HAL
open science

Multipolar Consensus for Phylogenetic Trees

Cécile Bonnard, Vincent Berry, Nicolas Lartillot

To cite this version:

Cécile Bonnard, Vincent Berry, Nicolas Lartillot. Multipolar Consensus for Phylogenetic Trees. Systematic Biology, 2006, 55 (5), pp.837-843. 10.1080/10635150600969880 . lirmm-00126475

HAL Id: lirmm-00126475

https://hal-lirmm.ccsd.cnrs.fr/lirmm-00126475v1

Submitted on 25 Jan 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Running head: Multipolar Consensus

Multipolar consensus for phylogenetic trees

Cécile Bonnard, Vincent Berry and Nicolas Lartillot[1]

Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier

UMR 5506, CNRS-Université de Montpellier 2

161, rue Ada, 34392 Montpellier Cedex 5, France

{cbonnard,vberry,lartillo}@lirmm.fr

tel: 00 33 4 67 41 86 55

fax: 00 33 4 67 41 85 00

[1]Corresponding author

Abstract

Collections of phylogenetic trees are usually summarized using consensus methods. These methods build a single tree, supposed to be representative of the collection. However, in the case of heterogeneous collections of trees, the resulting consensus may be poorly resolved (strict consensus, majority-rule consensus...), or may perform arbitrary choices among mutually incompatible clades, or splits (greedy consensus).

Here, we propose an alternative method, which we call the Multi-Polar Consensus (MPC). Its aim is to display all the splits having a support above a pre-defined threshold, in a minimum number of consensus trees, or poles. We show that the problem is equivalent to a graph coloring problem, and propose an implementation of the method.

Finally, we apply the MPC to real datasets. Our results indicate that, typically, all the splits down to a weight of 10% can be displayed in no more than 4 trees. In addition, in some cases, biologically relevant secondary signals, that would not have been present in any of the classical consensus trees, are indeed captured by our method, indicating that the MPC provides a convenient exploratory method for phylogenetic analysis. The method was implemented in a package freely available at `http://www.lirmm.fr/~cbonnard/MPC.html`.

Keywords : consensus, phylogeny, secondary signal, conciseness, graph-coloring.

Introduction

Two main methods have been proposed to account for uncertainty in phylogenetic reconstructions, both of which rely on resampling procedures. The first is the non-parametric bootstrap of (Efron, 1979), first applied to phylogenetics by Felsenstein (1985). More recently, a Bayesian alternative was proposed which samples trees directly from their posterior probability distribution (Huelsenbeck and Ronquist, 2001; Larget and Simon, 1999). These two methods have their respective advantages (Erixon et al., 2003; Huelsenbeck and Rannala, 2004), but on the other hand, they also have many features in common. In particular, both give as a raw output a collection of trees, instead of a single one, and the problem is then to summarize this collection into a synthetic, yet informative, picture.

By far the most common methods to do such a summary are the consensus methods. Their aim is simply to build a single tree displaying the most frequent splits (or clades) seen in the collection. Several variants exist reviewed in Bryant (2003): first, the strict consensus (McMorris et al., 1983) displays only the branches shared by all the trees. However, in most cases, this consensus is not informative enough. As an alternative, the majority-rule consensus (Margush and McMorris, 1981) aims at displaying the branches shared by at least 50% of the trees. This yields a much more satisfactory result than the strict consensus, and in fact, it is currently the method proposed by default by most phylogenetic softwares (Huelsenbeck and Ronquist, 2001; Swofford, 1998).

An issue about these consensus methods is when a split should be considered as reliable. A traditional statistical threshold of 95% is often used. Alternatively, for trees generated by bootstrap, a threshold of 70% has been proposed (Hillis and Bull, 1993). In both cases, assuming these thresholds are valid, a single consensus tree is sufficient, since branches appearing in more than

50% of the trees are always compatible. In practice, however, models used to build phylogenetic trees are simplistic compared to the true evolutionary processes. This often results in artifacts such as long-branch attraction, which can hide real phylogenetic signal (Brinkmann et al., 2005; Gaut and Lewis, 1995; Sullivan and Swofford, 1997). In the worst cases, the artifacts completely dominate, and there is no hope recovering the true corresponding relationships, except by using other models or methods. In most cases, however, phylogenetic signal in favor of the true relationships is still present in the data, as secondary signals. These signals are revealed to some extent in some of the trees built from bootstrap replicates or sampled by a Bayesian process. Therefore, it might be useful to have a method displaying these signals, and more generally, enabling one to explore the information contained in a collection of trees, including signals of lesser intensity.

