

The potential and peril of the supertree approach: A response to van der Linde and Houle

PATRICK M. O'GRADY, RICHARD T. LAPOINT and GORDON M. BENNETT

Insect Syst. Evol. O'Grady, P. M., Lapoint, R. T., Bennet, G. M.: The potential and peril of the supertree approach: A response to van der Linde and Houle. *Insect Syst. Evol.* 39: 269-280. Copenhagen, October, 2008. ISSN 1399-560X.



P. M. O'Grady^{1,2}, R. T. Lapoint¹, G. M. Bennett¹

¹University of California, Department of Environmental Science, Policy and Management, 137 Mulford Hall, Berkeley, CA 94720. tel: (510) 642-0662, email: ogrady@nature.berkeley.edu

²To whom correspondence should be addressed.

"False facts are highly injurious to the progress of science, for they often endure long; but false views, if supported by some evidence, do little harm, for every one takes salutary pleasure in proving their falseness; and when that is done, one path towards error is closed and the road to truth is often at the same time opened."

Charles Darwin (1871)

Introduction

The family Drosophilidae currently contains over 3500 described species distributed in most habitat types (Bächli 2008). Members of the genus *Drosophila* have been the subject of genetic and developmental research for the past century and this group also serves as a model system for evolutionary and comparative studies (Markow & O'Grady 2007). Phylogenetic relationships within Drosophilidae, with particular emphasis on the largest genus, *Drosophila*, have been examined since Sturtevant's taxonomic work in the 1920s (reviewed in Markow and O'Grady 2006). Throckmorton (1962, 1966, 1975, 1982) formalized morphological relationships in this family in a series of landmark papers. The first molecular studies on this group (Hubby 1963; Hubby and Lewontin 1966; Hubby and Throckmorton 1965; Lewontin and Hubby 1966; Throckmorton 1962) led the way to extensive allozyme work in the 1960s and 1970s (reviewed in Powell 1997). This work was followed by restriction fragment polymorphism studies in the early 1980s (e.g., DeSalle et al. 1987), widespread use of DNA sequences in the 1990s (DeSalle 1992; Pelendakis et al. 1991; Pelendakis and Solignac 1993; Remsen and O'Grady 2002; Tatarenkov et al. 1999) and full genome sequences in the early 21st century (Adams et al. 2000;

Pollard et al. 2006; *Drosophila* 12 genome Consortium 2007). These studies have continually refined our notions of phylogenetic relationships within this group. Early molecular phylogenetic work (Pelendakis et al. 1991; and Solignac 1993) suggested that the genus *Drosophila* was not a monophyletic group and subsequent studies have corroborated these findings (e.g., Remsen and O'Grady 2002).

Coincident with these studies was the widespread acceptance of phylogenetic methods and their applications to evolutionary biology. Systematists began revising para- and polyphyletic groups in favor of monophyletic assemblages in order to generate taxonomic hierarchies that are concordant with evolutionary relationships. The bulk of these revisionary efforts have been driven by primary character information (i.e., morphological or molecular characters) and follow traditional taxonomic and nomenclatural process. Recently supertree methods, the most common of which codes source trees into a presence-absence matrix represented using parsimony (MRP), have been introduced to summarize the results of primary data and generate hypotheses based on secondary analyses (phylogenetic trees). While this may be acceptable for character analysis in a comparative framework, it has some serious drawbacks includ-

ing lack of a statistical framework (but see Moore et al. 2006), reliance on secondary sources of information, and non-independence between studies that partially overlap in taxon and character sampling.

Our critique of the van der Linde and Houle (2008) paper is based on three main points: (1) issues with supertree methods in general, (2) issues with the specific methodology employed by van der Linde and Houle, and (3) their approach of using of supertree data as the basis of taxonomic and nomenclatural revisions.

Supertrees vs. supermatrices

Supertrees are generated by coding multiple input tree topologies into a single character matrix and analyzing these data using a parsimony algorithm (Sanderson et al. 1998; Bininda-Emonds et al. 2002). The strength of such an approach is that a single composite phylogeny can be generated from many studies with different gene and taxon sampling strategies. The supermatrix approach (Gatesy et al. 2002) is an alternate method to summarize multiple phylogenetic studies. In a supermatrix analysis primary character information from different studies is combined into a large matrix and analyzed using phylogenetic methods. The benefits of this approach are that analyses can be placed in a statistical context, analytical methods that accommodate rate heterogeneity and different rates of evolution can be used, and conflicts between characters can be resolved in a total evidence framework. If researchers do decide to pursue a supertree approach, an understanding of the basic assumptions, and potential pitfalls of these methods is required.

Non-independence. – Supertree methods, while they provide an adequate summary of previous studies may not constitute independent corroboration of research findings. This is particularly true if the input trees are based on overlapping data matrices. This is an issue common to many studies and can lead to bias in resultant supertrees topology due to the effective overweighting of some data and underweighting of others. Gatesy and Springer (2004) state “all redundant source data are unacceptable and must be eliminated to avoid the non-independence problem.” While others (e.g., Bininda-Emonds et al. 2004) consider non-independence to be a major issue in tree combina-

tion, they point out that it is usually impossible to eliminate. Bininda-Emonds et al. (2004) present an explicit set of guidelines for minimizing the impact of overlapping data including selecting different genes that do not share a common evolutionary history, including non-overlapping taxa sets derived from the same gene, coding total evidence analyses generated by unique or partially overlapping data sets, picking the most recent study or most taxonomically inclusive analysis for a given locus, or using the hypothesis explicitly preferred by the authors. While not all authors may agree with these criteria, it is important that they be explicitly stated in their methods so that non-independence can readily be assessed and supertree analyses can be repeated by subsequent workers.

