A genus-level supertree of Adephaga (Coleoptera)

Rolf G. Beutel\textsuperscript{a,}\textsuperscript{*}, Ignacio Riberab, Olaf R.P. Bininda-Emonds\textsuperscript{a}

\textsuperscript{a}Institut für Spezielle Zoologie und Evolutionsbiologie, FSU Jena, Germany
\textsuperscript{b}Museo Nacional de Ciencias Naturales, Madrid, Spain

Received 14 October 2005; accepted 17 May 2006

Abstract

A supertree for Adephaga was reconstructed based on 43 independent source trees – including cladograms based on Hennigian and numerical cladistic analyses of morphological and molecular data – and on a backbone taxonomy. To overcome problems associated with both the size of the group and the comparative paucity of available information, our analysis was made at the genus level (requiring synonymizing taxa at different levels across the trees) and used Safe Taxonomic Reduction to remove especially poorly known species. The final supertree contained 401 genera, making it the most comprehensive phylogenetic estimate yet published for the group. Interrelationships among the families are well resolved. Gyrinidae constitute the basal sister group, Haliplidae appear as the sister taxon of Geadephaga + Dytiscoida, Noteridae are the sister group of the remaining Dytiscoida, Amphizoididae and Aspidytidae are sister groups, and Hygrobiidae forms a clade with Dytiscidae. Resolution within the species-rich Dytiscidae is generally high, but some relations remain unclear. Trachypachidae are the sister group of Carabidae (including Rhysodidae), in contrast to a proposed sister-group relationship between Trachypachidae and Dytiscoida. Carabidae are only monophyletic with the inclusion of a non-monophyletic Rhysodidae, but resolution within this megadiverse group is generally low. Non-monophyly of Rhysodidae is extremely unlikely from a morphological point of view, and this group remains the greatest enigma in adephagan systematics. Despite the insights gained, our findings highlight that a combined and coordinated effort of morphologists and molecular systematists is still required to expand the phylogenetic database to enable a solid and comprehensive reconstruction of adephagan phylogeny.

\textcopyright{} 2007 Gesellschaft für Biologische Systematik. Published by Elsevier GmbH. All rights reserved.

See also Supplementary material in the online edition at doi:10.1016/j.ode.2006.05.003

Keywords: Adephaga; Cladistics; Phylogenetic systematics; Safe Taxonomic Reduction; Supertree construction

Introduction

Adephaga is the second largest suborder of Coleoptera and comprises ca. 30,000 species in three terrestrial and eight aquatic or hygropetric families (Meruidae, which was first described by Spangler and Steiner 2005, is not considered here). Numerous efforts have been undertaken to resolve the phylogeny of this group, and themselves show an evolution in terms of the methods and data used. Older phylogenetic studies, such as Crowson (1960), were not based on Hennigian (Hennig 1950, 1966) or cladistic methods, but rather on intuition (see, for example, Wheeler 1995: “Crowsonian phylogenetics”). A Hennigian approach was used in most studies carried out in the later decades of the last
century. Some of these studies were very detailed comparative investigations focused on a specific body part (e.g. ovipositor of Hydradephaga, Burmeister 1976; ovipositor of Geadeaphaga, Bils 1976; prothorax, Baehr 1979) and with Hennigian character evaluation, but without explicit use of the outgroup comparison method (see Maddison et al. 1984). A large increase in the number of larval studies started in the 1980s (e.g. Beutel 1986a, 1992a, 1993; Ruhnau 1986; Arndt 1993; Alarie 1997, 1998; Alarie and Larson 1998; Alarie and Balke 1999), with the first comprehensive cladistic analyses of morphological data being published in the 1990s (Beutel and Haas 1996; Beutel 1997, 1998). Analyses of DNA sequence data followed shortly thereafter. The first studies were largely or exclusively based on 18S rRNA sequences (Vogler and Barraclough 1998; Maddison et al. 1999; Shull et al. 2001; Ribera et al. 2002b), but sequences of several genes have been used more recently (Ribera et al. 2002a; Balke et al. 2005).

The main purpose of constructing supertrees (sensu Sanderson et al. 1998), like total evidence, is to combine the available phylogenetic information to derive an estimate that is more comprehensive than one based on any single information source. In a supertree context, this information amounts to trees with different but jointly overlapping taxon samples and trees restricted to different subgroups of the more inclusive taxon under consideration.

