Many hexapod groups originated earlier and withstood extinction events better than previously realized: inferences from supertrees

Robert B. Davis^{*,†}, Sandra L. Baldauf[‡] and Peter J. Mayhew

Department of Biology, University of York, Heslington, York YO10 5YW, UK

Comprising over half of all described species, the hexapods are central to understanding the evolution of global biodiversity. Direct fossil evidence suggests that new hexapod orders continued to originate from the Jurassic onwards, and diversity is presently higher than ever. Previous studies also suggest that several shifts in net diversification rate have occurred at higher taxonomic levels. However, their inferred timing is phylogeny dependent. We re-examine these issues using the supertree approach to provide, to our knowledge, the first composite estimates of hexapod order-level phylogeny. The Purvis matrix representation with parsimony method provides the most optimal supertree, but alternative methods are considered. Inferring ghost ranges shows richness of terminal lineages in the order-level phylogeny to peak just before the end-Permian extinction, rather than the present day, indicating that at least 11 more lineages survived this extinction than implied by fossils alone. The major upshift in diversification is associated with the origin of wings/wing folding and for the first time, to our knowledge, significant downshifts are shown associated with the origin of species-poor taxa (e.g. Neuropterida, Zoraptera). Polyneopteran phylogeny, especially the position of Zoraptera, remains important resolve because this influences findings regarding shifts in diversification. Our study shows how combining fossil with phylogenetic information can improve macroevolutionary inferences.

Keywords: fossil record; key innovations; insect diversity; macroevolution; mass extinction; supertree

1. INTRODUCTION

One of the major aims of the field of macroevolutionary biology is to describe and explain differences in species richness across taxa (Barracough *et al.* 1998; Nee 2001). Because over half of described species belong to just one class of organisms, the hexapods (insects and their six-legged kin such as springtails) (Grimaldi & Engel 2005), understanding the evolution of species richness in this taxon is a substantial contribution to understanding the forces that shape global biodiversity (Labandeira & Eble 2002; Mayhew 2002). A central part of explaining the diversity of speciose groups involves identifying when major clades originated, and changes in the rate of production of new lineages. Knowing when major clades originated can help to identify potential environmental or other causes of adaptive radiation. The same goes for identifying shifts in the rate of diversification, which can help test ideas about potential key evolutionary innovations (Barracough *et al.* 1998; Davis *et al.* 2009).

The hexapods comprise 46 orders, living and extinct, many of which are familiar groups such as the Coleoptera (beetles), Diptera (true flies) and Lepidoptera

(butterflies/moths) (Grimaldi & Engel 2005). Some morphological or developmental traits are shared by several, but not all orders, through common descent (synapomorphies), and are hypothesized to have increased the rates of diversification in those groups that share them, potentially making these traits key evolutionary innovations. Such innovations include external mouthparts, wings, wing folding and complete metamorphosis (Mayhew 2007).

Comparing species richness between sister groups provides potentially powerful tests of such key innovation hypotheses (Nee *et al.* 1994). However, because comparisons at deep nodes can depend on the relationships towards the tree tips (the trickle-down effect, see §2), such analyses require a phylogeny containing all the higher taxa in the group under consideration (Davies *et al.* 2004; Hunt *et al.* 2007). Yet, phylogenies containing all hexapod orders are rare and are based on only subsets of available evidence such as a particular gene (Yoshizawa & Johnson 2005), or specific morphological characters (Beutel & Gorb 2001), with different studies suggesting different relationships, complicating the analysis considerably (Mayhew 2002). A useful tool would therefore be a single tree of hexapod relationships summarizing all the existing phylogenetic evidence. Such trees are now routinely produced for similar evolutionary analyses using supertree methods, which use existing phylogenetic topologies as their input data (Bininda-Emonds 2004).

Previous work on the times of origin of the hexapod orders has focused on documenting the appearances of

* Author for correspondence (davis@ut.ee).

† Present address: Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu, 46 Vanemuise Street, Tartu 51014, Estonia.

‡ Present address: Department of Evolutionary Biology, Uppsala University, Uppsala, Sweden.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2009.2299> or via <http://rspb.royalsocietypublishing.org>.

the first fossil of each (Jarzembowski & Ross 1996). The oldest hexapod fossil is a collembolan from the Early Devonian (approx. 400 Ma) (Whalley & Jarzembowski 1981), and some molecular dating matches the fossil evidence for this (Gaunt & Miles 2002). Many orders are first documented from the Palaeozoic, and, following the apparent extinction of several orders at the ends of the Permian and Triassic, a number of other orders are then first documented from the Jurassic, Cretaceous and Cenozoic, suggesting a continuing diversification of major body plans, perhaps driven by the diversification of other terrestrial taxa that dominate modern ecosystems such as angiosperms, birds and mammals (Labandeira & Eble 2002). However, evidence also suggests that the fossil record of arthropods is relatively poor (Wills 2001), and it is currently unclear how much this continuous rise in order richness reflects the true continuing diversification pattern rather than an incomplete fossil record. Phylogenies provide a useful tool for detecting gaps in the fossil record because sister clades should have identical origination dates (Norrell 1992). Discrepancies between the dates of first fossil appearances of sisters therefore either suggest that the phylogeny is inaccurate (which is refutable) or that the clade that appeared later actually existed before this but has not been discovered (Norrell & Novacek 1992), hence that the number of lineages rose earlier than the fossil data alone suggest.

Here we use the first (to our knowledge) composite estimates of the phylogeny of all 46 hexapod orders to identify significant shifts in the net rate of cladogenesis, and to find to what extent the first fossil finds of several hexapod orders that date from after the Palaeozoic represent true lineage originations rather than an incomplete fossil record.

