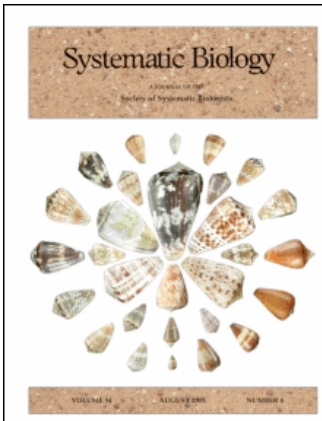

Systematic Biology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713658732

Filtered Z-Closure Supernetworks for Extracting and Visualizing Recurrent Signal from Incongruent Gene Trees

James B. Whitfield [a]; Sydney A. Cameron [a]; Daniel H. Huson [b]; Mike A. Steel [c]
[a] Department of Entomology, University of Illinois, Urbana, Illinois, USA [b] Center for Bioinformatics, Tübingen University, Tübingen, Germany [c] Allan Wilson Centre, University of Canterbury, Christchurch, New Zealand

First Published on: 01 December 2008

To cite this Article Whitfield, James B., Cameron, Sydney A., Huson, Daniel H. and Steel, Mike A.(2008)'Filtered Z-Closure Supernetworks for Extracting and Visualizing Recurrent Signal from Incongruent Gene Trees',Systematic Biology,57:6,939 — 947
To link to this Article: DOI: 10.1080/10635150802552849
URL: http://dx.doi.org/10.1080/10635150802552849

Points of View

Syst. Biol. 57(6):939–947, 2008

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150802552849

Filtered Z-Closure Supernetworks for Extracting and Visualizing Recurrent Signal from Incongruent Gene Trees

JAMES B. WHITFIELD,[1] SYDNEY A. CAMERON,[1] DANIEL H. HUSON,[2] AND MIKE A. STEEL[3]

[1]*Department of Entomology, University of Illinois, Urbana, Illinois 61801, USA; E-mail: jwhitfie@life.uiuc.edu (J.B.W.); scameron@life.uiuc.edu (S.A.C.)*
[2]*Center for Bioinformatics, Tübingen University, 72075 Tübingen, Germany; E-mail: huson@informatik.uni-tuebingen.de*
[3]*Allan Wilson Centre, University of Canterbury, Christchurch, New Zealand; E-mail: m.steel@math.canterbury.ac.nz*

Most modern phylogenetic inference methods assume that data should be fitted to a set of possible trees, from which the optimal tree is to be identified. A bifurcating tree-like diagram is thought to best represent evolutionary history. This expectation follows largely from the view that (i) simultaneous divergence of multiple organismal lineages from the same common ancestor is likely to be rare (Hennig, 1966), hence the bifurcating attribute; and (ii) reticulate patterns of relationship caused by hybridization, horizontal gene transfer, and gene conversion and recombination are relatively minor deviations from an underlying single bifurcating pattern of evolution (Brown et al., 2001; Daubin et al., 2003).

Whether most of the pattern of evolutionary history is truly bifurcating is still uncertain, but it has become increasingly evident that complications such as hybridization, horizontal gene transfer, and lineage sorting of ancestral polymorphisms are not as rare as once supposed (Takahashi et al., 2001; Pääbo, 2003; Zhaxybayeva et al., 2004; Gogarten and Townsend, 2005). Under these conditions, individual gene histories may not be congruent with the general pattern of species or higher taxon relationships, and a phylogenetic tree estimated from the combined (concatenated) gene sequence represents an oversimplified version of the genetic history (Huber and Moulton, 2005; Morrison, 2005; Huson and Bryant, 2006). An usual alternative to simultaneous analysis of combined sequences from multiple gene partitions is to present a consensus tree as a representation of relationships common to each individual gene tree, but this also suffers from oversimplification (Swofford, 1991) and can be performed only when each gene tree has the same taxon representation. A solution to the oversimplification and taxon constraints to topological congruence would allow visualization of both the principal conflicting patterns among individual gene trees and the phylogenetic patterns common among trees that may not overlap completely in taxon representation (partial trees).

Huson et al. (2004) presented the Z-closure method for constructing a supernetwork from a set of partial trees. In cases where gene trees contain numerous different incongruent relationships, the resulting supernetwork can, however, be too complex to visualize and interpret. Huson et al. (2006), therefore, extended the supernetwork approach to enable filters to remove relationships (splits) from the supernetwork that are represented only once or sporadically among the source trees. The result is a network that summarizes the relationships found repeatedly among the source trees, even when they have somewhat different taxon representation. This new method has been used recently to explore plant and animal relationships (Huson et al., 2006; Whitfield and Kjer, 2008; Murphy et al., 2008), but its use has not yet been fully illustrated for a biological audience. Here we describe the basic underlying methods of the filtered supernetwork and illustrate it with two biological examples from parasitoid wasps and social bees, demonstrating applications of the method to common phylogenetic problems arising from gene tree incongruence.

Filtering Supernetworks: A Primer

A collection of phylogenetic trees $T_1, T_2, \ldots, T_K$ can be summarized by a consensus tree if all the trees are defined by exactly the same set of taxa. Consensus methods work by extracting all splits (bipartitions of the taxon set) from the set of input trees and then returning a tree that represents a subset of the splits, such as the strict consensus comprising the splits contained in all input trees or the majority consensus consisting of those splits contained in the majority of all input trees.

In practice, the taxon content of trees from different data partitions usually differs slightly due to the absence of data for identical taxa across all partitions. This is often

addressed by removing taxa that do not occur in all trees under consideration, which leads to a reduced data set. To avoid taxon reduction, a supertree method (Bininda-Emonds, 2004) can be used to obtain a consensus. Supertree methods, however, employ a heuristic rather than an exhaustive search for solutions to resolve incompatibilities among trees and thus may discard well-supported splits.