Although supertrees are appearing increasingly in the literature (see Bininda-Emonds 2004a), the method remains highly controversial. The key points of contention involve what effects the analysis of tree topologies rather than the data upon which they are based have on the accuracy of the resultant supertree (e.g. due to the loss of information inherent in the former as compared to the latter) and, more generally, whether or not this procedure represents a legitimate method of phylogenetic inference. As important as this debate is, a full summary of it is beyond the scope of the present work. We refer the reader to the relevant literature instead, particularly to the exchange between Gatesy et al. (2002, 2004) and Bininda-Emonds et al. (2003) and Bininda-Emonds (2004b). However, it is worth pointing out that even the staunchest critics of the supertree approach (e.g. Gatesy and Springer 2004) admit that such a framework will ultimately be necessary to reconstruct the Tree of Life.

What we attempt here is the use of supertree construction to combine presently available adephagan phylogenies to reflect the state-of-the-art of adephagan studies, particularly areas of the adephagan tree that remain poorly resolved due to a lack of consensus and/or of sufficient research effort. We also point out potential methodological and technical problems to supertree construction as highlighted in this study, and present and discuss potential solutions to these problems.

### Material and methods

#### Source data

We compiled trees from manuscripts published or in press by December 2004 that were found using a combination of searches of the available databases (e.g. Web of Science, Zoological Record) as well as from a thorough examination of the relevant literature. To be included, a source tree had to meet the following criteria: (1) relationships of all terminal taxa had to be unambiguously established from a figure or from the text; (2) characters used to build the tree had to be clearly specified; and (3) trees had to contain some novel data and/or be derived from novel analyses with respect to other trees to avoid any pseudoreplication in which a given data set is represented in more than one source tree (see Gatesy et al. 2002).

As a result, a considerable number of studies (e.g. Forsyth 1968, 1969, 1972; Hlavac 1975; Kavanaugh 1986; Ruhnau 1986; Deuve 1988, 1994; Jaglarz 1998) were not included here. These works contain valuable data, but the systematic conclusions were not presented in a form suitable for the inclusion in our supertree (there was no clear specification of the characters used, the terminal taxa, the relationships of all groups or of the criteria used for the elaboration of the tree). The nomenclature of all source trees was updated and standardized according to the recent catalogues of Nilsson (2001) and Löbl and Smetana (2003), and completed with other source references where necessary. Because of the large number of species of Adephaga, many of which are poorly studied at best, we used genera as the terminal taxa in this study. In source trees where species formed the terminal taxa and the genus in which they are currently placed was not reconstructed as being monophyletic, either some of the species were considered to belong to a different, ‘informal’ genus (e.g. ‘Stictotarsus gr.’, corresponding to a group of species within the genus Stictotarsus; Ribera 2003), or all relevant taxa were collapsed to a polytomy for that node.

The generally poor overlap between the published phylogenetic trees made it necessary to use a backbone taxonomy (see Appendix A in the online edition at: doi:10.1016/j.ode.2006.05.003) as an additional source tree, particularly to guide the placement of those genera with insufficient representation in the source trees. This procedure has been shown in simulation to improve the efficacy and accuracy of supertree construction (Bininda-Emonds and Sanderson 2001), albeit at the cost of including a source tree that, in this specific case, might not fulfil the above criteria 2 and 3. However, the poorly resolved nature of the seed taxonomy means that it will be easily overruled by any of the other, more robust source trees (in contrast to using the taxonomy as
Analytical methods – supertree construction

We used the matrix representation with parsimony (MRP; Baum 1992; Ragan 1992) supertree method to combine the source trees listed in Table 1. Although numerous supertree methods exist, MRP represents by far the most popular method due to its universal applicability (it can combine any set of source trees) and its use of the well-known parsimony as an optimization criterion (see Bininda-Emonds 2004a). In addition, MRP has shown good performance in simulation, with accuracy about on a par with a total-evidence analysis of the analogous data set (Bininda-Emonds and Sanderson 2001).

MRP operates by encoding the topology of the source trees as a partial binary matrix. For each node in a source tree, taxa descended from that node are scored as 1; taxa not descended from that node but present elsewhere on the source tree are scored as 0; all other taxa (i.e., those not present on that source tree) are scored as? An all-zero outgroup is added to the ‘matrix representation’ for rooting purposes during the analysis (but see Bininda-Emonds et al. 2005), but is pruned from the tree thereafter. Analysing the matrix representation for any single source tree using parsimony will recover that source tree exactly, whereas analysing the combined matrix representations of the set of source trees will yield a tree (the supertree) with the best fit to all source trees in that set.