2. MATERIAL AND METHODS

(a) *Taxonomy*

The taxonomic nomenclature of Grimaldi & Engel (2005) was used for consistency (see hexapod order names and traditional taxonomy in the electronic supplementary material). The orders Blattaria (cockroaches), Mecoptera (e.g. scorpionflies) and Psocoptera (e.g. barklice) may be paraphyletic to Isoptera (termites), Siphonaptera (fleas) and Phthiraptera (lice), respectively (Lyal 1985; Whiting 2002; Grimaldi & Engel 2005; Inward *et al.* 2007), but were treated here for convenience as monophyletic, as in the majority of order-level phylogenetic and fossil treatments. Regardless, whether these are treated as paraphyletic or not, our overall conclusions are unaffected. The Hexapoda are also treated as monophyletic. While some recent evidence suggests they may be polyphyletic with respect to crustaceans (Nardi *et al.* 2003) or mutually paraphyletic with crustaceans (Cook *et al.* 2005), there is still much evidence that hexapods are monophyletic (Regier *et al.* 2005; Mallatt & Giribet 2006). For molecular studies, with terminal taxa at the species level, these species were assigned to their respective orders.

(b) *Input trees*

Input trees were searched for online (Web of Science, Google Scholar) using the following keywords: insect*, hexapod*, holometabol*, endopterygota*, neoptera*, paraneoptera*, apterygota*, pterygota*, orthopter*, hemimetabola*, dictyoptera*, neuropter*, dicondylia*, entognatha*,

ectognatha*, phylogen*, cladistic, cladogram. (*where any word begins with those letters). The reference sections of these studies were also searched for more potential source studies.

Because Hennig (1969) is considered a major landmark on the phylogeny of hexapod orders, only studies carried out post-1969 were considered. A cut-off date of November 2007 was used. We discuss more recent studies below. All the phylogenies analysed attempted primarily to assess hexapod order-level relationships and include at least three ingroup taxa. We also omitted taxonomies (Richards & Davies 1977), reviews (Kristensen 1981) and trees that authors state should not be used as a reliable phylogeny (Carmean *et al.* 1992). A total of 76 potential input trees met these criteria (see input trees in the electronic supplementary material).

It is an inevitable downside of supertree analysis that more phylogenies will be published after analysis has begun. As an addition to this study, the literature was scoured post-diversification analysis for more recent papers, the input trees taken from these and data non-independence accounted for. The supertree analysis was re-run with the method producing the most optimal supertree (§2c) to see whether there is a change in the topology (see updated Purvis matrix representation with parsimony (MRP) supertree in the electronic supplementary material).

(c) *Supertree methods*

In the absence of compelling reasons to choose a single supertree method, we applied several to assess the sensitivity of the results. As the best-characterized and most widely used method, standard MRP (Baum 1992; Ragan 1992) was used for a series of initial analyses, in which the dataset was refined (see dataset refinement in the electronic supplementary material). Five other supertree methods were also implemented: Purvis MRP (Purvis 1995), matrix representation with compatibility (Ross & Rodrigo 2004), matrix representation with flipping (Eulenstein *et al.* 2004), the most similar supertree method or distance fit (Creevey *et al.* 2004), and the average consensus (Lapointe & Cucumel 1997). We avoided using agreement supertree methods (Bininda-Emonds 2004) because they only identify relationships common to all input trees, providing little information when the many competing hypotheses of hexapod phylogeny are taken into account. Software and settings for supertree analysis are provided in the electronic supplementary material. The initial dataset of 76 input trees was then refined to minimize data non-independence, leaving 58 trees for the final analyses (see dataset refinement in the electronic supplementary material).

The *V* index (Wilkinson *et al.* 2005) was used to assess support for supertree relationships, as was the more liberal *V*+ modification of the *V* index, which considers polytomies as support. These indices work on a scale from -1 to 1, where negative values indicate more disagreement than agreement between a relationship in the supertree and its constituent input trees, and positive values indicate more agreement. *V* scores were calculated using 'stsupport' (<http://taxonomy.zoology.gla.ac.uk/%7Ejcotton/software.html>). Supertree bootstrapping (Burleigh *et al.* 2006) is not applicable to our dataset given the low representation of some taxa in the input tree set (Moore *et al.* 2006).

Trees produced by the different methods were assessed by their overall consistency with the underlying input trees. Diversification analyses were applied to the Purvis MRP

Figure 1. Trees used in diversification analyses and significant shifts in diversification. (a) Majority rule (50%) Purvis MRP super-tree topology; (b) 50% consensus tree of four matrix-based supertrees. Grey circles, topological differences. Open stars, positive shift; black stars, negative shift; grey stars, shift both ways. The five broad groups of hexapods are indicated; (1) Apterygota, (2) Palaeoptera, (3) Polyneoptera, (4) Paraneoptera and (5) Holometabola. Internal branch thickness relates to V scores. Thick solid lines, positive V (including O) and $V+$; medium solid lines, negative V but positive $V+$ (including O); thin solid lines, negative V and $V+$; broken lines, uncertain placement. Asterisk denotes extinct order.

tree as the best-supported phylogeny, and to explore alternative but well-supported topologies, also to a consensus of the four matrix-based methods (figure 1), which all produce relatively well-supported trees in contrast to the two distance-based methods. As the method producing the most optimal phylogeny, Purvis MRP was used to analyse the updated set of input trees. The resulting supertree is shown in figure 2.