Statistical support. – Statistical support in phylogenetic analyses, including supertree analyses, is critical in expressing confidence in tree topology and in subsequent inference. Recent studies describe ways in which uncertainty can be expressed on a supertree using bootstrapping procedures (Burleigh et al. 2006; Moore et al. 2006). Three approaches have been described, each differing in the way in which they deal with issues of supertree reconstruction. The first, referred to as bootstrap-weighted MRP, weights data columns in the matrix proportionally to nodal support in source trees. This method, while it does not provide a direct estimate of nodal support, incorporates support in analyses of primary data into the supertree analysis (Moore et al. 2006; Ronquist 1996; Bininda-Emonds and Bryant 1998). This is preferable, because treating all nodes as equally supported, regardless of whether support at that node in a phylogenetic study was 100% or 51%, may have the effect of downweighting well supported nodes and upweighting those that are not statistically supported.

Two other methods allow for direct approximation of nodal support from the MRP matrix using source-tree bootstrapping and hierarchical bootstrapping (Moore et al. 2006; Burleigh et al. 2006; Page 2004). Source tree bootstrapping is similar in spirit to standard bootstrapping procedures, where resampled data columns are representative of input trees rather than amino acid or nucleotide sites. Hierarchical bootstrapping draws trees from original bootstrap sets derived from primary data in the creation of a supertree bootstrap matrix (Page

2004). In empirical and simulation studies supertree support measures generally produce lower support, more trees, and more ambiguity (Burleigh et al 2006). Hierarchical bootstrapping methods generally outperform source-tree bootstrapping (Moore et al. 2006). Regardless, bootstrapping supertrees provides some measure of support for hypotheses and conclusions and should be performed. Each of the three methods described here are implemented in the program tREeBOOT (Smith and Moore, submitted).

The van der Linde and Houle approach

van der Linde and Houle's supertree methods are flawed in at least three major ways: (1) the criteria for how a study was selected for inclusion in the MRP matrix was not stated in the paper, (2) several analytical errors were made in how searches were conducted on their MRP matrix and (3) there was no effort made to estimate statistical support on their supertree topology.

Input tree bias. – Supertree methods in general suffer from a “garbage in – garbage out” problem in that the quality of studies coded into the MRP matrix are directly related to the quality of the resultant supertree (Bininda-Emonds et al. 2004). Most supertree analyses must, therefore, winnow down the phylogenies that are coded into the supertree matrix to those that contain trees based on different genes, unique combinations of genes from published phylogenies, and/or trees constructed with rigorous phylogenetic analyses (Bininda-Emonds et al 2004). The van der Linde and Houle (2008) study contains no explicit statement concerning the selection of input trees.

Outgroup selection is critically important in order to rigorously test the monophyly of the ingroup and to polarize characters on a phylogeny (Maddison et al 1984). van der Linde and Houle (2008) employ a composite outgroup strategy in order to root their supertree. Aside from issues with composite taxa (see below), this can create a bias, depending on how outgroups were designated in the original analyses. Many drosophilid phylogenies use outgroups that are nested within the genus *Drosophila* as it is currently circumscribed. For example, Tatarenkov et al. (1999) used the Hawaiian *Drosophila* as an outgroup while Kambysellis et al (1995) used *D. melanogaster*. Both taxa are in distinctly different parts of the phyloge-

ny and neither is outside of the genus *Drosophila*. Combining these with outgroups from outside the family Drosophilidae (e.g., *Ceratitis*; Kwiatowski & Ayala 2001) can influence the signal at the root node and may create bias in the resultant supertree.

Bininda-Emonds and others suggest that, when a single study publishes multiple phylogenies, only those preferred by the original authors should be included in the supertree reconstruction (Bininda-Emonds et al. 2004; Gatesy et al. 2002). This is particularly important when including studies that analyze several independent loci and then combine them into a single “total evidence” analysis. van der Linde and Houle (2008) do not follow this suggestion. Instead, they code all individual trees, even when the original authors cited the combined analysis as preferred. While we agree that authors taking a supertree approach should be free to select their input trees, we also strongly feel that some discussion of how these were selected should be included. Such discussion is especially important when the supertree authors may disagree with the original authors. No such discussion is evident in the van der Linde and Houle (2008) paper.

When selecting input trees, several authors (e.g., Gatesy et al. 2002) have advocated only using published phylogenies. Restricting source trees to the published literature means that some sort of peer review process has vetted the phylogeny and the methods employed to generate source trees are available for the community to assess. One or more studies used to generate the present supertree (e.g., van der Linde et al. *flexa* and van der Linde et al. *quadrilineata*) are unpublished and violate this basic tenet of supertree construction.

Many of the datasets utilized by van der Linde and Houle (2008) overlap with respect to the genes employed. Although an effort is made to exclude some studies, weighting is applied to reduce the problem of non-independence of data. Decisions about exclusion and weighting should be clearly defined and based on logically defensible criteria. This is lacking in the submitted paper. They used a total of 117 source trees, drawn from “over 120 studies” to build their supertree. There is no explicit statement of how they selected these trees. Together, these 117 trees included over 620 species [they state “624” (line 67), “623” (line 146), and “about 623” (line 120)]. A total of 185 spp. were in a single tree, 141 were in two trees and 75 were present in three trees. The summary of their

supertree analysis includes 293 taxa (they say 297 on line 147 and their matrix includes 296 taxa).

