Molecular Phylogenetics and Evolution 94 (2016) 447–462

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Review

Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics ☆

Scott V. Edwards [a,*], Zhenxiang Xi [a], Axel Janke [b], Brant C. Faircloth [c], John E. McCormack [d], Travis C. Glenn [e], Bojian Zhong [f], Shaoyuan Wu [g], Emily Moriarty Lemmon [h], Alan R. Lemmon [i], Adam D. Leaché [j], Liang Liu [k], Charles C. Davis [a]

[a] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
[b] Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25, D-60325 Frankfurt am Main, Germany
[c] Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA
[d] Moore Laboratory of Zoology, Occidental College, Los Angeles, CA 90041, USA
[e] Department of Environmental Health Science, University of Georgia, Athens, GA 30602, USA
[f] College of Life Sciences, Nanjing Normal University, Nanjing 210023, China
[g] Department of Biochemistry and Molecular Biology & Tianjin Key Laboratory of Medical Epigenetics, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China
[h] Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA
[i] Department of Scientific Computing, Florida State University, Tallahassee, FL 32306, USA
[j] Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA
[k] Department of Statistics and Institute of Bioinformatics, University of Georgia, Athens, GA 30602, USA

ARTICLE INFO

Article history:
Available online 27 October 2015

ABSTRACT

In recent articles published in *Molecular Phylogenetics and Evolution*, Mark Springer and John Gatesy (S&G) present numerous criticisms of recent implementations and testing of the multispecies coalescent (MSC) model in phylogenomics, popularly known as "species tree" methods. After pointing out errors in alignments and gene tree rooting in recent phylogenomic data sets, particularly in Song et al. (2012) on mammals and Xi et al. (2014) on plants, they suggest that these errors seriously compromise the conclusions of these studies. Additionally, S&G enumerate numerous perceived violated assumptions and deficiencies in the application of the MSC model in phylogenomics, such as its assumption of neutrality and in particular the use of transcriptomes, which are deemed inappropriate for the MSC because the constituent exons often subtend large regions of chromosomes within which recombination is substantial. We acknowledge these previously reported errors in recent phylogenomic data sets, but disapprove of S&G's excessively combative and taunting tone. We show that these errors, as well as two nucleotide sorting methods used in the analysis of *Amborella*, have little impact on the conclusions of those papers. Moreover, several concepts introduced by S&G and an appeal to "first principles" of phylogenetics in an attempt to discredit MSC models are invalid and reveal numerous misunderstandings of the MSC. Contrary to the claims of S&G we show that recent computer simulations used to test the robustness of MSC models are not circular and do not unfairly favor MSC models over concatenation. In fact, although both concatenation and MSC models clearly perform well in regions of tree space with long branches and little incomplete lineage sorting (ILS), simulations reveal the erratic behavior of concatenation when subjected to data subsampling and its tendency to produce spuriously confident yet conflicting results in regions of parameter space where MSC models still perform well. S&G's claims that MSC models explain little or none (0–15%) of the observed gene tree heterogeneity observed in a mammal data set and that MSC models assume ILS as the only source of gene tree variation are flawed. Overall many of their criticisms of MSC models are invalidated when concatenation is appropriately viewed as a special case of the MSC, which in turn is a special case of emerging network models in phylogenomics. We reiterate that there is enormous promise and value in recent implementations and tests of the MSC and look forward to its increased use and refinement in phylogenomics.

☆ This paper was edited by the Associate Editor Derek Wildman.

* Corresponding author.
E-mail address: sedwards@fas.harvard.edu (S.V. Edwards).

http://dx.doi.org/10.1016/j.ympev.2015.10.027
1055-7903/

Contents

1. Introduction

In their article "The gene tree delusion" (Springer and Gatesy, 2016) and other recent articles in Molecular Phylogenetics and Evolution (i.e., Gatesy and Springer, 2014; Simmons and Gatesy, 2015; hereafter S&G when referring to this ensemble of articles), Mark Springer and John Gatesy present a wide-ranging critique of recent phylogenomic analyses employing models based on the multispecies coalescent. The multispecies coalescent (MSC) model is a relatively new and arguably successful approach to phylogenomics in which individual gene trees are estimated simultaneously or separately with a species tree as a means of estimating phylogenetic relationships. Species tree methods are diverse, and now include parsimony, maximum likelihood, Bayesian, and nonparametric approaches, most of which take advantage of some condition of the MSC model. S&G critique the data sets analyzed by several recent papers, including McCormack et al. (2012), Song et al. (2012), and Xi et al. (2014), and state that these papers violate aspects of the MSC and therefore misapply it, resulting in unreliable phylogenetic trees. They also identify a number of errors in data sets used by Song et al. (2012) and Xi et al. (2014), and conclude that these errors compromise the major conclusions of these papers.

S&G make many additional criticisms of the MSC models and coalescent computer simulations in phylogenomics, stating that:

(1) There are deficiencies and errors in the Song et al. (2012) and Xi et al. (2014) data sets that, when cleaned up, lead to different results from those published by those authors.
(2) The frequency of recombination in mammalian and other genomes violates fundamental assumptions of MSC methods and that a "recombination ratchet" ensures that only very short (~12 bp) segments of DNA of mammals can be analyzed by MSC methods.
(3) Transcriptome data fundamentally violates MSC analyses by concatenating exons that span large segments of chromosomes, much larger than the scale on which recombination takes place.
(4) The coalescent simulations used to test the MSC model in the critiqued papers and in earlier works are circular, are designed to protect MSC methods by focusing on unrealistic parameter values and other "illogical…shell games", and that the MSC model explains 0–15% of the variation in gene trees observed in empirical data sets, which apparently are caused by gene tree estimation error.
(5) MSC methods fail when gene tree variation is due to gene tree error or any source other than incomplete lineage sorting (ILS).
(6) By assuming neutrality, the MSC model is violated by many kinds of molecular data, particularly those influenced by natural selection.
(7) S&G recommend MSC models employing SNPs to get around the problem of recombination in MSC data sets using linked loci.
(8) Authors of papers favorable to MSC methods (e.g., Lemmon and Lemmon, 2013) have turned a "blind eye" to the deficiencies noted above, relegating standard concatenation methods as "not recommended".

In many more subtle claims too numerous to mention, S&G disapprove of MSC methods and their application to phylogenomic data. This reply counters many of those claims and shows that most them are either invalid, misapplied, or represent fundamental misunderstandings of the MSC model. Although the authors of the main papers critiqued by S&G acknowledge errors in their data sets, many of these errors had already been pointed out in the literature and do not imply fundamental changes in the conclusions of the respective papers. Moreover, S&G fail to suggest any sustainable approach for identifying such errors in phylogenomic data sets going forward. We show that the "recombination ratchet" and other concepts introduced by S&G are flawed, causing more confusion than clarity in thinking about phylogenomic data. Similar misconceptions plague their simulations of gene tree heterogeneity; we show that ILS indeed accounts for a substantial fraction of gene tree heterogeneity observed in the critiqued data sets. We emphasize that MSC models are not predicated on the existence of ILS,

whose presence or absence, ultimately, is irrelevant to the application of the MSC model. Rather, it is the conditional independence of loci used to make gene trees, not the presence of ILS, that is the fundamental assumption of MSC methods. This conditional independence follows from first principles of genetics, such as recombination and the chromosomal structure of genomes, a field from which phylogenetics is derived and with which it must be consistent. Although S&G have performed an admirable service by identifying errors in phylogenomic data sets, their recommendations for phylogenomics and the provocative language used to proffer them represent fundamental throwbacks that do not advance the field.

2. First principles of (phylo)genetics: model hierarchies and recombination

2.1. Flawed appeal to "first principles" of phylogenetics

S&G make an appeal to "first principles" of phylogenetics in their critique of recent coalescent approaches to an angiosperm phylogeny (Simmons and Gatesy, 2015). They write (p. 120): "Enthusiastic application of novel tools is not a substitute for rigorous application of first principles [of phylogenetics], and new 'sophisticated methods' may be novel sources of previously under-appreciated, systematic errors…". Among the "first principles" they mention are practices such as rigorous taxon sampling, accurate alignments, character coding, "total evidence with inclusion of hidden support", and rigorous tree searches. We point out that some of these "first principles" are legacies of the gene tree era of phylogenetics, in which concatenation methods were tantamount to assuming an average gene tree across all genes represents the species tree and when data sets were very modest in size. They are also, of course, first principles that should be used when building gene trees as a first step to constructing a species tree, in so-called "two-step" methods of species tree inference (Liu et al., 2015b). Although we acknowledge their general merit in phylogenetic inference, it is not yet clear the extent to which these first principles apply directly to species tree inference. For example, taxon sampling is now generally agreed to improve phylogenetic reconstruction (of gene trees). However, it is unclear whether this principle applies to species tree inference independently of its importance for gene trees. Of course, a species tree will usually contain the same species found in its constituent gene trees, so dense taxon sampling at the level of gene trees will perforce imply dense taxon sampling at the level of the species tree. Yet recent studies have found that some species tree methods may be more robust than concatenation methods to taxon sampling and consequences of poor taxon sampling such as long-branch attraction (Liu et al., 2015a). Additionally, some species tree methods appear to be robust to issues such as mis-rooting of gene trees, which may be common in some phylogenomic data sets (e.g., Fig. 3 of Liu et al. [2009]; Fig. 4 of Liu et al. [2010]). For example, Liu et al. (2010:314) found in one simulation on the MSC method MP-EST that "We observed that nearly half (49%) of the gene trees generated from the species tree across replicates were rooted incorrectly ….Yet in these cases and in the simulation generally, the correct species tree was always consistently estimated." Thus, whereas we agree with S&G that a lax approach to the first principles would only degrade the quality of phylogenomic data sets, S&G's appeal to "first principles" of phylogenetics is outdated insofar as it does not acknowledge new levels of robustness of phylogenetic inference that may arise from species tree inference and may in some cases ameliorate challenges of accurate gene tree inference and loose application of first principles on gene tree estimation. In the age of phylogenomics, new principles arise when considering the additional levels above the gene tree brought about by the species tree approach, and consideration of microevolution enhances our approach to macroevolutionary problems such as phylogenomics (Lynch, 2007). For example, McCormack et al. (2009) found that dense sampling of alleles within species – never a strong first principle of gene tree phylogenetics – can improve species tree inference in some cases. Thus, in the age of phylogenomics, the MSC requires first principles to be re-evaluated and novel principles to be incorporated into standard practice of phylogenomic analysis.

Another example of an outdated "first principle" by S&G involves the concept of synapomorphy and hidden support, especially for molecular data. By our reading, hidden support in Gatesy and Springer (2014) refers to instances in which individual gene trees may not support a given clade in the species tree but concatenation of those genes nonetheless supports that clade. They use this definition here despite the fact that hidden support was not originally rationalized primarily in the context of coalescent heterogeneity in early articulations of the concept (Gatesy et al., 1999; Gatesy and Baker, 2005). Gatesy and Springer (2014:235ff) show that analyzing only those loci that contradict a given clade with coalescent methods such as STAR and MP-EST often yields species trees that are unexpected. However, such a result is entirely expected when one picks and chooses a biased set of gene trees as they have done, since 'two-step' coalescent methods require an accurate estimate of the gene tree distribution. The fact that concatenation analysis of those loci yields the expected topology in these cases does not imply that concatenation is better than coalescent approaches in these cases, because Gatesy and Springer (2014) have not studied branch support or branch length error, and certainly have not highlighted those cases where concatenation gives an unexpected result. We predict that the hidden support in concatenation will often overestimate branch support and lead to biased branch lengths. But the perceived poor performance of coalescent methods in these cases is irrelevant, because they are presenting a biased set of gene trees to these methods; it is the data here, not the method, that is flawed in their analysis.

We also challenge S&G to study hidden support in the context of phylogenomic subsampling, and to demonstrate the consistency of concatenation over different phylogenomic subsamples of data (see below). Both simulations (Edwards, 2009a) and empirical analyses (Song et al., 2012) suggest MSC models display consistency superior to concatenation when analyzed with multiple subsamples of phylogenomic data. The recovery of 10 "uncontroversial" nodes in the mammalian tree via hidden support (Fig. 3 of Springer and Gatesy [2014]) is indeed intriguing. But without subsampling or analysis via coalescent simulations with known levels of ILS, it is unclear whether such hidden support is consistently recovered across replicates and data subsamples.