(d) Taxon age

For each order, a minimum first appearance date from the fossil record was obtained (see minimum first appearance dates for fossils of each hexapod order in the electronic supplementary material) using Ross & Jarzembowski (1993), and the online fossil insect database 'EDNA' (Mitchell 2007). A minimum first appearance date corresponds to the latest date in the first rock stratum a taxon is identified in. They are conservative estimates and are used for assessing order lineage richness over time (§2e). To provide a sensitivity analysis, maximum first appearance dates (i.e. the earliest date in the first rock stratum a taxon is identified in) are also taken and the effect of this is considered.

The age of the oldest family of each order was used as a proxy for that of the order, removing doubt over specimens that are not confidently assigned to families. Ghost ranges were added for orders with older sister taxa. The time scale of Harland *et al.* (1990), as used by Ross & Jarzembowski

(1993), was used. Recent changes to this time scale only have a significant effect on periods earlier than those considered here.

The fossil record of three orders, Blattaria (cockroaches), Mecoptera (scorpionflies) and Trichoptera (caddisflies), may actually represent the stem of more inclusive clades (Grimaldi & Engel 2005). For these orders, the minimum first appearance date for the 'stem' and 'crown' were recorded. The problematic fossil order Miomoptera was considered in two different positions within the Panorpida; first as the most basal panorpidan order. This is most stratigraphically congruent and has no effect on the ranges of the other panorpidan orders. Second, it was also placed where it causes the greatest effect on the extension of ranges, as sister to Strepsiptera. This results in eight possible scenarios based on the supertree topology (i.e. Purvis MRP or four-method consensus and alternative positions of Miomoptera) and whether or not stem or crown dates are used for Blattaria, Mecoptera and Trichoptera.

(e) Diversification analyses

Using the fossil record, order-level richness over geological time is assessed, in the same manner as previous studies (Jarzembowski & Ross 1996; Jarzembowski 2003). We also investigate the effects of including ghost range information by assigning each lineage the age of the earliest fossil of it or its sister group. There are several ways to define an

Figure 2. Updated Purvis MRP supertree including input trees up until 2009. Shifts in diversification indicated. Open stars, positive shift; black stars, negative shift; grey stars, shift both ways. The five broad groups of hexapods are indicated; (1) Apterygota, (2) Palaeoptera, (3) Polyneoptera, (4) Paraneoptera and (5) Holometabola. Internal branch thickness relates to V scores. Thick solid lines, positive V (including O) and $V+$; medium solid lines, negative V but positive $V+$ (including O); thin solid lines, negative V and $V+$; broken lines, uncertain placement/grafted in post-analysis. Asterisk denotes extinct order.

order based on how we attribute taxonomic rank. One can simply attribute taxonomic rank as a result of a sister-group relationship regardless of the level of morphological distinctiveness or species richness. Alternatively, it can be argued

that new orders should only be established when a group of organisms is sufficiently distinct in morphology from other groups (Smith 1994). Since the ranges of some lineages here are extended back before the age of their first fossil, it

is possible that the missing representatives of these lineages would not be sufficiently distinct under the latter definition, and therefore would only meet the former definition. Essentially, we consider stem lineages of orders here when we consider the origination of an order, and to avoid confusion do not refer to these simply as orders, but rather as lineages or branches leading to orders.

To assess where significant upshifts and downshifts in diversification have occurred across the hexapod tree, comparisons of sister-group species richness were made for each node across both phylogenies. Assuming that sister taxa radiated (allowing for extinction) at equal, but not necessarily constant, rates over time (Nee *et al.* 1994), all possible splits of N species between these two clades are equally likely (Farris 1976). The two-tailed probability of an equal or greater magnitude of split under the null model is given by $2[N_{\text{small}}/(N_{\text{small}} + N_{\text{large}} - 1)]$. Care must be taken when attributing significant differences in species richness between sister clades to shifts in diversification, because seemingly significant differences at a basal node may be attributable to such occurrences towards the crown (i.e. a 'trickle-down' effect). The approach of Davies *et al.* (2004; also Hunt *et al.* 2007) is used to remove such effects (see the method of Davies *et al.* (2004) for countering the trickle-down effect in the electronic supplementary material). The adjustment of species richness in this way has proved to be statistically robust in simulation studies (Moore *et al.* 2004). Alternative methods for detecting diversification exist, but are not applicable to this study: for example that of Chan & Moore (2005) requires the phylogeny to be relatively complete at species level, while that of Rabosky *et al.* (2007) is designed to detect the position at which a single shift is most likely, rather than multiple shifts. Described species numbers for each order were taken from Grimaldi & Engel (2005) (table 1).

3. RESULTS

(a) *Origin, extinction and diversity of orders*

Using the fossil record alone, order richness rises rapidly during the Carboniferous and Permian peaking at 26, with extinctions at the end of the Permian and Triassic. Richness then rises through the Mesozoic and Cenozoic, peaking at present with the 33 extant orders (figure 3).