An alternative consensus approach to maintaining all taxa across multiple data partitions is the supernetwork method known as Z-closure, introduced in Huson et al. (2004). The Z-closure method takes as input a set of trees based on taxonomically overlapping but nonidentical data sets and uses the so-called Z-closure rule to infer a collection of splits for the full taxon set. Z-closure is an operation that takes the splits that appear in the input trees and combines them in pairs to provide splits of progressively larger sets of taxa. The process is applied to pairs of splits that intersect (the Z refers to this pattern of intersection) such that if A_1 and B_1 are the two taxon sets defined by one split, and A_2 and B_2 are the two taxon sets defined by a second split, A_1 and A_2 must overlap, A_2 and B_1 must overlap, and B_1 and B_2 must overlap, whereas A_1 and B_2 do not necessarily overlap (Huson et al., 2004). The justification for combining pairs of splits in this way originates with Meacham (1983), who showed that any evolutionary tree that is consistent with such a pair of splits for subsets of taxa must necessarily be consistent with the corresponding pair of splits for the more inclusive subsets of taxa. The Z-closure operation is applied repeatedly (and in randomized order as the results of the operation are order dependent) to obtain sets of splits for the entire set of taxa.

In the resulting set of splits, called the Z-closure of the trees, all of the trees are not necessarily compatible and thus cannot necessarily be represented by a single tree, but rather by a more general split network (Huson and Bryant, 2006), which uses bands of parallel lines to represent splits in the presence of incompatibilities (e.g., Figs. 3 and 5). Edge lengths in the network are derived from the branch lengths in the original gene trees.

A split network derived from a set of splits arising from the Z-closure method is called a supernetwork because it contains all taxa found in any of the trees being summarized. It also has the important property that any split S contained in the network is present in every input tree that contains at least one taxon from each part of the split S. However, for algorithmic reasons there is no guarantee that every split that occurs in an input tree also occurs in the network. The user supplies two parameters for inclusion of splits in the supernetwork: maxdist, the maximum distortion under which S is still considered to be close to T, and mintrees, the minimum number of trees to which S must be close. In this article we do not consider the distortion filter, focusing instead on the simple mintrees filter. If one sets maxdist to 0 (essentially removing the distortion filter) and mintrees to a chosen number t, then only those splits that are contained in t or more trees pass the filter, and one can obtain the strict or majority consensus by setting t appropriately. Note, however, that the resulting majority consensus is not necessarily compatible with a single bifurcating tree. In practice, the mintrees value set by the user should be guided by biological insight rather than explicit rules. The resulting filtered supernetworks are therefore best viewed as exploratory tools for visualizing conflicting splits and assessing the impact of removing splits that occur infrequently among the different input trees.

Identifying Underlying Species Trees from Conflicting Gene Trees: An Example from Parasitoid Wasps

One important application of filtered supernetworks is the identification of the most likely species tree from a collection of gene trees. We demonstrate this with an example from parasitoid wasps.

Lapointe et al. (2007) compared the genomes of 10 polydnaviruses living in association with parasitoid wasps. As a reference for their examination of polydnavirus genome differentiation and gene duplication, they inferred a phylogeny of the 10 corresponding parasitoid host wasps from DNA sequences of multiple wasp genes (Fig. 1). The wasps are members of two families: Braconidae, represented by exemplars from the subfamilies Microgastrinae (*Cotesia*, *Glyptapanteles*, and *Microplitis*), Cardiochilinae (*Toxoneuron*), and Cheloninae (*Chelonus*); the family Ichneumonidae is represented by Campopleginae and Banchinae (*Glypta*). *Cotesia* and *Glyptapanteles* are sister taxa relative to *Microplitis* in the microgastrine clade (Fig. 1), Cardiochilinae is sister group to Microgastrinae, and Cheloninae attaches at the base of the braconid clade. Four of the five Ichneumonidae belong to Campopleginae (relationships among them uncertain), with Banchinae as sister group.

Relationships within and between these wasp families are supported by other diverse morphological and molecular studies (Gauld, 1985; Wahl, 1991; Whitfield, 1997, 2002; Dowton and Austin, 1998; Belshaw et al., 1998; Banks and Whitfield, 2006); thus, the backbone phylogeny in Figure 1 serves as a realistic framework upon which to map polydnaviral genome evolution. Yet, not all gene trees inferred from the six individual and combined data sets for these 10 wasps are compatible with this expected phylogeny. For instance, the single-partition Bayesian analyses misplace several of the subfamilies or root the braconid clade incorrectly (Fig. 2a). Although the mixed-model Bayesian analysis of the combined partitions (Lapointe et al., 2007) recovers the expected phylogeny, maximum parsimony results in a tree that misplaces *Chelonus* as sister group to *Toxoneuron* (Fig. 2b) and maximum likelihood (uniform GTR+I+Γ model across the six gene partitions) places the root of the braconid clade within the Microgastrinae (Fig 2b). Other models can lead to other misplaced taxa, such as *Glypta* falling inside the monophyletic Campopleginae. The instability of these different results with respect to model assumptions reduces confidence in the underlying species phylogeny for these six genes.