An extreme solution to this problem would be to keep, and visualize separately, each split observed in the tree collection. However, this can be unreadable in practice, in particular for analyses with many taxa. Another possibility is to build the greedy consensus (Felsenstein, 1993), which consists in the majority-rule consensus tree complemented with all the branches that one can add to it by scanning the list of the remaining branches sorted by decreasing frequency. However, this only displays branches that happen to be compatible with those displayed by the majority-rule consensus tree. Therefore, it does not allow one to detect all biologically interesting alternatives to the splits having the highest weight. The same can be said from the asymetric median tree consensus of Phillips and Warnow (1996): although it aims at maintaining as much evolutionary information as possible contained in the input set of trees, it can only retain branches that can be combined into a single tree. Yet another way to summarize a tree collection is to build a consensus network (Holland et al., 2005), which displays all input splits in a single *graph*. Though this type

of structure requires a little bit of training to be interpreted, it allows one to spot the splits that conflict with one another provided there is a reasonable number.

An alternative to all these methods is to modify the consensus paradigm used traditionally, and summarize the tree collection by several trees instead of a single one. This idea has been explored by several previous works, which we now detail. The *Reduced Consensus* method of Wilkinson (1994) outputs a profile of trees representing *all* positive statements of relationships that are common to a set of input trees. In the present work, in contrast, we aim at representing all splits that appear *sufficiently frequently*. In this direction, several clustering methods have also been proposed. The *Phylogenetic Islands* method (Maddison, 1991) first clusters the input trees into several tree islands, i.e. trees more closely related in terms of tree rearrangements, and then computes a distinct consensus tree for each island. More generally, alternative clustering methods were investigated by Stockham et al. (2002) in an extensive study on several data sets. Their results show the validity of the approach in that the information content of the trees output for each cluster is significantly greater than a single consensus of the initial collection, while only adding a small amount of complexity. Their results also show that the complete linkage method (well-known in classification) outperforms the other clustering methods, including Phylogenetic Islands, according to several measures of the information content.

The clustering methods retain more splits in their output than single-tree consensus methods. However, by their mere principle, i.e. representing each cluster of trees by a consensus tree, it is still possible that splits present in a large proportion of the input trees impede alternative, less frequent, splits to be displayed.

Here, we propose an alternative method, called Multi-Polar Consensus (MPC), which explicitly

aims at representing all splits above a pre-defined threshold in as few trees as possible. In the following, we first formally define the MPC and show that the resulting optimization problem reduces to one of graph coloring. Then, we describe a simple algorithm implementing the method. Lastly, an application to several protein datasets with long branch attraction problems shows that the method enables secondary phylogenetic signal to emerge and to be displayed together with the primary signal by allowing very few extra trees in the output.

Materials and Methods

Formalisms

Splits and trees

Each branch of a given phylogenetic tree divides the set of taxa into two subsets: it is therefore equivalent to a *split*. A tree, whether rooted or unrooted, can always be characterized by a set of splits. Two splits are *compatible* if they can be displayed by the same tree. A collection of splits is compatible if every pair of splits is compatible. To build a multipolar consensus, we start from a collection of trees on the same set of taxa. We collect the list of the weighted splits displayed by the internal branches of the trees, where the weight assigned to a split S_i, denoted $f(S_i)$, represents its frequency in the collection (Bryant, 2003). The split list is a synthetic representation of the collection. We discard splits displayed in no more than a proportion α ($0 < \alpha \leq 1$) of the trees and call $\mathscr{L}_\alpha$ the *reduced split list*.