With ghost ranges implied, important differences emerge (figure 3 and table 2). The greatest difference sees the Protura branch of the tree, which lacks any fossil record, extending back to the Devonian when their sister order Collembola originate. The Embiodea, Mantophasmatodea and Phthiraptera branches have their ranges extended from the Cenozoic to the Palaeozoic, and those of Dermaptera, Hymenoptera, Phasmatodea, Siphonaptera, Titanoptera, Zoraptera and Zygentoma have their ranges extended back from the Mesozoic to Palaeozoic. When stem dates of the appearance of the first fossils for all orders are used over crown dates, the same is the case for the Lepidoptera, Isoptera and Mantodea lineages, and if Miomoptera is considered in a crownward position, the ranges of Diptera and Strepsiptera branches also extend back to the Palaeozoic. In total, 20–23 terminal branches of the order tree (depending on the use of stem/crown dates and the inferred position of Miomoptera) have their ranges pulled back beyond their appearance date of their first fossil (see

Table 1. Described numbers of living species.

extant order	living described species numbers
Archaeognatha	500
Blattaria	4000
Collembola	9000
Coleoptera	350 000
Dermaptera	2000
Diplura	1000
Diptera	120 000
Embiodea	500
Ephemeroptera	3100
Grylloblattodea	26
Hemiptera	90 000
Hymenoptera	125 000
Isoptera	2900
Lepidoptera	150 000
Mantodea	1800
Mantophasmatodea	15
Mecoptera	600
Megaloptera	270
Neuroptera	6010
Odonata	5500
Orthoptera	20 000
Phasmatodea	3000
Phthiraptera	4900
Plecoptera	2000
Protura	600
Psocoptera	4400
Raphidioptera	220
Siphonaptera	2500
Strepsiptera	550
Trichoptera	11 000
Thysanoptera	5000
Zoraptera	32
Zygentoma	400

lineage richness curves of lineages leading to orders under different phylogenetic hypotheses and dates in the electronic supplementary material). The largest increase in richness occurs in the Carboniferous (23–30 branches emerge, not just the 14–16 that have first fossils of this age). Richness of order branches peaks at the end of the Permian, instead of the present, at 37–44, of which 11–16 are inferred to have originated where no fossil evidence is known (figure 3 and table 2). The effect of using maximum first and last appearance dates does not alter our findings greatly. Although the end-Permian extinction appears less severe still (a net reduction of three lineages), five lineages are shown to go extinct here, but this coincides with the origination of Mecoptera and Siphonaptera (figure 3). Furthermore, using the updated tree (figure 2) does not affect the overall conclusions either, with 20 terminal lineages that would have their ranges pulled back based on sister-group relationships.

(b) *Significant shifts in the net rate of cladogenesis*

The two supertrees show considerable similarities in the shifts identified but also some interesting differences (figure 1 and sister-group species richness comparisons in the electronic supplementary material). Both show net-diversification shifts associated with Diptera and Strepsiptera (probably an upshift in Diptera and a

Figure 3. (a) Lineage richness curves constructed in (black, minimum first/last appearance dates and dark grey maximum first/last appearance dates lines) and out (light grey line minimum first/last appearance dates only) of a phylogenetic context. Phylogeny used in this example is the Purvis MRP supertree, (Miomoptera in a basal position, crown group dates for Blattaria, Mecoptera and Trichoptera). Key extinctions indicated. (b) Timing of origins in a phylogenetic context using minimum first appearance dates. (c) Timing of origins outside a phylogenetic context using minimum first appearance dates. (d) Timing of extinctions using minimum first appearance dates. Geological time scale provided at bottom of figure (O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; T, Tertiary/Cenozoic). See table 2 for details of origins and extinctions.

downshift in Strepsiptera), downshifts associated with the origin of Neuropterida (i.e. Neuroptera + Megaloptera + Raphidioptera), and with Zoraptera, one or more downshifts in the clade containing Mantophasmatodea and

Grylloblattodea, and one or more upshifts associated with deep nodes. Different positions of some of these shifts are owing to differences in tree topology. In the four-method consensus, the deepest upshift is detected at the origin of

Table 2. Originations and extinctions of lineages leading to orders using minimum first/last appearance dates. The phylogeny-based originations correspond to figure 3. (Note, there is a difference between extinctions (i.e. last appearances) and first absences. For example, at the end of the Permian, in a phylogenetic context, 33 orders are recorded as existing immediately prior to the extinction. This includes those five orders that go extinct. They have to have existed at this time point in order to go extinct. Their absence is first noted in the first stage of the Triassic by the richness curve in figure 3.)

date (Ma)	originations (phylogeny-based)	originations (fossil-based)	extinctions (i.e. last appearances)
390.4	Collembola, Protura	Collembola	
322.8	Archaeognatha, Diplura, Ephemeroptera, Geroptera, Palaeodictyoptera, Zygentoma	Geroptera, Palaeodictyoptera	
320.6	Odonata, Protodonata	Protodonata	
311.3	Caloneuroidea, Diaphanopteroidea, Dicliptera, Glosselytroidea, Hymenoptera, Megasecoptera, Miomoptera, Zoraptera	Caloneuroidea, Diaphanopteroidea, Megasecoptera, Miomoptera	Geroptera
305		Ephemeroptera	
303		Diplura	
290	Grylloblattodea, Hemiptera, Mantophasmatodea, Orthoptera, Phasmatodea, Phthiraptera, Psocoptera, Titanoptera	Archaeognatha, Grylloblattodea, Hemiptera, Orthoptera, Psocoptera	
281.5	Coleoptera, Dermaptera, Embiodea, Lophioneurida, Plecoptera, Protelytroptera, Protoperlaria, Thysanoptera	Coleoptera, Glosselytroidea, Lophioneurida, Odonata, Protelytroptera, Protoperlaria, Thysanoptera	
256.1	Neuroptera	Dicliptera, Neuroptera, Plecoptera	Dicliptera
255	Megaloptera, Raphidioptera	Megaloptera, Raphidioptera	
245 (end-Permian)	Mecoptera, Siphonaptera	Mecoptera	Caloneuroidea, Diaphanopteroidea, Megasecoptera, Palaeodictyoptera, Protelytroptera
241.1	Diptera, Strepsiptera	Diptera, Phasmatodea, Titanoptera	
223.4	Lepidoptera, Trichoptera	Hymenoptera, Trichoptera	
208 (end-Triassic)			Miomoptera, Protodonata, Protoperlaria, Titanoptera
194.5		Dermaptera, Lepidoptera	
145			Glosselytroidea
140.7	Blattaria, Isoptera, Mantodea	Blattaria, Isoptera, Mantodea	
112		Siphonaptera	
90.4		Zygentoma	
65		Zoraptera	Lophioneurida
50		Strepsiptera	
35.4		Embiodea, Phthiraptera	
23.3		Mantophasmatodea	
0		Protura	

