Supertrees

Noe de la Sancha
Principles and Practice of Phylogentic
What is a super tree?

- A supertree is a tree that amalgamates, synthesizes, or otherwise represents the phylogenetic relationships included in a set of input trees (Wilkinson et al., 2004).

- Supertree construction is the generation of one or more output trees (the supertrees) from a set of source trees that process fully or partially overlapping sets of taxa (Bininda-Emonds et al., 2002).

Gordon 1986

- Problems
- Published in Mathematical journal
- Not computationally feasible until the last 10 years or so.
- Method limited to overlapping trees which were compatible.
Indirect Methods

- MiniCut Supertrees (Semple and Steel, 2000).
- MinFlip Supertrees (Chen et al., 2001).
- Average consensus (Lapointe and Cucumel, 1997).
**Matrix representation with parsimony**

**MRP**
- Baum, 1992; Doyle, 1992; and Ragan, 1992
- MRP - makes use of additive binary coding to represent a given tree in matrix format.
- The “matrix representations” of the different source trees are then combined into a single matrix that can be analyzed using any desired optimization criterion (but usually parsimony).
- This removes Gordon’s 1986 strict supertree method

Trees can be overlapped even if they conflict.

---

**MRP**
- Nodes of each source tree are encoded as follows:
  - Taxa descended from the focal node score 1.
  - Those not presented elsewhere in source tree score 0 (assuming this is a primitative state).
  - All other taxa scored missing.
  - A fictitious all-zero outgroup is added to the matrix to polarize the subsequent parsimony analysis (Bininda-Edmonds, 2002).
Purvis 1995

- MRP supertree of all 203 extant species of primates represented the first supertree of a clade.

- It large size, unprecedented completeness, high amount of resolution should supertrees could be acheived

- Showed biological applications
- Purvis et al, 1995 reference for super trees
  - Used to test other methods (Moore et al., 2004; Vos and Mooers, 2004)
Problems Supertrees

- Taxon sampling.
- Computer power.
- Cumulative problems from source trees.

Critiques

- Super trees lose contact with the primary data.
- Present a useful summary of the source trees, rather than accurate phylogenetic reconstruction.
- Do not provide strong tests of previous phylogenetic hypotheses because they are not based on new data sets.
- How do you interpret in terms of biological meaning of any homoplasy.
Table 1: Examples of supertrees and their applications

<table>
<thead>
<tr>
<th>Supertree</th>
<th>Taxon</th>
<th>Level of type (size)</th>
<th>Application</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purvis et al. (1999)</td>
<td>Primates</td>
<td>Species (201)</td>
<td>Evolutionary rates (life histories, brain size)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exclusion risk</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Immune system function</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exclusion and speciation</td>
<td>1, 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Brain size and behavior</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Allometry and base range size</td>
<td>9</td>
</tr>
<tr>
<td>Bininda-Emonds et al. (1999)</td>
<td>Carnivora</td>
<td>Species (271)</td>
<td>Evolutionary rates (life histories)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Body size and species richness</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Taxonomic differences</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Extinction risk</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exclusion rates</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Species richness</td>
<td>superree study</td>
</tr>
<tr>
<td>Kreft et al. (1997)</td>
<td>Marsupiala</td>
<td>Species (51)</td>
<td>Extinction and range size</td>
<td>4</td>
</tr>
<tr>
<td>Jones et al. (2002b)</td>
<td>Chiroptera</td>
<td>Species (925)</td>
<td>Extinction risk</td>
<td>13</td>
</tr>
<tr>
<td>Liu et al. (2001)</td>
<td>Mammals</td>
<td>Families (95)</td>
<td>Phylogeny</td>
<td>superree study</td>
</tr>
<tr>
<td>Welhins et al. (2000)</td>
<td>Monocotyledons</td>
<td>Species (918)</td>
<td>Evolution of breeding systems</td>
<td>superree study</td>
</tr>
<tr>
<td>Welh (2000)</td>
<td>Trees</td>
<td>Species (324)</td>
<td>Structure of ecological communities</td>
<td>superree study</td>
</tr>
</tbody>
</table>

*1, Purvis et al. (1999); 2, Benton (1996); 3, Gittleman & Purvis (1991); 4, Johnson (1991); 5, Parrish (1991); 6, Desper & Nunn (1999); 7, Bininda-Emonds & Goldman (2000); 8, Nunn & Benton (2000); 9, Tate et al. (2003); 10, Purvis et al. (2004a); 11, Purvis et al. (2008); 12, Purvis et al. (2000b); 13, Jones et al. (2002).*

### Bininda-Emonds et al., 1999

- **Supertree of order Carnivora**

- **Problems too taxonomic for evolutionary journals too evolutionary for taxonomic journals.**

- **271 species**
A

B

Fig. 12. The composite tree for Virocarinae. Node numbers refer to Table 15; other details are as in Fig. 1.