With the widespread incidence of ILS, we now know that a clade in species trees can sometimes be supported by DNA sites, or other genomic features such as transposable elements (e.g., Suh et al., 2015), that appear to contradict that clade, such as in the simulation in Fig. 2 of Edwards (2009a). Such sites will occur when gene trees exhibit discordance, yet support and are consistent with a species tree with a different topology. Such phenomena compel us to revise the strict interpretation of synapomorphy in phylogenetics, particularly for molecular data. Whereas we see a clear biological explanation for support of sites for a discordant clade in the framework of ILS and the MSC, hidden support in a concatenation framework appears more reminiscent of the erratic behavior of concatenation when applied to rapidly radiating clades and polytomies with abundant ILS (e.g., Fig. 4 of Edwards [2009a]). In such cases, we know that concatenation can yield high yet erroneous support for unexpected clades that are not encountered in the trees of individual genes, whereas MSC models appropriately yield muted support for an array of possible species trees.

2.2. Concatenation is a special case of the MSC

S&G's appeal to "first principles" is worrisome in another very important aspect. Many of their critiques have, at their core, an assumption that concatenation and species tree methods represent alternatives, and that phylogenetics is witnessing a battle of sorts for each of these to achieve supremacy of the field. Concatenation, as an outgrowth of "total evidence", is generally cast as the founding approach of the modern incarnation of the field that should not be carelessly abandoned in favor of newer, different models. However, this perspective is flawed for the key reason that phylogenetics itself is a particular instance of the broader field of genetics, with its focus on inheritance, the chromosomal structure of genomes and its focus on explaining phenotypic variation within and among species (Rosenberg and Nordborg, 2002; Fig. 1A). Conditional independence of loci – when loci are considered stochastically independent of each other due to genetic drift, conditional on the underlying demographic or phylogenetic history – is built into this world-view of genetics. Whereas conditional independence is inherent in the MSC, it is patently ignored by concatenation methods. More particularly for this discussion, concatenation is best viewed not as an alternative to species tree inference but as a specific case of the MSC (Fig. 1B; Edwards, in press). Concatenation represents that case of the MSC in which all gene trees are identical in topology and branch lengths. This suggestion may seem radical, in part because it casts a 30-year old tradition (concatenation) as a specific case of a much newer model in phylogenetics (the MSC). However, it was recently pointed out that the MSC can indeed subsume the model assumed by concatenation (Liu et al., 2015b), in the same way that parsimony can be viewed as a specific instance of the broader model inherent in maximum likelihood (Goldman, 1990; Lewis, 1998). Coalescent theory itself has proceeded in this way, from specific instances to more general models (Kingman, 2000). This updated perspective on concatenation raises serious problems for the repeated criticisms of the MSC by S&G, which in this new light can be viewed also as criticisms of concatenation. This twist becomes particularly problematic when they discuss population genetic phenomena such as recombination as if they were a problem only for the MSC and not for concatenation. Thus, in criticizing the perceived weaknesses of the MSC in the face of forces like recombination, S&G necessarily draw attention to the possibility that concatenation may also fail under these conditions, because concatenation is a special case of the MSC. This represents a logical flaw in their arguments that is distinct from whether or not recombination does indeed represent a problem for phylogenetic analysis (Posada and Crandall, 2002; see below).

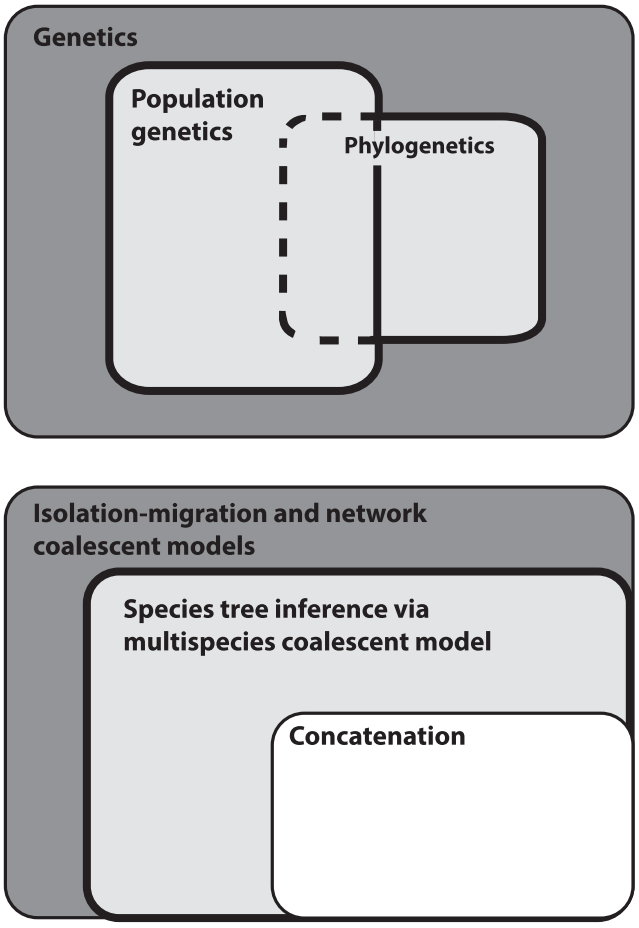


Fig. 1. Hierarchy of domains and models in the fields of genetics and phylogenetics. Phylogenetics is viewed as a particular instance of the broader field of genetics in which phylogenetics is nested. (Top) A view of the relationship among the domains of genetics, phylogenetics and population genetics encourages approaches to phylogenetics that are consistent with the major tenets of genetics, such as the chromosomal structure of genomes, random assortment of alleles during meiosis, independent transmission of alleles at unlinked loci and other mainstays of genetics. Most sampling schemes for phylogenetics (sampling multiple alleles per species, or single alleles from multiple loci from multiple species) demand consideration of population genetic principles. (Bottom) Relationship among models in phylogenetics, including the MSC and concatenation models. Concatenation is best viewed as a particular case of the broader model inherent in the MSC, which is itself a particular case of models that incorporate gene flow and other reticulations, such as recently proposed MSC network models.

We are convinced that recent inquiries into the influence of population genetic phenomena such as recombination, gene flow and horizontal gene transfer (HGT) on species trees and concatenation approaches to phylogenetic inference are positive trends for the field and should be continued (e.g., Eckert and Carstens, 2008; Liu et al., 2009; Leaché et al., 2014; Davidson et al., 2015). Yet in focusing on the problems of recombination for MSC models, S&G somehow assume that concatenation itself is immune to such genetic forces. We note that there were very few discussions of the effect of recombination or gene flow on phylogenetic inference prior to the introduction of the MSC (but see Posada and Crandall, 2002; reviewed in Posada et al., 2002), and can only conclude that the introduction of the MSC has spurred additional interest in this area. But it did not have to proceed this way. Population genetic forces were just as much an issue for concatenation before the MSC was introduced, but the discussion did not trend in this direction, we believe, because systematists generally did not think about population genetics, and vice versa (Felsenstein, 1988). We reject the argument that concatenation is somehow immune to population genetic forces in ways that the MSC is not. Arguing that "hidden support" is a strength of concatenation is problematic, ultimately vague and confusing and, if anything, also likely appropriate for other ways of combining data in phylogenomics, such as the continuum of MSC models.

2.3. Recombination: not a severe problem for species tree inference

The issue of recombination is an example of S&G's misdirected approach to first principles. Recombination is a very important issue for phylogenetics (both concatenation and the MSC) because the MSC does indeed make the assumption of "recombination between genes but not within genes". In fact, because concatenation is a particular instance of the MSC, but assumes the same gene tree for all genes, concatenation makes the even stronger assumptions of "no recombination between genes" compared to the MSC, although it is rarely discussed. A brief response to S&G about the issue of recombination (Gatesy and Springer, 2013) has already been published in Wu et al. (2013), yet they raise it again in their

more recent papers. It is therefore important to detail this issue again, while acknowledging that more research is required.

As pointed out already by Wu et al. (2013) and Zhong et al. (2014), Lanier and Knowles (2012) conducted to our knowledge the only focused inquiry into the effects of recombination on phylogenetic reconstruction using species tree methods. Their overall conclusion is summarized by the opening sentences in their discussion: "Species-tree estimates under the multispecies coalescent are robust to uncertainty introduced by unrecognized recombination within a locus . . . at least within the natural range of recombination observed in empirical data across different methods of analysis and for different diversification histories (including both recent and deep radiations) . . . In no cases was recombination identified as a primary factor influencing the accuracy of species-tree estimates. . .." Their conclusion is unambiguous, and their analysis shows that only under extremely short internodes will recombination be a problem for species tree analysis.

As a response to this perceived flaw, S&G devise a "recombination ratchet" based on lengths of DNA sequence alignments within which there is no evidence of recombination as judged by homoplasy at different phylogenetic levels. We find numerous problems with the concept of the recombination ratchet and discourage its use. S&G have conflated homoplasy across any level in the phylogeny – whether across genera, families or taxonomic orders – with recombination, which, based on first principles of genetics, necessarily takes place only within species (both current and ancestral). In attempting to trace the longest genomic regions showing no evidence of recombination across the tree for mammals, S&G mistakenly attribute homoplasy across the mammal tree with recombination, thereby invalidating their arguments. We cannot find any biologically meaningful entity that corresponds with their recombination ratchet, since, as they suggest, there is likely no DNA segment except for individual SNPs that has not experienced recombination at some point across the history of life. Put another way, recombination occurs along one branch of the species tree, but S&G present the recombination ratchet as producing a set of recombination events that together affect the entire tree. That is very misleading.

But in some ways the predicted effects of recombination on phylogenetic inference, whether via species trees or concatenation, does not require simulations if one adheres to basic first principles of genetics (Fig. 2). Whereas phenomena such as horizontal gene transfer do involve multiple, sometimes long diverged, species, recombination by definition involves exchange of genetic material between alleles found in single individuals, which, under panmixia, is tantamount to exchanges between individuals of the same species. For this reason, recombination takes place almost entirely among gene lineages of a locus while those lineages are within the branches (internodes) of a species tree. (Some studies have used simulations of a generalized recombination mechanism that does not mimic the special properties of genetic recombination and includes, for example, swapping of DNA sequences between species, which is more germane to horizontal gene transfer than to classical recombination (Posada and Crandall, 2002;

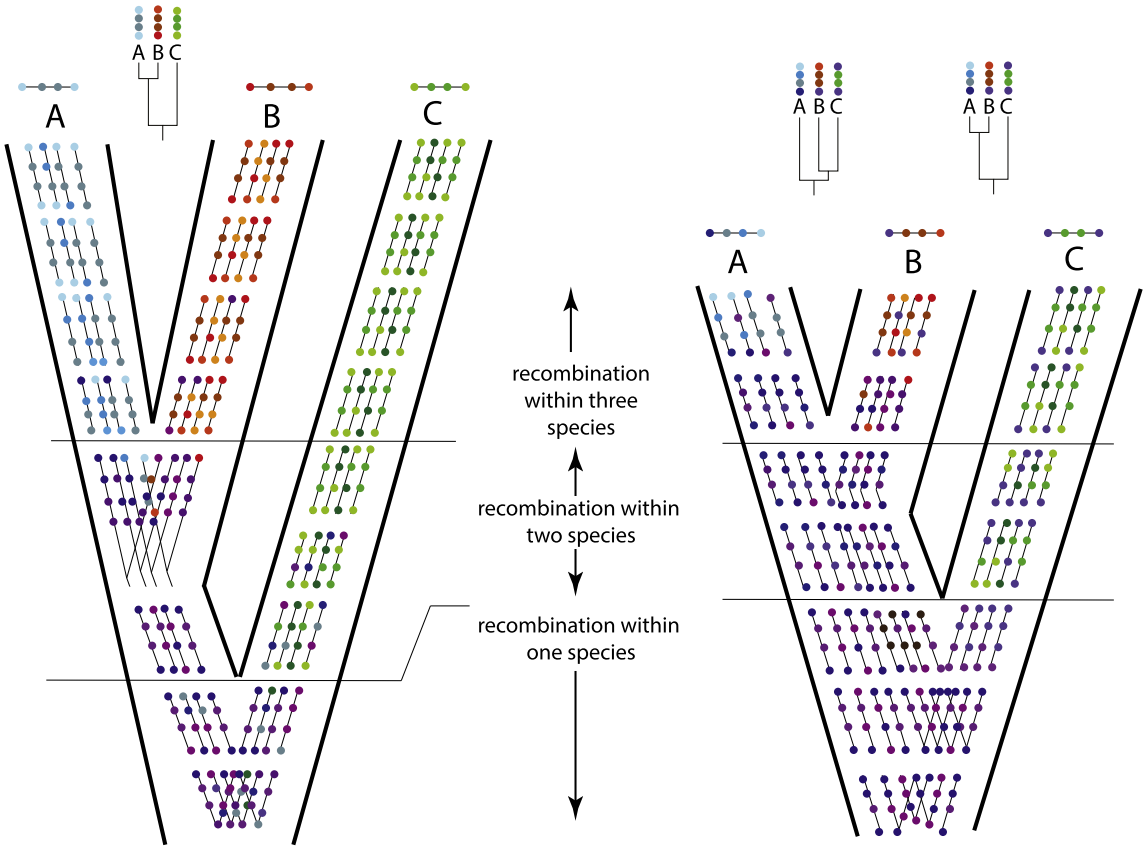


Fig. 2. Schematic of the recombination process along lineages of a species tree. Left, species tree with no incomplete lineage sorting. Right, species tree with all loci experiencing ILS. In both trees, colored balls represent alleles at four loci along a chromosome, each marked by a thin black line with four haplotypes per species. Colors indicate phylogenetic affinities of alleles within species. At the top are examples of possible reconstructed gene trees after a history of recombination. On the left, recombination does not compromise gene tree or species tree inference because recombining alleles, which must occur in the same species (current or ancestral) are generally more closely related to each other than to alleles from other species. On the right, recombination will only compromise gene tree inference when ILS is high, in this example 100%. Via drift and recombination, alleles maintain phylogenetic coherence along long branches; incongruences occurring at short internodes are to a certain extent erased by recombination along long branches. Recombination will tend to shorten branch lengths of gene trees in less severe cases or in more severe cases, lead to gene tree misestimation, as described in Posada and Crandall (2002). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Castillo-Ramírez et al., 2010)). If we imagine that most alleles at a locus are more similar to one another within a given species than are alleles found in different species, recombination will generally involve alleles that are closely related to one another than are alleles found in sister species or ancestral lineages. Thus, when taking place within branches of a species tree, recombination will generally not combine alleles that are divergent from one another or discordant from the species tree, and the net result will be consistent with the species tree (Fig. 2).