Representing split interaction with a graph

A graph G is composed of a set of *vertices*, V, and a set of *edges*, E, linking pairs of vertices. A *clique* is a subset of vertices all pairwise linked by an edge. A *maximal clique* is a clique not included in another clique. A *maximum clique* is the largest maximal clique of the graph. The *complementary graph*, $\overline{G}$ of G represents the same set of vertices, V, but two vertices are linked in $\overline{G}$ if and only if they are not linked in G.

The *compatibility graph* $G(\mathscr{L}_\alpha)$ of the split list $\mathscr{L}_\alpha$ is defined as follows: each split S_i with $f(S_i) > \alpha$ is represented by a vertex weighted by $f(S_i)$, and two vertices are linked by an edge if and only if they are compatible. Since a tree is a set of compatible splits, any tree of the collection is trivially represented by a clique in the graph $G(\mathscr{L}_0)$. Reciprocally, any clique of the compatibility graph represents a (possibly multifurcated) tree (Nelson, 1979). For instance, the greedy consensus tree is formed by searching a maximal clique in the graph $G(\mathscr{L}_0)$. This clique is first composed of all splits S_i such that $f(S_i) > 0.5$ and then considering splits in decreasing order of $f(S_i)$ and adding them to the clique only if they are linked to all the vertices already chosen (subgraph represented in bold in figure 1-C(3)).

The aim of the MPC is to represent all the splits of $G(\mathscr{L}_\alpha)$ in as few trees as possible. In the present graph-theoretic formulation, this is equivalent to searching for a *minimum-sized set of cliques covering* $G(\mathscr{L}_\alpha)$. Note that, when $\alpha \geq 0.5$, the whole graph is a clique and in this case, the MPC outputs just one tree, which is just the majority-rule consensus tree ($\alpha = 0.5$) or the strict consensus tree ($\alpha = 1$). In contrast, when $\alpha < 0.5$, $G(\mathscr{L}_\alpha)$ is not always a clique, and then, we need several cliques to cover all its vertices. Each corresponds to a tree, called a pole, and the poles form the MPC.

Kernel splits

In the graph $G(\mathscr{L}_\alpha)$, the vertices that are linked to every other vertex form a particular clique, which we call the *kernel clique*, and denote $kernel(\mathscr{L}_\alpha)$. The splits represented by the vertices in this clique, called *kernel splits*, form the *kernel tree* which corresponds to the loose consensus tree (Bremer, 1990) made on $\mathscr{L}_\alpha$. Note that the kernel splits can be included in all poles. More generally, some splits will be compatible with several poles, which raises the issue of how many times a split should appear in the MPC.

In the present paper, we use the following "*all-or-none*" rule: (1) any split in the graph $G(\mathscr{L}_\alpha)$, except the kernel splits, belongs to only one pole, and (2) all kernel splits are included into all poles as a common backbone. The poles of the MPC are thus obtained by first removing kernel splits from $G(\mathscr{L}_\alpha)$, then computing a set of cliques partitioning the vertices of $G(\mathscr{L}_\alpha)$ not in $kernel(\mathscr{L}_\alpha)$ and last, adding $kernel(\mathscr{L}_\alpha)$ to each clique. In this way, the previous problem, covering a graph with cliques, reduces to one of partitioning the graph with a minimum number of cliques. The MPC shares some similarities with the Nelson consensus as described in Page (1990), which seeks *a* clique of maximal weight in the compatibility graph (the weight of the clique is here defined as the sum of the $f(S_i)$ values on all vertices S_i of the clique). The main difference with the MPC is that all splits are not included in the output.

Partitioning a graph with cliques is a classical problem in graph theory, but traditionally, this problem is translated into the complementary graph $\overline{G}(\mathscr{L}_\alpha)$, which we call *incompatibility graph*. It then becomes a minimum coloration problem in the graph $\overline{G}(\mathscr{L}_\alpha)$, i.e. an assignment of colors to the vertices of the graph such that two vertices linked by an edge do not have the same color and a minimum number of color is used. Each pole of the MPC is composed by all vertices of the

same color.