Analytical biases. – There are also several clear analytical biases in this study that may have a significant impact on the phylogenetic hypothesis proposed by van der Linde and Houle (2008). These include an incomplete search of tree space and a non-repeatable weighting scheme. When performing a phylogenetic analysis it is critically important to thoroughly and aggressively search tree space in order to discover all equally parsimonious solutions. It is known that heuristic algorithms are not guaranteed to discover the shortest trees (Felsenstein 2004) and Maddison (1991) has clearly demonstrated that multiple islands of equally parsimonious trees can exist for a single data matrix. Failure to search tree space adequately erroneously represents information provided by the data. van der Linde and Houle (2008) constructed starting trees using closest addition. Closest, or simple, additions are not sufficient because the starting trees will always be the same and only a single area of tree space, containing very similar trees, will be explored. Furthermore, van der Linde and Houle (2008) only performed a single replicate. If only some of the most parsimonious trees are discovered in an analysis, resolution of the resulting phylogenetic hypothesis can be overestimated. In order to more fully explore tree space multiple, random additions should be performed (Maddison 1991). While they don't specifically detail most of the methods, it is clear that their heuristic search strategy was less than exhaustive and additional equally parsimonious phylogenies likely exist in the "tree space" for their matrix (Maddison 1991).

Another issue with this study is that they do not specify the exact number of trees discovered. Instead, they state that "over 20,000" trees were found. This suggests that their search was not allowed to go all the way to completion. If this is the case, the trees presented may also be more resolved than the data actually support, meaning that some nodes in the van der Linde and Houle (2008) phylogeny are not actually supported by the data.

Finally, van der Linde and Houle (2008) make a compelling case for including only those analyses that use species as terminal taxa in their supermatrix. Some studies (e.g., Remsen and O'Grady 2002) are excluded because they use composite

taxa. We agree with the rationale behind such an approach but are confused by van der Linde and Houle's composite outgroup approach. Using composite taxa, whether they are in the ingroup or outgroup, makes an assumption of monophyly and should be avoided unless it is clearly justified.

Weighting bias. – Supertree analyses often weight nodes in a phylogeny based on how well supported the relationships are in the original studies (Ronquist 1996; Bininda-Emonds and Sanderson 2001). This allows for some incorporation of statistical support (see above) and brings the supertree analysis in closer correspondence with primary data information (Cotton and Page 2002; Page 2004). Advocates for weighting approaches also consider the amount of data in source trees, suggesting weighting schemes consider total sequence length or the number of variable characters (Burleigh et al. 2006).

van der Linde and Houle used a novel weighting scheme in an attempt to ameliorate the problem of non-independence between studies employing the same gene or sets of genes to infer phylogeny. They apply weights to the entire matrix based on the number of genes involved. A weight of 1 is given to an analysis that uses 1 gene or morphology, 2 for 2 genes and so on. Studies with more extensive gene sampling were given higher weights, up to a maximum weight of 5. Such an approach is subjective because a considerable gap in weighting exists for those studies using between 5 genes and a large portion of the entire genome (~9000 genes). In fact, in the van der Linde and Houle paper, specifically what constitutes a large enough dataset to warrant a weight of 5 relative to one with a weight of 4 is never explicitly discussed. Furthermore, giving a genome wide analysis a weight of 5 and a morphological analysis a weight of 1 is problematic when one considers the difference in taxon sampling between these studies. The largest genome study included a total of 12 species; the largest morphological study had over 180 taxa. This approach may lead to datasets with dense taxon sampling being down weighted relative to those where many genes and few taxa were sampled. This may lead to an amplification of issues involved with poor taxon sampling (long branch attraction, incorrect placement of taxa whose genomes haven't been sequenced, etc.).

A weighting scheme simply based on the number of genes employed in source studies is prob-

lematic, however, as the number of genes used is not necessarily a proxy for the number of characters used in the primary data sets or for how well those studies may be supported based upon primary character data. Finally, the authors mention also doing an unweighted comparison, but the differences between these two supertrees, if any, are never discussed.

Statistics. – Despite several possible statistical approaches, van der Linde and Houle (2008) do not generate any confidence intervals or statistical support values for nodes on the resultant supertree. Without support values for taxonomic relationships in their supertree hypothesis, the authors are limited in their ability to draw conclusions from their results. Support methods, described above, are easily implemented using programs like PAUP for source-tree bootstrapping, or tREeBOOT, which can perform all three described methods (Burleigh et al. 2006, Moore et al. 2006). Instead of assessing support, van der Linde and Houle (2008) suggest that supertrees present an “average” of the nodes supported in the literature. This is not precisely the case. Under such a scenario, highly supported nodes are effectively given the same weight as those with little or no statistical support. This may weight nodes that are present in all trees, but not statistically supported in any, over those with strong support in some, but not all analyses. Thus, readers and researchers still have no way of interpreting the validity of taxonomic conclusions presented in their paper’s conclusion.

Reanalysis of van der Linde & Houle’s MRP matrix.

We repeated van der Linde and Houle’s (2008) weighted analysis using their original data matrix. We ran a single random addition replicate using TBR branch swapping and discovered 106,659 equally parsimonious trees (4155 steps) before terminating the search after only 24 hours. This suggests that there are far more than 20,000 trees, as stated by van der Linde and Houle (2008), and may mean that the phylogeny they present is more resolved than the data warrant.

Even though we discovered more than ten times more equally parsimonious trees than van der Linde and Houle (2008), it is clear that tree space has still not been exhaustively searched. In fact, the 106,659 trees we did find are likely to be very

similar to one another. In order to avoid having the analysis stuck on a single large plateau of equally parsimonious trees and to explore tree space for more dissimilar, yet equally parsimonious, phylogenies we performed a series of random addition replicates where the maximum number of trees was set to either 1000 or 5000. Each replicate was allowed to swap completely on all trees in memory in an effort to find shorter trees. In addition, we performed random addition searches swapping on only 100 trees and saving a maximum of 10000 trees if trees of 4155 steps or shorter were discovered. These analyses are summarized in Table 1. Out of a total of 556 random addition searches, only 24 (4.3%) discovered trees of 4155 steps, the most parsimonious in our reanalysis. Of the 239,659 total trees found, 225,803 (94.2%) were unique. This search strategy was by no means comprehensive, as the percentage of unique trees discovered in each search was over 87%. This suggests that many more equally parsimonious trees await discovery for this data matrix and that the phylogeny presented by van der Linde and Houle (2008) was far more resolved than their data would support.