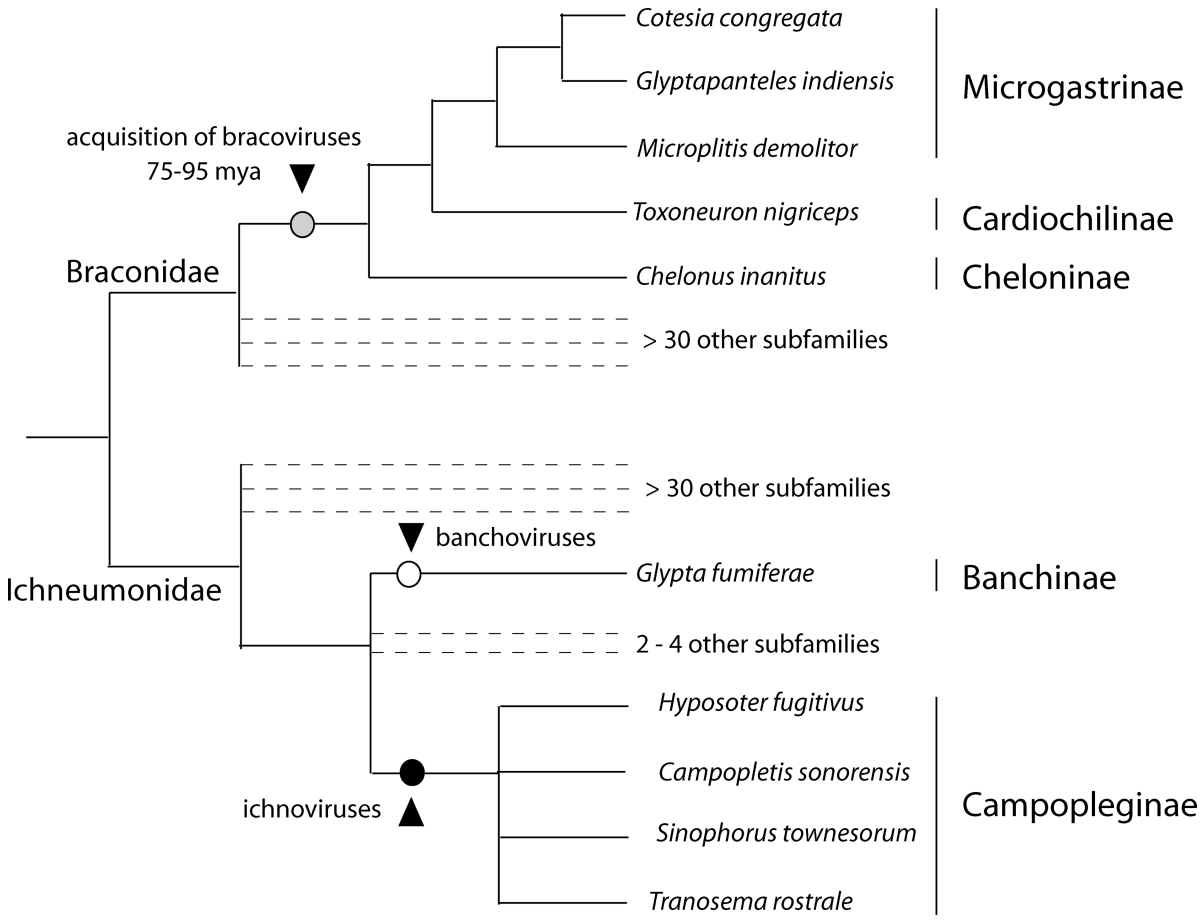

FIGURE 1. Expected phylogeny of 10 parasitoid wasp species from which associated polydnavirus genomes have been sequenced and compared. The phylogeny shows the evolutionary scenario proposed for the polydnaviruses (adapted from Lapointe et al., 2007). Numerous lines of evidence, including molecular data, comparative morphology, and fossils, support this pattern of relationships.

If instead we use a Z-closure supernetwork to depict the alternative splits found in the six gene trees (Fig. 3a), we obtain the expected backbone tree with a web of reticulations, a supernetwork, representing topological differences (uncertainty) among the gene trees. Applying a conservative filter to this supernetwork, for instance, allowing only the splits that occur consistently in four or more of the six gene trees, yields a tree that not only loses the reticulation but also loses some resolution (tree not shown). A more "lenient" filter, allowing all splits found in only three (or more) of the gene trees, yields the expected tree (Figure 3b). This procedure highlights those relationships that recur consistently among the source trees.

The filtered supernetwork method is advantageous in that it allows extraction of the most common recurrent phylogenetic signal from a set of taxonomically overlapping gene trees. It thus presents a solution to the common problem that arises when a combined-partition tree, in which all available data sources are pooled for analysis, is overly sensitive to one or two data partitions that have strong signal for an incorrect relationship, such as might result from systematic biases in a certain data set.

Identifying the Root of an Ingroup Tree When Outgroups Disagree: An Example from the Corbiculate Bees

The filtered supernetwork approach can be useful in clarifying the rooting of a species tree when different gees trees misplace some outgroup taxa. A prominent example of this is the controversy surrounding the relationships among the corbiculate bees.

The corbiculate bees comprise four tribes, three of which represent the major groups of social bees, including the stingless bees (Meliponini), bumble bees (Bombini), and honey bees (Apini). Orchid bees (Euglossini), the fourth tribe within the clade, are solitary or communal (see Michener, 1974, for definitions of sociality). Understanding the historical pattern of relationships among these taxa is important for a full understanding of the evolution of highly social behavior in bees, yet their phylogeny has been a difficult problem to resolve. Incompatibility exists between morphological (Schulz et al., 1999) and DNA data (Cameron and Mardulyn, 2001; Cameron, 2003; Kawakita, 2008). The principal finding from DNA is that the so-called primitively (intermediately) social Bombini and highly social Meliponini form a strongly supported clade, to the exclusion of