As an example, in figure 1.C, the graph (4) displays the incompatibility graph $\overline{G}(\mathscr{L}_0)$ for the collection $\mathscr{T}$ of figure 1.A, where S_1 and S_2 are the kernel splits. We can assign two colors to the rest of the vertices (one to S_3, S_6 and one to S_4, S_5, S_7), leading to two poles in the MPC, respectively containing splits $\{S_1, S_2, S_3, S_6\}$ and $\{S_1, S_2, S_4, S_5, S_7\}$.

Algorithm

The graph colouring problem of a general graph is NP-complete. Note that there is a well-known linear-time algorithm to know whether a graph can be two-coloured. However two colours are often not enough for reasonable values of α, and unfortunately, this algorithm does not generalize to more than two colours. Additionally, the practical graphs we obtained from real data sets had no obvious property affiliating them to specific classes of graphs for which the coloration problem is polynomial. Therefore, to find a MPC we rely on a heuristic coloration scheme called the Greedy Coloring Algorithm.

Algorithm Greedy coloring

1. `Create an order on the vertices.`

2. `Consider vertices one by one in this order, assigning to a vertex the first color not assigned to an already colored vertex linked to it.`

The efficiency of the greedy algorithm in proposing a small number of colors critically depends on the order defined on the set of vertices. A possible order on the vertices is that of decreasing weight. In fact, the greedy consensus tree is composed by the set of vertices to which the Greedy coloring scheme assigns the first color when the splits are sorted by decreasing frequency. Alternatively, the

vertices can be sorted by decreasing degree (Welsh-Powell algorithm, Welsh and Powell (1967)).

We implemented both versions of the algorithm and compared the MPC to other consensus methods, as described in the next paragraph.

Software

An implementation of the MPC was written in C++ using the STL. The source code and static libraries can be freely downloaded from `http://www.lirmm.fr/~cbonnard/MPC.html`. The program allows the user to control both the threshold α and the number of poles. It also includes other options not introduced in this paper.

Data and experimental validation

We used several datasets, some consisting of primary data from which we obtained collections of input trees, and others consisting of lists of trees obtained from other studies. The primary data is composed of 133 genes spanning 44 species and including, on average, around 200 amino acid positions per gene (Brinkmann et al., 2005). The datasets covers the eukaryotic phylogeny and contains 6 archaea (used as the outgroup) and 38 eukaryotes. Because of long branch attraction between fast evolving eukaryotes and archaea, the consensus trees may show some problematic groups of taxa. Based on an sequence selection procedure, Brinkmann et al. (2005) propose a eukaryotic phylogeny devoid of most of these artifacts, which we used as our reference tree.

For each gene, we performed 100 replicates of non-parametric bootstrap, which we analyzed by the Maximum Parsimony (MP) criterion, using the PAUP software (Swofford, 1998), and by the Maximum Likelihood criterion, using PHYML (Guindon and Gascuel, 2003). In this way, we

obtained 2 sets of 133 tree collections, to which we individually applied various consensus methods: the strict, majority-rule and greedy consensus methods, as well as the MPC.

Additionally, we used the four collections of trees that served as a basis for the experimental study of Stockham et al. (2002). The *Caesal* dataset is a collection of 450 trees on 51 taxa of *Caesalpinia*. The *Camp* dataset contains 216 trees on 13 taxa of *Campanulaceae*. The *Pevcca1* and *Pevcca2* datasets span 129 taxa and respectively contain 216 and 654 trees. *Pevcca* stands for *Porifera* (sea sponges), *Echinodermata* (sea urchins, sea cucumbers), *Vertebrata* (fish, reptiles, mammals), *Cnidaria* (jellyfish), *Crustacea* (crabs, lobsters, shrimp), and *Annelida* (annelid worms).

Results

The MPC aims at representing, in a minimum number of trees, all splits whose frequency in the tree collection is higher than a given threshold α. In a first experiment, we measured the number of trees contained in the MPC, as a function of α.