Similar results were recovered in the reanalysis of the unweighted supertree matrix (Table 1). A single random addition replicate using TBR branch swapping discovered 550,236 trees after only 20 hours. Combining these trees with those derived from subsequent searches setting maxtrees to either 1000, 5000, or 10000 trees yielded 729,584 novel trees. Asterisks on Figure 1 indicate nodes collapsing in the unweighted strict consensus.

The strict consensus of the 225,803 (weighted; Figure 1) and 729,584 (unweighted; Figure 1, asterisks) unique most parsimonious trees discovered in our analyses are far less well resolved than presented in van der Linde and Houle (2008). This includes the *immigrans-tripunctata* and *virilis-repleta* clades, the genus *Scaptomyza*, and large portions of the *melanogaster* and *obscura* species groups. While the basal nodes are still resolved, and still support van der Linde and Houle’s assertion that the genus *Drosophila* is paraphyletic, we would caution against taking this as support for their taxonomic redesignation. Based on the large number of equally parsimonious trees that likely exist, the biased analytical methods employed (e.g., weighting), and the lack of statistical support, their topology should be considered tentative at best.

Table 1. Summary of maximum parsimony reanalyses

Weighted Analysis							
# Char (# Trees) ¹	# Taxa ²	Tree Length ³	Maxtrees ⁴	# Reps (# MPT) ⁵	% MPTs ⁶	# Trees (# Unique) ⁷	% Unique ⁸
3343 (181)	296	4155	None	1 (1) ⁹	100	106,659 (106,659)	100
			10,000	11 (6)	54.5	60,000 (58,484)	97.5
			5,000	40 (14)	35	70,000 (63,664)	90.9
			1,000	514 (3)	0.6	3,000 (2,635)	87.8
			Total	556 (24)	4.3%	239,659 (225,803) ¹⁰	94.2%
Unweighted Analysis							
1919 (115)	296	2543	None	1 (1) ⁹	100	550,236 (550,236)	100
			10,000	10 (9)	90	90,000 (89,096)	99
			5,000	85 (18)	21.2	90,000 (89,196)	99.1
			1,000	394 (16)	4.1	16,000 (13,576)	84.9
			Total	490 (44)	9%	746,236 (729,584) ¹⁰	97.7%

1. Number of characters in MRP (matrix representation with parsimony) matrix; number of trees used to generate matrix.
2. Number of taxa analyzed.
3. Length of most parsimonious tree(s) discovered.
4. Maximum number of shortest trees (Maxtrees) held and swapped on each replicate.
5. Number of random addition sequence replicates (# Reps) performed under each maxtree setting; If trees equal to length of MPT were not found, then non-optimal trees were discarded and the next replicate was started.
6. Percent of replicates finding an island equal to the most parsimonious trees (# MPT).
7. Total number of MPTs discovered using each maxtree setting (# Trees); total number of unique MPTs (# Unique). Unique trees were identified using the "condense trees" command in PAUP*, version 4.0, (Swofford, 2008).
8. Percent of unique trees (% unique) discovered under each maxtree setting.
9. This search was terminated prematurely because of the large number of most parsimonious trees in memory. Allowing branch swapping to continue would have led to the discovery of many more equally parsimonious trees.
10. The number of unique trees is not additive because of overlap between trees from different searches.

