

Congruence of Morphological and Molecular Phylogenies

Davide Pisani · Michael J. Benton · Mark Wilkinson

Received: 10 March 2007 / Accepted: 1 July 2007 / Published online: 27 July 2007
© Springer Science+Business Media B.V. 2007

Abstract When phylogenetic trees constructed from morphological and molecular evidence disagree (i.e. are incongruent) it has been suggested that the differences are spurious or that the molecular results should be preferred a priori. Comparing trees can increase confidence (congruence), or demonstrate that at least one tree is incorrect (incongruence). Statistical analyses of 181 molecular and 49 morphological trees shows that incongruence is greater between than within the morphological and molecular partitions, and this difference is significant for the molecular partition. Because the level of incongruence between a pair of trees gives a minimum bound on how much error is present in the two trees, our results indicate that the level of error may be underestimated by congruence within partitions. Thus comparisons between morphological and molecular trees are particularly useful for detecting this incongruence (spurious or otherwise). Molecular trees have higher average congruence than morphological trees, but the difference is not significant, and both within- and between-partition incongruence is much lower than expected by chance alone. Our results suggest that both molecular and morphological trees are, in general, useful approximations of a common underlying phylogeny and thus, when molecules and morphology clash, molecular phylogenies should not be considered more reliable a priori.

D. Pisani (✉)

Laboratory of Evolutionary Biology, Department of Biology, The National University of Ireland, Maynooth, Co. Kildare, Ireland
e-mail: davide.pisani@nuim.ie

D. Pisani · M. Wilkinson

Department of Zoology, The Natural History Museum, London SW7 5BD, UK

D. Pisani · M. J. Benton

Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK

Keywords Congruence · Morphology · Molecules · Phylogenetic accuracy · Tree distance metrics

1 Introduction

For most of its history phylogenetics has depended primarily upon morphological data. Apart from a few rare cases (e.g. Yang et al. 1996) the study of extinct species is still the realm of morphology. However, since Zuckerkandl and Pauling's (1965) intuition that molecules may document evolutionary history, protein and DNA sequences ('molecular data') have been increasingly used to investigate the phylogeny and divergence times of extant organisms (Pagel 1999).

The remarkable growth of molecular phylogenetics (Pagel, 1999) has been accompanied by sometimes acrimonious debate on the relative utility of molecular and morphological data (Gura 2000). Fuelled by observations that molecular and morphological phylogenies often seem to be at odds with each other, this debate continues to produce very disparate opinions; compare Scotland et al. (2003) with both Jenner (2004) and Wiens (2004), or Benton (1999) with Easteal (1999). It continues partly because we do not know what proportion of alternative molecular and morphological trees are incongruent and to what extent any incongruence is spurious or real (see also Scotland et al. 2003). With the single notable exception of Bledsoe and Raikow (1990), who performed a statistical analysis of the congruence of a total of 48 (mostly molecular) trees, all studies of the congruence of molecules and morphology have focussed on only one or a few case studies (see Hillis and Wiens 2000 for a recent survey). While case studies have yielded important insights, they may lack generality, and sometimes have further inflamed debate; compare for example Lee (1997) with Hedges and Maxson (1996, 1997)).

We performed the largest survey to date of the congruence of molecular and morphological phylogenies. Essentially we use tree-to-tree distance metrics to quantify the incongruence between pairs of trees (see Sect. 2), with the aims of evaluating (I) whether molecular and morphological trees are significantly incongruent, and (II) whether there is empirical evidence suggesting greater utility of one data type over the other.

2 Materials and Methods

2.1 The Data Base

For three groups of amniotes, mammals, birds, and turtles, phylogenetic trees were culled from the literature using key word searches of Web of Science (<http://wos.mimas.ac.uk/>) over the whole time span covered at the time this study was initiated (1981–2001) and by manual searching of selected journals (*Systematic Biology*, *Molecular Biology and Evolution*, *Molecular Phylogenetics and Evolution*, *The Auk*, *Journal of Molecular Evolution*, and the *Zoological Journal of the Linnean Society*). All publications that were likely to include a phylogenetic tree were examined, and the bibliographies of the retained articles (those presenting one

or more phylogenetic trees) were screened for additional trees. Trees were retained irrespective of the phylogenetic method used to derive them. Where different trees were inferred using different methods each was retained. For studies that yielded multiple equally optimal trees we used a published consensus if available. Otherwise we used RadCon (Thorley and Page 2000) to compute the strict reduced consensus tree (Wilkinson 1994) with the highest cladistic information content (Thorley et al. 1998). We used consensus trees as a proxy for the (often unavailable) optimal trees that they represent and the reduced consensus to minimize number of polytomies (which flatten tree-to-tree distances) in the consensus. Trees were stored in separate files.