As Lanier and Knowles (2012) illustrate through simulations, the only situation in which recombination will be a problem for phylogenetic analysis is when recombination takes place in ancestral nodes where ILS is very high. We acknowledge that in a typical phylogeny there are many ancestral nodes whose gene trees may be comprised by recombination, yet ILS is often resolved along deep branches of phylogenies after brief periods of gene tree incongruence spanning such ancestral nodes. The scenarios illustrated in Figs. 3–5 of Springer and Gatesy (2016) may illustrate conceivable events, but it remains unclear how common they are. We also find their discussion of recombination confusing and their conclusion that 12-bp defines the largest length of DNA that is suitable for phylogenetic analysis (p. 29) unsubstantiated. For example, they write of the recombination ratchet, "As more taxa are added to a phylogenetic analysis, the number of recombination breakpoints can only increase" (p. 6), but we view recombination breakpoints as an intrinsic property of individual species, not entire clades, and thus adding taxa to a problem involving recombination does not make sense. Additionally, they conflate homoplasy with recombination when the two are not synonymous. Finally, it is not clear that the appropriate units for phylogenetic analysis need not have recombined at all during the entire history of the clade being studied.

Recent studies of topologically uniform tract lengths in eukaryotic genomes are more relevant to this discussion, as Springer and Gatesy (2016) point out. But our interpretation of the paper by Hobolth et al. (2007), which they rely on heavily, is different from theirs, and other workers have estimated much longer recombination-free tract lengths in natural populations. Depending on the scale over which Hobolth et al. (2007) applied their algorithm, they often estimated much longer tracts of topologically homogeneous DNA, sometimes on the order of tens of kilobases (e.g., their Fig. 3), than the 109 bp in primate Tract 122 that Springer and Gatesy (2016) highlight. Additionally, it is not clear to us that the appropriate tract length for phylogenomic analysis must be the minimum length within which there is no topological incongruence; after all, phylogenetics has for decades been comfortable analyzing gene sequences that exhibit some level of homoplasy. Using an algorithm for estimating breakpoints of phylogenetically coherent tract lengths in genomes (Ané and Sanderson, 2005), White et al. (2009) used genome wide data to estimate a median length of nearly 100 kb for phylogenetically coherent genomic regions in mice within which recombination was estimated regions of mouse genomes. Similarly, using computer simulations and a coalescent tree that could be interpreted to span many millions of years, Ané (2011) found topologically homogeneous tract lengths of several thousand SNPs, which translate into even larger genomic regions when one considers intervening invariant sites. Although the *Mus* species studied by White et al. (2009) diverged much more recently than the mammals studied by Song et al. (2012), this result suggests that the no-homoplasy principle used in Springer and Gatesy (2016)'s recombination ratchet to delimit such tract lengths is flawed. Additionally, even transcriptome data will sample the genome only sparsely, and to estimate breakpoints on individual genes, even those with widely dispersed exons, will not adequately capture genome-wide patterns, especially when those genes are not analyzed in a spatially appropriate context.

Finally, we reiterate that S&G imply that concatenation is somehow immune to whatever negative effects recombination may have on phylogenetic analysis. Peculiarly, recombination is viewed only as a problem for the MSC and not for concatenation. We fail to see how a problem for the more general model of phylogenetics can be avoided by the specific case of concatenation. At the end of their paper, Springer and Gatesy (2016), withdraw to a vague mainstay, pointing out the importance of partitioning genome scale data and rephrasing the well-known question "... how should these basic units [of phylogenetic analysis] be delimited and analyzed?" This question has been asked and partially answered by several workers (Posada and Crandall, 2001, 2002; Martin et al., 2005; Ruths and Nakhleh, 2005; White et al., 2009; Ané, 2011). Recombination is also an issue for models in phylogeography, where it is likely an even more severe problem than for phylogenetics. Researchers have developed extensions of basic phylogeographic models to accommodate recombination (Becquet and Przeworski, 2007; Naduvilezhath et al., 2011) and the same could be done for phylogenetics. We do not deny there is more work needed in this area but reiterating the question, pointing out potential violations of the MSC model, yet not demonstrating an effect on the phylogenetic analyses does not move the field forward. All models are simplifications, the MSC included, and it is acceptable, often necessary, to relax these assumptions when applying the models, especially when simulations and empirical analysis do not suggest deleterious consequences (Lanier and Knowles, 2012).

3. Data types in phylogenomics: model violations and sensitivity analyses

3.1. Transcriptomes: powerful tools for phylogenomics using the MSC

The issue of recombination bears directly on the use of transcriptome data in phylogenomic studies. S&G have repeatedly criticized the use of transcriptomes in phylogenomic studies using the MSC, an approach that has now been used in several studies (e.g., Chiari et al., 2012; Song et al., 2012; Liang et al., 2013; Tsagkogeorga et al., 2013; Wickett et al., 2014; Xi et al., 2013, 2014; Jarvis et al., 2014). They suggest, illustrating with several putative examples, that because of the dispersed nature of exons in the genome, phylogenomic analyses using exonic data, such as those from transcriptomes, likely violate the MSC assumption of lack of within-locus recombination, because exons of a single gene can in some cases be spread over tens or even hundreds of thousands of base pairs of genomic DNA. At one point, the specter of recombination looms so large that they recommend SNP approaches in higher level phylogenetics, which, by avoiding within-locus recombination, "merit further pursuit in both empirical systematic research and simulations." (Springer and Gatesy, 2016:2), a point also raised by Edwards (2009a) among others. Although we are sympathetic with the suggestion that SNP data deserves more study in phylogenomics, and note that new SNP-based methods are emerging (Bryant et al., 2012; Chifman and Kubatko, 2014), we find their arguments against transcriptome data flawed and unnecessarily alarmist, principally for the same reasons stated above that recombination is not a severe problem for phylogenomic studies using the MSC. It is hard to imagine that recombination severely compromises MSC analyses when, overall, the results of MSC and concatenation analyses are concordant for most nodes and lineages (Song et al., 2012; Liang et al., 2013; Tsagkogeorga et al., 2013; Wickett et al., 2014; Xi et al., 2013, 2014; Jarvis et al., 2014).

Gatesy and Springer (2014) suggest that the simulations performed by Lanier and Knowles (2012) are not relevant to the prob-

lems pointed out by S&G because of the limited lengths of genes they studied (1000 bp). They suggest that simulations must span the genomic lengths that individual genes are observed to span, lengths that can exceed 100,000 bp or more. We reject this argument on the grounds that the correlation between the histories of two segments within a gene is determined by the number of recombination events occurring between two segments (Slatkin and Pollack, 2006). Therefore, the relevant parameter in such simulations is the number of recombination events per locus over the relevant phylogenetic history – regardless of the actual number of base pairs over which the simulation is conducted. Using short loci in a simulation only requires increasing the rate of recombination in order to mimic the effects of recombination across 100,000 bp of a genome. Had Lanier and Knowles (2012) simulation resulted in fewer than one recombination event on average within their loci, then this would have compromised their ability to address the recombination problem. Lanier and Knowles (2012) used a population recombination rate ($\rho = 4Nc$, where N is the effective population size and c the per site recombination rate) spanning over three orders of magnitude; although we cannot reconstruct the number of recombination events per locus in their study, we expect it is commensurate with what is observed in real genomes.

Gatesy and Springer (2014) suggest that concatenating loci or exons that reside in the same genomic neighborhood results in a hybrid approach to phylogenetic analysis that "is confronted with the same critique that proponents of coalescence have leveled against the supermatrix approach – a confusing mixture of different historical signals." However, some studies have shown that using individual exons as loci yields species trees very similar or identical to those produced by use of entire multi-exonic genes in transcriptomic data (Tsagkogeorga et al., 2013). Contrary to S&G's concerns, the genomic proximity of concatenated loci may turn out to be an important aspect of methods that concatenate non-adjacent loci as in transcriptome data (Liu and Edwards, 2015). We acknowledge that there is an arbitrariness to the use of transcriptomes and other types of data in which genomic segments-even individual SNPs-are concatenated into individual loci for analysis using the MSC. However, as shown by Lanier and Knowles (2012), we believe that factors other than recombination play a more important role in the failure of species tree methods to reconstruct history accurately. Additionally, the robustness of transcriptome data to the effects of recombination is distinct from their utility in terms of information content for estimating gene trees, which will depend on a variety of factors including the mutation rate of the loci being used (Lanier et al., 2014).

3.2. Subsampling as a method for assessing phylogenetic robustness

The robustness of data used in MSC analyses to recombination has been demonstrated in simulations as well as in empirical data where phylogenetic resolution of short, deep internodes is concerned. For example, Song et al. (2012) showed via subsampling experiments that results of species tree methods were much more congruent across different subsampling protocols (subsampling both taxa and genes) than was concatenation, which tended to show highly erratic results from subsample to subsample (we believe this result was not an artifact of the compromised alignments in that study). The increased robustness of species tree methods compared to concatenation with regard to taxon and gene sampling, as well as to other challenges to phylogenetic analysis, has also been demonstrated extensively (Song et al., 2012; Xi et al., 2014; Liu et al., 2015a). Therefore, we recommend routinely subsampling the data in phylogenomic analyses to document the robustness of the methods to different data subsamples. Based on simulations of phylogenetic analyses across short internodes and polytomies in species trees (Edwards, 2009a), we predict that concatenation will behave erratically when faced with different subsamples of the same data, a phenomenon demonstrated empirically by Song et al. (2012). The erratic behavior of concatenation in subsampling studies has not been addressed by S&G, yet is one of the most conspicuous and unsettling differences in performance between the two approaches.

4. Phylogenetic consequences of alignment errors and misrooting of gene trees

4.1. Identification of errors in data sets

The attention to detail and exhaustive documentation by Springer and Gatesy (2016) of the myriad consequences of gene tree estimation error for species tree inference is admirable, but many of the same errors had already been pointed out by Mirarab et al. (2014a). Springer and Gatesy (2016) reviewed every alignment from the data set assembled by Song et al. (2012), and found that the alignments of 51 of the 447 genes were flawed. After curating, 413 genes remained for their reanalysis. Although it does not exonerate the present case, flawed datasets are unfortunately common in phylogenomics. In addition, after manual review of the 310 nuclear genes assembled by Xi et al. (2014), Simmons and Gatesy (2015) claimed that "ambiguously aligned and/or highly divergent regions were identified for 184 of the 310 nuclear genes (in Xi et al. [2014]), including 8772 positions entirely excluded from the matrices", which corresponded to 3.7% of the total 239,763 nucleotide sites analyzed by Xi et al. (2014). It was indeed careless of Song et al. (2012) and Xi et al. (2014) not to have checked their alignments more carefully. Still, Simmons and Gatesy (2015) lack a criterion for identification of "ambiguously aligned regions"; their assertion must be taken on faith, even though the burden is on them to prove erroneous alignments. For the mammal data, Song et al. (2012) did in fact include an error-finding step in their pipeline, flagging gene trees with excessively long branch lengths as a proxy for alignment accuracy. Additionally, what errors made it into their alignments were presumably propagated from the OrthoMaM database (Ranwez et al., 2007), which Song et al. (2012) used as their data source. Moreover, reducing errors in these data sets does not change many of conclusions of these papers, yet still S&G imply substantial consequences to these errors. Despite their effort, most of the reanalysis by Springer and Gatesy (2016) results in changes of bootstrap support values in the range of 54–95%. This discussion is not particularly helpful, even if the topology changes in a few instances. Branches with support values in this range are not well-resolved in any case, and do not warrant a lengthy discussion. Likelihood ratio tests (LRTs) of alternative topologies around such branches, one of the best ways of comparing phylogenetic hypotheses that is still undeveloped for MSC models, will nonetheless likely be non-significant. Even in phylogenomic analyses of the placental mammalian tree, using over 3300 protein coding genes, there are branches with 100% support that do not withstand a LRT (Hallström and Janke, 2008).