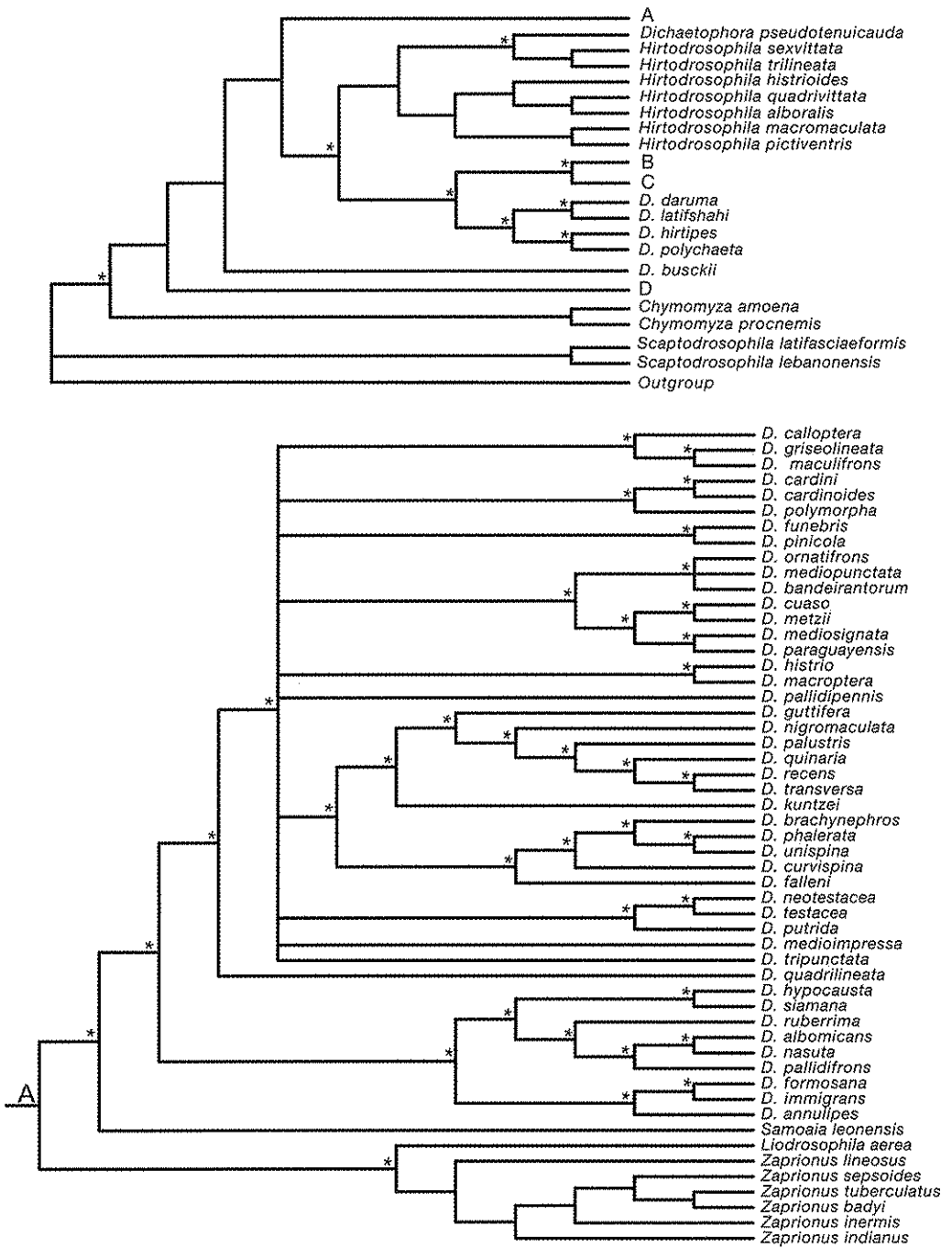
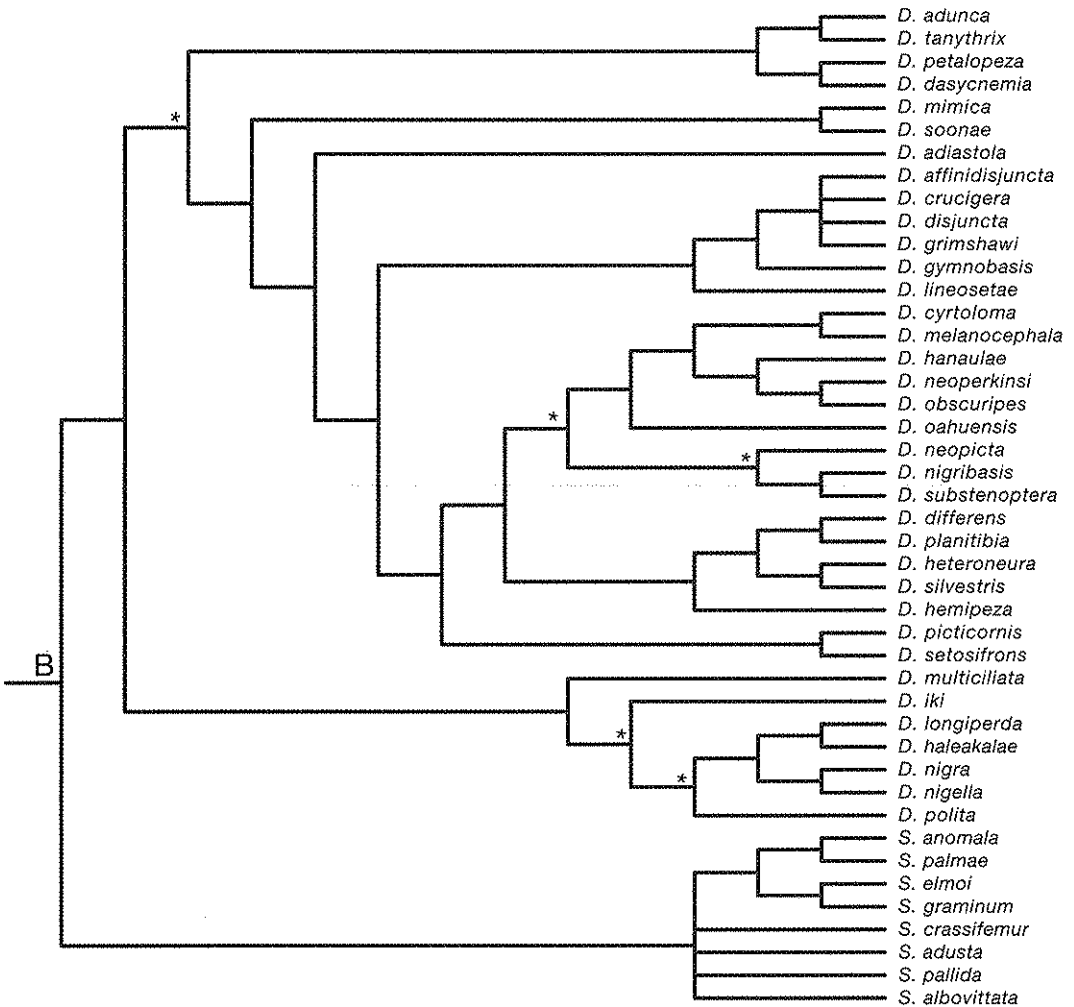


Figure 1: Strict consensus of 225,803 unique trees discovered in the reanalysis of van der Linde and Houle's weighted supertree matrix. While the backbone relationships presented by van der Linde and Houle are still evident, many equally parsimonious trees likely remain undiscovered. Asterisks indicate nodes that collapse in the strict consensus of 729,584 trees discovered in the unweighted supertree analysis. Little support for the genera proposed by van der Linde and Houle is evident in these analyses.

