Uncertainty in divergence time estimation

Tom Carruthers[1*] **and Robert W. Scotland**[2]

Affiliations:

[1] Royal Botanic Gardens Kew, Richmond, London, TW9 3AE, United Kingdom

[2] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, United Kingdom

[*] Correspondence to be sent to: t.carruthers@kew.org

ABSTRACT

Understanding and representing uncertainty is crucial in academic research, because it enables studies to build on the conclusions of previous studies, leading to robust advances in a particular field. Here, we evaluate the nature of uncertainty and the manner by which it is represented in divergence time estimation, a field that is fundamental to many aspects of macroevolutionary research, and where there is evidence that uncertainty has been seriously underestimated. We address this issue in the context of methods used in divergence time estimation, and with respect to the manner by which time-calibrated phylogenies are interpreted. With respect to methods, we discuss how the assumptions underlying different methods may not adequately reflect uncertainty about molecular evolution, the fossil record, or diversification rates. Therefore, divergence time estimates may not adequately reflect uncertainty, and may be directly contradicted by subsequent findings. For the interpretation of time-calibrated phylogenies, we discuss how the use of time-calibrated phylogenies for reconstructing general evolutionary timescales leads to inferences about macroevolution that are highly sensitive to methodological limitations in how uncertainty is accounted for. By contrast, we discuss how the use of time-calibrated phylogenies to test specific hypotheses leads to inferences about macroevolution that are less sensitive to methodological limitations. Given that many biologists wish to use time-calibrated phylogenies to reconstruct general evolutionary timescales, we conclude that the development of methods of divergence time estimation that adequately account for uncertainty is necessary.

Keywords: uncertainty, divergence time estimation, macroevolution

Uncertainty describes a situation where knowledge of something is incomplete, and it is expressed by the degree of precision with which a statement is made (Fischhoff and Davis 2014). Ideally, scientific research will adequately express uncertainty whilst moving to a state of less uncertainty, such that statements become more precise and more accurate (Fig 1a). By contrast, it may be the case that there is a misleading perspective of uncertainty, such that statements are overly precise, inaccurate, and potentially contradicted by subsequent findings (Fig. 1b). Adequately expressing uncertainty is important for estimates that are derived from analyses, and also for the assumptions that are made about variables prior to performing analyses.

Accounting for uncertainty is especially pertinent in macroevolutionary studies which are dependent on understanding complex interactions between biology, climate, and geology - variables about which knowledge is very incomplete (Hughes and Eastwood 2006; Hoorn et al. 2010; Hughes et al. 2013; Lagomarsino et al. 2016; Perrigo et al. 2020; Rahbeck et al. 2019). Here, we discuss the nature of uncertainty in divergence time estimation – a fundamental step in many macroevolutionary studies (eg. Magallón et al. 2015; Lagomarsino et al. 2016; Cardillo et al. 2017; Folk et al. 2019), and a field in which uncertainty may be considerably greater than expressed in previous studies (Fig. 1b) (Wilf and Escapa 2015; Wilf et al. 2017; Brown and Smith 2018).

Methodological and theoretical developments in divergence time estimation have previously been discussed in detail (Donoghue and Yang 2007; Donoghue and Yang 2016; Bromham et al. 2018). The purpose of this short opinion piece is not to provide another detailed discussion of the merits of different methods. Instead, we focus specifically on how the field of divergence time estimation incorporates uncertainty, both with respect to the methods that are used in divergence time estimation, and the manner by which divergence time estimates are interpreted when discussing macroevolution. We hope that this can stimulate a more widespread consideration of the type of information that can actually be conveyed in time-calibrated phylogenies. Although our discussion is limited to divergence time estimation, the general principles that we highlight are relevant to other fields that use phylogenies to study macroevolution.