Fig. 13. The composite tree for all 275 extant species of coriaceae, including estimated times of divergence and showing which lineages have given rise to significantly more extant descendants than expected (hashed squares, P < 0.05). Filled squares, P < 0.01). Within major taxa, species are presented in the same order as in Figs 1-12. Dates were estimated from the UTPS model (Table 15) or net a pure birth model. Negative branch lengths are shown as having zero length. See text for further details.
A phylogenetic supertree of the bats
(Mammalia: Chiroptera)

KATE E. JONES²,³,³, ANDY PURVIN², ANN MacLARNON³, OLAF R. P. BININDA-EMONDS³ and NANCY B. SIMMONS³¹

¹Department of Biology, Imperial College at Silwood Park, Ascot, Berkshire, SL7 1TA, UK
²School of Life Sciences, University of Surrey, Guildford, West Sussex GU2 7XH, UK
³Institute of Evolutionary and Ecological Sciences, Leeds University, Leeds LS6 9JS, UK

(RECEIVED 28 June 2001; revised 21 November 2001; accepted 12 December 2001)

ABSTRACT

We present the first estimate of the phylogenetic relationships among all 926 extant and nine recently extinct species of bats (Mammalia: Chiroptera), a group that accounts for almost one-quarter of existent mammal diversity. This phylogeny was derived by combining 105 estimates of bat phylogenetic relationships published since 1970 using the supermatrix construction technique of matrix representation with Parsimony (MRBP). Despite the explosive growth in the number of phylogenetic studies of bats since 1960, phylogenetic relationships in the order have been studied rate-randomly. For example, over 96% of all bat systematic studies to date have focused on relationships within Phyllostomidae, whereas relationships within clades such as Rhinolophidae and Vespertilionidae have been studied less extensively. Resolution in the supermatrix is lower than among clades overall (resolution is poor 46.4%, it’s a fully bifurcating solution) but reaches 100% in some groups (e.g. relationships within Microchiroptera). The supermatrix analyses are less adequate for a more common solution, Micropipitea is paraphyletic with respect to Microchiroptera, as the majority of sources topologies support microchiropteran monophyly. Although this is not a substitute for comprehensive phylogenetic analyses of primary molecular and morphological data, the bat supermatrix provides a useful tool for future phylogenetic comparative and macroevolutionary studies. Additionally, it identifies clades that have been little studied, highlights groups within which relationships are controversial, and like all phylogenetic studies, provides preliminary hypotheses that can form starting points for future phylogenetic studies of bats.

Key words: bats, evolution, matrix representation, parsimony, phylogeny, supertree construction.

(1) Taxonomic coverage and resolution

Fig. 1. Cumulative numbers of usable sources published since 1970 (N = 105). The solid line represents the total number of sources; solid boxes represent cumulative number of sources based on analyses of morphological data only; open boxes represent sources based on analyses of molecular data only; and crosses represent sources based on analyses combining both morphological and molecular data in a total evidence approach.

A. The supertree for Phyllostomidae excluding Phyllotominae and Streomiasminidae. Relationships within studied clades are shown in B and C, respectively.
Bat supertree

Fig. 3. The supertree for higher-level relationships. The number below each node represents the node number.
Molecular and Morphological Supertrees for Eutherian (Placental) Mammals

Fu-Guo Robert Liu,1 Michael M. Miyamoto,1* Nicole P. Freire,1
Phong Q. Ong,1 Michele R. Tennant,2 Timothy S. Young,1
Kikumi F. Gugel1

A large body of diverse comparative data now exists for a major phylogenetic synthesis of the higher-level relationships among eutherian (placental) mammals. We present such a phylogenetic synthesis using the composite trees or supertrees from the combined and separate analyses of their published molecular and morphological source phylogenies. Our combined and separate supertrees largely support the same superfamilial taxa and orders, but different interordinal clades. These similarities and differences reinforce the continuing contributions of morphological studies, while highlighting the growing influence of molecular information on the field. As current summaries of past research, our supertrees emphasize opportunities for future work, while providing a step toward the eventual integration of the data and characters themselves.
Molecular phylogenetics and the origins of placental mammals

William J. Murphy 1,2, Eduardo Izarzugaza 1,2, Warren E. Johnson 3, Ta Ping Zhang 4, Oliver A. Ryder 4 & Stephen J. O’Brien 4

1 Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702, USA
2 Key Laboratory of Cellular and Molecular Evolution, Knowledge Institute of Zoology, Chinese Academy of Sciences, Beijing, China
3 Center for Reproduction of Endangered Species, Zoological Society of San Diego, San Diego, California 92113, USA
4 Department of Biology, University of Maryland, College Park, Maryland 20742, USA

* These authors contributed equally to this work

The precise hierarchy of ancient divergence events that led to the present assemblage of modern placental mammals has been an area of controversy among morphologists, paleontologists and molecular evolutionists. Here we address the potential weaknesses of limited character and taxon sampling in a comprehensive molecular phylogenetic analysis of 64 species sampled across all extant orders of placental mammals. We examined sequence variation in 18 homologous gene segments (including nearly 16,000 base pairs) that were selected for maximal phylogenetic informativeness in resolving the hierarchy of early mammalian divergence. Phylogenetic analyses identify four primary superordinal clades: (i) Afrotheria (elephants, manatees, hyraxes, tarsiers, aardvarks and elephant shrews); (ii) Xenarthra (armadillos, anteaters and sloths); (iii) Glires (rodents and lagomorphs), as a sister taxon to primates, flying lemurs and tree shrews; and (iv) the remaining orders of placental mammals (artiodactyls, perissodactyls, caniforms, pangolins, bats and cetea inactinomorphs). Our results provide new insight into the pattern of the early placental mammal radiation.