Figure 2 displays the result obtained on the data sets of Brinkmann et al. (2005). As expected, the number of poles in the MPC is a decreasing function of the threshold. More interestingly, for small values of α, the decrease is very steep, so that for a threshold of 10%, the MPC already contains less than 5 poles in average. This means that if we relax the constraint of having a unique tree to form the consensus, we can represent a substantial amount of additional information (namely, all the splits with frequencies between 10% and 50% not represented in the greedy consensus tree). These splits represent an average fraction of 25% of the total number of splits of the graph $G(\mathscr{L}_\alpha)$. The latter result was obtained using the Welsh-Powell greedy coloring algorithm. We also tried the greedy coloring method on the split list sorted by decreasing weight: this almost always resulted

in a higher number of poles in the MPC (one additional pole on the average, not shown).

In some cases, these extra splits displayed by the MPC may represent biologically relevant signals. For example, the MP analysis of the bootstrap replicates generated for one of the proteins of the dataset (proteasome α-subunit on 35 taxa) yields the following picture (figure 3): the fungi are not correctly resolved neither by the strict consensus nor by the majority-rule consensus. The greedy consensus displays a phylogeny where ascomycetes are polyphyletic, which is incongruent with current knowledge. In contrast, a pole of the MPC (which contains only 2 poles in this case) displays a more satisfactory configuration, including monophyly of the ascomycetes, with a bootstrap support equivalent to splits included by the greedy consensus.

To quantify the latter phenomenon, we measured how many among all the correct splits, according to the tree of Brinkmann et al. (2005), are displayed by each consensus tree (figure 4). We first note that increasing the threshold α, i.e. lowering the number of splits considered by the methods, leads the loose consensus (KE) to contain more and more correct splits. This paradoxal observation is explained by the fact that, in this study, numerous correct splits are present in a large proportion of input trees. Progressively discarding incorrect alternative splits of lower support decreases the degree of the vertices corresponding to correct splits S_i in the incompatibility graph $\overline{G}(\mathscr{L}_\alpha)$. For a high enough α (and provided $\alpha < f(S_i)$) such a vertex ultimately reaches degree zero and S_i is then included in the loose consensus tree. Otherwise, for thresholds $\alpha \leq 50$ (and by the mere principle of the method), the MPC allows a higher proportion of correct splits to be displayed than the majority-rule and the greedy consenses. Specifically, the majority-rule consensus tree contains on average only 28% of the correct splits. Considering splits appearing with a frequency between 10% and 50% in tree collections enables to discover an additional 20% of correct splits. Among

those, slightly more than a third are displayed by the greedy consensus, the remaining two thirds appearing in the MPC only. Thus, compared to the MPC, the greedy consensus misses a significant amount of relevant information. Yet, it is the method including the most information among those proposing a single tree to summarize a tree collection. We also note that, interestingly, the majority of the correct splits with a low support and ignored by the greedy consensus have however a higher support than the split with lowest support displayed in the greedy consensus tree.

We also applied the MPC to lists of trees analyzed previously Stockham et al. (2002). Once again, setting a threshold $\alpha = 10\%$ is sufficient on average for four poles to represent all retained input splits (the *Camp* dataset needing up to $\alpha = 10\%$). Three poles were obtained on average when setting $\alpha = 15\%$, and two poles (i.e. cases when $\overline{G}(\mathscr{L}_\alpha)$ can be bicolored) were achieved for values of α between 30% and 35%. These results confirm the findings of figure 2. As expected, the computing time required to obtain the four poles differs from one dataset to the other depending on their size: the *Camp* dataset is processed in 1.2 seconds (on average over 10 runs), while *Pevca2*, the largest dataset, is processed in 43 seconds on average.

Discussion

As an alternative to the classical consensus methods such as the majority-rule or greedy consensus, we propose to build a multipolar consensus (MPC), which consists in a small set of trees displaying all the splits with a support greater than a pre-defined threshold. Given the splits to display, the number of trees in the MPC is to be minimized. Our main motivation is to allow for more secondary evolutionary signal to be displayed than what is proposed by the majority-rule consensus, without making the kind of somewhat arbitrary choices underlying the greedy consensus. Our last

experiment illustrates that such arbitrary decisions, where some splits are discarded to the profit of less supported splits because of incompatibility with a split of higher support, are very common.