Most of the figures in Springer and Gatesy (2016) illustrate the myriad consequences of the same errors in alignments from Song et al. (2012) on diverse aspects of species tree inference, including spurious inference of the extent of ILS on individual gene trees (their Fig. 13). They correctly point out that misaligments and missing data will cause errors in gene tree inference, which in turn will influence gene tree topologies (their Figs. 7–11 and 13; Tables 2–4). However, the diverse consequences of gene tree error are well known and it is well documented that such errors, even those caused by factors other than poor alignments, can mislead species tree inference (e.g., Thomson et al., 2008; Xi et al., 2015).

Unfortunately, S&G offer no solution for efficiently identifying errors in phylogenomic data sets and cleaning them up. Moreover, we assert that manual gene-by-gene curation is unsustainable in the phylogenomics era, and can lead to a lack of repeatability if alignments are curated in such a manner. Authors have used a variety of approaches to quantify the reliability of alignments (e.g., Wang et al., 2012; Notredame et al., 2000); further automated metrics for identifying errors in alignments, orthology, gene tree outliers and other critical aspects of phylogenetic analysis are needed (e.g., Philippe et al., 2011; Dunn et al., 2013; Weyenberg et al., 2014).

4.2. The position of Amborella in the plant tree of life

Simmons and Gatesy (2015) concluded that the placement of *Amborella trichopoda* Baill. as sister to water lilies (i.e., order Nymphaeales) inferred from 310 nuclear genes by Xi et al. (2014) was flawed, claiming that gene-tree-based coalescent methods such as MP-EST (Liu et al., 2010) and STAR (Liu et al., 2009) were not robust to the sometimes highly divergent and occasionally misrooted gene trees that were used by Xi et al. (2014). In addition, Simmons and Gatesy (2015) concluded that two nucleotide sorting methods used by Xi et al. (2014) were biased in favor of sites with highly asymmetrical distributions of character states that misled the analyses of Xi et al. (2014), in particular the placement of *Amborella*. Xi et al. (2014) used two tree-independent methods (Observed Variability [OV; Goremykin et al., 2010] and Tree Independent Generation of Evolutionary Rates [TIGER; Cummins and McInerney, 2011]) to estimate the relative evolutionary rate for each nucleotide, and examined the placement of *Amborella* for both slow and fast rate partitions. These analyses demonstrated that fast-evolving sites had a particularly strong influence on the spurious placement of *Amborella* alone as the sister to angiosperms. This was especially pronounced when using concatenation, but Simmons and Gatesy (2015) claim these analyses were flawed. Here we show that they are mistaken in this conclusion.

The OV method calculates the total number of pair-wise mismatches at a given site. One example given by Simmons and Gatesy (2015) was that an asymmetrically distributed site for which 98 terminals had an adenine and two terminals had a thymine would be identified by the OV method as more conserved than a symmetrically distributed site for which 50 terminals had an adenine and 50 terminals had a thymine. We think this example actually supports the validity of the OV method. Fig. 3 presents a simplified case. Suppose that the data set includes four taxa *A*, *B*, *C*, and *D*. For the first site, taxon *A* has an adenine while the other three taxa have a thymine; for the second site, taxa *A* and *B* have an adenine while taxa *C* and *D* have a thymine. Thus, the character state of the first site is considered to be asymmetrically distributed *sensu* Simmons and Gatesy (2015). Since the total number of pair-wise mismatches is three for the first site and four for the second site, the first site is identified by the OV method as more conserved. To demonstrate this, we mapped character state changes manually onto phylogenies for these two sites under parsimony (Fig. 3). Since the OV method is tree-independent, we consider all three possible unrooted binary trees for four taxa. For the first site, all three trees require at least one character state change. In contrast, for the second site, one tree requires at least one character state change while others require at least two character state changes (Fig. 3). Therefore, even though both sites include adenine and thymine, the site with an asymmetrical distribution of character states requires fewer character state changes. We thus conclude that the major criticism of the OV method by Simmons and Gatesy (2015) is unwarranted.

To further evaluate the effect of elevated substitution rates on the placement of *Amborella*, we used the branch lengths and nucleotide substitution parameters estimated from each of 310 nuclear genes assembled by Xi et al. (2014) to simulate nucleotide sequences based on gene trees representing varying percentages of the two alternative placements of *Amborella* (Fig. 4). For each simulation, "*X*" percent of the 310 nuclear genes (where "*X*" ranges from 0 to 100 in increments of 10) were randomly assigned topology 1 (i.e., *Amborella* + *Nuphar* as the first lineage of angiosperms; Fig. S1A), and the remaining genes were assigned topology 2 (i.e., *Amborella* alone as the first lineage of angiosperms; Fig. S1B). For each nuclear gene, the branch lengths of the assigned topology and parameters of the GTR + Γ model were estimated from the original nucleotide sequences using RAxML v8.1.3 (Stamatakis, 2014) ("-f e -m GTRGAMMAX -u --no-bfgs"). The resulting optimized gene tree and model parameters were then utilized to simulate nucleotide sequences using Seq-Gen v1.3.3 (Rambaut and Grassly, 1997) with the GTR + Γ model. The concatenated nucleotide matrix was generated from these 310 simulated genes using Phyutility v2.2.6 (Smith and Dunn, 2008). Sites were then sorted using the OV method and divided into OV-slow and OV-fast rate partitions following Xi et al. (2014). Next, species trees were inferred for each rate partition using both concatenation and gene-tree-based coalescent methods. Each simulation was repeated 100 times. For concatenation analyses, the best-scoring maximum likelihood trees were estimated from concatenated gene sequences using both unpartitioned (i.e., a single GTR + G model) and partitioned (i.e., a separate GTR + G model for each gene) models. Optimal tree searches were conducted using RAxML with five independent searches starting from random trees ("-d -f o -m GTRGAMMAX -u --no-bfgs"). For gene-tree-based coalescent analyses, gene trees were first estimated using RAxML with the GTR + Γ model ("-d -f o -m GTRGAMMAX -u --no-bfgs"), and rooted with the lycophyte *Selaginella*. These estimated gene trees were then used to construct species trees using ASTRAL v4.7.1, MP-EST v1.4, and the STAR method as implemented in Phybase v1.3 (Liu and Yu, 2010) (default settings were used for ASTRAL, MP-EST, and STAR). The average number of species and nucleotide sites for each of the 310 nuclear genes are 33 and 773, respectively. The final concatenated matrix included 46 species, 239,763

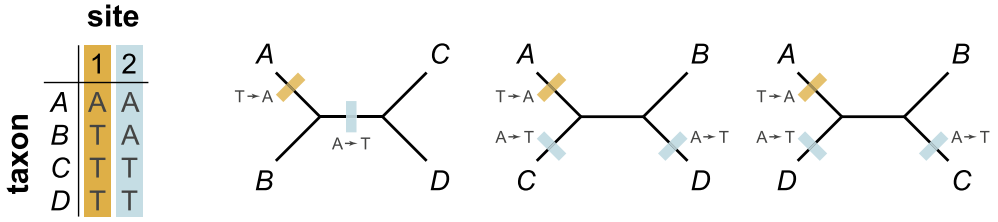
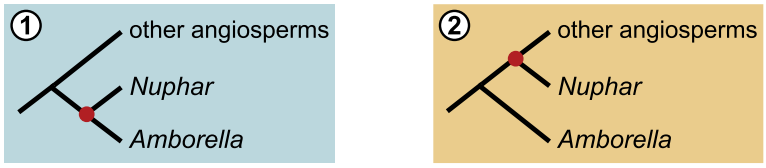


Fig. 3. Four-taxon case to evaluate the OV method. Manual character mapping using parsimony demonstrates that all three unrooted trees require at least one character state change for the first site. In contrast, for the second site, one tree requires at least one character state change while the others require at least two character state changes. Therefore, even though both sites include adenine and thymine, the site with an asymmetrical distribution of character states (i.e., site 1) requires fewer character state changes contra Simmons and Gatesy (2015).



Gene trees (n = 310)	OV - slow				OV - fast			
① : ②	RAxML	ASTRAL	MP-EST	STAR	RAxML	ASTRAL	MP-EST	STAR
0%:100%	1.0/1.0	1.0	1.0	1.0	1.0/1.0	1.0	1.0	1.0
10%:90%	1.0/1.0	1.0	0.99	1.0	1.0/1.0	1.0	0.99	1.0
20%:80%	1.0/1.0	1.0	0.99	1.0	1.0/1.0	1.0	0.94	1.0
30%:70%	1.0/1.0	1.0	1.0	1.0	1.0/1.0	1.0	0.77	0.89
40%:60%	1.0/1.0	0.99	0.99	0.96	1.0/1.0	1.0	0.52	0.67
50%:50%	0.99/1.0	0.69	0.56	0.68	1.0/1.0	0.99	0.83	0.74
60%:40%	0.81/0.75	0.90	0.95	1.0	1.0/1.0	0.82	1.0	0.96
70%:30%	1.0/1.0	1.0	1.0	1.0	0.96/1.0	0.61	1.0	0.98
80%:20%	1.0/1.0	1.0	1.0	1.0	0.58/0.89	0.94	1.0	1.0
90%:10%	1.0/1.0	1.0	1.0	1.0	0.83/0.74	0.99	1.0	1.0
100%:0%	1.0/1.0	1.0	1.0	1.0	1.0/0.98	1.0	1.0	1.0

Fig. 4. Proportions of the two alternative placements of *Amborella trichopoda* recovered from simulated nuclear genes using coalescent (ASTRAL, MP-EST, and STAR) and concatenation (unpartitioned/partitioned RAxML) methods. For each placement, the red dot highlights the node of interest, and for which the proportion of each rival topology was estimated (for concatenation using RAxML, proportions are presented as unpartitioned/partitioned). Nucleotide sequences were simulated based on 310 nuclear gene trees representing varying percentages of the two alternative placements of *Amborella* (Fig. S1) as indicated in the "Gene trees" column (blue = *Amborella* + *Nuphar*, orange = *Amborella* alone). Sites in each data set were sorted by evolutionary rates determined using the OV method, and divided into two equal partitions (i.e., slow and fast) following Xi et al. (2014). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

nucleotide sites (142,590 parsimony informative sites), and 29.9% missing data (including gaps). These simulations are similar to those presented by Xi et al. (2014) but are here expanded to address the criticisms of Simmons and Gatesy (2015). Namely, the new MSC method ASTRAL (Mirarab et al., 2014b) is used here, as well as both unpartitioned and partitioned concatenation analyses.

Simmons and Gatesy (2015) suggested that much of the gene-tree incongruence observed by Xi et al. (2014) could simply be explained by mis-rooting artifacts. In these cases, the placement of *Amborella* may be mis-rooted by the long branches leading to the outgroup *Selaginella* and the gymnosperms *Picea*, *Pinus*, and *Zamia*. Our simulation analyses of these 310 genes indeed show that the long branches separating *Selaginella* and the gymnosperms from the angiosperms could compromise gene tree estimation, especially for the OV-fast rate partition. When all 310 genes were assigned topology 1, on average 44.3% and 12.7% of the inferred gene trees recovered the correct phylogenetic relationships regarding basal angiosperms *Amborella* and *Nuphar* (Nymphaeales), three gymnosperms, and *Selaginella* for the OV-slow and OV-fast rate partitions, respectively. When all 310 genes were assigned topology 2, on average 50.2% and 21.7% of the inferred gene trees recovered the correct relationships regarding *Amborella*, *Nuphar*, *Selaginella*, and three gymnosperms for the OV-slow and OV-fast rate partitions, respectively. Despite these incorrect placements of *Amborella* in gene tree estimation, the proportions of the correct placement of *Amborella* recovered by two gene-tree-based coalescent methods (MP-EST and STAR) were high for OV-slow rate partitions (0.95–1.0) and moderate to high for OV-fast rate partitions (0.77–1.0; Fig. 4). The lone exception for MP-EST and STAR was when 60% of genes were simulated with the *Amborella* alone topology enforced. Here, the proportions of the correct placement of *Amborella* recovered by MP-EST and STAR were low–only 0.48 and 0.67, respectively (Fig. 4). A newly developed gene-tree-based coalescent method (ASTRAL) similarly recovered the correct placement of *Amborella* with a high proportion for OV-slow rate partitions (i.e., 0.90–1.0; Fig. 4). However, for OV-fast rate partitions, the proportion of the *Amborella* alone placement recovered by ASTRAL was 0.82, even though 60% of genes were simulated with the *Amborella* + *Nuphar* topology enforced (Fig. 4). These results suggest that MP-EST and STAR are robust to mis-rooted gene trees, consistent with simulations observed in those methods (Liu et al., 2009, 2010). In addition, our simulation analyses of these 310 genes do not support the claim by Simmons and Gatesy (2015) that ASTRAL is more robust to mis-rooted gene trees and is overall "superior" to MP-EST (Springer and Gatesy, 2016:1). We know of no paper that has claimed ASTRAL to be globally superior to MP-EST, and indeed ASTRAL remains untested in the anomaly zone, whereas MP-EST has been known to perform consistently in the anomaly zone (Liu et al., 2010). Furthermore, these results suggest that for recent re-analyses of these data, in which *Amborella* alone is inferred as sister to all angiosperms, that ASTRAL may be more prone to spurious results due to saturated nucleotide sites (Mirarab and Warnow, 2015). This may apply similarly to a separate analysis of the placement of *Amborella* using transcriptomes by Wickett et al. (2014; see also Liu et al., 2015b).