MRP encoding was performed using the Perl script SuperMRP and resulted in an initial MRP matrix of 409 taxa (including the hypothetical MRP outgroup) and 781 MRP ‘characters’. However, initial exploratory analyses of this matrix revealed large numbers of equally most parsimonious solutions and generally low resolving power (results not shown). In many cases, both factors result from the presence of ‘floating taxa’ that are poorly known (i.e., occur on few source trees in a supertree context) and hence can cluster equally parsimoniously with a large number of taxa with which they share the little information that is known for them. The end result is a local reduction in resolution. Often, the selective removal of such poorly known taxa will recover much of the resolution that was lost, revealing relationships between the remaining taxa that were otherwise obscured. This is the principle behind the method of Safe Taxonomic Reduction (STR; Wilkinson 1995).

We identified potential floating taxa using STR as implemented in the Perl scripts PerlEQ (Jeffery and Wilkinson, unpubl.) and STRindexer. In particular, we focused on taxa for which the known information was identical to that for one or more better-known index taxa (‘potential asymmetric equivalents’) and which had to originate from the same node as the index taxon (category C* taxa). Combining both factors often allows the removed taxa to be unambiguously re-included on the tree. STR revealed 100 category C* taxa, which we pruned from the source trees (see Pisani and Wilkinson 2002) to yield a final MRP matrix of 309 taxa and 691 MRP characters. In doing so, one source tree (that of Beutel and Ruhnau 1990) was lost despite meeting the three criteria above, because the pruning procedure caused it to contain fewer than three taxa and become uninformative. The full matrix of 409 taxa has been deposited in TreeBASE (http://www.treebase.org; study accession number S1682, matrix accession number M3040), together with instructions to derive the final matrix of 309 taxa.

Analysis of the final MRP matrix was performed in PAUP* v4.0b10 (Swofford 2002) using the Perl script PerlRat v1.0.9 to encode the search instructions for a parsimony ratchet analysis (Nixon 1999). Ratchet searches here employed 200 individual sets of 500 reweighting iterations, where 25% of the characters were randomly selected to receive a weight of 3, with no final ‘brute-force’ search on the set of 100–500 trees that were found. The supertree was held to be the strict consensus of all equally most parsimonious trees from the maximal possible set of 100–500 trees. When possible, taxa removed using STR were re-included on the tree using the Perl script reverseSTR.

In addition to an analysis of the MRP matrix in which all characters were equally weighted, differential weighting was also used to account for differential robustness between source trees. Ideally, differential weighting would make use of some direct measure of support from the source trees (e.g. bootstrap frequencies or Bremer support), a procedure that has been demonstrated in simulation to improve the accuracy of MRP supertree construction (Bininda-Emonds and Sanderson 2001). Because no single such metric was present globally among the set of source trees, we used study size as a proxy for support at the level of entire trees. Study size was quantified by calculating the number of characters per taxon in each study. The average study size was determined for morphological and molecular studies separately, and each study was compared to the appropriate average using Z-scores. MRP characters from studies with a Z-score < -1 and the backbone taxonomy were given a weight of 1, those with a Z-score > 1 were given a weight of 4, and the remaining studies were given a weight of 2. For studies with mixed data types (e.g. Vogler and Barraclough 1998; Ribera et al. 2002a; Miller 2003), the study size was calculated separately for each of the morphological and molecular
Table 1. Selected characters of source trees used