Neoptera. The Purvis MRP tree, offers an alternative scenario showing two shifts along the backbone of the hexapod tree: at the origin of Pterygota, and at the origin of Eumetabola + Zoraptera. Zoraptera universally registers as a significant downshift despite differing in position on the two trees. The same is true of Mantophasmatodea and Grylloblattodea. In the Purvis MRP tree, the latter orders form a clade with Dermaptera, and together these are sister to the Dictyoptera (i.e. Blattaria + Isoptera + Mantodea). Consequently, significant downshifts are inferred within the ancestor of the three orders as well as Grylloblattodea. In the four-method consensus, a downshift is only inferred among the ancestor of Mantophasmatodea and Grylloblattodea. In the updated Purvis MRP tree the same major shifts occur as in the original Purvis MRP

tree, but there is also a possible upshift identified at the origin of Eumetabola alone in addition to the downshift in Zoraptera (figure 2).

4. DISCUSSION

In contrast to previous, purely fossil-based, studies, which show order-level diversity reaching an all time high in the present day, richness of terminal branches in the order phylogeny peaks in the Permian. Therefore, many first fossil appearances of orders later than this do not imply later lineage origination but rather reflect an incomplete fossil record. Consequently, many more such lineages survived the end-Permian extinction than fossils alone suggest. We also detect reductions (downshifts) in the

net rate of cladogenesis, as well as increases (upshifts), of which several are identified for the first time.

(a) *Order lineage richness through time*

The great advantage of placing the fossil record in a phylogenetic context is that ghost ranges are implied, and for hexapods they push the ranges of at least 20 terminal branches of the order-level tree earlier than appearances of their first fossils. Hexapod order lineage richness increased substantially through the Carboniferous. During this period 23–30 lineages leading to orders originated. Lineage richness was highest immediately before the Permian mass extinction. Before this extinction 37–44 such lineages had some representation: 4–11 more (12–15%) than at present. This is contrary to previous findings that used only fossil first appearance dates (Jarzembowski & Ross 1996; Jarzembowski 2003), but our finding agrees with recent molecular studies that places the origin of Holometabola in the Carboniferous (Wiegmann *et al.* 2009). Placing diversity in the phylogenetic context helps eliminate the varying effects of taphonomy between strata and poor preservation potential of certain groups, which plague studies of fossil diversity (Ponomarenko & Dmitriev 2009). Although it is obviously a more serious issue for taxa, like insects, with relatively poor fossil records, our study suggests that other studies of biodiversity through time could benefit from combining fossil with phylogenetic information.

The two major extinction events we note agree with previous findings, but previous studies have probably underestimated how many lineages survived the mass extinction events of the Permian and Triassic. Our trees suggest that 11–16 more lineages survived the Permian mass extinction than implied by first fossils of orders alone, including those leading to the Hymenoptera and possibly Lepidoptera and Diptera, and that 10–12 of these, previously not shown to exist pre-Triassic (including Lepidoptera), would have survived through the Triassic too.

Thus our data suggest that this extinction affected proportionally fewer hexapod lineages than the fossil record alone suggests, a fact that may be important in future studies of extinction-selectivity, or in general when studying the environmental causes of extinction severity (e.g. Mayhew *et al.* 2008). The same is true of the Triassic mass-extinction, because many more lineages are here shown to survive this event than the fossil evidence alone suggests. A small post-Permian recovery is identified in the scenario when crown-group dates are used, and with Miomoptera in a position basal to the other Panorpida. These are more conservative assumptions than applying stem dates, or placing Miomoptera higher up the tree without any hard evidence. Based on the more conservative assumptions that we are more comfortable with, we suggest a small recovery occurred at the beginning of the Triassic, and this relates to the origin of crown-group Mecoptera and Siphonaptera (figure 3 and table 2). Additionally, because the dating method used here is conservative (because in reality both, not just one, of two sister lineages have incomplete fossil records), the discovery of earlier fossils for particular orders or using more liberal dates will only strengthen the hypothesis presented here that diversity was higher before the Permian and Triassic mass extinctions.

(b) *Shifts in net cladogenesis*

Of the individual hexapod orders often considered to be highly diverse, our analyses suggest that only the Diptera represent a significant upshift in diversification (figure 1). The Diptera have arguably a greater ecological diversity than any other order (Grimaldi & Engel 2005). Any upward shift here could be assigned to mouthpart and feeding strategy diversity within this order (Labandeira 1997; Grimaldi & Engel 2005). Key innovations could include brachyceran larval mouthparts adapted for predation and pseudotracheae in brachyceran adult mouthparts, enhancing liquid-feeding (Grimaldi & Engel 2005). Diptera are also recognized by their halteres (club-like hindwing modifications, aiding flight stabilization), but haltere-like structures have evolved in other orders too (Grimaldi & Engel 2005). Low diversity in Strepsiptera could be a result of their move to a specialized parasitic lifestyle (Grimaldi & Engel 2005).