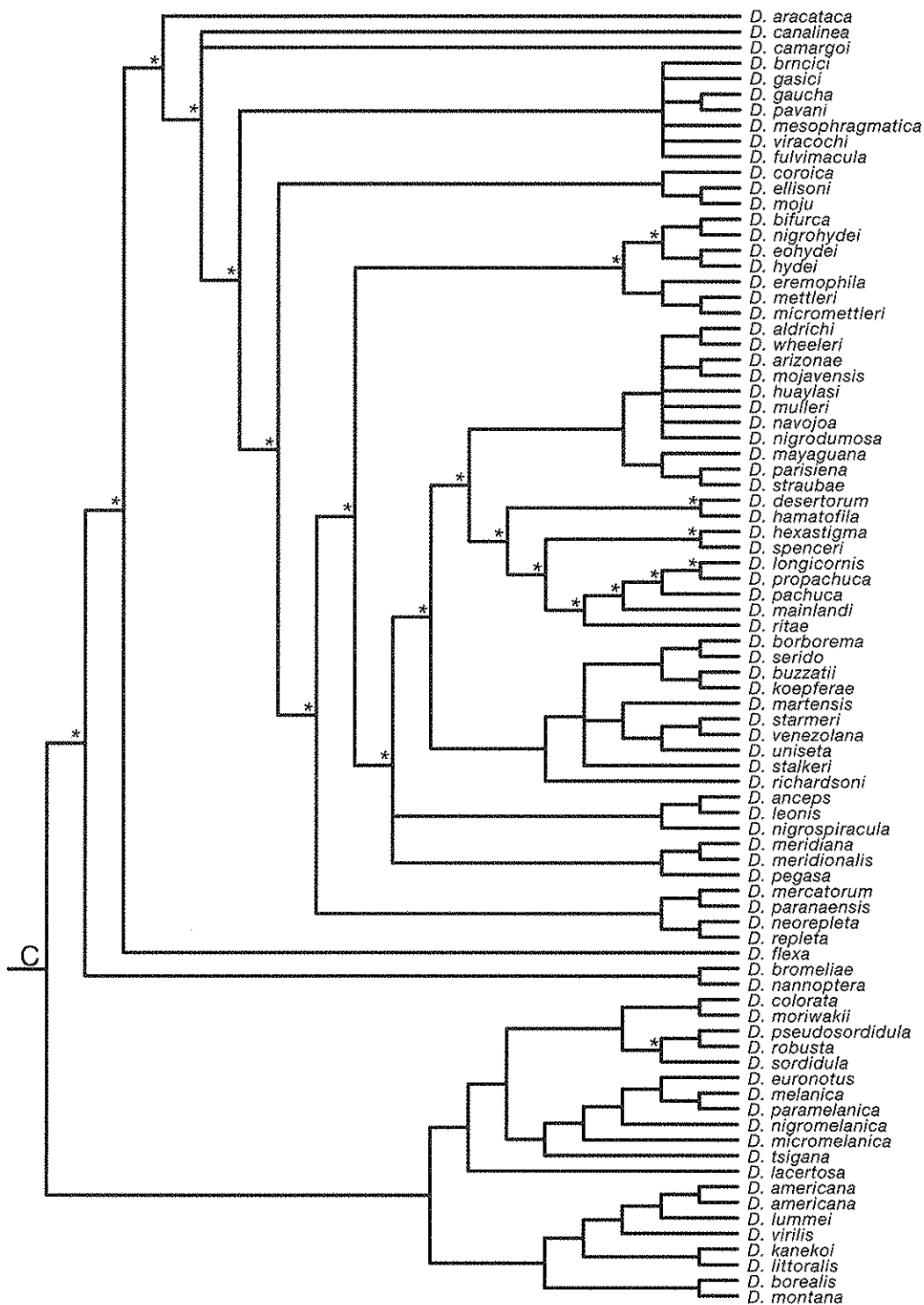


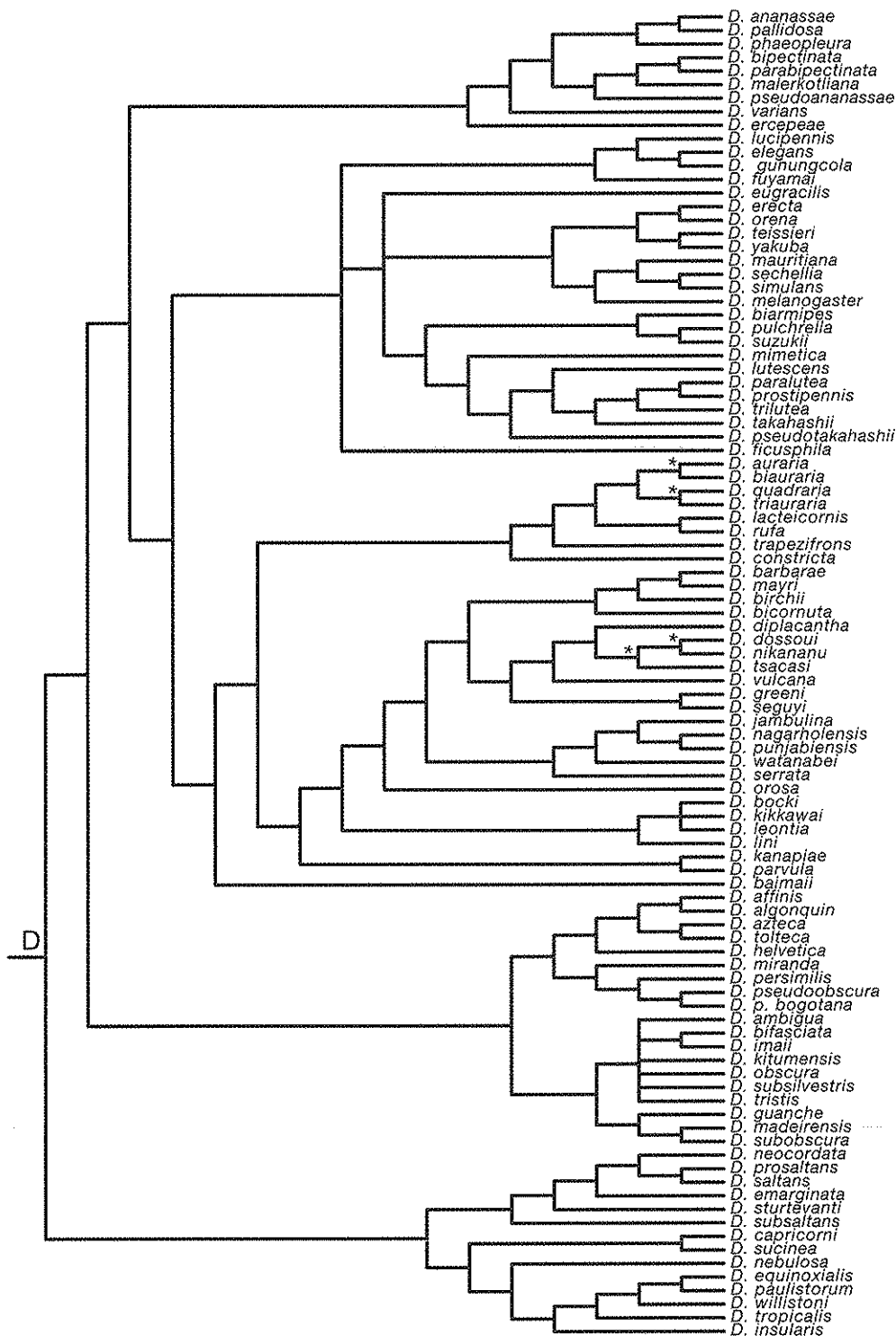
Supertrees as guides for taxonomy

One of the main goals of taxonomy is to provide a predictable, stable framework in which to describe new species and classify existing taxa. Taxonomic hypotheses, like those proposed in other scientific disciplines, must be based upon well-supported, repeatable analyses of primary data using philosophically rigorous and published methods. Taxonomic progress, in the sense of creating stable, natural groupings of monophyletic taxa, is not served by revisions of existing nomenclature based on poorly executed or incomplete studies, particularly when these do not add any new data relevant to the problem. Phylogenetic analysis can

be used as an aid to taxonomy in order to place an unknown species within a certain group and sort out higher-level taxonomy (e.g., APG 1998; 2005; Ward 2007), or to delineate cryptic species (e.g., Herbert et al 2004). Supertree methods may also be useful in taxonomy, but care needs to be taken because supertree approaches are highly susceptible to biases in phylogenetic method (long branch attraction, etc.). It is, therefore, critical that source trees be selected discerningly and objectively to minimize problems. The methodological implementation should likewise be consistent, objective, repeatable, and statistically supported.

In the van der Linde and Houle (2008) study,





there is no explicit statement of how trees were selected or weighted in the supertree analysis. In addition, we discovered severe methodological biases in our reanalysis, suggesting that their results are effectively non-repeatable. Given the status of the current data, there is no question that a comprehensive taxonomic revision of the genus *Drosophila* will be necessary in the future if the taxonomy is to reflect phylogenetic relationships. However, basing taxonomic revisions on a supertree analysis that only discovers a fraction of the equally parsimonious trees possible for the MRP matrix is unacceptable if one's goal is to create a stable taxonomic framework. Rather than invest time and effort in a poorly justified supertree approach, the *Drosophila* systematics community should collaborate to construct phylogenetic relationships based on primary data and use these hypotheses to drive thorough taxonomic revisions.