<table>
<thead>
<tr>
<th>References</th>
<th>Group</th>
<th>No. taxa</th>
<th>Terminal taxa</th>
<th>Data type</th>
<th>Character system</th>
<th>Method</th>
<th>No. char.</th>
<th>Inf. char.</th>
<th>Char./taxon</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Morph.</td>
<td>Mol.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[5] Arndt (1993)</td>
<td>Caraboidea</td>
<td>T.2, 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>41</td>
<td>0.41</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[12] Beutel (1992a)</td>
<td>Metriinae, Paussinae</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>0.63</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[16] Beutel and Roughley (1994)</td>
<td>Gyrinidae</td>
<td>18</td>
<td>G</td>
<td>f</td>
<td>L</td>
<td></td>
<td>22</td>
<td>2.75</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>[18] Bils (1976)</td>
<td>Caraboidea</td>
<td>1</td>
<td>F, T</td>
<td>f</td>
<td>Gf</td>
<td></td>
<td>47</td>
<td>1.07</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[20] Dettner (1979)</td>
<td>Hydradephaga</td>
<td>3</td>
<td>F, sF, G</td>
<td>f</td>
<td>pylgidial gland constituents</td>
<td></td>
<td>H &lt;10</td>
<td>1.11</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[22] Erwin (1985)</td>
<td>Caraboidea</td>
<td>3</td>
<td>sF, G</td>
<td>f</td>
<td>A</td>
<td></td>
<td>25</td>
<td>1.41</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[24] Hammond (1979)</td>
<td>– (text)</td>
<td>3</td>
<td>F</td>
<td>f</td>
<td>wing folding and venation</td>
<td></td>
<td>H &lt;15</td>
<td>0.37</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[26] Liebherr and Will (1998)</td>
<td>Caraboidea</td>
<td>57</td>
<td>T, G</td>
<td>f</td>
<td>Gf</td>
<td></td>
<td>20</td>
<td>0.36</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>[33] Oygur and Wolfe (1991)</td>
<td>Gyrinidae</td>
<td>160</td>
<td>G</td>
<td>f</td>
<td>A</td>
<td></td>
<td>80</td>
<td>7.27</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Taxa</td>
<td>Characters</td>
<td>Data type</td>
<td>Character system</td>
<td>Method</td>
<td>No. char.</td>
<td>Inf. char.</td>
<td>Weight</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>--------</td>
<td>----------</td>
<td>-----------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[34] Ribera (2003)</td>
<td>2b Hydroporini ((Graptodytes \text{ gr.}))</td>
<td>6 S</td>
<td>COI, 16S</td>
<td>C</td>
<td></td>
<td>1271</td>
<td>352</td>
<td>58.67</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>[35] Ribera (2003)</td>
<td>3 Hydroporini ((Deronectes \text{ gr.}))</td>
<td>9 S</td>
<td>COI, 16S</td>
<td>C</td>
<td></td>
<td>1274</td>
<td>451</td>
<td>50.11</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>[36] Ribera et al. (2002b)</td>
<td>3 Hydradephaga</td>
<td>64 S</td>
<td>18S</td>
<td>C</td>
<td></td>
<td>409</td>
<td>2759</td>
<td>736</td>
<td>6.39</td>
<td>1</td>
</tr>
<tr>
<td>[37] Ribera et al. (2003)</td>
<td>3 Hydroporini</td>
<td>12 S</td>
<td>COI, 16S</td>
<td>C</td>
<td></td>
<td>1316</td>
<td>388</td>
<td>32.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>[38] Ribera et al. (2004)</td>
<td>2 Agabinae</td>
<td>23 S</td>
<td>COI, 16S</td>
<td>C</td>
<td></td>
<td>1284</td>
<td>531</td>
<td>23.09</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>[40] Shull et al. (2001)</td>
<td>4 Adepata</td>
<td>35 S</td>
<td>18S</td>
<td>C</td>
<td></td>
<td>2480</td>
<td>834</td>
<td>23.83</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>[41] Vogler and Barraclough (1998)</td>
<td>1 Cicindelinae</td>
<td>18 S</td>
<td>f+1</td>
<td>L, 18S, 16S</td>
<td>C</td>
<td>34</td>
<td>906</td>
<td>320</td>
<td>1.89</td>
<td>17.78</td>
</tr>
<tr>
<td>[42] Wolfe (1985)</td>
<td>45 Hydroporinae</td>
<td>17 G</td>
<td>A</td>
<td>H</td>
<td></td>
<td>18</td>
<td>1.06</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[43] Wolfe (1988)</td>
<td>3 Hydroporinae</td>
<td>13 S</td>
<td>A</td>
<td>C</td>
<td></td>
<td>19</td>
<td>1.46</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[44] Backbone taxonomy</td>
<td>– Adepata</td>
<td>389 S</td>
<td>F</td>
<td>L, A</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

For bibliographic citations see References; for the backbone taxonomy used see Appendix A in the online version of this paper (doi:10.1016/j.ode.2006.05.003).

Terminal taxa: F = family; sF = subfamily, T = tribe, G = genus, S = species.

Data type: f = morphological; l = molecular.

Character system: morphological: A = adult, Gf = female genitalia, Gm = male genitalia, L = larva; molecular: COI = cytochrome oxidase I, CytB = cytochrome B, H3 = histone 3, 16S = large mt ribosomal unit, 18S = small nuclear ribosomal unit, 28S = large nuclear ribosomal unit; wing = wingless.