For concatenation analyses (RAxML), when there was a single placement of *Amborella* in the simulated gene trees (i.e., "X" equals 0 or 100), despite rate heterogeneity across genes, the proportions of the correct placement of *Amborella* recovered by the concatenation method were very high ($\geqslant$0.98) for both rate partitions (Fig. 4). In contrast, when 60–80% of genes were simulated with

the *Amborella* + *Nuphar* topology enforced, the concatenation analyses produced incongruent placements of *Amborella* across the two rate partitions (Fig. 4). Here, the OV-slow rate partitions corroborated results from the gene-tree-based coalescent analyses: the proportion of the correct placement of *Amborella* + *Nuphar* recovered by the concatenation method was high, ranging from 0.81 to 1.0 and 0.75 to 1.0 in unpartitioned and partitioned RAxML analyses, respectively. For OV-fast rate partitions, however, the concatenation method inferred the incorrect placement of *Amborella* alone at a high rate, i.e., 0.58–1.0 and 0.89–1.0 in unpartitioned and partitioned RAxML analyses, respectively. This observation that the concatenation analyses of OV-fast rate partitions support the placement of *Amborella* alone, despite the fact that up to 80% of the genes are simulated with the alternative *Amborella* + *Nuphar* topology, indicates that the concatenation analyses of OV-fast rate partitions are biased toward the placement of *Amborella* alone, even when it is incorrect. Therefore, our simulation corroborates previous results from Xi et al. (2014) and indicates that analyzing this 310-gene data set using gene-tree-based coalescent methods, or only the OV-slow rate partitions, is more likely to recover the correct placement of *Amborella*.

5. Testing the multispecies coalescent model using simulations

5.1. Previous coalescent simulations of species trees are not circular

In phylogenomic data analysis, MSC is the null model. The data simulated from the coalescent estimate of the species tree under MSC can be used to approximate the null distribution of data generated from the true species tree, assuming that the coalescent estimate of the species tree is close to the true tree. Alternatively, the true tree may be replaced by the trees estimated from other independent analyses, as long as there is evidence that the estimates from other studies are more accurate (i.e., closer to the true tree) than the tree estimated in the present study. S&G emphasized that the estimates of the species tree topology, branch lengths, and population sizes from other sources be used to simulate data, but there is no quantitative evidence that the estimates from previous studies are more accurate than the estimates in Song et al. (2012). In fact, the mammal data set in Song et al. (2012) contains 447 genes for 37 taxa, much larger than the phylogenetic data from which S&G obtained the estimates of the species tree, branch lengths, and population sizes in their analysis. Thus, using the estimates from Song et al. (2012) to simulate data is not circular, and in fact, it produces more accurate simulation under the null hypothesis.

5.2. Gene tree heterogeneity observed in empirical data sets is real

The accuracy of gene tree estimation is critical to the performance of gene-tree-based coalescent methods for species tree estimation. Previous simulation and empirical studies suggest that biased gene trees can mislead species tree estimation and produce wrong species trees with a high probability (e.g., Knowles et al., 2012; Chiari et al., 2012; Xi et al., 2015). A number of factors, including GC content, rate heterogeneity, codon usage bias, weak phylogenetic signal, biased resolution of polytomies and model misspecification, can lead to biased gene trees. Thus, gene tree estimation should be given more attention even though the ultimate goal is to estimate species trees. Although gene tree misestimation may introduce large errors in species tree accuracy, this is not legitimate evidence for advocating concatenation, which fundamentally oversimplifies the handling gene tree heterogeneity. We suggest that the tremendous amount of gene tree variation observed in the phylogenomic data by Song et al. (2012) and many other studies cannot simply be explained by gene tree estimation error. To study this, we re-estimated gene trees for the curated set of 413 genes by a newer version of PhyML v3.1 (Guindon et al., 2010) than that used in Song et al. (2012), with the GTR + Γ model and the NNI + SPR search scheme, and by RAxML using the GTRGAMMA model. The median RF distance (i.e., symmetric difference of Robinson and Foulds [1981]) between the estimated gene trees and the species tree is 16, regardless of the program (i.e., PhyML or RAxML) used for estimating individual gene trees. Then, we selected the first 413 trees (the number analyzed by Springer and Gatesy [2016]) from the 1000 gene trees simulated from the timetree with 1CU = 0.8 and the mutation rate = 1e−9 from S&G's study (Springer and Gatesy, 2016). The median RF distance between the 413 simulated gene trees (i.e., true gene trees) and the species tree was 2. We further simulated DNA sequences from 413 trees with the GTR + Γ model. The simulated sequences have the same length as in the real data set. The parameters of the GTR + Γ model for simulating sequences were estimated by RAxML from the 413-gene data set. Gene trees were estimated from the simulated sequences by RAxML with the GTRGAMMA model. In this case, the median RF distance between the gene trees estimated from the simulated sequences and the species tree is 4. The distances between the estimated gene trees and the species tree are the cumulative results of the distances between the true gene trees (i.e., gene trees simulated from the species tree) and the species tree and the distances between the estimated and true gene trees (i.e., gene tree estimation error). Since the median RF distance between the true gene trees and the species tree is 2, gene tree estimation errors account for only $(4-2)/16 = 12.5\%$ of the observed gene tree variation, indicating that gene tree estimation errors cannot adequately explain observed gene tree variation. Note that this simulation is internally consistent and does not depend on highly questionable parameters for branch lengths and ancestral population sizes such as those used by Springer and Gatesy (2016). Thus, gene tree variation is real and should be considered when modeling the evolution of individual genes.

Yet even if ILS explained 0% of gene tree variation, as Springer and Gatesy (2016) suggest in some of their simulations, it is a misconception that this situation invalidates the MSC. A common statement in the literature is that ILS is an assumption of the MSC model. However, ILS is not the key assumption of the MSC model, but rather a possible outcome of the MSC. The MSC in principle might produce gene trees identical in topology and branch lengths, completely devoid of ILS. This will happen when internal branch lengths of the species are large and/or population sizes are small. The fundamental assumption of the MSC is that gene trees are independently evolved within the species tree, assuming that an infinite number of recombination events have occurred between genes. In contrast, the concatenation model assumes no recombination between genes and thus no variation among gene trees. This assumption has rarely been stated, and, just as defenders of parsimony used to claim that the method was assumption-free, so S&G suggest that concatenation is immune to recombination.

S&G simulated gene trees using the timetree from dos Reis et al. (2012) with branches in millions of years (Fig. 1 of Springer and Gatesy [2016]) as the true species tree. The branch lengths in years were converted to coalescent units using the effective population sizes estimated from other sources, i.e., 1CU = 1.5 based on the estimate of effective population size from Carstens and Dewey (2010), 1CU = 2.75 based on Hobolth et al. (2007), 1CU = 0.4 based on Patel et al. (2013), and 1CU = 0.8 based on Rannala and Yang (2003). S&G suggested that the median RF distance between the simulated gene trees and the species tree can only explain a small portion of the observed gene tree variation. To demonstrate the effect of ancestral population size on gene tree variation, we

simulated gene trees with the same species tree used in S&G but with a larger CU = 3, 3.5, 4. The median RF distance becomes 8, 10, 12 for 1CU = 3, 3.5, 4, respectively. When 1CU = 3, 3.5, 4, the MSC accounts for 8/16 = 50%, 10/16 = 62.5%, and 12/16 = 75%, respectively, which is close to the proportion in Song et al. (2012). Our simulation does not take into account the variation in branch lengths of the species tree, which may further increase the proportion of the observed gene tree variation explained by the MSC.

In general, goodness of fit of a mathematical model derived from a real biological process, such as the coalescence process, does not necessarily indicate goodness of fit of the corresponding biological process, unless all other biological processes cannot produce similar patterns described by the same mathematical model. For example, 75% refers to the proportion of observed gene tree variation explained by the MSC. If we further exclude other biological phenomena that may lead to gene tree heterogeneity, then we can interpret 75% as the proportion of observed gene tree variation explained by ILS. Our simulation analysis indicates that gene tree heterogeneity observed in Song et al. is real, because it cannot be adequately explained by gene tree estimation error alone. The simulations by Springer and Gatesy (2016) make the highly unrealistic assumption that population sizes were constant along the entire tree for mammals and that a single value (mean/median) of the population size was adequate to simulate data. The estimates of the timetree branch lengths and population sizes used by Springer and Gatesy (2016) were generated under the MSC. It is ironic, and indeed circular, that Springer and Gatesy (2016) use parameters estimated from the MSC model as the evidence against the MSC model. Finally, although the mathematical model of the MSC explains a large portion of the observed gene tree variation, it does not necessarily imply that the observed gene tree variation is primarily caused by ILS. Identifying the biological origin of discordant gene trees requires collection of data sets containing information for distinguishing different biological origins.

5.3. Previous simulations do not suggest that concatenation is generally better than species tree inference

In terms of phylogenetic accuracy, recent published simulations testing comparing coalescent versus concatenation approaches in phylogenomics overwhelmingly favor coalescent methods. Where they are comparable in performance, or when concatenation methods appear to outperform MSC methods, the authors usually have not studied the statistical confidence of the resulting trees and therefore miss possible overestimation and instability of that confidence by concatenation methods (e.g., Bayzid et al., 2015; Chou et al., 2015). Concatenation and MSC methods are classic examples of the 'bias-variance' dilemma in statistics, in which one achieves reduced variance, and therefore higher support (as in concatenation methods) only at the cost of considerable bias in the estimates produced by the model (Liu et al., 2015b). In general, the situations in which concatenation is more accurate than coalescent methods are overstated by S&G. For example, the performance of the coalescent method STEM, one of the first implementation of the MSC model in a maximum likelihood framework (Kubatko et al., 2009), is highlighted repeatedly as an instance of the superiority of concatenation. However, S&G fail to point out an important property of STEM, namely, that the method is not misleading under the wide range of parameters assessed by previous simulations (Leaché and Rannala, 2011; Patel et al., 2013), but instead is uninformative regarding the species tree when there is a lack of information in the gene trees. When analyzing multilocus data with low information content such as those used in these simulations, empiricists must decide if they would prefer to use a method that has a strong tendency to be positively misleading (e.g., concatenation), or an approach that will accurately reflect the uncertainty in the data (e.g., the MSC).

One clear area where coalescent methods outperform concatenation is in the anomaly zone, as well as regions of tree space outside the anomaly zone that still exhibit substantial ILS. The anomaly zone is characterized by the presence of gene tree topologies that are more probable than the true species tree (termed anomalous gene trees = AGTs), which are the inevitable outcome of consecutive rapid speciation events in the species tree (Degnan and Rosenberg, 2006). Concatenation has no chance of escaping the anomaly zone, and will always favor the wrong tree with definitive support (Kubatko and Degnan, 2007). Coalescence methods accommodate ILS among genes instead of ignoring this process, and therefore provide a natural solution to the anomaly zone (Liu and Edwards, 2009).

S&G assert that the anomaly zone simulation studies were flawed because they focused on small asymmetric trees containing only four or five species. Unfortunately, S&G failed to realize that these logical simulation study designs reflected the scenarios where the anomaly zone has a firm theoretical basis. AGTs can occur on species trees containing four or more species, but in the four-taxon case, only the pectinate tree contains a set of internodes that can be in the anomaly zone, which can produce up to three AGTs (Degnan and Rosenberg, 2006). Scaling up to just five species introduces multiple sets of branches that can lead to as many 45 AGTs (Rosenberg and Tao, 2008). Computing the number of AGTs in trees larger than five-taxa quickly becomes impractical, but there are approaches for simplifying the problem (Rosenberg, 2013). Previous simulations used species trees with attributes that are well characterized under the anomaly zone. These simulations demonstrate that concatenation is positively misleading with as few as four taxa. Concatenating more genes will amplify the problem, and adding species provides more opportunities for concatenation to fail.