The collected trees were partitioned into subsets that overlapped in their included taxa. Trees sometimes differed in the Linnean rank of the included taxa and sometimes employed different taxa (e.g. genera) as representatives of a common higher rank taxon (e.g. family). To facilitate tree-to-tree comparisons the leaf labeling of the trees within each profile was homogenised to include only genera, or, more rarely, families. Family to genus transformations were carried out by replacing families with a polytomy subtending all the genera that are present in the other trees of the same profile and considered members of the family. Genus to family transformations were also performed, and where there was a lack of general agreement among authors, assignment to family followed electronic taxonomy databases. Mammals were homogenised with reference to the Smithsonian National Museum of Natural History, Mammal Species of the World Database (MSW database; <http://www.nmnh.si.edu/msw>), and birds were homogenized with reference to the Field Museum Birds Collection Database (<http://fm1.fieldmuseum.org/collections/search.cgi?dest=birds>). The resulting data base contained 230 (49 morphological and 181 molecular) trees permitting some 929 and 879 potentially meaningful pairwise comparisons of overlapping trees in the rooted and unrooted cases respectively (see below). References for the original articles where the trees used in this study were published can be downloaded (Table S1) from the web site: <http://www.bmnh.org/~mw/Data/>.

2.2 Tree Distance Metrics

Incongruence between trees was estimated as their tree-to-tree distance, using three tree distance metrics: Partition Metric (PM; Robinson and Foulds 1981) which is the symmetric difference on full splits (bipartitions of the full leaf set), and the quartet and triplet Do not Conflict (DCq, and DCt; Estabrook et al. 1985; Critchlow et al. 1996) which are the number of quartets (subtrees on four leaves) and triplets (a quartet in which one leaf is the root) that are resolved and different respectively. DCt and DCq differ in treating the trees as rooted or unrooted respectively and comparing them informs on the impact of rooting upon congruence. Each full split in a tree entails a set of triplets/quartets and two trees that do not share any full splits may share some triplets/quartets. As a consequence, triplet and quartet based metrics have a greater range than full split based metrics like PM, and trees that are maximally distant using PM can still be discerned and ranked using DCt and DCq (Day 1983). All overlapping pairs of trees were compared and in cases where two

trees only partially overlapped, the trees were pruned and only the subtrees induced by the shared taxa were compared. Tree to tree distances for all possible pairwise comparison between the trees in our data set can be downloaded (Table S2) from the web site: <http://www.bmnh.org/~mw/Data/>.

To avoid problems due to data non-independence, not all distances in Table S2 were used for statistical analyses (see below the “*Data Selection and Statistical Analyses*” section for details). All tree-to-tree distances were calculated using Component Lite (Page 1997) and were normalised to allow meaningful comparison of results for differently sized leaf sets. For each metric considered, regression analysis was performed on the normalized tree-to-tree distances to evaluate whether the number of taxa in the compared trees was influencing their incongruence (results not shown; but see Pisani 2002). These analyses demonstrated that the normalised tree-to-tree distances were independent of the number of taxa (i.e. the cardinality) of the compared trees and thus the existence of a possible “cardinality effect” on the congruence of the pairwise compared trees was ruled out and not considered further.

2.3 Classes of Data

Tree-to-tree distances were classified according to the data underpinning the compared trees, methods of analyses and date of publication as follows. With trees inferred from either molecules (M) or morphology (m) there are three categories for distances between two molecular trees (MM), two morphological trees (mm), and a molecular and a morphological tree (Mm). Some data, e.g. caryological, can be considered either molecular or morphological. Here the very few (see Table S2) caryological and behavioural data sets considered were assumed to be molecular and morphological respectively. Alg and Opt distances are those between pairs of trees obtained using an algorithmic clustering method (e.g. UPGMA and Neighbor Joining) or an optimality criterion based method (*sensu* Swofford et al. 1996; e.g. Maximum Parsimony, Maximum Likelihood and Minimum Evolution) respectively. Lastly, Post 95 distances are those from pairs of trees published from 1995 onwards. This is an arbitrary temporal division of the data, and was chosen because 1995 has a special relevance in the history of molecular biology, marking the release of the first complete genomic sequence (Fleishmann et al. 1995) and the beginning of the “genomic era”.

