearctic species of *Pandivirilia* Irwin and Lyneborg (Diptera: Therevidae: Therevinae). *Annals of the Entomological Society of America* **96**, 369–402. doi: 10.1603/0013-8746(2003)096[0369:TNSOPI]2.0.CO;2

Webb, D. W., and Metz, M. A. (2006). A revision of the New World genera *Brachylinga* Irwin and Lyneborg and *Lysilinga* Irwin and Lyneborg (Diptera: Therevidae: Therevinae) with the description of a new genus, *Elcaribe* Webb. *Zootaxa* **1288**, 1–241.

Webb, D. W., and Metz, M. A. (2008). A revision of the New World genus *Penniverpa* Irwin and Lyneborg (Diptera: Therevidae: Therevinae). *Zootaxa* **1720**, 1–45.

Wiegmann, B. M., Yeates, D. K., Thorne, J. L., and Kishino, H. (2003). Time flies, a new molecular time-scale for brachyceran fly evolution without a clock. *Systematic Biology* **52**, 745–756.

Wiens, J. J. (2003). Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology* **52**, 528–538. doi: 10.1080/10635150390218330

Wiens, J. J. (2006). Missing data and the design of phylogenetic analyses. *Journal of Biomedical Informatics* **39**, 34–42. doi: 10.1016/j.jbi.2005.04.001

Wiens, J. J., and Reeder, T. W. (1995). Combining data sets with different number of taxa for phylogenetic analysis. *Systematic Biology* **44**, 548–558.

Wilkinson, M. (1995). Coping with abundant missing entries in phylogenetic inference using parsimony. *Systematic Biology* **44**, 501–514.

Wilkinson, M., Cotton, J. A., and Thorley, J. L. (2004). The information content of trees and their matrix representations. *Systematic Biology* **53**, 989–1001. doi: 10.1080/10635150490522737

Winterton, S. L. (2000). Phylogenetic revision of *Acupalpa* Kröber (Diptera: Therevidae). *Insect Systematics & Evolution* **31**, 225–240.

Winterton, S. L. (2006). New species of *Eupsilocephala* Kröber from Australia (Diptera: Therevidae). *Zootaxa* **1372**, 17–25.

Winterton, S. L. (2007a). New species of *Laxotela* Winterton & Irwin from Australia (Diptera: Therevidae: Agapophytinae). *Zootaxa* **1407**, 57–64.

Winterton, S. L. (2007b). New species of *Nanexila* Winterton & Irwin and *Taenogera* Kröber from Australia (Diptera: Therevidae). *Zootaxa* **1413**, 55–64.

Winterton, S. L., and Irwin, M. E. (1999). *Laxotela* – a new genus of Therevidae (Diptera) from Australia. *Entomologica Scandinavica* **30**, 299–310.

Winterton, S. L., and Irwin, M. E. (2001). Phylogenetic revision of *Agapophytus* Guérin (Diptera: Therevidae: Agapophytinae). *Invertebrate Taxonomy* **15**, 467–526. doi: 10.1071/IT00027

Winterton, S. L., Irwin, M. E., and Yeates, D. K. (1999a). Phylogenetic revision of the *Taenogera* Kröber genus-group (Diptera: Therevidae) with descriptions of two new genera. *Australian Journal of Entomology* **38**, 274–290. doi: 10.1046/j.1440-6055.1999.00126.x

Winterton, S. L., Irwin, M. E., and Yeates, D. K. (1999b). Systematics of *Nanexila* Winterton & Irwin, gen. nov. (Diptera: Therevidae) from Australia. *Invertebrate Taxonomy* **13**, 237–308. doi: 10.1071/IT97029

Winterton, S. L., Skevington, J. H., Irwin, M. E., and Yeates, D. K. (2000). Phylogenetic revision of *Bonjeania* Irwin & Lyneborg (Diptera: Therevidae). *Systematic Entomology* **25**, 295–324. doi: 10.1046/j.1365-3113.2000.00107.x

Winterton, S. L., Yang, L., Wiegmann, B. M., and Yeates, D. K. (2001). Phylogenetic revision of the Agapophytinae subf. n. (Diptera: Therevidae) based on molecular and morphological evidence. *Systematic Entomology* **26**, 173–211. doi: 10.1046/j.1365-3113.2001.00142.x

Wüster, W., Ferguson, J. E., Quijada-Mascareñas, J. A., Pook, C. E., da Graça Salomão, M., and Thorpe, R. S. (2005). Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). *Molecular Ecology* **14**, 1095–1108. doi: 10.1111/j.1365-294X.2005.02471.x

Yang, L., Wiegmann, B. M., Yeates, D. K., and Irwin, M. E. (2000). Higher-level phylogeny of the Therevidae (Diptera: Insecta) based on 28S ribosomal and elongation factor-1α gene sequences. *Molecular Phylogenetics and Evolution* **15**, 440–451. doi: 10.1006/mpev.1999.0771

Yeates, D. K., Irwin, M. E., and Wiegmann, B. M. (2003). Ocoidae, a new family of asiloid flies (Diptera: Brachycera: Asiloidea), based on *Ocoa chilensis* gen. and sp.n. from Chile, South America. *Systematic Entomology* **28**, 417–431. doi: 10.1046/j.1365-3113.2003.00224.x

Yeates, D. K., Wiegmann, B. M., Courtney, G. W., Meier, R., Lambkin, C., and Pape, T. (2007). Phylogeny and systematics of Diptera: two decades of progress and prospects. *Linnaeus Tercentenary: Progress in Invertebrate Taxonomy. Zootaxa* **1668**, 565–590.