S&G routinely equate strong branch support and resolution with phylogenetic accuracy, and this mistake undermines their perceptions of the studies by Song et al. (2012) and Zhong et al. (2013). Higher support values and complete resolution are admirable qualities for any phylogeny, but these attributes should not be misinterpreted as signs of accuracy. For example, long-branch attraction leads to decisive phylogenetic trees when using parsimony (Swofford et al., 2001), but practitioners of phylogenetics are not fooled by this outcome. Similarly, as we stated above, concatenation will return strong branch support under the anomaly zone, but simulations have shown that concatenation is positively misleading in the anomaly zone. The inability of coalescence methods to produce resolved species trees with strong support from inaccurate gene trees is not a fundamental flaw of the approach (or with respect to the data by Song et al. [2012]). Rather, it is the desired outcome when analyzing data that lack information (e.g., a small number of loci or poorly supported gene trees). The incorporation of gene tree uncertainty into the species tree analysis is a benefit of using the MSC, not a detriment.

S&G contend that the benefits of concatenation are to decrease sampling error, find hidden support, and to offset homoplasy that might be present in some data partitions. These might be benefits from the perspective of a parsimony-based approach, but attempting to evaluate the performance of coalescent methods using parsimony rhetoric exposes some basic misunderstandings about MSC models shared by S&G. All genetic loci contain useful information regarding the species tree in a coalescent framework, although fast-evolving loci are now known to improve species tree analysis more so than loci with little phylogenetic information (Lanier et al., 2014; Liu et al., 2015b). Gene tree discordance provides valuable information regarding topology, population sizes, and divergence times, three important dimensions of species trees that are

estimated by many coalescent methods (Heled and Drummond, 2010). The information content in gene trees is not limited to their topology, branch support, or homoplasy metric. Each gene tree provides an independent source of evidence for estimating parameters of biological interest (Edwards, 2009a).

6. Biological realism of the multispecies coalescent model

6.1. Effects of selection on analyses using the MSC

S&G, in particular Springer and Gatesy (2016), raise the issue that many loci analyzed using the MSC are likely non-neutral, and therefore constitute a violation of the MSC. We agree that more work needs to be done in this arena, but casual thought experiments suggest that selection will rarely compromise the MSC so strongly that it misleads estimation of phylogenetic topologies. Selection often occurs at a minority of loci in the genome, in which case any spurious signals would likely be swamped out by the many neutral loci that are sampled. In such cases, methods such as gene tree outlier analysis could prove very effective (Weyenberg et al., 2014), analogous to F_{st} outlier analysis in phylogeography. Rather than influence tree topologies, selection will most likely influence branch lengths and estimates of ancestral effective population sizes that are yielded by some MSC methods. Extreme cases of selection-driven convergence in DNA or amino acid sequences are known to compromise estimation of gene trees (e.g., Castoe et al., 2009) and in such cases, phylogenetic analyses using the MSC will also be compromised, but not out of any deficiency of MSC methods per se. Such extreme cases of selection are thought to be rare and indeed, most forms of natural selection will not greatly compromise phylogenetic analysis (Edwards, 2009b). To our knowledge, balancing selection, such as occurs at genes of the major histocompatibility complex (MHC), and extreme cases of molecular convergence (e.g., Castoe et al., 2009), are the only types of selection that will distort gene trees in such a way that they no longer reflect the underlying species tree from which they were generated, and can be expected to grossly mislead phylogenetic analysis (Edwards, 2009b). Positive Darwinian or directional selection is likely to speed up substitution rates in one or more lineages, but all but the most extreme variation in substitution rates among lineages are easily accommodated by most phylogenetic methods for building gene trees. Positive selection will also have the effect of reducing variation within species, and in some cases selective sweeps have been known to reduce the incidence of ILS in genomic data sets (Scally et al., 2012). Such an event could actually be a boon for phylogenetic analysis, in so far as it reduces ILS, making gene trees more congruent with their species trees. Such events will, however, alter the effective population size of the constituent branches and thereby alter branch lengths for those methods that yield species tree branch lengths. As with discussions of recombination or gene flow, we view the introduction of natural selection into discussions about phylogenetic analysis as another positive consequence of population thinking that was brought about by the increasing use of MSC methods. But we caution that concatenation is no less immune to these effects just because they have not been discussed in the context of those approaches.

6.2. Networks as explanations for phylogenomic data

Putatively full resolution of trees based on genome-scale data (bootstrap percentages 100%, posterior probabilities 1.0) is a commonly encountered in many recent phylogenomic studies (e.g., Jarvis et al., 2014; Misof et al., 2014), but a closer look at many such studies reveals some poorly supported branches. Such poorly resolved branches may be an indication of reticulate or other non-phylogenetic processes in operation. For example, the placement of the tree shrew on the tree for mammals has been a contentious issue and remains unresolved even when analyzing the curated dataset by Song et al. (2012). This lack of resolution suggests conflicts in the data that place tree shrews either with Glires (rodents and Lagomorpha) or with primates. Such conflicts may not be reconciled by ILS or by any patterns generated from a bifurcating tree, and indeed may be better explained by processes evolving along a phylogenetic network (Kumar et al., 2013).

The era of genomics has unleashed unprecedented amounts of phylogenetic heterogeneity in data sets. This heterogeneity is poorly captured by concatenation approaches and the high certainty often yielded by concatenation may be one symptom of ignoring this heterogeneity. Thus interpreting high phylogenetic support as phylogenetic certainty is becoming increasingly untenable, and really is better thought of as a relict from traditional single gene analyses and the concatenation of mitochondrial (Janke et al., 1994) and later nuclear genes (Murphy et al., 2001). Concatenation of mitochondrial and sex chromosome sequences (Bidon et al., 2015) makes sense because they represent sequences that behave like single loci; they often yield a well-resolved gene tree. However, recent developments in phylogenetics have shown that for many clades a bifurcating species tree may not be the best explanation of the data, either because of ILS or, increasingly, because of introgressive hybridization, which increases the heterogeneity of phylogenetic signals in the data. Such phenomena may be particularly prevalent for plant systems (Zwickl et al., 2014; Stenz et al., 2015). Some have argued that, even for eukaryotes, the paradigm of a strictly bifurcating species tree may not exist given the mosaic nature of genomes and interacting clades (Whitfield and Lockhart, 2007). Sequence analyses and, in particular, analysis of genomic insertion of mobile genetic elements have shown that some divergences may not be resolvable as a bifurcating tree (Hallström and Janke, 2010; Suh et al., 2015) and that, instead, evolutionary networks are needed as more flexible explanations of the data (Bapteste et al., 2013). Biologists should therefore be more aware that "phylogenetic incongruence [is] a signal, rather than a problem" (Nakhleh, 2013) and treat it accordingly. In the case of the tree shrew, and many other lineages in vertebrate phylogenetics, different algorithms may yield different trees because of the mosaic nature of the data (Kumar et al., 2013) and the inability of a bifurcating tree to explain the patterns. Just as concatenation is a specific instance of the MSC, the MSC itself can be interpreted as a special case of network methods that model introgression and horizontal gene transfer and employing the MSC (Fig. 1). In that framework, traditional MSC models would then be those specific cases of network MSC models in which there is no introgression or other reticulations. We expect such network models to have increasing utility going forward.

Both S&G and the creators of recent MSC models have focused primarily on MSC models in which there is no gene flow or introgression. The field in general has overlooked introgression as another major source for phylogenetic conflict at higher taxonomic levels, even as several emerging models have appeared (Than et al., 2008; Durand et al., 2011; Mailund et al., 2012; Park and Nakhleh, 2012; Smith et al., 2013; Liu et al., 2014; Yu and Nakhleh, 2015; reviewed in Nakhleh, 2013). The short internal branches of the mammal tree that are the focus of most phylogenetic conflict usually span only 1–3 million years in time (Hallström and Janke, 2010). This is a typical time interval during which hybridization between different mammal species typically occurs (Jónsson et al., 2014; Kutschera et al., 2014). We affirm that evolutionary reticulations due both to ILS and hybridization are common and should not be ignored in an attempt to produce a "fully resolved" bifurcating tree. Even for deep mammalian divergences these

two processes can now be distinguished by available algorithms and may better explain some of the patterns observed in existing phylogenomic data sets.

7. Serious scientific exchange deserves a civil tone

Finally, we disapprove of the language and tone of S&G critiques, particularly that in Springer and Gatesy (2016). In using inflammatory language, S&G exaggerate many assertions about errors in the papers critiqued and the conclusions once those errors are rectified. The title of their paper, "The gene tree delusion", captures the unnecessarily taunting, provocative, and excessively combative language throughout this work. Phylogenetics, and apparently now phylogenomics, has a long history of combative, exaggerating, and ultimately unhelpful language in what are otherwise academic exchanges in the literature. Those readers who actually experienced those exchanges may chuckle at the present episode, thinking that we got off lightly. But, however caustic the written exchanges were in the cladist/pheneticist/likelihood wars of the 1970s and 1980s, we respectfully suggest that the world has changed since then, and that such language and tone are misplaced in phylogenetics or any other field of science. In particular, the phylogenetics community today is substantially more diverse than it was then, and in the intervening years, factors both internal and external to science have moved it to a realm of greater accountability and responsibility.

Along these lines, it is unscientific and exaggerating to refer to MSC models as "delusional", simulations as "illogical", or data sets as "error-ridden" with "wholesale" mistakes. This needlessly ridicules the earnest attempts by numerous scientists in the field to advance phylogenomics and improve the models on which they are based. Ultimately authors of such papers and the editors handling them are responsible for such language and must appreciate that it undercuts their authority and ultimately works against the long-term advancement of science. We trust that most readers of the S&G papers see through the smoke-and-mirrors of such language; any other conclusion would be an insult not only to the authors of S&G's targeted critiques, but to the entire phylogenomics community.

8. Concluding remarks

Much of the confusion clarified above is likely a result of a failure by S&G to appreciate that concatenation occurs at one end of a spectrum encompassed by MSC models. S&G portray coalescent methods as a new scientific paradigm that should be subjected to critical scrutiny before abandoning concatenation, what they refer to as the *status quo* of phylogenetics. We agree that there is merit in comparing and contrasting new phylogenetic approaches against existing methods in a rigorous manner. But coalescent theory has firm conceptual roots in population genetics, and early work on the subject described the nature of gene tree discordance mathematically (Tajima, 1983; Takahata and Nei, 1985; Pamilo and Nei, 1988). Like phylogenetic theory, coalescent theory began with simpler models and was later expanded to more general models (Kingman, 1982a,b, 2000), such as the MSC. However, during the 20-year gap between the development of coalescent theory and the implementation of an MSC model (Rannala and Yang, 2003), phylogenetic analysis using a biologically realistic model for multilocus data was not feasible due to lack of theory and algorithms as well as lack of large multilocus data sets. After considerable debates in the literature about "total evidence" (reviewed in Edwards, 2009a), concatenation was the *status quo* in phylogenetics by default, not because it attempted to model any critical aspect of genetic inheritance, but because it was in many ways a stop gap approach with few alternatives, and was the best way to maximally resolve a phylogeny with a small number of genes. It has priority over the MSC only temporally, not conceptually. We view concatenation as a method of convenience that carried molecular systematics for several decades and still does. We also suggest that "total evidence" does not discriminate between concatenation and MSC models, which we view as equally consistent with that philosophical mandate in so far as they both provide means of combining multiple sources of data. But now that computational tools and availability of data have caught up with theory and first principles of genetics, the time has come to adopt and refine methods that can leverage genomic data to estimate phylogeny using approaches that are motivated by biological principles.

A key means by which S&G argue for concatenation is by suggesting that recombination within loci (e.g., between exons in coding regions), violates the MSC model. Although intra-locus recombination does technically violate the MSC model, researchers who have actually studied the problem assert that its effects are minimal, occur only in extreme situations, and are dwarfed by other issues germane to species tree reconstruction. The concept of the recombination ratchet advanced by S&G is not helpful because recombination events are most often localized to individual branches (as opposed to nodes) and the smallest non-recombining unit in the history of a clade is not synonymous with the smallest unit of phylogenetic analysis. Implementation of all models incur model violations, and we contend that recombination within loci has not been demonstrated to drastically compromise recent phylogenomic studies. We contend, and recent research shows, that transcriptomes and other type of data sets involving linked loci, remain reasonable and viable data sets for analysis by the MSC, even if recombination may compromise their use. Finally, any negative effects or recombination also violate the concatenation model, in which all sites are assumed to belong to one non-recombining unit. Unwittingly, S&G construct arguments that build a stronger case against the concatenation approach that they champion than against the more general MSC model, because concatenation lies at the end of the spectrum at which the negative effects of recombination should be most pronounced.