2.4 Data Selection and Statistical Analyses

Not all the inferred tree-to-tree distances (Table S2) are independent (see also above). Non-independence arises when (i) trees are inferred from the same data with multiple methods, (ii) trees are inferred from overlapping data (e.g. separate and combined analyses of partitions or a historical sequence in which data is expanded through addition of taxa or characters) and (iii) a single tree contributes to multiple tree-to tree distances. To avoid non-independence in our statistical analyses, we (i) excluded any distances between trees inferred from identical data, (ii) used only the tree-to-tree distances based on the most inclusive of any overlapping data sets, and (iii) randomly selected a single tree-to-tree distance from those involving the same tree. Further, for comparisons involving a single (e.g. morphological) tree to multiple

(e.g. molecular) trees that were inferred from the same data, we used an average tree-to-tree distance.

For each data class, several transformations to normalise their distribution and homogenise their variances were attempted, and, in general, the best results were obtained using the square root transformation. However, data normalisation and homogenisation of variances could not always be achieved, and thus non-parametric statistical methods: the U Test of Mann–Whitney (Sokal and Rohlf 1995), the Kruskal–Wallis test (Sokal and Rohlf 1995; Siegel and Castellan 1989) and non-parametric multiple comparisons (Siegel and Castellan 1989) were used to test for significant differences in the congruence of various data partitions. This is an important departure from Bledsoe and Raikow (1990), which performed only limited non-parametric statistical analyses, and based their main conclusions from the results they obtained from standard ANOVA although their data violated its normality and homoscedasticity assumptions (Bledsoe, personal communication). All statistical analyses were performed using the statistical data analysis software “R” (www.r-project.org/).

A first analysis investigated the potential impact of method of analysis upon congruence. The distribution of the pairwise tree-to-tree distances within the Alg and Opt classes was investigated generating a boxplot representation of the data, and the Mann–Whitney U test was used to test the null hypothesis that the median (θ) Alg distance (θ -Alg) and the median Opt distance (θ -Opt) are not significantly different ($H_0: \theta$ -Alg = θ -Opt). Rejection of the null hypothesis would imply that the method of analysis used to recover the compared trees significantly affected their congruence.

Congruence of molecular and morphological trees was investigated by generating a boxplot representation of the distances within the three defined groups (MM, Mm, and mm). The Kruskal–Wallis test was used to test the null hypothesis of no significant differences between the median distances of the three groups ($H_0: \theta$ -MM = θ -Mm = θ -mm). Because this test cannot distinguish which, if any, are the significantly different medians, non-parametric multiple comparisons (Siegel and Castellan 1989) were performed to test for significant differences between the medians of each pair of these groups: (1) $H_0: \theta$ -MM = θ -Mm, (2) $H_0: \theta$ -mm = θ -Mm, and (3) $H_0: \theta$ -MM = θ -mm. For the distances in the Mm class to be significantly different from those in the MM and/or mm classes, i.e. for molecular and morphological trees to be significantly incongruent, the first and/or the second null hypotheses must be rejected. Rejection of the third null hypothesis would indicate that the molecular trees are significantly more similar to each other than morphological trees or vice versa. Parallel analyses were performed using only (i) Opt class distances in order to exclude any possible effect of the use of algorithmic clustering methods upon congruence and (ii) Post -95 Opt class distances to investigate whether changes in phylogenetics during the 21 year window we have studied affect congruence.