References

- Adams, M.D., et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287: 2185-2195.
- Angiosperm Phylogeny Group (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399-436.
- Angiosperm Phylogeny Group. (1998) An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden*. 85(4): 531-553.
- Bächli, G. (2008) Taxodros, the database on taxonomy of Drosophilidae. Available from: <http://www.taxodros.unizh.ch/>. (Accessed May, 2008).
- Bininda-Emonds, O. R. P. & Bryant, H. N. (1998) Properties of matrix representation with parsimony analyses. *Systematic Biology* 47: 497-508.
- Bininda-Emonds, O. R. P., Gittleman, J. L., & Steel, M. A. (2002) The (Super)tree of Life: Procedures, Problems and Prospects. *Annual Review of Ecology and Systematics*. 33: 265-289.
- Bininda-Emonds, O. R. P., Jones, K. E., Price, S. A., Cardillo, M., Grenyes, R., & Purvis, A. (2004) Garbage in, garbage out: Data issues in supertree construction. Pp 267-277. in: Bininda-Emonds, O. R. P.: *Computational Biology Series, Volume 4, Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*. 550 pp. Kluwer Academic Publishers, Boston, MA.
- Bininda-Emonds, O. R. P., & Sanderson, M. J. (2001) Assessment of the accuracy of matrix representation with parsimony supertree construction. *Systematic Biology* 50: 565-579.
- Burleigh, J. G., Driskell, A., & Sanderson, M. J. (2006) Supertree bootstrapping methods for assessing phylogenetic variation among genes in genome scale data. *Systematic Biology* 55: 426-440.
- Cotton, J.A., & Page, R.D.M. (2002) Going nuclear: gene family evolution and vertebrate phylogeny reconciled. *Proceedings of the Royal Society of London B* 269(1500): 1555-1561.
- Darwin, C. (1871) *The Descent of Man and Selection in Relation to Sex*. 1031 pp. John Murray, London.
- DeSalle, R. 1992. The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. *Molecular Phylogenetics and Evolution* 1: 31-34.
- DeSalle, R., Freedman, T., Prager, E. M., & Wilson, A. C. (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of Molecular Evolution* 26: 157-164.
- Drosophila* 12 Genome Consortium. (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450: 203-218.
- Felsenstein, J. (2004) *Inferring Phylogenies*. 664 pp. Sinauer Associates, Inc., Sunderland MA.
- Gatesy, J. E., & Springer, M. S. (2004) A critique of matrix representation with parsimony supertrees. Pp 369-388. in: Bininda-Emonds, O. R. P.: *Computational Biology Series, Volume 4, Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*. 550 pp. Kluwer Academic Publishers, Boston, MA.
- Gatesy, J. E., Matthee, C., DeSalle, R., & Hayashi, C. Y. (2002) Resolution of a supertree / supermatrix paradox. *Systematic Biology* 51: 652-664.
- Herbert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., & Hallwachs, W. (2004) Ten species in one: DNA barcoding reveals cryptic species in Neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*. 101: 14812-14817.
- Hubby, J. L. (1963) Protein differences in *Drosophila*. I. *Drosophila melanogaster*. *Genetics* 48: 871-879.
- Hubby, J. L., & Lewontin, R. C. (1966) A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles and different loci in *Drosophila pseudoobscura*. *Genetics* 54: 577-594.
- Hubby, J.L., & Throckmorton, L. H. (1965) Protein differences in *Drosophila*. II. Comparative species genetics and evolutionary problems. *Genetics* 52: 203-215.
- Kambysellis, M. P. Ho, K-F, Craddock, E. M., Piano, F., Parisi, & Cohen, J. (1995). Pattern of ecological shifts in the diversification of Hawaiian *Drosophila* inferred from a molecular phylogeny. *Current Biology* 5: 1129-1139.
- Kwiatowski, J., & Ayala, F. J. (2001) Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. *Molecular Phylogenetics and Evolution* 13: 319-328.
- Lewontin, R. C., & Hubby, J. L. (1966) A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* 54: 595-609.
- Maddison, D. R. (1991) The discovery and importance of multiple islands of most parsimonious trees. *Systematic Zoology* 40: 315-328.
- Maddison, W. P., Donoghue, M. J., & Maddison, D. R. (1984). Outgroup analysis and parsimony. *Systematic Zoology* 33: 83-103.
- Markow, T. A., & O'Grady, P. M. (2006) *Drosophila*: A

- Guide to Species Identification and Use*. 259 pp. Academic Press, London.
- Markow, T. A., & O'Grady, P. M. (2007) *Drosophila* biology in the genomic age. *Genetics* 177: 1269-1276.
- Moore, B. R., Smith, S. A., & Donoghue, M. J. (2006) Increasing data transparency and estimating phylogenetic uncertainty in supertrees: approaches using non-parametric bootstrapping. *Systematic Biology* 55: 662-676.
- Page, R. D. M. (2004) Taxonomy, Supertrees, and the Tree of Life. Pp 247-265. in: Bininda-Emonds, O. R. P.: *Computational Biology Series, Volume 4, Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*. 550 pp. Kluwer Academic Publishers, Boston, MA.
- Pélandakis, M., Higgins, D. G. & Solignac, M. (1991) Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. *Genetica* 84: 87-94.
- Pélandakis, M. & Solignac, M. (1993) Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *Journal of Molecular Evolution* 37: 525-543.
- Pollard, D. A., Iyer, V. N., Moses, A. M., & Eisen, M. B. (2006) Whole genome phylogeny of the *Drosophila melanogaster* species subgroup: Widespread discordance with species tree and evidence for incomplete lineage sorting. *PLoS Genetics* 2(10): e173.
- Powell, J. R. (1997) *Progress and Prospects in Evolutionary Biology: the Drosophila Model*. 562 pp. Oxford University Press, New York.
- Remsen, J. & O'Grady, P. (2002) Phylogeny of Drosophilinae (Diptera; Drosophilidae), with comments on combined analysis and character support. *Molecular Phylogenetics and Evolution* 24: 249-264.
- Ronquist, F. (1996) Matrix representation of trees, redundancy, and weighting. *Systematic Biology* 45: 247-253.
- Smith, S.A., & Moore, B.R. (submitted) tREeBOOT: Assessing phylogenetic uncertainty in supertrees. *Bioinformatics*.
- Sanderson, MJ, Purvis, A, & Henze, C. (1998) Phylogenetic supertrees: assembling the trees of life. *Trends in Ecology and Evolution* 13: 105-109.
- Tatarenkov, A., Kwiatowski, J., Skarecky, D., Barrio, E., & Ayala, F. J. (1999) On the evolution of Dopa decarboxylase (Ddc) and *Drosophila* systematics. *Journal of Molecular Evolution* 48: 445-462.
- Throckmorton, L. H. (1962) The use of biochemical characteristics for the study of problems of taxonomy and evolution in the genus *Drosophila*. *University of Texas Publications* 6205: 415-488.
- Throckmorton, L. H. (1966) The relationships of the endemic Hawaiian Drosophilidae. *University of Texas Publications* 6615: 335-396.
- Throckmorton, L. H. (1975) The phylogeny, ecology and geography of *Drosophila*. Pp. 421-469 in: King, R. C.: *Handbook of genetics Vol 3*. 874 pp Plenum, New York.
- Throckmorton, L. H. (1982) The *virilis* species group. Pp. 227-296. in: Ashburner, M. *The Genetics and Biology of Drosophila*. xvi + 429. Academic Press, New York.
- van der Linde, K. & Houie, D. (2008) supertree analysis and literature review of the genus *Drosophila* and closely related genera (Diptera, Drosophilidae). *Journal of Insect Systematics and Evolution* (in this issue).
- Ward, P.S. (2007) Phylogeny, classification, and species-level taxonomy of ants (Hymenoptera: Formicidae). *Zootaxa* 1668: 549-563.

Accepted for publication August 2008