The overall robustness of phylogenetic models employing the MSC to recombination and other model violations still needs to be fully evaluated, but their robustness compared to the more restricted concatenation model has already been demonstrated in several contexts (Lanier and Knowles, 2012; Liu et al., 2015a). S&G unnecessarily highlight the potential conflicts between concatenation and species tree methods without acknowledging that in most cases, concatenation and species tree methods will in fact recover the same relationships – albeit with different branch lengths and levels of support. There are important conditions, however, in which estimates from the two methods may strongly differ, conditions such as short internodes or the anomaly zone, which we believe have been encountered in several recent phylogenomic data sets. It is precisely under these conditions in which MSC models have been shown to result in better estimates than concatenation approaches. S&G should not take inadvertent errors in recent papers as reasons to discard MSC models or make claims that model violations undermine their utility. Conspicuously, S&G have little to say about the use of subsampling phylogenomic data as a test for the consistency and robustness of phylogenetic methods (Song et al., 2012), possibly because under these conditions the deficiencies of concatenation become even more glaring.

As we approach an increasingly better understood Tree of Life, the phylogenomics community will be faced with an increasing proportion of short, difficult branches for which full resolution requires more accurate and precise modeling. However, reticulation in the evolutionary tree may make it impossible to resolve some branches as bifurcations with any amount of data or

improvement of traditional bifurcating models. A realization that our models have limitations, however, is no justification for refusing to account for the biological processes that we know exist and know should not be ignored. Instead, we should strive to strike the balance between model complexity and tractability that maximizes our ability to accurately elucidate evolutionary history. We believe MSC models achieve such a balance, while acknowledging that even more complex models, such as those involving networks, may prove not only biologically plausible but even better at explaining the diversity of signals we now routinely encounter in phylogenomics data.

Acknowledgments

We thank MPE Editor Derek Wildman for allowing us to write this piece, Vadim Goremykin for valuable discussion, one anonymous reviewer and Ed Braun for constructive criticism on the manuscript, Joe Felsenstein, Monty Slatkin, and Michael Lynch for discussion of Fig. 1, and John Gatesy and Mark Springer for sharing data and details of their work. The writing of this comment was supported by US National Science Foundation (NSF) grants DEB-1120243 to CCD, DEB-1120516 to EML, DEB-1145978 to D. Rokyta, ARL, and EML, and EAR-1355343 to SVE; National Natural Science Foundation of China grant 31570219 and funding from the Priority Academic Program Development of Jiangsu Higher Education Institutions to BZ.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.10.027.

References

Ané, C., Sanderson, M., 2005. Missing the forest for the trees: phylogenetic compression and its implications for inferring complex evolutionary histories. Syst. Biol. 54, 146–157.
Ané, C., 2011. Detecting phylogenetic breakpoints and discordance from genome-wide alignments for species tree reconstruction. Genome Biol. Evol. 3, 246–258.
Bapteste, E., van Iersel, L., Janke, A., Kelchner, S., Kelk, S., McInerney, J.O., Morrison, D.A., Nakhleh, L., Steel, M., Stougie, L., Whitfield, J., 2013. Networks, expanding evolutionary thinking. Trends Genet. 29, 439–441.
Bayzid, M.S., Mirarab, S., Boussau, B., Warnow, T., 2015. Weighted statistical binning: enabling statistically consistent genome-scale phylogenetic analyses. Plos One 10, e0129183.
Becquet, C., Przeworski, M., 2007. A new approach to estimate parameters of speciation models with application to apes. Genome Res. 17, 1505–1519.
Bidon, T., Schreck, N., Hailer, F., Nilsson, M., Janke, A., 2015. Genome-wide search identifies 1.9 megabases from the polar bear Y chromosome for evolutionary analyses. Genome Biol. 7, 2010–2022.
Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., RoyChoudhury, A., 2012. Inferring species trees directly from Biallelic genetic markers, bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29, 1917–1932.
Castoe, T.A., de Koning, A.P.J., Kim, H.-M., Gu, W., Noonan, B.P., Naylor, G., Jiang, Z.J., Parkinson, C.L., Pollock, D.D., 2009. Evidence for an ancient adaptive episode of convergent molecular evolution. Proc. Natl. Acad. Sci. USA 106, 8986–8991.
Carstens, B.C., Dewey, T.A., 2010. Species delimitation using a combined coalescent and information-theoretic approach: an example from North American *Myotis* bats. Syst. Biol. 59, 400–414.
Castillo-Ramírez, S. et al., 2010. Bayesian estimation of species trees: a practical guide to optimal sampling and analysis. In: Knowles, L.L., Kubatko, L.S. (Eds.), Estimating Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell, New Jersey, pp. 15–33.
Chiari, Y., Cahais, V., Galtier, N., Delsuc, F., 2012. Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). BMC Biol. 10, 65.
Chifman, J., Kubatko, L., 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics 30, 3317–3324.
Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., Warnow, T., 2015. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. BMC Genom. 16 (Suppl. 10), S2.
Cummins, C.A., McInerney, J.O., 2011. A method for inferring the rate of evolution of homologous characters that can potentially improve phylogenetic inference, resolve deep divergence and correct systematic biases. Syst. Biol. 60, 833–844.
Davidson, R., Vachaspati, P., Mirarab, S., Warnow, T., 2015. Phylogenomic species tree estimation in the presence of incomplete lineage sorting and horizontal gene transfer. BMC Genom. 16 (Suppl. 10), S1.
Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2, 762–768.
dos Reis, M., Inoue, J., Hasegawa, M., Asher, R.J., Donoghue, P.C., Yang, Z., 2012. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. Proc. Biol. Sci. 279, 3491–3500.
Dunn, C.W., Howison, M., Zapata, F., 2013. Agalma, an automated phylogenomics workflow. BMC Bioinform. 14, 330.
Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–2252.
Eckert, A.J., Carstens, B.C., 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. Mol. Phylogenet. Evol. 49, 832–842. http://dx.doi.org/10.1016/j.ympev.2008.09.008.
Edwards, S.V., 2009a. Is a new and general theory of molecular systematics emerging? Evolution 63, 1–19.
Edwards, S.V., 2009b. Natural selection and phylogenetic analysis. Proc. Natl. Acad. Sci. USA 106, 8799–8800.
Edwards, S.V., 2015. Inferring species trees. In: Richard Kliman (Ed.), Encyclopedia of Evolutionary Biology. Elsevier, Amsterdam (in press).
Felsenstein, J., 1988. Phylogenies and quantitative characters. Annu. Rev. Ecol. Syst. 19, 445–471.
Gatesy, J., O'Grady, P., Baker, R., 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15, 271–313.
Gatesy, J., Baker, R.H., 2005. Hidden likelihood support in genomic data: can forty-five wrongs make a right? Syst. Biol. 54, 483–492.
Gatesy, J., Springer, M.S., 2013. Concatenation versus coalescence versus "concatalescence". Proc. Natl. Acad. Sci. USA 110, E1179.
Gatesy, J., Springer, M.S., 2014. Phylogenetic analysis at deep timescales, unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80, 231–266.
Goldman, N., 1990. Maximum likelihood inference of phylogenetic trees, with special reference to a Poisson process model of DNA substitution and to parsimony analyses. Syst. Biol. 39, 345–361.
Goremykin, V., Nikiforova, S., Bininda-Emonds, O., 2010. Automated removal of noisy data in phylogenomic analyses. J. Mol. Evol. 71, 319–331.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.
Hallström, B.M., Janke, A., 2008. Resolution among major placental mammal interordinal relationships with genome data imply that speciation influenced their earliest radiations. BMC Evol. Biol. 8, 162.
Hallström, B.M., Janke, A., 2010. Mammalian evolution may not be strictly bifurcating. Mol. Biol. Evol. 27, 2804–2816.
Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biology Evol. 27, 570–580.
Hobolth, A., Christensen, O.F., Mailund, T., Schierup, M.H., 2007. Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. PLoS Genet. 3, e7.
Janke, A., Feldmaier-Fuchs, G., Thomas, W.K., Von Haeseler, A., Pääbo, S., 1994. The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137, 243–256.
Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y.W., Faircloth, B.C., Nabholz, B., Howard, J.T., Suh, A., Weber, C.C., da Fonseca, R.R., Li, J.W., Zhang, F., Li, H., Zhou, L., Narula, N., Liu, L., Ganapathy, G., Boussau, B., Bayzid, M.S., Zavidovych, V., Subramanian, S., Gabaldon, T., Capella-Gutierrez, S., Huerta-Cepas, J., Rekepalli, B., Munch, K., Schierup, M., Lindow, B., Warren, W.C., Ray, D., Green, R.E., Bruford, M.W., Zhan, X.J., Dixon, A., Li, S.B., Li, N., Huang, Y.H., Derryberry, E.P., Bertelsen, M.F., Sheldon, F.H., Brumfield, R.T., Mello, C.V., Lovell, P.V., Wirthlin, M., Schneider, M.P.C., Prosdocimi, F., Samaniego, J.A., Velazquez, A.M.V., Alfaro-Nunez, A., Campos, P.F., Petersen, B., Sicheritz-Ponten, T., Pas, A., Bailey, T., Scofield, P., Bunce, M., Lambert, D.M., Zhou, Q., Perelman, P., Driskell, A.C., Shapiro, B., Xiong, Z.J., Zeng, Y.L., Liu, S.P., Li, Z.Y., Liu, B.H., Wu, K., Xiao, J., Yinqi, X., Zheng, Q.M., Zhang, Y., Yang, H.M., Wang, J., Smeds, L., Rheindt, F.E., Braun, M., Fjeldsa, J., Orlando, L., Barker, F.K., Jonsson, K.A., Johnson, W., Koepfli, K.P., O'Brien, S., Haussler, D., Ryder, O.A., Rahbek, C., Willerslev, E., Graves, G.R., Glenn, T.C., McCormack, J., Burt, D., Ellegren, H., Alstrom, P., Edwards, S.V., Stamatakis, A., Mindell, D.P., Cracraft, J., Braun, E.L., Warnow, T., Jun, W., Gilbert, M.T.P., Zhang, G.J., 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346, 1320–1331.
Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., Albrechtsen, A., Petersen, B., Korneliussen, T.S., Vilstrup, J.T., Lear, T., Myka, J. L., Lundquist, J., Miller, D.C., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Stagegaard, J., Strauss, G., Bertelsen, M.F., Sicheritz-Ponten, T., Antczak, D.F., Bailey, E., Nielsen, R., Willerslev, E., Orlando, L., 2014. Speciation with gene flow in equids despite extensive chromosomal plasticity. Proc. Natl. Acad. Sci. USA 111, 18655–18660.
Kingman, J.F.C., 1982a. The coalescent. Stoch. Proc. Appl. 13, 235–248.
Kingman, J.F.C., 1982b. On the genealogy of large populations. J. Appl. Probab., 27–43
Kingman, J.F.C., 2000. Origins of the coalescent: 1974–1982. Genetics 156, 1461–1463.