Method (of analysis): C = computer based; H = Hennigian (by direct estimation).

No. char.: Total number of characters in the data matrix.

Inf. char.: Number of informative characters (for molecular data). In square brackets, estimated values where exact data not provided in original publication.

Char./taxon: Number of informative characters per taxon.

Weight: Weight assigned to source tree in weighted ratchet analysis. For mixed morphological and molecular studies, weights for each partition are separated by a slash; the average of these values is the weight assigned to the tree.

*morphological weight = 1; molecular weight = 2; average = 1.5.*
partitions in the study, with the final weight for the study being the average of the weights given to each partition.

All Perl scripts used in this study were written by the third author (unless otherwise noted), and are freely available from http://www.uni-jena.de/~b6biol2/(under “Programs”).

**Results**

We found 43 independent source trees matching our criteria (Table 1). Eight source trees were found to fulfill criteria 1 and 2, but not 3 (i.e., their full data matrix was included in a subsequent, more inclusive analysis), and thus were not considered (Appendix 1). The source trees, with taxonomically equivalent taxa removed according to the rules of STR, included a total of 309 genera in all extant families of Adephaga (excluding the newly described Meruidae). The trees derived from a ratchet search (with either weighted or unweighted source trees) had very similar topologies, with only 42 clades (6.9% of the total possible) appearing on one tree or the other but not both. Both trees remained poorly resolved despite the use of STR (resolution of unweighted and weighted trees 35.5% and 46.7%, respectively). However, the generally high goodness-of-fit measures (Table 2) indicate that the lack of resolution derives more from a lack of information than from conflict between the source trees.

It was possible to unambiguously re-include 93 of the 100 removed STR taxa, resulting in a tree with 401 genera in total (Figs. 1 and 2; the strict and majority-rule consensus supertrees of both unweighted and weighted analyses have also been deposited in TreeBASE; study accession number S1682, matrix accession number M3040). The seven taxa that could not be re-included were *Platyderus*, *Pristonychus*, *Calathus*, *Dolichus*, *Antisphodrus*, *Sphodropsis*, and *Thermoscelis*, all belonging to the extremely species-rich family Carabidae. This procedure increased resolution slightly (resolution of unweighted trees 41.4%, of weighted trees 49.9%). After the inclusion of the removed taxa, the topology of the two supertrees remained very similar to one another, with the same 42 clades present in just one of them.

**Discussion**

**Resolution of the supertrees**

Most nodes were well resolved in the final supertrees for the aquatic families (for which there were more source trees available; see Table 1), but resolution was very low for the basal Carabidae and in Harpalinae. This reflects both the lack of robust phylogenies for the former group, the most species-rich of all Adephaga (see, for example, Ober 2002), and also several groups of poorly known STR taxa in Harpalinae that all share the same index taxon. In the only other published supertree of Coleoptera, Hughes and Vogler (2004) obtained a similar lack of resolution for the species of the acorn weevil genus *Curculio*. In that study, however, the supertree was constructed from only four gene trees, so that the small sample size compounded the loss of information in going from DNA sequence data to tree topologies in a supertree analysis. In the present case, the sample of source trees was much larger and from a wider variety of data types. Thus, the lack of resolution suggests real lack of sufficient information rather than insufficient sample size.

**Phylogeny of Adephaga**

(Figs. 1, 2; for undivided version of Fig. 2 see online edition at: doi:10.1016/j.ode.2006.05.003)

**Inter-family relationships**

In all the resulting supertrees, Gyrinidae were placed as the sister group of the remaining Adephaga, reflecting the prevalence of trees resulting from the analysis of morphological data (e.g. Beutel and Roughley 1988; Beutel 1993; Beutel and Haas 1996). In the available molecular studies (based on 18S rRNA), Gyrinidae were placed as sister group of a clade comprising Haliplidae + Dytiscoidea (Shull et al. 2001; Ribera et al. 2002b).

The position of Haliplidae as sister taxon of Geadephaga + Dytiscoidea reflects morphological studies by Beutel and Roughley (1988) and Beutel (1993), but is in contrast to Beutel and Haas (1996) and analyses of molecular data. Haliplidae have been variously placed within a clade Hydradephaga in Shull et al. (2001), and as sister group of Dytiscoidea in Beutel and Haas (1996) and Ribera et al. (2002b).