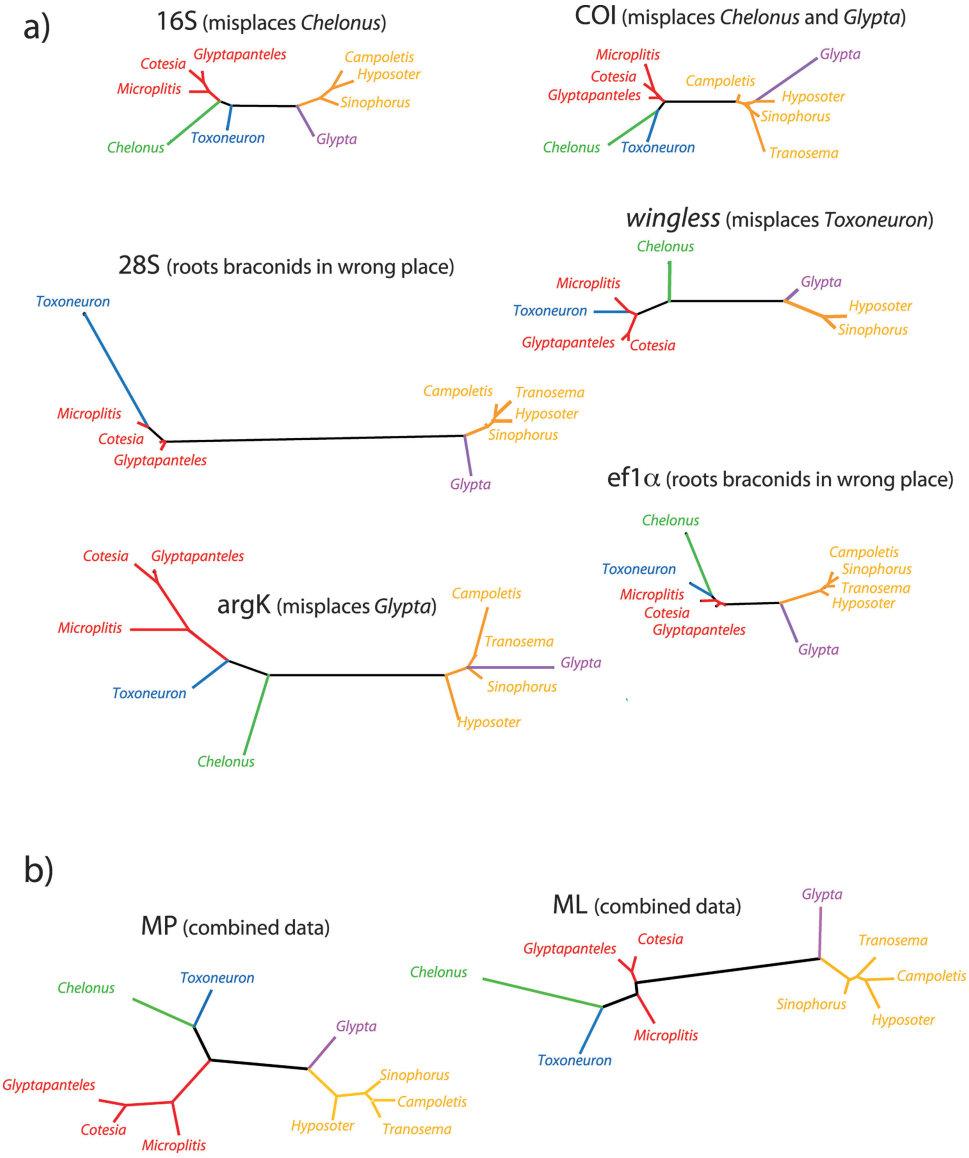

FIGURE 2. (a) Individual gene trees of parasitoid wasps from the families Braconidae and Ichneumonidae inferred from independent Bayesian analyses of six gene partitions. The sequence data are from Lapointe et al. (2007); (b) maximum parsimony (MP) and maximum likelihood (ML) trees inferred from the combined six-gene partitions indicated in (a). Relative to the expected tree (Fig. 1), MP misplaces *Chelonus*, and ML (GTR+I+Γ model) roots the braconid lineage incorrectly within the Microgastrinae.

highly social Apini. This was a surprise in the early stages of molecular research on these bees, because it was conventionally thought from morphological characters that highly social behavior arose once, in the common ancestor of Apini and Meliponini (Roig-Alsina and Michener, 1993). Yet nucleotide data obtained from multiple independent genes over the last decade has given a mostly congruent pattern of relationships, repeatedly grouping Bombini and Meliponini, regardless of the optimality criterion applied to tree estimation (Cameron and Mardulyn, 2001; Michel-Salzat et al., 2004; Cameron et al., 2007; Rasmussen and Cameron, 2007; Kawakita et al., 2008). Unrooted trees are mostly congruent, although conflicting signals for the placement of the root have not been fully resolved.

Rooting the corbiculate tree has been difficult even with the large number of nucleotide characters and multiple outgroups (Lockhart and Cameron, 2001; Cameron and Mardulyn, 2001). The phylogeny represents a classic example of the evolutionary scenario described by Hendy and Penny (1989) in which the juxtaposition of long external branches and short internal branches makes it difficult to place outgroups correctly (Lockhart and Cameron, 2001). With this tree shape, the root and direction of evolution are difficult to determine because homoplasy is expected to result in conflicting signals for root placement—outgroups tend to be drawn toward the long external branches in the tree, irrespective of the true placement (Holland et al., 2003; Shavit et al., 2007).