The Coleoptera, have long been regarded as significantly species rich (Farrell 1998). However, our analyses suggest otherwise. The Coleoptera are not significantly more species rich than the remaining Holometabola, which form the sister group to them and their sister group Neuropterida. Downshifts have been reported within Coleoptera in addition to upshifts (Hunt *et al.* 2007). Their richness is owing to events or innovations experienced at the origin of a much more inclusive clade. Instead it is the Neuropterida which have undergone a significant downshift in radiation. Neuropterida are a relict lineage of holometabolous insect, which had much greater diversity in the past (Grimaldi & Engel 2005), suggesting that higher extinction rates may have contributed to this downshift. Furthermore, none of the other hyper-speciose orders of hexapods (Hemiptera, Hymenoptera, Lepidoptera) are shown to have undergone a significant increase in diversification (figures 1 and 2).

The Zoraptera experienced a significant reduction in diversification. Why this enigmatic group is so species-poor is uncertain: their biology is still poorly known (Grimaldi & Engel 2005). The phylogenetic position of Zoraptera is also important to resolve. It influences our findings relating to the deeper shifts in hexapod diversity: in the four-method consensus tree (figure 1b), the deepest significant upshift occurs at the origin of Neoptera (wing folding). Wing folding has been cited previously as a key innovation in the evolution of hexapods, allowing better use of an environment while storing and protecting their wings, especially perhaps in the case of tight spaces for example under bark or leaf litter (Grimaldi & Engel 2005). However, in the Purvis MRP tree (figures 1a and 2), the origin of Eumetabola + Zoraptera represents a significant upshift in radiation and consequently the deepest increase is shifted back to the origin of Pterygota (wings). The evolution of wings is another key innovation cited for the success of hexapods and the ability to fly would have enabled higher dispersal, the use of novel micro-habitats and enhanced ability to escape a predator (Grimaldi & Engel 2005), all of which might raise speciation rates or lower extinction rates (Mayhew 2007).

The character matrix of Wheeler *et al.* (2001) indicates there are two morphological characteristics shared between the Zoraptera and Eumetabola: fusion of the media and radia in the wing, and cryptosterny

(invagination of the thoracic sternal area). However, the effects of these on speciation and extinction rates are unknown. Mayhew (2002) suggested a shift somewhere between the origins of Neoptera and Eumetabola, which is consistent with our findings.

A significant downshift relating to the orders Mantophasmatodea and Grylloblattodea has also probably occurred. Their phylogenetic positions are not well constrained, but both are relict orders, having a previously wider distribution (Vrsansky *et al.* 2001; Engel & Grimaldi 2004).

(c) Hexapod phylogeny

The phylogenies presented here (figure 1) provide, to our knowledge, the first estimates of the relationships of all hexapod orders combining current information. Many areas of hexapod phylogeny are well supported by super-tree analysis, particularly basal apterygote, palaeopteran and paraneopteran relationships. Phylogenetic uncertainty still remains within Polyneoptera, and especially regarding the position of Zoraptera, which affects our findings concerning comparisons of species richness of sister groups. Holometabolan relationships are mostly well supported, but a notably poorly supported relationship is that of Hymenoptera to the Panorpida (i.e. Mecoptera, Siphonaptera, Lepidoptera, Trichoptera, Diptera, Strepsiptera), requiring further study.

In the updated analysis (figure 2), while Diptera remain as the sister group to Strepsiptera, the Hymenoptera are recovered as the basal holometabolan lineage. Furthermore, the 'Polyneoptera' emerge in this analysis as four clades in a paraphyletic lineage, although Zoraptera remain sister to Eumetabola, and Diplura fall within a monophyletic Entognatha, and are no longer sister to Insecta. These changes in the topology have no effect on what we identify to be key innovations relating to significant shifts in diversification.

(d) Conclusions

Our analyses predict that the fossils records of several terminal branches of the order-level phylogeny could potentially be extended backwards considerably in the future by new fossil finds, at least in the stem-group form, and correspondingly that survival of hexapod lineages across the Permian–Triassic and Triassic–Jurassic boundaries were greater than implied by studies that use fossils without an explicit phylogenetic context. This may have important implications for future extinction and origination studies. Our analyses support the suggestion that the origin of wings or wing flexion played an important role in hexapod diversification. Regarding individual orders, we identify the Diptera, Neuropterida, Zoraptera, Strepsiptera, Grylloblattodea and Mantophasmatodea as worthy of more in-depth macroevolutionary analysis. The shifts in diversification rates associated with these higher taxa may well have occurred anywhere within these clades, and only analyses at finer taxonomic scales can identify their exact timing.

R.B.D. was supported by BBSRC PhD studentship BB/D527026/1. We thank James McInerney, Davide Pisani and Mark Wilkinson for advice regarding supertree methods, Andrew Ross for providing the unpublished

supplements to Ross & Jarzembowski (1993), and Andy Purvis and two anonymous referees for comments.