HOW METHODS OF DIVERGENCE TIME ESTIMATION INCORPORATE UNCERTAINTY

Strict clocks, relaxed clocks, fossil calibrations, and branching processes: fundamental uncertainty about rate and time

Divergence time estimation, whereby the branching events in a molecular phylogeny are placed on an absolute timescale, is based on the premise that the number of substitutions (*n*) that have occurred along a branch can be used as a basis for estimating the temporal duration of that branch (*t*) (Fig. 2). *n*, which is a product of the substitution rate (*r*) and *t*, can be inferred directly from molecular sequence data, and sampling more data will usually result in more accurate inferences of this parameter (Britton 2005). However, it is important to note that only *n* can be inferred directly from molecular sequence data. Divergence time estimation therefore requires important assumptions about *r* and *t* (Fig. 2) (Gillespie 1991; Sanderson 1997, 2002; Thorne et al. 1998; Britton 2005, Yang and Rannala 2006; Rannala and Yang 2007; dos Reis and Yang 2013; Zhu et al. 2015).

Ideally, assumptions would reflect current levels of uncertainty about *r* and *t*. As knowledge of *r* and *t* improves, assumptions can become more precise and accurate, leading to increasingly precise and accurate divergence time estimates (Fig. 1a).

Methodological developments in divergence time estimation have not followed this trajectory. An example of this is with the implementation of models for *r*. Early divergence time analyses used strict clock models, that assume *r* is the same on every branch (Langley and Fitch 1974; Britten 1984; Gillespie 1989, 1991; Bromham 2006), because they were considered to reflect the fact that *t* and the degree of divergence between a pair of sequences are correlated (Zuckerkandl and Pauling 1962, 1965; Margoliash 1963; Miyata et al. 1980). However, this assumption does not reflect uncertainty about the degree of variation in *r* for each branch. The subsequent development of relaxed clock models that allow *r* to vary among branches (Fig. 2) (Sanderson 1997, 2002; Thorne et al. 1998; Kishino et al. 2001; Drummond et al. 2006) better reflects this aspect of uncertainty. Therefore, in this case, methods have developed from a state of misleading certainty, when *r* was assumed to be the same on every branch, to increasing uncertainty about *r* for each branch (Fig. 1b).

Even though relaxed clock models accommodate more uncertainty about *r*, they are still underpinned by important assumptions. In correlated models, *r* is inherited between ancestral and descendant branches (Sanderson 1997, 2002; Thorne et al. 1998; Kishino et al. 2001), in uncorrelated models *r* for each branch is drawn from a parametric distribution (Drummond et al. 2006), whilst random local clocks assume discreet jumps in *r* that occur at a given probability (Drummond and Suchard 2011). Importantly, there is very rarely any direct evidence to support the implementation of any of these models in a given dataset, and studies which implement these models often do not include any explicit justification (eg. Jarvis et al. 2014; Magallón et al. 2015; Cardillo et al. 2017; Folk et al. 2019). As such, relaxed clock models often do not reflect the true level of uncertainty about *r* for each branch (Fig. 1b).

Likewise, for assumptions about *t* (Fig. 2), uncertainty about the relationship between fossil ages and clade ages is often such that it is difficult to meaningfully implement fossil calibrations (Carruthers and Scotland 2020). Therefore, methods often make misleadingly precise assumptions that do not reflect true levels of uncertainty (Carruthers and Scotland 2020). Additional uncertainty is also inherent in the phylogenetic placement of fossil calibrations, and such uncertainty is rarely incorporated into divergence time analyses that implement each fossil calibration at a single node. To further complicate matters, multiple fossil calibrations are often used as a means by which to constrain estimates of r derived from relaxed clock models, whilst in Bayesian methods a further underlying assumption is that t for each branch is controlled by a branching process model with constant diversification rates, an assumption that in no way reflects uncertainty about diversification rates (and by extension the expected t for each branch). Given that these methods for implementing assumptions about t and r can fail to adequately reflect uncertainty about t and r, they can lead to misleadingly precise divergence time estimates (Carruthers and Scotland 2020).