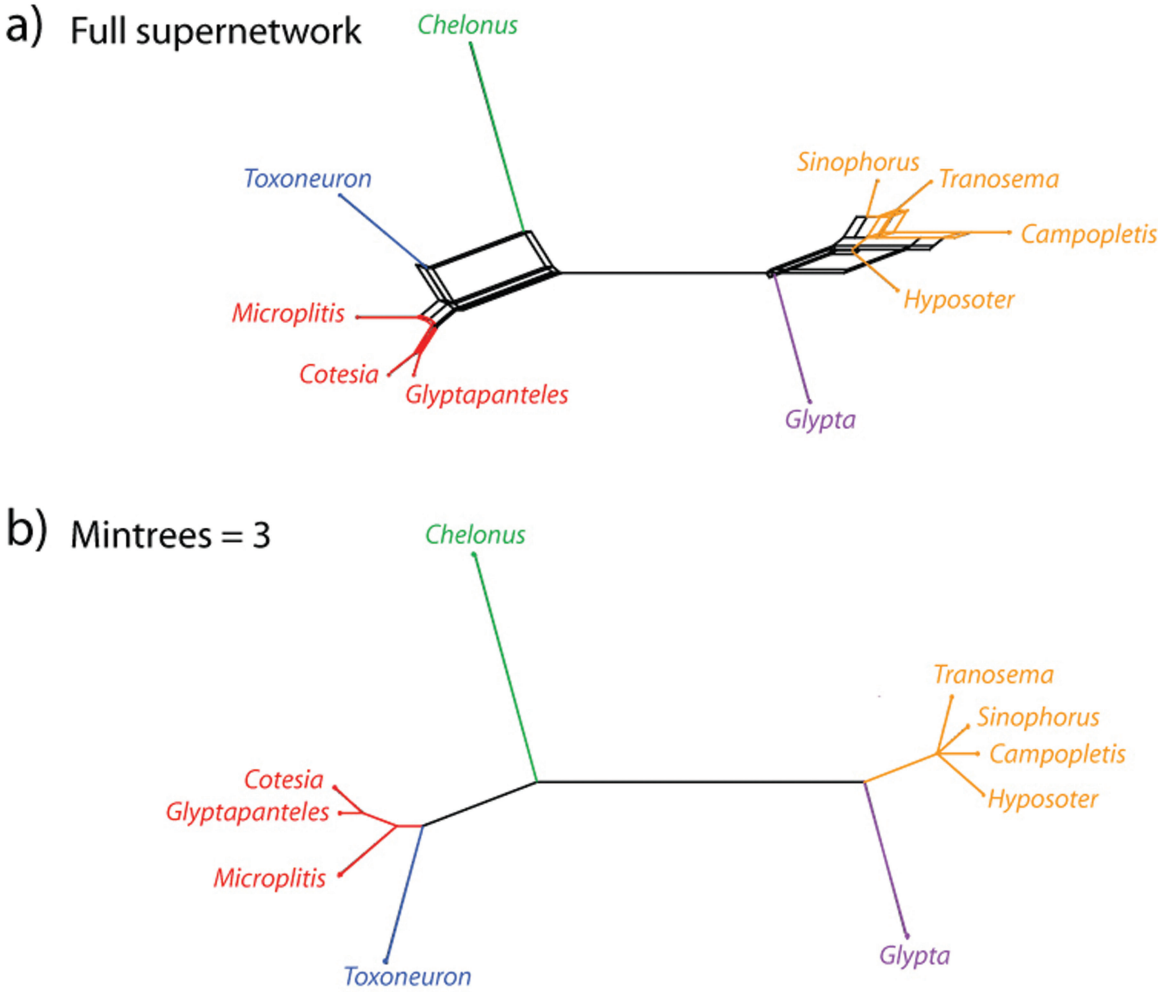

FIGURE 3. (a) Z-closure supernetwork of six parasitoid wasp gene trees (Fig. 2), prior to filtering. The supernetwork resembles the expected wasp phylogeny (Fig. 1) with internal conflicts displayed. (b) Z-closure supernetwork after implementing the filter (mintrees = 3), displaying only the splits found in (or found to be fully compatible with) at least three of the gene trees. This method recovers the expected wasp phylogeny.

In a recent study, Kawakita et al. (2008) analyzed data from 12 nuclear genes for representatives of the four corbiculate tribes plus four outgroup taxa. We conducted Bayesian analyses of each of their 12 nuclear gene fragments plus two mtDNA gene fragments (substitution models provided in Table 1). Seven of the resulting gene trees in Figure 4 (rh, Nak, cytb, GlyK, COI, EF, and Pol II) unite Bombini and Meliponini, five (CamK, ArgK, white, CAD, and Dnk) neither support nor contradict that grouping due to lack of resolution, and two (wg and bub) contradict it. Overall, there appears to be strong support for a Bombini + Meliponini ingroup split. The more serious problem lies with the unstable placement of the outgroups, which in some cases render the apparently monophyletic corbiculate clade paraphyletic or even polyphyletic or lead to an inconsistent placement of the ingroup root.

The Z-closure supernetwork of the 14 gene trees helps to visualize the conflict among the trees (Fig. 5), but it is difficult to interpret with respect to intertribal relationships and rooting. We applied a filter to the supernetwork such that a split had to be contained in, or be

TABLE 1. Substitution models for Bayesian analysis of each of the 14 genes used in the corbiculate bee example. Nst: number of substitution rate types; Statefreqpr: assumption of stationary base composition frequencies as equal or unequal; Inv: proportion of invariant sites estimated; Gamma: among-site rate variation modeled using a gamma distribution.

Gene	Nst	Statefreqpr	Inv	Gamma
ArgK (arginine kinase)	6	Equal	No	Yes
Bub (mitotic checkpoint control protein)	6	Unequal	Yes	No
CAD (carbamoyl-phosphate synthetase–aspartate transcarbamoylase–dihydroorotase)	6	Equal	Yes	Yes
CamK (calmodulin-dependent protein kinase II)	6	Unequal	No	Yes
DnK (deoxyribonucleoside kinase)	6	Unequal	Yes	Yes
EF-1α (elongation factor 1α F2)	6	Unequal	Yes	Yes
GlyK (glycerol kinase)	6	Unequal	No	Yes
NaK (sodium-potassium ATPase)	6	Unequal	No	Yes
polII (RNA polymerase II)	6	Unequal	Yes	Yes
Rh (long-wavelength rhosopsin)	2	Unequal	Yes	Yes
white	6	Equal	No	Yes
Wg (*wingless*)	6	Unequal	No	Yes
COI (cytochrome oxidase I)	6	Unequal	Yes	Yes
Cytb (cytochrome *b*)	6	Unequal	Yes	Yes