REFERENCES

- Barracough, T. G., Vogler, A. P. & Harvey, P. H. 1998 Revealing the factors that promote speciation. *Phil. Trans. R. Soc. Lond. B* **353**, 241–249. (doi:10.1098/rstb.1998.0206)
- Baum, B. R. 1992 Combining trees as a way of combining datasets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**, 3–10. (doi:10.2307/1222480)
- Beutel, R. G. & Gorb, S. N. 2001 Ultrastructure of attachment specializations of hexapods (Arthropoda): evolutionary patterns inferred from a revised ordinal phylogeny. *J. Zoolog. Syst. Evol. Res.* **39**, 177–207. (doi:10.1046/j.1439-0469.2001.00155.x)
- Bininda-Emonds, O. R. P. 2004 The evolution of supertrees. *Trends Ecol. Evol.* **19**, 315–322.
- Burleigh, J. G., Driskell, A. C. & Sanderson, M. J. 2006 Supertree bootstrapping methods for assessing phylogenetic variation among genes in genome-scale data sets. *Syst. Biol.* **55**, 426–440. (doi:10.1080/106351505000541722)
- Carmean, D., Kimsey, L. S. & Berbee, M. L. 1992 18S rDNA sequences and the holometabolous insects. *Mol. Phylogenet. Evol.* **1**, 270–278. (doi:10.1016/1055-7903(92)90002-X)
- Chan, K. M. A. & Moore, B. R. 2005 SYMMETREE: whole-tree analysis of differential diversification rates. *Bioinformatics* **21**, 1709–1710. (doi:10.1093/bioinformatics/bti175)
- Cook, C. E., Yue, Q. & Akam, M. 2005 Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proc. R. Soc. B* **272**, 1295–1304. (doi:10.1098/rspb.2004.3042)
- Creevey, C., Fitzpatrick, D. A., Philip, G. A., Kinsella, R. J., O'Connell, M. J., Travers, S. A., Wilkinson, M. & McInerney, J. O. 2004 Does a tree-like phylogeny only exist at the tips in the prokaryotes? *Proc. R. Soc. Lond. B* **271**, 2551–2558. (doi:10.1098/rspb.2004.2864)
- Davies, T. J., Barracough, T. G., Chase, M. W., Soltis, P. S. & Savolainen, V. 2004 Darwin's abominable mystery: insights from a supertree of angiosperms. *Proc. Natl Acad. Sci. USA* **101**, 1904–1909. (doi:10.1073/pnas.0308127100)
- Davis, R. B., Baldauf, S. L. & Mayhew, P. J. 2009 Eusociality and the success of the termites: insights from a supertree of dictyopteran families. *J. Evol. Biol.* **22**, 1750–1761. (doi:10.1111/j.1420-9101.2009.01789.x)
- Engel, M. S. & Grimaldi, D. A. 2004 A new rock crawler in Baltic amber, with comments on the order (Mantophasmatodea: Mantophasmatidae). *Am. Mus. Novit.* **3431**, 1–11. (doi:10.1206/0003-0082(2004)431<0001:ANRCIB>2.0.CO;2)
- Eulenstein, O., Chen, D., Burleigh, J. G., Fernández-Baca, D. & Sanderson, M. J. 2004 Performance of flip supertree construction with a heuristic algorithm. *Syst. Biol.* **53**, 299–308. (doi:10.1080/10635150490423719)
- Farrell, B. D. 1998 'Inordinate fondness' explained: why are there so many beetles? *Science* **25**, 196–198.
- Farris, J. S. 1976 Expected asymmetry of evolutionary rates. *Syst. Zoolog.* **25**, 196–198. (doi:10.2307/2412748)
- Gaunt, M. W. & Miles, M. A. 2002 An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographical landmarks. *Mol. Biol. Evol.* **19**, 748–761.
- Grimaldi, D. A. & Engel, M. S. 2005 *Evolution of the insects*. Cambridge, UK: Cambridge University Press.