Dytiscoidea was recovered in all analyses, with Noteridae placed as the sister group of the remaining families, as in most morphological and molecular studies (e.g. Beutel 1993; Beutel and Haas 1996; Ribera 2002).

Table 2. Statistics relating to parsimony ratchet analyses of MRP matrix with source trees either unweighted or weighted according to size of study

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unweighted</th>
<th>Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPTs</td>
<td>22,709</td>
<td>218</td>
</tr>
<tr>
<td>Length</td>
<td>1,014</td>
<td>1,559</td>
</tr>
<tr>
<td>CI</td>
<td>0.6815</td>
<td>0.7223</td>
</tr>
<tr>
<td>RI</td>
<td>0.9076</td>
<td>0.9101</td>
</tr>
<tr>
<td>RC</td>
<td>0.6185</td>
<td>0.6573</td>
</tr>
</tbody>
</table>

MPTs = number of equally most parsimonious trees.
Fig. 1. Summary tree based on majority-rule consensus of all equally most parsimonious solutions from weighted supertree analysis (see Tables 1 and 2) and after re-inclusion of 93 (of 100) STR taxa. Monophyletic clades pooled according to backbone taxonomy (see online edition of this paper at: doi:10.1016/j.ode.2006.05.003). Taxon names followed by a number indicate number of genera pooled into the clade. For detailed relationships at genus level see Fig. 2.
et al. 2002a, b; Balke et al. 2005). This is in contrast to the traditional treatment of Noterinae as a subfamily of Dytiscidae (e.g. Franciscolo 1979) or of Noteridae and Dytiscidae as sister groups (e.g. Miller 2001). Within the remaining Dytiscoidea, Amphizoidae and Aspidytidae were placed as sister groups, in agreement with Balke et al. (2005) but contrary to Ribera et al. (2002a) (although results of the latter study were not included in the final analyses because all data were subsumed within the former). Hygrobiidae was placed as the sister group of Dytiscidae, in agreement with most morphological data (e.g. Beutel 1986b, 1998; Beutel and Haas 1996) and some molecular studies (Ribera et al. 2002a), but contrary to Balke et al. (2005), who found support for a sister relationship of Hygrobiidae with the clade (Dytiscidae + (Aspidytidae + Amphizoidae)).
In our supertrees, Trachypachidae were the sister group of Carabidae (incl. Rhysodidae), as proposed by, for example, Erwin (1985) and Beutel and Haas (1996) from morphological evidence, and by Shull et al. (2001) from molecular data. This is in conflict with a sister-group relationship between Trachypachidae and Dytiscidae, which in turn was not monophyletic. The non-monophyly of Rhysodidae, with the clade of Clinidium and Omoglymmius grouping with Cicindelinae+Scaritinae, and with Rhysodes placed as sister to Psydrus (= Psydrinae partim) (Fig. 2d), is extremely unlikely considering the numerous larval and adult autapomorphies of this family (e.g. Bell and Bell 1962, 1978; Bell 1991; Beutel 1997, 2005). It is most likely an artefact of how the information for members of this family is distributed among the data partitions (source trees). Specifically, the family is represented by different genera in different source trees (Rhysodes in Bils 1976 and in Baehr 1979; Clinidium and Omoglymmius in Beutel and Haas 1996 and in Maddison et al. 1999; Omoglymmius in Ober 2002), and these genera are placed in different positions in these trees. Together, these two factors outweigh rhysonid monophyly, which is only indicated for the three genera simultaneously in phylogenies that did not meet the criteria for inclusion in the present study. It is worth noting that the backbone taxonomy also specified rhysonid monophyly, thereby supporting our contention that the seed tree can be overruled by other, more robust source trees (if perhaps wrongly in
this case). We would also note that because the root cause of this problem stems from the distribution of the phylogenetic information among the partitions, similar errors are likely to arise even in a combined analysis of the character data.

**Intra-family relationships**

The interrelationships within the family Gyrinidae were well resolved, and did not show any major conflict with published trees (Beutel 1989a,b, 1990 from morphological data; Ribera et al. 2002a,b from molecular data). Gyrinidae, Gyrinini (excl. *Heterogyrus*), and a clade comprising Enhydriini + Orectochilini are monophyletic. However, in Beutel (1990) the genus *Enhydrus* was placed as the sister taxon of the remaining enhydrine-orectochiline tribe, which implies enhydrine paraphyly.