The MPC is meant as an exploratory device to analyze the output of a given phylogenetic study: often, analyses of real data may lead to artifacts and/or contradictory signals, due for instance to model-misspecification problems. Sometimes the artifacts can have a greater support than some real splits, and therefore, examining a larger part of the collection of splits, including those with lower support, reveals additional correct splits, omitted by classical consensus methods. Including splits with lower support also adds incorrect splits to the proposed picture. However, they are in relatively small number (on average, 3 or 4 poles are enough to represent all conserved splits, including correct and incorrect ones), so that the MPC really fulfills its role in giving a concise exploratory picture.

Two other approaches intending to display more input splits than can fit into a single tree are the the consensus network (Holland et al., 2005) and clustering methods (Maddison, 1991; Stockham et al., 2002). Compared to the MPC, the consensus network has the advantage of yielding only one structure as output. On the other hand, the complexity of such a structure can make it difficult to be interpreted by people not acquainted with the method. Besides these differences, the conciseness of the results of both MPC and consensus networks are dependent on the size of the maximal clique in the incompatibility graph. A first consequence is that the worst case situations will be the same for both methods. In the most extreme cases, and for large values of α, the MPC outputs as many poles as input trees and the consensus networks contain a large number of vertices compared to the number of initial splits. In typical cases however, it is likely that the MPC and the consensus network give outputs of comparable size. For example, discarding splits present in less than 10%

of the input trees typically leads to a network containing 3-cubes as more complex parts (Barbara Holland, personnal communication). This echoes our findings that the MPC usually outputs less than 5 poles for $\alpha = 10\%$.

The MPC is also quite close in spirit to the clustering approach (Maddison, 1991; Stockham et al., 2002): namely, both methods propose several consensus trees, instead of a single one. The MPC differs from the clustering methods in at least two respects: first, it is more parsimonious, in that any non-kernel split present in an input tree, is represented only once. Second, it does not require any preliminary clustering of the input trees. As a consequence, the computing time of MPC is polynomial in the number of input splits but only linear in the number of input trees. This might be an advantage as practical collections of trees tend to be dense (e.g., the data sets of Stockham et al. (2002) contained on average three times less splits than trees).

Yet, at least in its present state, the MPC has a drawback compared to the clustering approach: the way in which splits are grouped in poles does not exactly reflect their cooccurence in trees of the input collection. However, it is possible to reduce this effect by giving a weight to the edges of the compatibility graph, corresponding to the number of cooccurrence of the splits (vertices) in the collection. This weight can then be taken into account when agglomerating splits into poles so that the splits composing a pole do not come from sources that are too heterogeneous.

We then face a multi-criterion optimization problem, where the best compromise between a small number of poles and a minimum intra pole heterogeneity is sought. As a first solution to this problem, the software we implemented enables the user to specify a minimum percentage of input trees in which splits should jointly appear in order to be output in a same pole. Researches on this topic are currently on progress to obtain a bicriterion algorithm for simultaneously optimizing the

number of poles and the co-occurrence of the splits in the poles.

Acknowledgment

We thank Li-Shang Wang for providing us with the data sets used in Stockham et al. (2002), as well Hervé Philippe, Nicolas Rodrigue, and the three referees for their useful comments on the manuscript. This work was funded by the Centre National de la Recherche Scientifique, the french funding program ACI IMP-BIO "Model Phylo".

References

Bremer, K. 1990. Combinable component consensus. Cladistics 6:369–372.

Brinkmann, H., M. van der Giezen, Y. Zhou, G. Poncelin de Raucourt, and H. Philippe. 2005. An empirical assessment of long branch attraction. Syst. Biol. 54:745–757.

Bryant, D. 1997. Building trees, hunting for trees, and comparing trees. Phd thesis University of Canterbury.

Bryant, D. 2003. A classification of consensus methods for phylogenetics. Bioconsensus, DIMACS Series in Discrete Mathematics and Theoretical Computer Science. Pages 163–184.

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

Efron, B. 1979. Bootstrap methods: another look at the jackknife. Ann. Statistics 7:1–26.

Erixon, P., B. Svennblad, T. Britton, and B. Oxelman. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52:665–673.