The phylogeny of morphological, ecological and genomic diversity among extant mammals offers considerable potential for studies of speciation, adaptation, molecular evolution, genome organisation and biogeography 1–3. Studies on morphology and both mitochondrial and nuclear genes have revealed several higher-level phylogenetic associations 4–9, but a full resolution of the earliest placental...
A species-level phylogenetic supertree of marsupials

Marcel Cardillo1,2,*, Olaf R. P. Bininda-Emonds3, Elizabeth Bouvier1,2 and Andy Purvis1

1 Department of Biological Sciences, Imperial College London, Silwood Park, Ascot SL5 7PY, U.K.
2 Institute of Biology, Zoological Society of London, Regent’s Park, London, NW1 4NY, U.K.
3 Lehrstuhl für Tierethologie, Technische Universität Munich, Albert-Ludwigs-University, Germany

(Accepted 26 January 2004)

Abstract

Comparative studies require information on phylogenetic relationships, but complete species-level phylogenetic trees of large clades are difficult to produce. One solution is to combine algorithmically many small trees into a single, larger supertree. Here we present a virtually complete, species-level phylogeny of the marsupials (Mammalia: Metatheria), built by combining 258 phylogenetic estimates published since 1990, using matrix representation with parsimony. The supertree is well resolved overall (73.7%), although resolution varies across the tree, indicating variation both in the amount of phylogenetic information available for different taxa, and the degree of conflict among phylogenetic estimates. In particular, the supertree shows poor resolution within the American marsupial taxa, reflecting a relative lack of systematic effort compared to the Australian taxa. There are also important differences in the supertrees based on source phylogenies published before 1995 and those published more recently. The supertree can be viewed as a meta-analysis of marsupial phylogenetic studies, and should be useful as a framework for phylogenetically explicit comparative studies of marsupial evolution and ecology.

Key words: comparative studies, matrix representation with parsimony, Metatheria, QS support

![Phylogenetic supertree of marsupials](image)

**Fig. 2.** Unweighted supertree relationships among the families and orders of marsupials recognized by Wilson & Reeder (1993). Note that Potomidae and Peramelemorphiae are paraphyletic. Branch lengths are arbitrary. Nodes are numbered sequentially.
Fig. 3. Family-level topologies for the mammal superfamilies weighted by method and data quantity, and for "old" (source trees pre-1993).

Fig. 4. Unrooted species relationships for the families Tetrapodidae and Heterocephala.

Fig. 5. Unrooted species relationships for the families Microsphenidae and Rhesodonta.
Fig. 5. Unweighted supertree relationships for the families Megalobatidae, Symposiidae, Leucorineae and Notoryctidae. One (Megalobates rufipes) has been omitted from this part of the supertree under soft taxonomic reduction.

Fig. 6. Unweighted supertree relationships for the families Perarycteridae and Peramidae. One species (Peramys eximius) has been omitted from this part of the supertree under soft taxonomic reduction.

Fig. 8. Unweighted supertree relationships for the families Peramidae and Perarycteridae. One species (Peramys eximius) has been omitted from this part of the supertree under soft taxonomic reduction.
The delayed rise of present-day mammals

Olaf R. P. Bininda-Emonds¹, Marcel Cardillo², Kate E. Jones¹, Ross D. E. MacPhee³, Robin M. D. Beck³, Richard Grenyer⁴, Samantha A. Price³, Rutger A. Vos⁵, John L. Gittleman⁶ & Andy Purvis⁷

Did the end-Cretaceous mass extinction event, by eliminating non-avian dinosaurs and most of the existing fauna, trigger the evolutionary radiation of present-day mammals? Here we construct, date and analyse a species-level phylogeny of nearly all extant Mammalia to bring a new perspective to this question. Our analyses of how extant lineages accumulated through time show that net per-lineage diversification rates barely changed across the Cretaceous/Tertiary boundary. Instead, these rates spiked significantly with the origins of the currently recognized placental superorders and orders approximately 93 million years ago, before fading and remaining low until accelerating again throughout the Eocene and Oligocene epochs. Our results show that the phylogenetic ‘fuses’ leading to the explosion of extant placental orders are not only very much longer than suspected previously, but also challenge the hypothesis that the end-Cretaceous mass extinction event had a major, direct influence on the diversification of today’s mammals.
Supertrees