As well as implementing assumptions that do not reflect current levels of uncertainty, the nature of divergence time estimation, which is dependent on the combination of several assumptions, has led to a dramatic increase in the complexity of divergence time analyses. This complexity is likely to explain in part why methodological studies of divergence time estimation address highly specific methodological problems, that relate for example to the correct implementation of relaxed clock models across multiple loci, or the integration of fossil calibrations with branching process models (for example dos Reis et al 2014; Zhu et al. 2015; Barba-Montoya et al. 2017; Foster and Ho 2017; Angelis et al. 2018). By contrast, studies typically fail to address the fundamental and far simpler point, that understanding of r and t is very limited, and that this uncertainty about r and t will inevitably lead to uncertain divergence time estimates, regardless of how much data is analysed (Britton 2005; Zhu et al. 2015; Carruthers et al. 2019).

Many of the issues relating to the incorporation of uncertainty in divergence time estimation can be broadly considered in the context of the bias-variance trade-off, which describes how simple models tend to result in inferences that are less accurate but highly precise (eg. strict molecular clocks), and complex and parameter rich models tend to lead to inferences that are more accurate but less precise (eg. relaxed molecular clocks) (Wertheim et al. 2010). However, we stress that the unique properties of divergence time estimation, whereby *r* and *t* are not inferred directly from the sequence data, means that inferences are uniquely prone to the inherent biases of different methods, and that this can lead to unexpected results when considering findings in the context of the bias-variance trade-off. For example, dos Reis et al. (2014); Zhu et al. (2015); Foster and Ho (2017); Angelis et al. (2018) discuss the value of partitioning molecular clock models in multi-locus datasets (which inevitably increases the complexity of the model) and describe how it is useful because it accounts for variation in *r* among loci. This approach may well be useful where lineage-specific rates are weak, and counterintuitively it leads to more precise inferences (Zhu et al. 2015; Foster and Ho 2017). However, this approach cannot account for lineage-specific-rates, and in the presence of lineage-specific-rates it leads to misleadingly precise inferences that do not adequately account for uncertainty (Carruthers et al. 2019).

Implementing imperfect methods and accounting for uncertainty

It is apparent that methods used in divergence time estimation often fail to adequately account for uncertainty in *r* and *t*. An intuitive solution to this problem is to implement methods that make far less precise assumptions about *r* and *t*, so that they better reflect uncertainty. For example, fossil calibrations can be implemented that make far less precise assumptions about clade ages in relation to fossil ages, whilst relaxed clock models can be implemented that have considerably higher variances (Magallón and Sanderson 2005; Morris et al. 2018; Carruthers and Scotland 2020). Although this is likely to lead to very imprecise parameter estimates, performing analyses in this manner can provide a basis for adequately reflecting the true level of uncertainty underlying divergence time estimates.

Nonetheless, the characteristics of many methods of divergence time estimation mean that it is challenging to configure analyses such that they adequately account for uncertainty. For example, a constant rate branching process model is an essential component of most

Bayesian analyses, despite the fact that it is likely to make unrealistic assumptions about the expected *t* for each branch.

An alternative focus of recent discussion, especially in a Bayesian context, has been on methods of model selection, model fit, and model averaging (Suchard et al. 2001; Huelsenbeck et al. 2004; Beale et al. 2013, 2014 Brown 2014a, b; Brown and Thomson 2018). Model selection and the assessment of model fit are both methods for determining the suitability of a model for a dataset, either by comparing the marginal likelihood among a set of models (model selection), or determining whether a model can be used as a basis for simulating data that is consistent with the dataset being analysed (model fit). By contrast, model averaging recognises that no single model is appropriate for a dataset, and therefore derives parameter estimates according to average values across several models. All three methods may provide a powerful means for avoiding the use of any single model that is inconsistent with a given dataset, and which does not therefore adequately account for uncertainty. However, the incorporation of these methods into divergence time analyses remains the exception rather than the rule (eg. Jarvis et al. 2014; Magallón et al. 2015; Cardillo et al. 2017; Folk et al. 2019), and they are often too computationally intensive to incorporate into large analyses (eg. Barba-Montoya et al. 2018). Further exploration of these methods is therefore required, alongside the development of computational resources that enable them to be implemented in larger datasets.