3 Results and Discussion

The boxplot representation of the data partitioned according to the method of analysis (i.e. the distribution of the distances within the Alg and Opt classes)

showed that trees obtained using optimality criteria were generally more similar to each other than are trees obtained using algorithmic methods (Fig. 1) when tree-to-tree distances were calculated using PM and DCt. The superiority of optimality criterion-based methods over clustering methods has been claimed (Swofford et al. 1996; Huelsenbeck and Hillis 1994), but is not universally accepted (Nei et al. 1998). Accepting congruence as evidence of accuracy (Miyamoto and Fitch 1995), our results suggest that optimality-based methods slightly outperform algorithmic clustering methods. However, assessed using the U test, the difference between the medians of the two groups (θ -Opt and θ -Alg) is not statistically significant (Table 1). In addition, no such difference is apparent using DCq (which treats trees as unrooted), suggesting that the incongruence revealed by the other measures may be due to incorrect rooting. Some clustering methods (e.g. UPGMA) root trees under dubious assumptions of rate homogeneity which we suspect may have inflated average incongruence in the Alg data when comparisons are of rooted trees (PM and DCt). In summary, the effect of method of analysis seems marginal, and more likely to impact upon congruence in rooted than in unrooted trees.

Comparing the congruence of the molecular and morphological phylogenies, the boxplot representation Fig. 2A) shows the θ -Mm distance is larger than both the θ -MM and θ -mm distances, that is, between partition incongruence is greater than within partition incongruence. Additionally, incongruence appears greater for morphological than for molecular data (θ -MM < θ -mm). The Kruskal–Wallis test shows the difference among all combinations of pairs of molecular and morphological trees (MM, Mm and mm) to be significant for all the three tree metrics considered (Table 2, Fig. 2A). These results hold also using Opt and post-95 Opt class distances (see Tables 2 and 3, Fig. 2B), and removing the pre-1995 trees did not result in an increase in the similarity of the compared molecular trees that might have been anticipated as a result of methodological advances (see Fig. 3). Non-parametric multiple comparisons (Table 3) showed θ -Mm to be significantly different (see also Fig. 2) from θ -MM for all considered distances and subsets of the data. Only with the most inclusive data is the difference between θ -Mm and θ -mm also significant (for PM and DCt but not for DCq, see Table 3)—a possible effect of incorrect rooting when trees inferred using clustering methods are included.

In their review, Hillis and Wiens (2000) suggested that much incongruence between morphological and molecular phylogenies may be spurious, reflecting incorrect rooting, poor choice of phylogenetic method, or conflict among only poorly supported relationships. Our results suggest that rooting has a marginal effect on incongruence in association with the use of algorithmic clustering methods. We also expect the inclusion of poorly supported relationships (noise) to increase incongruence. However, the combined results of the Kruskal–Wallis tests and of the non-parametric multiple comparisons show that the pairs of molecular and morphological phylogenies are significantly more incongruent than pairs of alternative molecular phylogenies, a pattern we would not expect if all incongruence were the result of randomly distributed noise. Either poorly supported relationships are non-randomly distributed with respect to data type or not all incongruence is due to poorly supported relationships. Morphological trees are much more similar to each other than they are to alternative molecular trees and less similar than pairs of

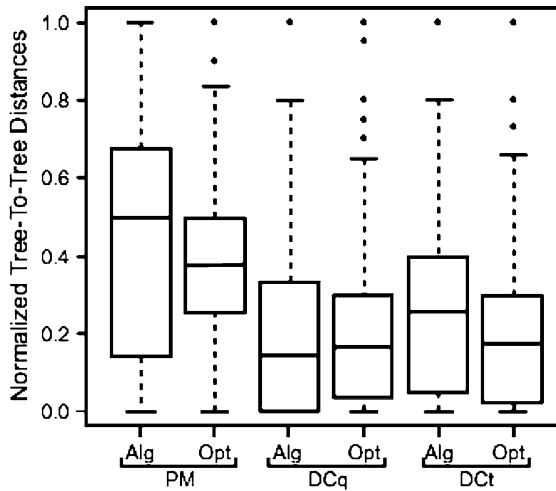


Fig. 1 Boxplot representation of the tree-to-tree distances within the Alg and Opt classes for all three metrics considered (PM, DCq, DCt). The boxes represent the values included between the lower and upper quartiles, with the medians (θ -Alg and θ -Opt) splitting the respective boxes in two. The whiskers extend to more extreme values (1.5 times the interquartile range from the box). Full circles represent outliers. This figure shows that trees obtained using methods based on the use of an optimality criterion are more congruent with each other than trees obtained using an algorithmic method, suggesting optimality criterion based methods tend to be more accurate. However, the difference between distances in the two classes is not critical. This is true for all the metrics considered (see also Table 1)