Knowles, L.L., Lanier, H.C., Klimov, P.B., He, Q., 2012. Full modeling versus summarizing gene-tree uncertainty, method choice and species-tree accuracy. Mol. Phylogenet. Evol. 65, 501–509.
Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.
Kubatko, L.S., Carstens, B.C., Knowles, L.L., 2009. STEM, species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25, 971–973.
Kumar, V., Hallström, B.M., Janke, A., 2013. Coalescent-based genome analyses resolve the early branches of the euarchontoglires. PLoS One 8, e60019.
Kutschera, V.E., Bidon, T., Hailer, F., Rodi, J.L., Fain, S.R., Janke, A., 2014. Bears in a forest of gene trees: phylogenetic inference is complicated by incomplete lineage sorting and gene flow. Mol. Biol. Evol. 31, 2004–2017.
Lanier, H.C., Knowles, L.L., 2012. Is recombination a problem for species-tree analyses? Syst. Biol. 61, 691–701.
Lanier, H.C., Huang, H., Knowles, L.L., 2014. How low can you go? The effects of mutation rate on the accuracy of species-tree estimation. Mol. Phylogenet. Evol. 70, 112–119.
Leaché, A.D., Rannala, B., 2011. The accuracy of species tree estimation under simulation, a comparison of methods. Syst. Biol. 60, 126–137.
Leaché, A.D., Harris, R.B., Rannala, B., Yang, Z., 2014. The influence of gene flow on species tree estimation: A simulation study. Syst. Biol. 63 (1), 17–30. http://dx.doi.org/10.1093/sysbio/syt049.
Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44, 99–121.
Lewis, P.O., 1998. Maximum likelihood as an alternative to parsimony for inferring phylogeny using nucleotide sequence data. Mol. Syst. Plants II. Springer, pp. 132–163.
Liang, D., Shen, X.X., Zhang, P., 2013. One thousand two hundred ninety nuclear genes from a genome-wide survey support lungfishes as the sister group of tetrapods. Mol. Biol. Evol. 30, 1803–1807.
Liu, L., Edwards, S.V., 2009. Phylogenetic analysis in the anomaly zone. Syst. Biol. 58, 452–460.
Liu, L., Edwards, S.V., 2015. Comment on "Statistical binning enables an accurate coalescent-based estimation of the avian tree". Science 350, 171. http://dx.doi.org/10.1126/science.aaa7343.
Liu, L., Yu, L., 2010. Phybase, an R package for species tree analysis. Bioinformatics 26, 962–963.
Liu, L., Yu, L., Pearl, D.K., Edwards, S.V., 2009. Estimating species phylogenies using coalescence times among sequences. Syst. Biol. 58, 468–477.
Liu, L., Yu, L., Edwards, S.V., 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol. Biol. 10, 302.
Liu, K.J., Dai, J., Truong, K., Song, Y., Kohn, M.H., Nakhleh, L., 2014. An HMM-based comparative genomic framework for detecting introgression in eukaryotes. Plos Comput. Biol. 10, e1003649.
Liu, L., Xi, Z., Davis, C.C., 2015a. Coalescent methods are robust to the simultaneous effects of long branches and incomplete lineage sorting. Mol. Biol. Evol. 32, 791–805.
Liu, L., Xi, Z., Wu, S., Davis, C.C., Edwards, S.V., 2015b. Estimating phylogenetic trees from genome-scale data. Ann. NY Acad. Sci. http://dx.doi.org/10.1111/nyas.12747. Article first published online: 14 Apr 2015.
Lynch, M., 2007. The frailty of adaptive hypotheses for the origins of organismal complexity. Proc. Natl. Acad. Sci. USA 104 (1), 8597–8604.
Martin, D.P., Posada, D., Crandall, K.A., Williamson, C., 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. Aids Res. Human Retrovir. 21, 98–102.
McCormack, J.E., Huang, H.T., Knowles, L.L., 2009. Maximum likelihood estimates of species trees: how accuracy of phylogenetic inference depends upon the divergence history and sampling design. Syst. Biol. 58, 501–508.
McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species tree analysis. Genome Res. 22, 746–754.
Mailund, T., Halager, A.E., Westergaard, M., Dutheil, J.Y., Munch, K., Andersen, L.N., Lunter, G., Prufer, K., Scally, A., Hobolth, A., Schierup, M.H., 2012. A new isolation with migration model along complete genomes infers very different divergence processes among closely related great ape species. PLoS Genet. 8, e1003125.
Mirarab, S., Bayzid, M.S., Boussau, B., Warnow, T., 2014a. Statistical binning enables an accurate coalescent-based estimation of the avian tree. Science 346, 1240463.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014b. ASTRAL, genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548.
Mirarab, S., Warnow, T., 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44–i52.
Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspock, U., Aspock, H., Bartel, D., Blanke, A., Berger, S., Bohm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B.M., Schutte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M., Zhou, X., 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science 346, 763–767.
Murphy, W.J., Eizirik, E., O'Brien, S.J., Madsen, O., Scally, M., Douady, C.J., Teeling, E., Ryder, O.A., Stanhope, M.J., de Jong, W.W., Springer, M.S., 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294, 2348–2351.
Naduvilezhath, L., Rose, L.E., Metzler, D., 2011. Jaatha: A fast composite-likelihood approach to estimate demographic parameters. Mol. Ecol. 20, 2709–2723. http://dx.doi.org/10.1111/j.1365-294X.2011.05131.x.
Nakhleh, L., 2013. Computational approaches to species phylogeny inference and gene tree reconciliation. Trends Ecol. Evol. 28, 719–728.
Notredame, C., Higgins, D.G., Heringa, J., 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.
Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583.
Park, H.J., Nakhleh, L., 2012. Inference of reticulate evolutionary histories by maximum likelihood, the performance of information criteria. BMC Bioinform. 13 (Suppl. 19), S12.
Patel, S., Kimball, R.T., Braun, E.L., 2013. Error in phylogenetic estimation for bushes in the tree of life. J. Phylogenet. Evol. Biol. 1, 110.
Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., et al., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9 (3), e1000602.
Posada, D., Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. USA 98, 13757–13762.
Posada, D., Crandall, K.A., 2002. The effect of recombination on the accuracy of phylogeny estimation. J. Mol. Evol. 54, 396–402.
Posada, D., Crandall, K.A., Holmes, E.C., 2002. Recombination in evolutionary genomics. Annu. Rev. Genet. 36, 75–97.
Rambaut, A., Grassly, N.C., 1997. Seq-Gen, an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Appl. Biosci. 13, 235–238.
Rannala, B., Yang, Z., 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164, 1645–1656.
Ranwez, V., Delsuc, F., Ranwez, S., Belkhir, K., Tilak, M.-K., Douzery, E.J., 2007. OrthoMaM: A database of orthologous genomic markers for placental mammal phylogenetics. BMC Evol. Biol. 7, 241. http://dx.doi.org/10.1186/1471-2148-7-241.
Robinson, D.F., Foulds, L.R., 1981. Comparison of phylogenetic trees. Math. Biosci. 53, 131–147.
Rosenberg, N.A., 2013. Discordance of species trees with their most likely gene trees: a unifying principle. Mol. Biol. Evol. 30, 2709–2713.
Rosenberg, N.A., Nordborg, M., 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nat. Rev. Genet. 3, 380–390.
Rosenberg, N.A., Tao, R., 2008. Discordance of species trees with their most likely gene trees, the case of five taxa. Syst. Biol. 57, 131–140.
Ruths, D., Nakhleh, L., 2005. Recombination and phylogeny: effects and detection. Int. J. Bioinform. Res. Appl. 1, 202–212.
Scally, A., Dutheil, J.Y., Hillier, L.W., Jordan, G.E., Goodhead, I., Herrero, J., Hobolth, A., Lappalainen, T., Mailund, T., Marques-Bonet, T., McCarthy, S., Montgomery, S.H., Schwalie, P.C., Tang, Y.A., Ward, M.C., Xue, Y.L., Yngvadottir, B., Alkan, C., Andersen, L.N., Ayub, Q., Ball, E.V., Beal, K., Bradley, B.J., Chen, Y., Clee, C.M., Fitzgerald, S., Graves, T.A., Gu, Y., Heath, P., Heger, A., Karakoc, E., Kolb-Kokocinski, A., Laird, G.K., Lunter, G., Meader, S., Mort, M., Mullikin, J.C., Munch, K., O'Connor, T.D., Phillips, A.D., Prado-Martinez, J., Rogers, A.S., Sajjadian, S., Schmidt, D., Shaw, K., Simpson, J.T., Stenson, P.D., Turner, D.J., Vigilant, L., Vilella, A.J., Whitener, W., Zhu, B.L., Cooper, D.N., de Jong, P., Dermitzakis, E.T., Eichler, E.E., Flicek, P., Goldman, N., Mundy, N.I., Ning, Z.M., Odom, D.T., Ponting, C.P., Quail, M.A., Ryder, O.A., Searle, S.M., Warren, W.C., Wilson, R.K., Schierup, M.H., Rogers, J., Tyler-Smith, C., Durbin, R., 2012. Insights into hominid evolution from the gorilla genome sequence. Nature 483, 169–175.
Simmons, M.P., Gatesy, J., 2015. Coalescence vs. concatenation, sophisticated analyses vs. first principles applied to rooting the angiosperms. Mol. Phylogenet. Evol. 91, 98–122.
Slatkin, M., Pollack, J.L., 2006. The concordance of gene trees and species trees at two linked loci. Genetics 172, 1979–1984.
Smith, S.A., Dunn, C.W., 2008. Phyutility, a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24, 715–716.
Smith, S.A., Brown, J.W., Hinchliff, C.E., 2013. Analyzing and synthesizing phylogenies using tree alignment graphs. Plos Comput. Biol. 9, e1003223.
Song, S., Liu, L., Edwards, S.V., Wu, S., 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proc. Natl. Acad. Sci. USA 109, 14942–14947.
Springer, M.S., Gatesy, J., 2014. Land plant origins and coalescence confusion. Trends Plant Sci. 19, 267–269.
Springer, M.S., Gatesy, J., 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94, 1–33.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Stenz, N.W.M., Larget, B., Baum, D.A., Ané, C., 2015. Exploring tree-like and non-tree-like patterns using genome sequences: an example using the inbreeding plant species *Arabidopsis thaliana* (L.) Heynh. Syst. Biol. 64, 809–823.

Suh, A., Smeds, L., Ellegren, H., 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. PLoS Biol. 13, e1002224.

Swofford, D.L., Waddell, P.J., Huelsenbeck, J.P., Foster, P.G., Lewis, P.O., Rogers, J.S., 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Syst. Biol. 50, 525–539.

Tajima, F., 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105, 437–460.

Takahata, N., Nei, M., 1985. Gene genealogy and variance of interpopulational nucleotide differences. Genetics 110, 325–344.

Than, C., Ruths, D., Nakhleh, L., 2008. PhyloNet, a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinform. 9, 322.

Thomson, R.C., Shedlock, A.M., Edwards, S.V., Shaffer, H.B., 2008. Developing markers for multilocus phylogenetics in non-model organisms, a test case with turtles. Mol. Phylogenet. Evol. 49, 514–525.

Tsagkogeorga, G., Parker, J., Stupka, E., Cotton, J.A., Rossiter, S.J., 2013. Phylogenomic analyses elucidate the evolutionary relationships of bats. Curr. Biol. 23, 2262–2267.

Wang, N., Braun, E.L., Kimball, R.T., 2012. Testing hypotheses about the sister group of the Passeriformes using an independent 30 locus dataset. Mol. Biol. Evol. 29, 737–750.

Weyenberg, G., Huggins, P.M., Schardl, C.L., Howe, D.K., Yoshida, R., 2014. KDETREES: non-parametric estimation of phylogenetic tree distributions. Bioinformatics 30, 2280–2287.

White, M.A., Ane, C., Dewey, C.N., Larget, B.R., Payseur, B.A., 2009. Fine-scale phylogenetic discordance across the house mouse genome. PLoS Genet. 5, e1000729.

Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22, 258–265.

Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., Ruhfel, B.R., Wafula, E., Der, J.P., Graham, S.W., Mathews, S., Melkonian, M., Soltis, D.E., Soltis, P.S., Miles, N.W., Rothfels, C.J., Pokorny, L., Shaw, A.J., DeGironimo, L., Stevenson, D.W., Surek, B., Villarreal, J.C., Roure, B., Philippe, H., dePamphilis, C.W., Chen, T., Deyholos, M.K., Baucom, R.S., Kutchan, T.M., Augustin, M.M., Wang, J., Zhang, Y., Tian, Z., Yan, Z., Wu, X., Sun, X., Wong, G.K., Leebens-Mack, J., 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc. Natl. Acad. Sci. USA 111, E4859–E4868.

Wu, S., Song, S., Liu, L., Edwards, S.V., 2013. Reply to Gatesy and Springer: the multispecies coalescent model can effectively handle recombination and gene tree heterogeneity. Proc. Natl. Acad. Sci. USA 110, E1180.

Xi, Z., Rest, J.S., Davis, C.C., 2013. Phylogenomics and coalescent analyses resolve extant seed plant relationships. PLoS One 8, e80870.

Xi, Z., Liu, L., Rest, J.S., Davis, C.C., 2014. Coalescent versus concatenation methods and the placement of *Amborella* as sister to water lilies. Syst. Biol. 63, 919–932.

Xi, Z., Liu, L., Davis, C.C., 2015. Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. Mol. Phylogenet. Evol. 92, 63–71.

Yu, Y., Nakhleh, L., 2015. A maximum pseudo-likelihood approach for phylogenetic networks. BMC Genom. 16 (Suppl. 10), S10.

Zhong, B., Liu, L., Yan, Z., Penny, D., 2013. Origin of land plants using the multispecies coalescent model. Trends Plant Sci. 18, 492–495.

Zhong, B., Liu, L., Penny, D., 2014. The multispecies coalescent model and land plant origins, a reply to Springer and Gatesy. Trends Plant Sci. 19, 270–272.

Zwickl, D.J., Stein, J.C., Wing, R.A., Ware, D., Sanderson, M.J., 2014. Disentangling methodological and biological sources of gene tree discordance on Oryza (Poaceae) chromosome 3. Syst. Biol. 63, 645–659.