In contrast to Gyrinidae, interrelationships within the family Haliplidae were not well resolved. The sister-group relationship between *Brychius* and *Haliplus* is clearly in contrast to the only morphology-based phylogenetic investigation of the group (Beutel and Ruhnau 1990), which pointed out that *Algosiphon* and *Apteraliplus* are likely subordinate subgroups of the genus *Haliplus* related to the subgenus *Liaphlus*.

The branching pattern within the family Noteridae reflects the only extensive analysis of noterid interrelationship by Belkaceme (1991) (see also Beutel and Roughley 1987), which is fully congruent with the available molecular data (Ribera et al. 2002a,b; Balke et al. 2005).

The high resolution within the species-rich Dytiscidae reflects the wide range of studies dealing with larval morphology (e.g. Nilsson and Hilsenhoff 1991; Alarie 1996; Alarie et al. 1997, 1998, 1999, 2000, 2001a, b, 2002a, b), adult characters (especially female genitalia) (Burmeister 1976; Miller 2000, 2001), molecular data (e.g. Ribera et al. 2002b, 2003, 2004; Ribera 2003; Balke and Ribera 2004; Balke et al. 2004), and combined data sets (Ribera et al. 2002a; Miller 2003; Balke et al. 2005). Matinae were placed as the sister group of the rest of the family, which is in agreement with Miller (2001) based on female genitalia, but in contrast to Ruhnau (1986) and Beutel (1994), who each proposed a basal position for Copelatinae (mostly because of the absence of mandibular sucking channel). A basal position of Copelatinae (along with Laccophilinae) was also found in a consensus tree based on cladograms obtained in the simultaneous analysis of all morphological and molecular data in Miller (2003, fig. 24). In our supertree, Hydroporinae was placed as the sister group of a large clade including all remaining Dytiscidae except Laccophilinae, Hydrodytinae and Matinae. Hydrodytinae was placed as sister to Laccophilinae, in agreement with the molecular data (Balke et al. 2004) but in contrast with morphological analyses, which placed it as sister to Hydroporinae (Miller 2001). Lancetinae was placed as sister to Dytiscinae, and the two combined as sister to Colymbetinae, which in turn were sister of Agabinae (Platynectes, currently placed in Agabinae, was placed as unresolved at the base of this lineage, in agreement with its undefined position in Ribera et al. 2002b). The genus *Coptotomus* (subfamily Coptotominae) was placed within Copelatinae.

Within Hydroporinae, the most diverse subfamily of Dytiscidae, Laccornini and Methlini were placed as sister to the remaining taxa, in agreement with published morphological (e.g. Wolfe 1985; Miller 2001) and some molecular data (Ribera et al. 2002b). The tribe Bidessini was recovered as monophyletic, in contrast to Hydroporini, which was found to be highly polyphyletic. The genus-level relationships of the other smaller groups reflect more or less directly the source trees used in our analyses.

Several major lineages within the family Carabidae, such as Carabinae, Paussinae, Cicindelinae, Trechinae, and Brachininae (placed with Harpalinae partim), were recovered as monophyletic. However, the resolution of this part of the tree is very low. Carabidae conjunctae (sensu Maddison 2004), Harpalinae, Brosicinae, Psydrinae, Scartininae, and Nebrinae (see, e.g., Maddison et al. 1999; fig. 1) do not appear as clades, and some relationships suggested by the branching pattern appear very unlikely. A placement of Gehringiini at the base of the family is not unlikely (e.g. Beutel 1992b), although not in combination with *Cymbionotum* (Melaeninae) as suggested by the supertrees.

**Main phylogenetic questions to be addressed**

As stated in the Introduction, the main objective of this study was to identify areas from the supertree of Adephaga for which the existing data were clearly insufficient (whether due to a lack of consensus or of data as such) and could not provide a well-resolved consensus phylogeny. Although the inter-family relationships were fully resolved, the fact remains that the indicated relationships tend to reflect the more numerous morphological studies, and are at least partly in contrast to molecular investigations. The crucial question as to whether or not the aquatic families (i.e., the ‘Hydradephaga’) form a clade should still be considered as unresolved (as should the position of the terrestrial Trachypachidae). Hydradephaga were postulated based on morphological features in pre-Hennigian studies (e.g. on laterally extended metacoxae in Crowson 1960), rejected based on cladistic analyses of morphological data (e.g. Beutel and Haas 1996), and re-erected in molecular studies (Shull et al. 2001; Ribera et al. 2002b). However, the relevant molecular analyses pointing towards ‘Hydradephaga’ monophyly were
based exclusively on the 18S rRNA gene (Shull et al. 2001; Ribera et al. 2002b), which has been found to be insufficient to reconstruct relationships at the deeper family level in beetles (e.g. Maddison et al. 1999; Vogler 2005; Whiting 2002). A coordinated combined effort of morphologists and molecular systematists is desirable here to obtain a molecular data set comprising several genes together with larval and adult morphological characters, and a taxon sampling scheme that comprises representatives of all adephagan families (including Rhyssodidae and Meruidae) plus a suitable outgroup with members from the other three coleopteran suborders.