Table 1 Probabilities and results of the Mann–Whitney U test

	Tree distance metric		
	PM	DCt	DCq
Probability	0.05596	0.1670	0.1102
U	31670	25967.5	21597.5

This test was used to evaluate whether θ -Alg and θ -Opt were significantly different. Results are clearly not significant for DCq and DCt, but they are nearly significant for PM (PPM \sim 0.05). See Sect. 2 for abbreviations

molecular trees but the differences are not significant. This lack of significance might reflect a lack of statistical power due to the smaller sample size for mm distances, a consequence of our attempt to compare only trees inferred from independent data. However, the very high P -values (mostly unity) for the null hypothesis that θ -mm = θ -MM suggest that this result may be unlikely to change with any additional sampling. All the median distances for the metrics considered, including those between morphological and molecular trees, are much lower than expected for pairs of random trees (see Fig. 2). This general non-random congruence of pairs of trees, including molecular and morphological trees, suggests both data types usefully approximate the same underlying species tree (see also Penny et al. 1991).

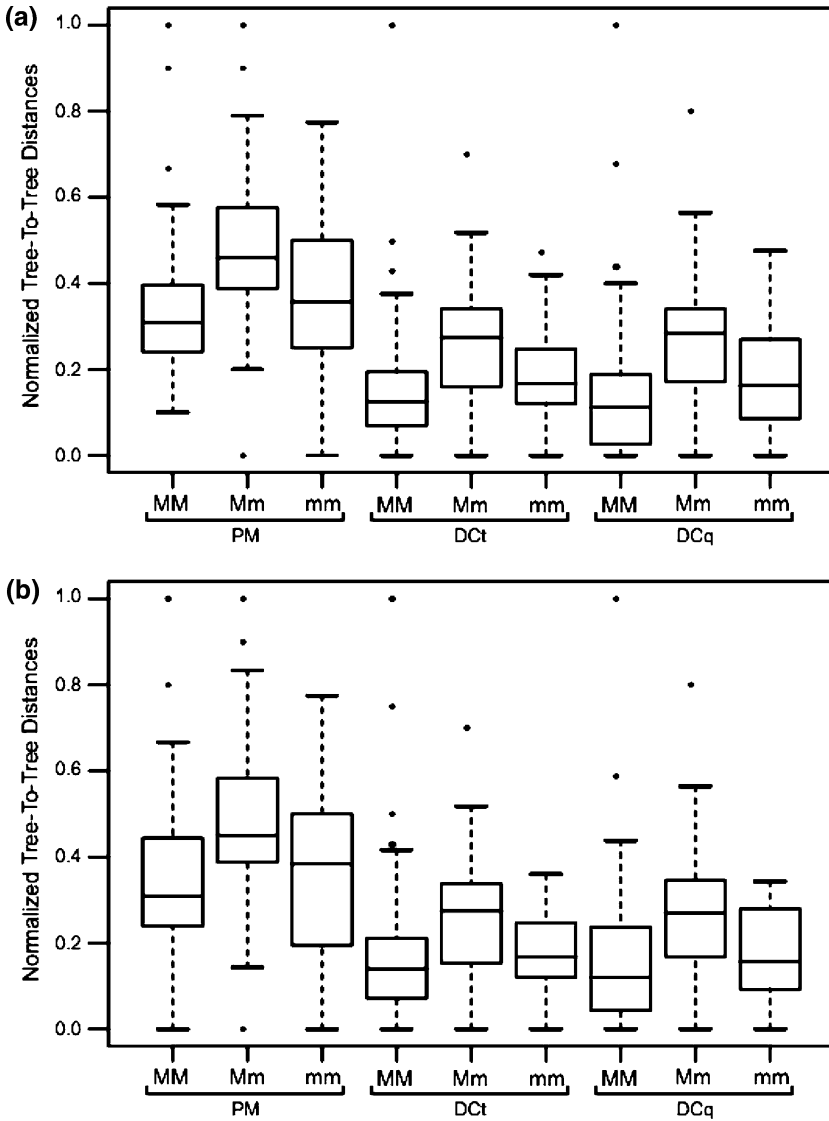


Fig. 2 Boxplot representation of the tree-to-tree distances in (a) the Full data set and (b) the Post-1995 data set. All the distances within the MM, Mm, and mm classes, for all three metrics considered (PM, DCq, DCt) are represented. The figure shows the Mm distances are higher than the MM and mm distances. That is, molecular and morphological phylogenies are incongruent (see also Table 2). Furthermore, θ -MM and θ -mm are very similar for all the metrics considered (see also Table 2 and text). Note also that within each of the considered classes (MM, Mm and mm) the compared trees were more similar to each other than random trees are expected to be. For example, two random trees are expected to have normalized DCq and DCt distance $\delta = 0.667$, i.e. even if incongruent molecular and morphological trees are still useful approximations of the same, underlying, unknown tree

Table 2 Probabilities and results of the Kruskal–Wallis test