However, these approaches also present important new challenges. First, it is likely that the central problem in divergence time estimation, that *r* and *t* are not inferred directly from the sequence data, will undermine comparisons between models. This is because it necessarily means that methods for comparing models of *r* are dependent on underlying assumptions about *t*, whilst methods for comparing models of *t* (branching process models) are dependent on underlying assumptions about *r*. For example, in their analysis of different branching process models for estimating divergence times in *Cycas*, Condamine et al. (2015) compared different branching process models in the context of a single model for *r*. In short, in the context of *r* and *t*, the selection of an appropriate model for one parameter is dependent on the assumptions that are implemented for the other parameter. Louca and Pennel (2020) also highlight the problem of model comparison when dealing with non-identifiable parameters, in this instance in the context of diversification rate estimation.

In contrast to model selection and model fit, model averaging has the advantage of incorporating uncertainty from several models. However, it necessarily incorporates a limited number of models, which at times may have contradictory and biologically unrealistic assumptions. For example, Betts et al. (2018) used a model averaging approach for a time-calibrated phylogeny of eukarya based on a range of different models for *r* and *t*. The divergence time estimates derived from this model averaging approach may better account for uncertainty compared to individual models. However, it is challenging to quantify how a necessarily limited number of contradictory and potentially unrealistic models adequately account for uncertainty.

An alternative option that may enable uncertainty in divergence time estimates to be expressed more effectively, is to extract parts of datasets that are likely to conform to the simplistic assumptions of many methods. Early versions of this approach involved pruning lineages that violate the assumption of the strict molecular clock (Li and Tanimura 1987; Takezki et al. 1995). However, this will often lead to the removal of a large number of branches, and determining whether shifts in *r* have actually occurred is in itself fundamentally problematic (Fitch 1976; Sanderson 1997; Lanfear et al. 2013). Alternatively, Near and Sanderson (2004) and Near et al. (2005) developed a cross-validation method for removing fossil calibrations that have an excessively large impact on divergence time estimates, and which may therefore be a source of error and lead to misleading inferences of uncertainty. However, it is unclear whether selecting fossil calibrations in this manner actually provides the most reliable temporal signal for divergence time analyses (Carruthers and Scotland 2020), and there is still a problem of accounting for uncertainty in the remaining fossil calibrations. Further, in many divergence time analyses, especially in plants, the very small number of fossil calibrations that are actually available will make it difficult to implement methods that focus on identifying suitable subsets of fossil calibrations.

Genomic scale datasets provide a new opportunity for selecting subsets of data that are congruent with simplistic methodological assumptions. The best example of this is “gene shopping” presented by Smith et al. (2018), who devised a protocol for selecting loci that conform to a strict molecular clock, and then incorporating only these loci in divergence time analyses. Ideally, divergence time estimates derived from these analyses would not be affected by variation in *r*, meaning that accounting for uncertainty becomes more straightforward. However, detecting the presence or absence of shifts in *r* is fundamentally difficult, and this method is of little value to the many divergence time analyses that are based on a limited number of markers (eg. Condamine et al. 2015; Magallón et al. 2015; Lagomarsino et al. 2016; Ramírez-Barahona et al. 2020).

INTERPRETING TIME-CALIBRATED PHYLOGENIES TO REDUCE UNCERTAINTY ABOUT MACROEVOLUTION

The points above highlight that despite important and novel methodological developments, there are significant limitations in how methods of divergence time estimation account for uncertainty. This has important implications for how divergence time estimates should be interpreted if they are to be used as a basis for reducing uncertainty about macroevolution.