- Harland, W. B., Armstrong, R. L., Cox, A. V., Craig, L. E., Smith, A. G. & Smith, D. G. 1990 *A geologic time scale*, 1989 edn. Cambridge, UK: Cambridge University Press.
- Hennig, W. 1969 *Die Stammesgeschichte der Insekten*. Frankfurt am Main, Germany: Kramer.
- Hunt, T. *et al.* 2007 A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* **318**, 1913–1916. (doi:10.1126/science.1146954)
- Inward, D., Beccaloni, G. & Eggleton, P. 2007 Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches. *Biol. Lett.* **3**, 331–335. (doi:10.1098/rsbl.2007.0102)
- Jarzembowski, E. A. 2003 Palaeoentomology: towards the big picture. *Acta Zoolog. Cracov.* **46**(Suppl.), 25–36.
- Jarzembowski, E. A. & Ross, A. J. 1996 Insect origination and extinction in the Phanerozoic. In *Biotic recovery from mass extinction events*, vol. 102 (ed. M. B. Hart), pp. 65–78. London, UK: Geological Society Special Publication.
- Kristensen, N. P. 1981 Phylogeny of insect orders. *Annu. Rev. Entomol.* **26**, 135–157. (doi:10.1146/annurev.en.26.010181.001031)
- Labandeira, C. C. 1997 Insect mouthparts: ascertaining the paleobiology of insect feeding strategies. *Annu. Rev. Ecol. Syst.* **28**, 153–193. (doi:10.1146/annurev.ecolsys.28.1.153)
- Labandeira, C. C. & Eble, G. J. 2002 Global diversity patterns of insects from the fossil record. Santa Fe Institute working paper, vol. 121, pp. 1–54.
- Lapointe, F.-J. & Cucumel, G. 1997 The average consensus procedure: combination of weighted trees containing identical or overlapping sets of taxa. *Syst. Biol.* **46**, 306–312.
- Lyal, C. H. C. 1985 Phylogeny and classification of the Psocodea with particular reference to the lice (Psocodea: Phthiraptera). *Syst. Entomol.* **10**, 145–165.
- Mallatt, J. & Giribet, G. 2006 Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Mol. Phylogenet. Evol.* **40**, 772–794. (doi:10.1016/j.ympev.2006.04.021)
- Mayhew, P. J. 2002 Shifts in hexapod diversification and what Haldane could have said. *Proc. R. Soc. Lond. B* **269**, 969–974. (doi:10.1098/rspb.2002.1957)
- Mayhew, P. J. 2007 Why are there so many insects? Perspectives from fossils and phylogenies. *Biol. Rev.* **82**, 425–454. (doi:10.1111/j.1469-185X.2007.00018.x)
- Mayhew, P. J., Jenkins, G. B. & Benton, T. G. 2008 A long-term association between global temperature and biodiversity, origination and extinction in the fossil record. *Proc. R. Soc. B* **275**, 47–53. (doi:10.1098/rspb.2007.1302)
- Mitchell, T. 2007 The EDNA fossil insect database. See <http://edna.palass-hosting.org>.
- Moore, B., Chan, K. M. A. & Donoghue, M. J. 2004 Detecting diversification rate variation in supertrees. In *Phylogenetic supertrees: combining information to reveal the tree of life* (ed. O. R. P. Bininda-Emonds), pp. 487–533. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Moore, B., Smith, S. A. & Donoghue, M. J. 2006 Increasing data transparency and estimating phylogenetic uncertainty in supertrees: approaches using nonparametric bootstrapping. *Syst. Biol.* **55**, 662–676. (doi:10.1080/10635150600920693)
- Nardi, F., Spinsanti, G., Boore, J. L., Carapelli, A., Dallai, R. & Frati, F. 2003 Hexapod origins: monophyletic or paraphyletic? *Science* **21**, 1887–1889.
- Nee, S. 2001 Inferring speciation rates from phylogenies. *Evolution* **55**, 661–668. (doi:10.1554/0014-3820(2001)055[0661:ISRFPP]2.0.CO;2)
- Nee, S., May, R. M. & Harvey, P. H. 1994 The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* **344**, 305–311. (doi:10.1098/rstb.1994.0068)
- Norrell, M. A. 1992 Taxic origin and temporal diversity: the effect of phylogeny. In *Extinction and phylogeny* (eds M. J. Novacek & Q. D. Wheeler), pp. 88–118. New York, NY: Columbia University Press.
- Norrell, M. A. & Novacek, M. J. 1992 Congruence between superpositional and phylogenetic patterns: comparing cladistic patterns with fossil records. *Cladistics* **8**, 319–337. (doi:10.1111/j.1096-0031.1992.tb00074.x)
- Ponomarenko, A. G. & Dmitriev, V. Y. U. 2009 Diversity curves revisited. *Paleontol. J.* **43**, 226–229. (doi:10.1134/S0031030109020154)
- Purvis, A. 1995 A modification to Baum and Ragan's method for combining phylogenetic trees. *Syst. Biol.* **44**, 251–255.
- Rabosky, D. L., Donnellan, S. C., Talaba, A. L. & Lovette, I. J. 2007 Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proc. R. Soc. B* **274**, 2915–2923. (doi:10.1098/rspb.2007.0924)
- Ragan, M. A. 1992 Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet. Evol.* **1**, 53–58. (doi:10.1016/1055-7903(92)90035-F)
- Regier, J. C., Schultz, J. W. & Kambic, R. E. 2005 Pancrustacean phylogeny: hexapods are terrestrial crustaceans and maxillopods are not monophyletic. *Proc. R. Soc. B* **272**, 395–401. (doi:10.1098/rspb.2004.2917)
- Richards, O. W. & Davies, R. J. 1977 *Imm's general textbook of entomology*. London, UK: Chapman & Hall.
- Ross, A. J. & Jarzembowski, E. A. 1993 Arthropoda (Hexapoda: Insecta). In *The fossil record 2* (ed. M. J. Benton), pp. 363–426. London, UK: Chapman & Hall.
- Ross, H. A. & Rodrigo, A. J. 2004 An assessment of matrix representation with compatibility in supertree construction. In *Phylogenetic supertrees: combining information to reveal the tree of life* (ed. O. R. P. Bininda-Emonds), pp. 35–63. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Smith, A. B. 1994 *Systematics and the fossil record*. Oxford, UK: Blackwell Science.
- Vrsansky, P., Storozhenko, S. Y., Labandeira, C. C. & Ihringova, P. 2001 *Galloisiana olgae* sp. nov. (Grylloblattodea: Grylloblattidae) and the paleobiology of a relict order of insects. *Ann. Entomol. Soc. Am.* **94**, 179–184. (doi:10.1603/0013-8746(2001)094[0179:GOSNGG]2.0.CO;2)
- Whalley, P. & Jarzembowski, E. A. 1981 A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature* **291**, 317. (doi:10.1038/291317a0)
- Wheeler, W. C., Whiting, M., Wheeler, Q. D. & Carpenter, J. M. 2001 The phylogeny of the extant hexapod orders. *Cladistics* **17**, 113–169. (doi:10.1111/j.1096-0031.2001.tb00115.x)
- Whiting, M. F. 2002 Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoolog. Scripta* **31**, 93–104. (doi:10.1046/j.0300-3256.2001.00095.x)
- Wiegmann, B. M., Trautwein, M. D., Kim, J.-W., Cassel, B. K., Bertone, M. A., Winterton, S. L. & Yeates, D. K. 2009 Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol.* **7**, 34. (doi:10.1186/1741-7007-7-34)
- Wilkinson, M., Pisani, D., Cotton, J. A. & Corfe, I. 2005 Measuring support and finding unsupported relationships in supertrees. *Syst. Biol.* **54**, 823–831. (doi:10.1080/1063515050950362)
- Wills, M. A. 2001 How good is the fossil record of arthropods? *Geol. J.* **36**, 187–210. (doi:10.1002/gj.882)
- Yoshizawa, K. & Johnson, K. P. 2005 Aligned 18S for Zoraptera (Insecta): phylogenetic position and molecular evolution. *Mol. Phylogenet. Evol.* **37**, 572–580. (doi:10.1016/j.ympev.2005.05.008)