Tree distance metric											
Opt and Alg Class Data											
Opt-Class Data			Opt-Class Data			Post-95 Opt-Class Data			Post-95 Opt-Class Data		
PM	DCt	DCq	PM	DCt	DCq	PM	DCt	DCq	PM	DCt	DCq
Probability	0.004523**	0.004086**	0.001528**	0.0013**	0.001013**	0.000029**	0.000022**	0.000047**	0.0002263**	0.000047**	0.0002263**
χ^2	10.797	11.0004	12.967	13.1729	13.7896	20.8359	21.4306	19.92	21.4306	19.92	16.7874
Degrees of freedom	2	2	2	2	2	2	2	2	2	2	2

The results show $\theta-MM$, $\theta-Mm$ and $\theta-mm$ to be significantly different for all three tree distance metrics considered. The level of significance is high (**) no matter whether the distances in the Alg class, and the oldest molecular trees are considered (see also Fig. 1) and main text for details

Table 3 Probabilities and results of the non-parametric multiple comparisons

Hypothesis tested	Tree distance metric								
	Opt and Alg Class Data			Opt-Class Data			Post-95 Opt-Class Data		
	PM	DCt	DCq	PM	DCt	DCq	PM	DCt	DCq
$\theta-MM = \theta-Mm$	0.00003	0.000001	0.0012	0.0012	0.0006	0.00001	0.004	0.0012	0.00013
$\theta-mm = \theta-Mm$	0.05	0.04	0.08	0.09	0.15	0.11	0.1	0.16	0.15
$\theta-MM = \theta-mm$	1	1	1	1	1	0.78	1	1	0.92

The results of these comparisons show $\theta-Mm$ to be significantly different than $\theta-MM$, that is, molecular and morphological trees are significantly incongruent. They also show that $\theta-MM$ and $\theta-mm$ are not significantly different, see also Fig. 1 and Table 2

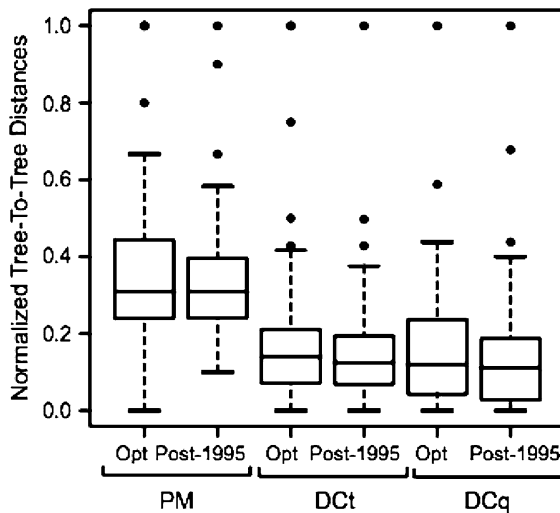


Fig. 3 Boxplot representation of the distances in the MM class in the Opt and Post- 95 data sets (see also Fig. 1 and 2 for details). The boxes represent the values included between the lower and upper quartiles, with the medians ($\theta-MM$) splitting the respective boxes in two. The whiskers extend to more extreme values (1.5 times the interquartile range from the box). Full circles represent outliers. This figure clearly shows that, for all the three distance metrics, if only comparisons between optimality based trees are considered, no significant increase in congruence can be observed when trees published before 1995 are excluded from the analysis