Felsenstein, J. 1985. Confidence Limites on Phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Felsenstein, J. 1993. Phylip http://evolution.genetics.washington.edu/phylip.html.

Felsenstein, J. 2004. Inferring Phylogenies. Sinauer Associates, Sunderland, Massachusetts.

Gaut, B. S. and P. O. Lewis. 1995. Success of the maximum likelihood phylogeny inference in the four taxon case. Mol. Biol. Evol. 12:152–162.

Guindon, S. and O. Gascuel. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst. Biol. 52:696–704.

Hendy, M. D., M. A. Steel, D. Penny, and I. M. Henderson. 1991. Families of trees and consensus. Syst. Zool. 40:315–328.

Hillis, D. M. and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assesing confidence in phylogenetic analysis. Syst. Biol. 42:182–192.

Holland, B. R., F. Delsuc, and V. Moulton. 2005. Visualizing conflicting evolutionary hypotheses in large collections of trees using consensus networks. Syst. Biol. 54:66–76.

Huelsenbeck, J. and B. Rannala. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53:904–913.

Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Larget, B. and D. Simon. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Mol. Biol. Evol. 16:750–759.

Lartillot, N. and H. Philippe. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 21:1098–1109.

Maddison, D. R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. Syst. Zool. 40:315–328.

Margush, T. and F. R. McMorris. 1981. Consensus n-trees. B. Math. Biol. 2:239–244.

McMorris, F. R., Meronk, and D. A. Neumann. 1983. A view of some consensus methods for trees. J. Felsenstein ed., Numerical taxonomy, Berlin.

Nelson, G. 1979. Cladistic analysis and synthesis: principles and definitions, with a historical note on Adanson's *famille des plantes (1763-1764)*. Syst. Zool. 28:1–21.

Page, R. D. M. 1990. Tracks and trees in the antipodes: A reply to Humphries and Seberg. Syst. Zool. 39:288–299.

Phillips, C. and T. Warnow. 1996. The asymetric median tree – a new model for building consensus trees. Disc. Appl. Math. 71:311–355.

Stockham, C., L.-S. Wang, and T. Warnow. 2002. Statistically based postprocessing of phylogenetic analyses by clustering. Bioinformatics 18:S285–S293.

Sullivan, J. and D. L. Swofford. 1997. Are guinea pigs rodents? the importance of adequate models in molecular phylogenetics. J. Mammal. Evol. 4:77–86.

Swofford, D. L. 1998. Paup*, phylogenetic analysis using parsimony (*and other methods) vol. 4. Sinauer Associates, Sunderland, Massachusetts.

Welsh, D. and M. Powell. 1967. An upper bound on the chromatic number of a graph and its applications to timetabling problems. Comput. J. 10:85–87.

Wilkinson, M. 1994. Common cladistic information and its consensus representation: Reduced Adams and reduced cladistic consensus trees and profile. Syst. Biol. 43:343–368.

Figure 1. Compatibility graph of a split list

(a) A collection of trees $\mathscr{T}$; (b) the corresponding split list $\mathscr{L}$; (c) and its compatibility graphs: $G(\mathscr{L}_{\frac{t-1}{t}})$ (1), $G(\mathscr{L}_{0.5})$ (2), $G(\mathscr{L}_0)$ (3), with in bold the clique corresponding to the greedy consensus; and the incompatibility graph $\overline{G}(\mathscr{L}_0)$ (4).

Figure 2. Number of trees in the MPC as a function of the threshold.

The MPC was built using the greedy algorithm on the Welsh-Powell order.

Figure 3. A case study using the α subunit of the proteasome on 35 taxa.

Phylogenetic relationships for a subset of 7 fungi as proposed by (a) the MPC, (b) the strict, majority-rule, and greedy consenses, and (c) the reference tree (Brinkmann et al., 2005).

Figure 4

Average proportion of the correct splits retrieved by the different methods (under ML criterion) for various thresholds α (splits appearing in less than α of the input trees are ignored by the consensus methods): majority-rule consensus (MRC), greedy consensus (GC), loose consensus (KE) and the MPC.