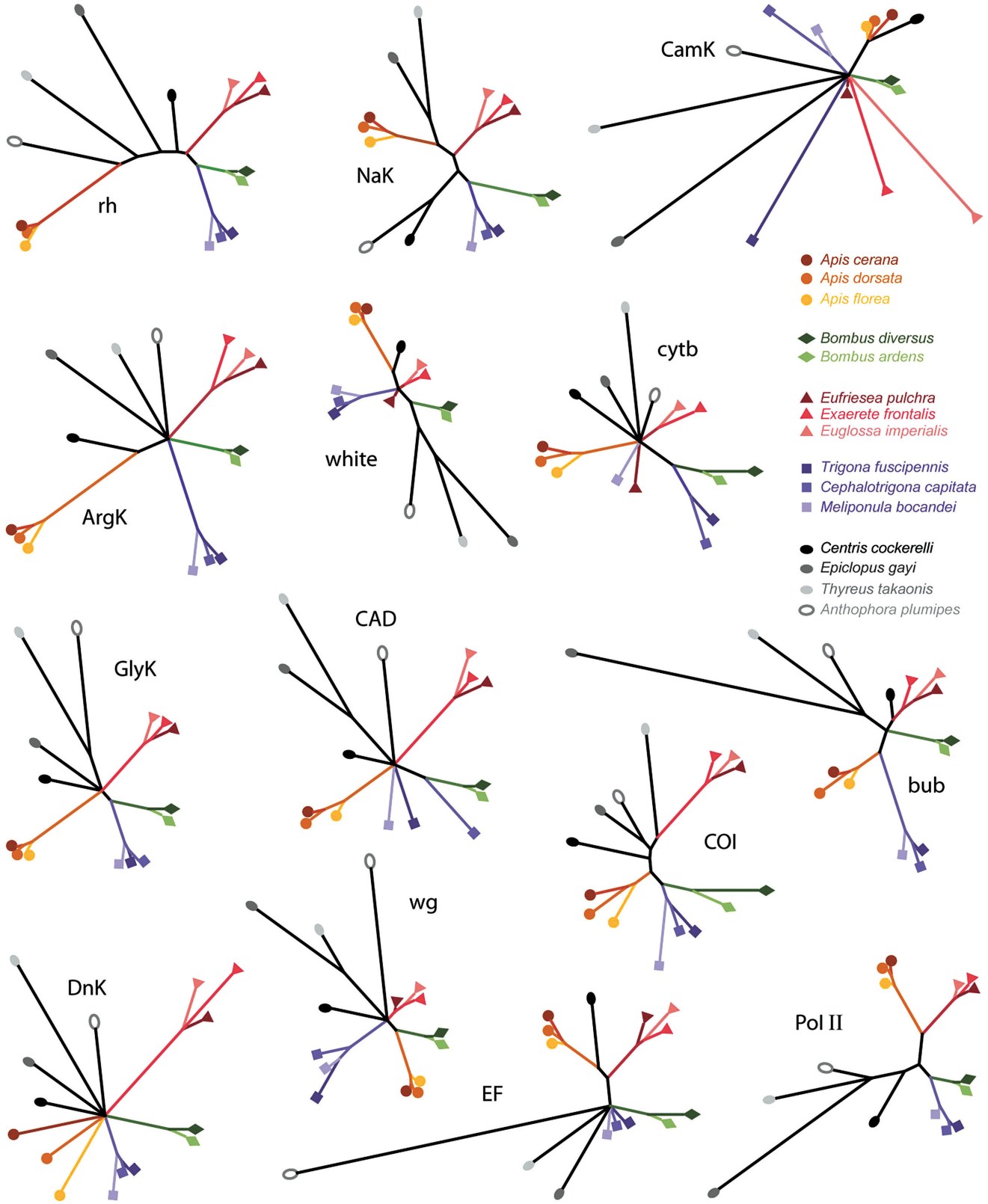

FIGURE 4. Corbiculate bee gene trees from Bayesian analyses of each of 14 gene partitions (sequence data, in part, from Kawakita et al., 2008); outgroups are represented by black and white shaded branches. Substitution models for each analysis are listed in Table 1.

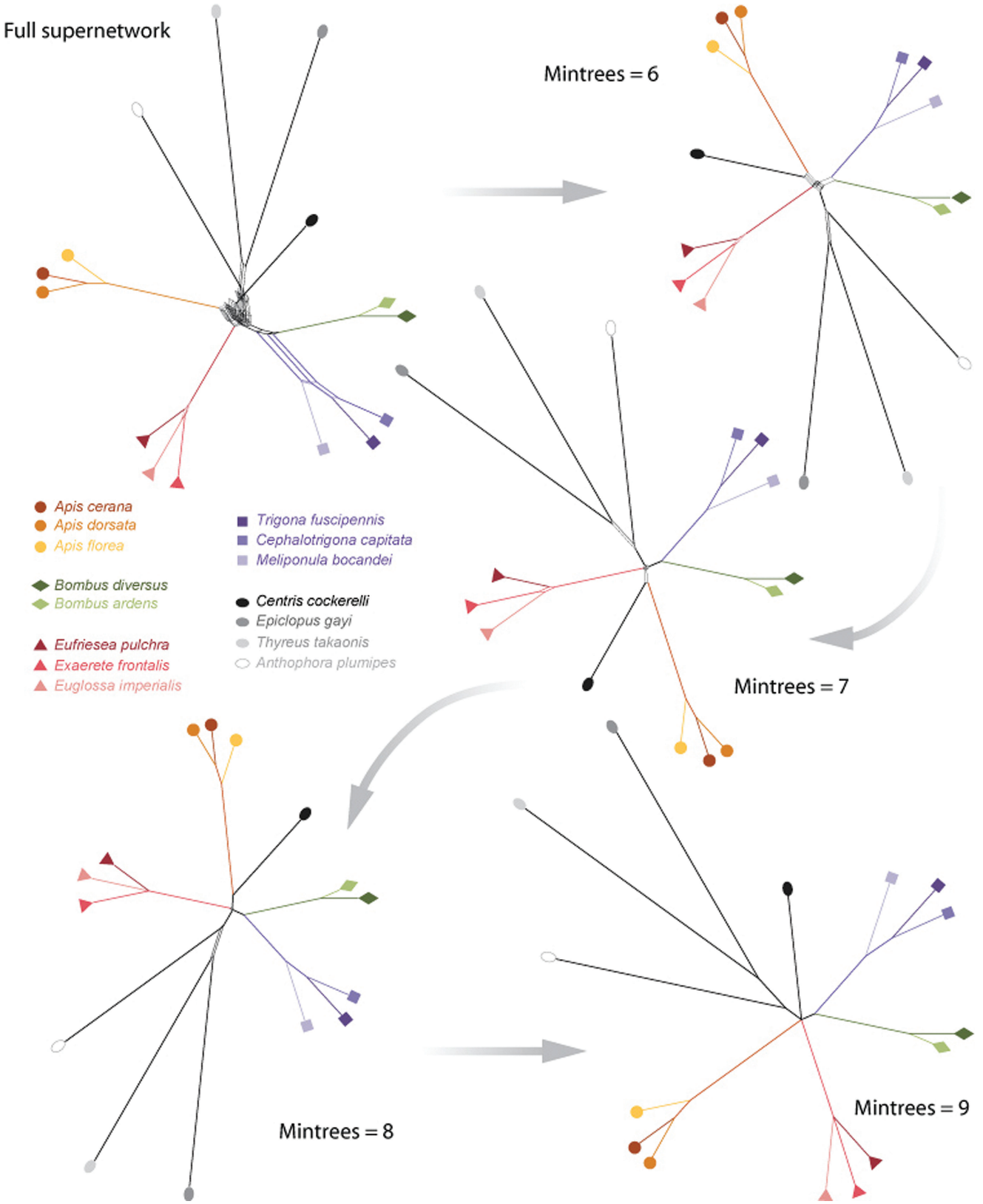

FIGURE 5. Corbiculate bee Z-closure supernetworks of the 14 gene trees of Figure 4, indicating filtering with mintrees = 6, 7, 8, and 9, respectively; each consecutive filtered supernetwork (arrows indicating the order of increasing stringency) shows only the splits found in (or fully compatible with) at least 6, 7, 8, or 9 source trees. The higher stringency filter (mintrees = 9) clarifies the position of the root (shaded as in Fig. 4) of the ingroup tree.

consistent with, a threshold number of trees (mintrees = 6, 7, 8, or 9) to be included in the network. The supernetwork is greatly simplified even with the less stringent filter, mintrees = 6; however, it contains some splits that place an outgroup (*Centris*) within the ingroup. With mintrees = 9, all reticulation within the network is gone, and the outgroups root the tree at the base of the Apini + Euglossini split. Thus, the most recurrent splits among the 14 gene trees, despite the outgroup "noise," are consistent with previous findings that Meliponini and Bombini are sister taxa, with the relationships still uncertain for Apini and Euglossini. In this example, the solution to the rooting problem is greatly simplified, if not fully solved.