One point that appears indubitable in our supertrees is the monophyly of Dytiscoidea, which is strongly supported by both larval and adult morphological features (e.g. Beutel 1997) and also by combined analyses (Ribera et al. 2002a, b; Balke et al. 2005). The recently discovered Aspidytidae clearly belongs to Dytiscoidea (like Meruidae; Spangler and Steiner 2005; Beutel et al. 2006), but its placement within this lineage is not yet unambiguously clarified (see Balke et al. 2005).

One of the greatest enigmas in adephagan phylogeny have been the affinities of the Rhyssodidae (or Rhysodini) (e.g. Bell 1998; Bell and Bell 1962; Beutel 2005), a highly aberrant group that has abandoned the typical predaceous lifestyle of Adephaga and switched to specialized xylobiontic habits. The placement of this group is clearly impeded by numerous autapomorphies of larvae and adults (e.g. Beutel 1997). Interestingly, a similar problem seems to occur in molecular investigations of this group (i.e., long branches; Maddison et al. 1999; Ober 2002). Our analyses were likewise unable to resolve this question satisfactorily, favouring the unlikely scenario of rhyssodid polyphyly, which is doubtlessly an artefact as pointed out above.

With respect to relationships within each of the families, the phylogenies of the smaller (aquatic) groups of Adephaga (Gyrinidae, Haliplidae, Noteridae) seem to be reasonably well established, although corroboration of these hypotheses in future studies based on a numerical approach (Beutel et al. 2006) and molecular data or combined data sets is always desirable. Resolution in the most species-rich aquatic family, Dytiscidae, is comparatively good, but it is evident that a certain amount of work remains to be done. The position of Copelatinae continues to be problematic, especially considering the conflicting results of recent studies (e.g. Miller 2000, 2001, 2003). The relationships within Hydroporinae are also highly unstable, in particular for the species-rich Hydroporini, which were not recovered as monophyletic in any of the trees.

The supertree shows the basal branching pattern within the megadiverse Carabidae (including Rhyssodidae) as well as within its largest subfamily, Harpalinae, to be largely unresolved. Apparently, the extreme complexity and species richness of this lineage in combination with a number of taxa being particularly poorly known makes the reconstruction of its phylogeny a difficult task at present, despite of the concerted efforts of numerous systematists and taxonomists. Again, a major goal for this group, if not also for Adephaga globally, should be an improved coordination (e.g. with respect to taxon and gene sampling) between systematists, especially between groups specialized on morphological and molecular data, respectively. The objective should be to create comprehensive and well-documented morphological matrices for a representative taxon sample, and to combine these characters with molecular data including several genes. Exactly such an integrative approach underlies the recently initiated Beetle Tree of Life project. But even with more comprehensive data sets such approaches will deliver, the sheer size of the phylogenetic problem would probably still require a supertree approach as part of a divide-and-conquer analytical framework (see Bininda-Emonds 2004a). Such a framework is also desirable in that it would allow each data partition to be analysed under the most appropriate model of evolution and optimization criteria, thereby enabling the final result to be built from the most robust source trees possible (Bininda-Emonds 2004a).

Acknowledgements

We thank two anonymous reviewers for their comments that helped to improve the manuscript. O.B.-E. was funded by the BMBF (Germany) through the project “Bioinformatics for the Functional Analysis of Mammalian Genomes”.

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ode.2006.05.003.

References


Alarie, Y., Larson, D.J., 1998. Larvae of Agabimus Crotch: characteristics, descriptions of A. glabellus (Motschulsky), and comparison with other genera of the subfamily.


Appendix B

References for potential source trees not included in the analyses due to redundant data (i.e., trees that fulfilled criteria 1 and 2, but not 3; see Methods). Square brackets at the end of each entry enclose a code number for a source tree derived from a more comprehensive version of the respective same data matrix; the reference for this latter source tree can be identified by finding the code number in the far left column in Table 1.