But why do molecular trees seem to be more similar to other molecular trees and morphological trees to other morphological trees than they are to each other? Clearly, some of this congruence is spurious and must result from some non-independence of alternative morphological and/or alternative molecular trees. The possibility of an historical burden affecting morphological phylogenies so that pre-existing phylogenetic theories affect the construction of new morphological characters has been discussed in the literature (e.g. Lee 1995; Rieppel and Kearney 2002; Harris et al. 2007). This may contribute to and inflate the congruence of

morphological trees. For molecular data, systematic errors resulting from, for example, long-branch attraction (Pisani 2004), patterns of codon usage, and base composition biases (Swofford et al. 1996) may affect different genes, sometime all the genes, of a given taxon (Driskell et al. 2004). This could potentially result in similarly incorrect, incompletely independent molecular trees, artificially inflating their congruence.

Scotland et al. (2003) argued that the superiority of molecules over morphology warrants concentrating on molecular phylogenetics and limiting the role of morphology in phylogenetic reconstruction. While our data are consistent with the superiority of molecules over morphology, our results do not support the view that morphological data are of insufficient utility to grant an independent role to morphological phylogenetics. To the contrary, the ability of comparisons between morphological and molecular trees to detect incongruence (and thus error) that would not be apparent from comparisons of trees inferred from molecular data only underlines the utility of such comparisons, for which both types of data are necessary.

Our inferences are not without substantial caveats. More powerful inferences would be possible with an increased sample size. Ideally, we would like to take differential support into account to determine if there is any relationship between patterns of incongruence and support, but here we have focussed only upon the nature of the principal signal (*sensu* Pisani and Wilkinson 2002) rather than its strength because information about levels of support was not reported for all the trees we used. Most importantly our conclusions are based on studies of amniotic vertebrates, which have relatively accessible, complex and potentially information rich morphologies. Obviously, where there is little morphology or it is relatively inaccessible, morphological data will be less important.

4 Conclusions

Bledsoe and Raikow (1990) performed the first statistical test of the congruence of molecular and morphological trees. Although they relied on a much smaller data set, their results were similar to those reported here in highlighting greater congruence between molecular trees than between molecular and morphological trees. However, this difference was not statistically significant, due perhaps to a lack of statistical power (their database included only 48, mostly molecular, trees) and inadequate statistical analyses. Here we build on Bledsoe and Raikow (1990) pioneering approach. Although still rather limited, our larger sample size and the use of non-parametric statistics reveal a significant difference in the congruence of molecular and morphological trees. In particular molecular trees are significantly more congruent than are molecular and morphological trees. Comparison of morphological and molecular trees is thus a good way to discover incongruence that may not be apparent using only one kind of data, but which is usefully diagnostic of non-independence and spurious congruence in molecular and/or morphological trees and data.

Acknowledgements This work was partially funded by a Leverhulme Trust grant to MJB, and by a BBSRC grant 40/G18385, and an NHM-MRF award to MW. DP was supported by a Marie Curie Intra European Individual Fellowship (MEIF-CT-2005-010022). We would like to thank Anthony Bledsoe, Andy Purvis and Clive Moncrieff for the useful advice on statistical data analyses, and Richard Olmstead and two anonymous reviewers for their helpful and insightful comments.