Discussion

Filtered supernetworks provide a tractable solution to the problem of visualizing the most recurrent phylogenetic pattern among a collection of gene trees for a set of taxa, while retaining the ability to represent the observed conflicts simultaneously. An additional advantage of the method is that it does not require identical taxon representation among the gene trees. Filtered supernetworks are easily obtained from sets of input gene trees using SplitsTree4 (Huson and Bryant, 2006; obtainable at http://www.splitstree.org).

Derived from a conceptually related but algorithmically dissimilar approach to that of filtered supernetworks, Holland et al. (2007) use a quartet imputation method to "add in" the "missing" taxa to partial trees, thus constructing a complete taxon set so that standard consensus network methods can be applied. An executable implementation of their method is available at http://awcmee.massey.ac.nz/downloads. The filtered Z-closure supernetwork method does not require extrapolation from the original gene trees but may require further exploration to gauge the efficiency in which it recovers and accurately summarizes splits. Some initial comparisons of Z-closure with Q-imputation were made by Holland et al. (2007), and the two methods appeared to have complementary strengths and weaknesses in terms of their accuracy in identifying sets of splits. These two supernetwork methods have yet to be extensively tested when splits are filtered.

We have described two applications for the filtered Z-closure supernetworks, each of which typifies a common problem encountered in phylogenetic analysis of multiple genes or data sources. It will be important to explore how the filtered supernetwork method performs with other real-data problems. We anticipate that it will be especially useful when relationships are controversial or when hybridization, lateral gene transfer, or lineage sorting could lead to incongruent gene trees.

Acknowledgments

We are grateful to the Cass Field Station of the University of Canterbury, where we first developed the filtered supernetwork methods. J.B.W. would like to thank the Allan Wilson Centre for Molecular Ecology and Evolution for a Sabbatical Leave Grant in 2006 that allowed him to join the other authors in New Zealand, and the National Science Foundation (NSF; DEB 0316566) for funding research leading to the wasp data. S.A.C. thanks Atsushi Kawakita for sharing the corbiculate bee sequence data, Janet Hanlon for graphical design assistance, and U.S. Department of Agriculture NRI CSREES grant 2002-35302-11553 and NSF grant DEB 0446325 for research support. D.H.H. would like to thank the Deutsche Forschungsgemeinschaft (DFG) and the Erskine Programme for funding. Thanks to P. Lockhart for many useful discussions and to D. Ford and an anonymous reviewer for helpful comments on the manuscript.

References

Banks, J. C., and J. B. Whitfield. 2006. Dissecting the ancient rapid radiation of microgastrine wasp genera using additional nuclear genes. Mol. Phylogenet. Evol. 41:690–703.

Belshaw, R., M. G. Fitton, E. Herniou, C. Gimeno, and D. L. J. Quicke. 1998. Molecular phylogeny of the Ichneumonoidea (Hymenoptera) based the on D2 expansion region of 28S ribosomal RNA. Syst. Entomol. 23:109–123.

Bininda-Emonds, O. 2004. Phylogenetic supertrees: Combining information to reveal the Tree of Life. Kluwer Academic Publisher, Dordrecht.

Brown, J. R., C. J. Douady, M. J. Italia, W. E. Marshall, and M. J. Stanhope. 2001. Universal trees based on large combined protein sequence data sets. Nat. Genet. 28:281–285.

Cameron, S. A. 2003. Data from the elongation factor-1a gene corroborates the phylogenetic pattern from other genes, revealing common ancestry of bumble bees and stingless bees. III Seminario Mesoamericano sobre Abejas sin Agijon 1:132–136.

Cameron, S. A., H. M. Hines, and P. H. Williams. 2007. A comprehensive phylogeny of the bumble bees (*Bombus*). Biol. J. Linn. Soc. 91:161–188.

Cameron, S. A., and P. Mardulyn. 2001. Multiple molecular data sets suggest independent origins of highly eusocial behavior in bees (Hymenoptera: Apinae). Syst. Biol. 50:194–214.

Daubin, V., N. A. Moran, and H. Ochman. 2003. Phylogenetics and the cohesion of bacterial genomes. Science 301:829–832.

Dowton, M., and A. D. Austin. 1998. Phylogenetic relationships among the microgastroid wasps (Hymenoptera: Braconidae): Combined analysis of 16S and 28S rDNA genes and morphological data. Mol. Phylogenet. Evol. 10:354–366.

Gauld, I. D. 1985. The phylogeny, classification and evolution of parasitic wasps of the subfamily Ophioninae (Ichneumonidae). Bull. Brit. Mus. (Nat. Hist.) Entomol. 51:61–185.

Gogarten, J. P., and J. P. Townsend. 2005. Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. 3:679–687.

Hendy, M. D., and D. Penny. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38:297–309.

Hennig, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana.

Holland, B., G. Conner, K. Huber, and V. Moulton. 2007. Imputing supertrees and supernetworks from quartets. Syst. Biol. 56:57–67.

Holland, B. R., D. Penny, and M. D. Hendy. 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock—A simulation study. Syst. Biol. 52:229–238.

Huber, K. T., and V. Moulton. 2005. Phylogenetic networks. Pages 178–204 *in* Mathematics of evolution and phylogeny (O. Gascuel, ed.). Oxford University Press, Oxford, UK.

Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254–267.

Huson, D. H., T. Dezulian, T. Kloepper, and M. A. Steel. 2004. Phylogenetic super-networks from partial trees. IEEE/ACM Trans. Comput. Biol. Bioinform. 1:151–158.

Huson, D. H., M. Steel, and J. B. Whitfield. 2006. Reducing distortion in phylogenetic networks. Lect. Notes Bioinform. 4175:150–161.

Kawakita, A., J. S. Ascher, T. Sota, M. Kato, and D. W. Roubik. 2008. Phylogenetic analysis of the corbiculate bee tribes based on 12 nuclear protein-coding genes (Hymenoptera: Apoidea: Apidae). Apidologie 9:163–175.

Lapointe, R., K. Tanaka, W. Barney, J. Whitfield, J. Banks, C. Beliveau, D. Stoltz, B. Webb, and M. Cusson. 2007. Genomic and morphological features of a banchine polydnavirus: A comparison with bracoviruses and ichnoviruses. J. Virol. 81:6491–6501.
Linder, C. R., and L. H. Rieseberg. 2004. Reconstructing patterns of reticulate evolution in plants. Am. J. Bot. 91:1700–1708.
Lockhart, P. J., and S. A. Cameron. 2001. Trees for bees. Trends Ecol. Evol. 16:84–88.
Lockhart, P. J., P. A. McLenachan, D. Havell, D. Glenny, D. H. Huson, and U. Jensen. 2001. Phylogeny, dispersal and radiation of New Zealand alpine buttercups: Molecular evidence under split decomposition. Ann. Missouri Bot. Gard. 88:458–477.
Meacham, C. A. 1983. Theoretical and computational considerations of the compatibility of qualitative taxonomic characters. Pages 304–314 *in* Numerical taxonomy, NATO ASI Series, volume 1 (J. Felsenstein, ed.). Springer, Berlin.
Michel-Salzat, A., S. A. Cameron, and M. L. Oliveira. 2004. Phylogeny of the orchid bees (Hymenoptera: Apinae: Euglossini): DNA and morphology yield equivalent patterns. Mol. Phylogenet. Evol. 32:309–323.
Michener, C. D. 1974. The social behavior of the bees: A comparative study. Harvard University Press, Cambridge, Massachusetts.
Morrison, D. A. 2005. Networks in phylogenetic analysis: New tools for population biology. Int. J. Parasitol. 35:567–582.
Murphy, N., J. C. Banks, J. B. Whitfield, and A. D. Austin. 2008. Phylogeny of the parasitic microgastroid subfamilies (Hymenoptera: Braconidae) based on sequence data from seven genes, with an improved time estimate of the origin of the lineage. Mol. Phylogenet. Evol. 47:378–395.
Pääbo, S. 2003. The mosaic that is our genome. Nature 421:409–412.
Rasmussen, C., and, S. A. Cameron. 2007. A molecular phylogeny of the Old World stingless bees (Hymenoptera: Apidae: Meliponini) and the non-monophyly of the large genus *Trigona*. Syst. Entomol. 32:26–39.
Roig-Alsina, A., and C. D. Michener. 1993. Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). Univ. Kansas Sci. Bull. 55:124–162.
Schultz, T. R., M. S. Engel, and M. Prentice. 1999. Resolving conflict between morphological and molecular evidence for the origin of eusociality in the corbiculate bees: A hypothesis-testing approach. Pages 125–138 *in* Entomological contributions in memory of Byron A. Alexander (G. W. Byers, R. H. Hagen, and R. W. Brooks, eds.). University of Kansas Natural History Museum Special Publication 24.
Semple, C., and M. A. Steel. 2003. Phylogenetics. Oxford University Press, Oxford, UK.
Shavit, L., D. Penny, M. D. Hendy, and B. R. Holland. 2007. The problem of rooting rapid radiations. Mol. Biol. Evol. 24:2400–2411.
Swofford, D. L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? Pages 295–333 *in* Phylogenetic analysis of DNA sequences (M.M. Miyamoto and J. Cracraft, eds.). Oxford University Press, Oxford, UK.
Takahashi, K., Y. Terai, M. Nishida, and N. Okada. 2001. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. Mol. Biol. Evol. 18:2057–2066.
Wahl, D. B. 1991. The status of *Rhymphoctona*, with special reference to the higher categories within Campopleginae and the relationships of the subfamily (Hymenoptera: Ichneumonidae). Trans. Am. Entomol. Soc. 117:193–213.
Whitfield, J. B. 1997. Molecular and morphological data suggest a common origin for the polydnaviruses among braconid wasps. Naturwissen. 84:502–507.
Whitfield, J. B. 2002. Estimating the age of the polydnavirus/braconid wasp symbiosis. Proc. Natl. Acad. Sci. USA 99:7508–7513.
Whitfield, J. B., and K. Kjer. 2008. Ancient rapid radiations of insects: Challenges for phylogenetic analysis. Annu. Rev. Entomol. 53:449–472.
Whitfield, J. B., P. Mardulyn, A. D. Austin, and M. Dowton. 2002. Phylogenetic analysis of relationships among microgastrine braconid wasp genera based on data from the 16S, COI and 28S genes and morphology. Syst. Entomol. 27:337–359.
Zhaxybayeva, O., P. Lapierre, and J. P. Gogarten. 2004. Genome mosaicism and organismal lineages. Trends Genet. 20:254–260.

First submitted 18 April 2008; reviews returned 27 June 2008; final acceptance 4 September 2008
Associate Editor: Thomas Buckley

Syst. Biol. 57(6):947–954, 2008

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150802562400

Consistent Estimation of Divergence Times in Phylogenetic Trees with Local Molecular Clocks

BODIL SVENNBLAD

Department of Mathematics, Uppsala University, Box 480, SE-751 06 Uppsala, Sweden; E-mail: bodil.svennblad@ucr.uu.se

Estimating divergence times in a phylogenetic tree without assuming a global molecular clock is a nontrivial task. In phylogenetic inference, branch lengths are a product of rates and times and therefore estimated divergence times cannot be extracted without additional assumptions or information about rates. If a global molecular clock is assumed and at least one time calibration node is known, then the rate can be estimated and hence also the divergence times of the internal nodes.

If the global molecular clock assumption is violated, a method of divergence times estimation assuming a molecular clock gives misleading results. No method can consistently estimate divergence times without assumptions about rate variation over the tree (Britton, 2005). There are methods that implement relaxed clocks, such as nonparametric rate smoothing (e.g., R8s; Sanderson, 2003) or local molecular clocks (e.g., BASEML; Yang, 1997; and QDATE; Rambaut and Bromham, 1998). For a review of methods estimating divergence times, see Rutschmann (2006).

A recent contribution in the spirit of local molecular clocks and rate smoothing is PATHd8 (Britton et al., 2007), a nonparametric method that smoothes