References

- Benton MJ (1999) Early origins of modern birds and mammals: molecules vs. morphology. *BioEssays* 21:1043–1051
- Bledsoe AH, Raikow RJ (1990) A quantitative assessment of congruence between molecular and nonmolecular estimates of phylogeny. *J Mol Evol* 30:247–259
- Critchlow DE, Pearl DK, Quian C (1996) The triplet distance for rooted bifurcating phylogenetic trees. *Syst Biol* 45:323–334
- Day WHE (1983) Distribution of distances between pairs of classifications. In: Felsenstein J (ed) *Numerical taxonomy*. NATO ASI Series, vol G1. Springer-Verlag, Berlin, pp 127–131
- Driskell AC, Ane C, Burleigh JG, McMahon MM, O'Meara BC, Sanderson MJ (2004) Prospects for building the tree of life from large sequence databases. *Science* 306:1172–1174
- Eastale S (1999) Molecular evidence for the early divergence of placental mammals. *BioEssays* 21:1052–1058
- Estabrook GF, McMorris FR, Meacham CA (1985) A Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. *Syst Zool* 34:193–200
- Fleischmann RD, Adams MD, White O et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512
- Gura T (2000) Bones, molecules...or both? *Nature* 406:230–233
- Harris SR, Pisani D, Gower DJ, Wilkinson M (2007) Investigating stagnation in morphological phylogenetics using consensus data. *Syst Biol* 56:125–129
- Hedges SB, Maxson LR (1996) Molecules and morphology in amniote phylogeny. *Mol Phylogenet Evol* 6:312–319
- Hedges SB, Maxson LR (1997) Complementary uses of molecules and morphology: a reply to Lee. *Mol Phylogenet Evol* 8:445
- Hillis DM, Wiens JJ (2000) Molecules versus morphology in systematics: conflicts, artefacts, and misconceptions. In: Wiens JJ (ed) *Phylogenetic analysis of morphological data*. Smithsonian Institution Press, Washington, pp 1–19
- Huelsbeck JP, Hillis DM (1994) Success of phylogenetic methods in the four-taxon case. *Syst Biol* 42:247–264
- Jenner RA (2004) Accepting partnership by submission? Morphological phylogenetics in a molecular millennium. *Syst Biol* 53:333–342
- Lee MSY (1995) Historical burden in systematics and the interrelationships of 'parareptiles'. *Biol Rev* 70:459–547
- Lee MSY (1997) Molecules, morphology, and phylogeny: a response to Hedges and Maxson. *Mol Phylogenet Evol* 7:394–395
- Miyamoto MM, Fitch WM (1995) Testing species phylogenies and phylogenetic methods with congruence. *Syst Biol* 44:64–76
- Nei M, Kumar S, Takahashi K (1998) The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotides or amino acids used is small. *Proc Natl Acad Sci USA* 95:12390–12397
- Page RDM (1997) *Component lite v.0.1*. Department of Zoology, University of Glasgow
- Page M (1999) Inferring the historical pattern of biological evolution. *Nature* 401:877–884
- Penny D, Hendy MD, Steel MA (1991) Testing the theory of descent. In: Miyamoto MM, Cracraft J (eds) *Phylogenetic analysis of DNA sequences*. Oxford University Press, New York, pp 155–183
- Pisani D (2002) Comparing and combining data and trees in phylogenetic analysis. Ph.D. Thesis. Department of Earth Sciences, University of Bristol
- Pisani D (2004) Identifying and removing fast-evolving sites using compatibility analysis: an example from the Arthropoda. *Syst Biol* 53:978–989

- Pisani D, Wilkinson M (2002) Matrix representation with parsimony, taxonomic congruence and total evidence. *Syst Biol* 51:151–155
- Rieppel O, Kearney M (2002) Similarity. *Biol J Linn Soc Lond* 75:59–82
- Robinson DF, Foulds LR (1981) Comparison of phylogenetic trees. *Math Biosci* 53:131–147
- Scotland RW, Olmstead RG, Bennett JR (2003) Phylogeny reconstruction: the role of morphology. *Syst Biol* 52:539–548
- Siegel S, Castellan NJ Jr (1989) Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York
- Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research, 3rd edn. W.H. Freeman and Company, New York
- Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK (eds) *Molecular systematics*. Sinauer, Sunderland, pp 407–514
- Thorley JL, Page RDM (2000) RadCon: phylogenetic tree comparison and consensus. *Bioinformatics* 16:486–487
- Thorley JL, Wilkinson M, Charleston M (1998) The information content of consensus trees. In: Rizzi A, Vichi M, Bock HH (eds) *Advances in data science and classification*. Springer-Verlag, Berlin, pp 91–98
- Wiens JJ (2004) The role of morphological data in phylogeny reconstruction. *Syst Biol* 53:653–661
- Wilkinson M (1994) Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. *Syst Biol* 43:343–368
- Yang H, Golenberg EM, Shoshani J (1996) Phylogenetic resolution within the Elephantidae using fossil DNA sequence from the American mastodon (*Mammuth americanum*) as an outgroup. *Proc Natl Acad Sci USA* 93:1190–1194
- Zuckerlandl E, Pauling L (1965) Molecules as documents of evolutionary history. *J Theoret Biol* 8:357–366