Using time-calibrated phylogenies to test specific macroevolutionary hypotheses

Often, time-calibrated phylogenies are used as a basis for testing explicit macroevolutionary hypotheses, with the falsification of a hypothesis acting as a basis for reducing uncertainty about that aspect of macroevolution (Lavin et al. 2004; Crisp et al. 2011; Muñoz-Rodríguez et al. 2018). In this instance, the conclusions that are derived from time-calibrated phylogenies are less sensitive to the methodological shortcomings inherent in many methods. For example, Lavin et al. (2004) tested the hypothesis that disjunct distributions within legume clades resulted from continental break-up. They were able to reject this hypothesis and support the alternative hypothesis - that disjunct distributions resulted from long-distance-dispersal - by showing that divergence times for disjunct taxa significantly post-date the break-up of Gondwana. Even though there may be methodological limitations in this study such that uncertainty is not adequately accounted for, the testing of such a specific hypothesis makes it possible to evaluate the implications of alternative divergence time estimates on their biological conclusion. Alternatively, Muñoz-Rodríguez et al. (2018) were able to test the hypothesis of whether the storage root of sweet potato evolved in pre-human times by biasing divergence time estimates to as young ages as possible. The authors explicitly recognise that the inferences from their analyses are likely to significantly underestimate the true age of the sweet potato (Muñoz-Rodríguez et al. 2018). However, this issue is irrelevant with respect to the conclusions of the study, because the purpose of the study was not to estimate the age of sweet potato, but rather to determine whether the sweet potato evolved in pre-human times.

Using time calibrated phylogenies to investigate general questions about when and why clades evolved

Alternatively, the aim of a time-calibrated phylogeny may be to estimate the age of a clade, or to provide a framework for re-constructing the macroevolutionary history of a clade (for example Magallón et al. 2015; Hohmann et al. 2015; Berger et al. 2016; Lagomarsino et al. 2016; Zhang et al. 2018). Many ground-breaking time-calibrated phylogenies have been constructed in this context that have provided important biological and methodological insights, and stimulated valuable discussion. However, when time-calibrated phylogenies are framed in this way, accurately reflecting the true degree of uncertainty in parameter estimates becomes increasingly important, otherwise the findings that are presented by the study are misleading. For example, the confidence interval that is presented for the age of a clade may not actually contain the true age of the clade.

An example of this is the angiosperm time-calibrated phylogeny presented by Magallón et al. (2015). Although the authors were clearly aware of the important assumptions underlying their approach, and provide a very useful discussion of this, their time-calibrated phylogeny has often been interpreted as a roadmap for the evolution of major angiosperm clades, and been used as a basis for estimating divergence times in hundreds of subsequent studies. This is despite the fact that this study considers a very limited set of assumptions, including one important assumption that the crown node of angiosperms must be younger than of 136 Myr. The nature of the assumptions that are implemented means that this study is unlikely to represent the true degree of uncertainty about the age of angiosperms, or major clades within angiosperms.

Nonetheless, this study can be considered in the context of the more specific hypothesis of whether molecular divergence time estimates for major angiosperm clades can be constrained such that they are congruent with the fossil record. In this case, the study shows that when there is a large number of fossil calibrations combined with an unusually young maximum constraint for the crown node of angiosperms, molecular divergence time estimates can be constrained such that they conform very closely to the fossil record. The manner by which a time-calibrated phylogeny is interpreted therefore plays a crucial role in how robust conclusions can be drawn from it. This reiterates our more general point, that when testing specific hypotheses, the conclusions that are drawn from time-calibrated phylogenies are less sensitive to the methodological limitations of a given method. This is because you are simply asking the question of whether, or in what context, the data may support or reject a given hypothesis. By contrast, attempts to adequately account for all sources of uncertainty, which can be underpinned by processes of which we have limited understanding, are more likely to be undermined by limitations in how methods account for uncertainty.

Underlying issues relating to the interpretation of time-calibrated phylogenies is the fact that users of time-calibrated phylogenies are often not fully aware of the fundamental problems in methods of divergence time estimation. This is not surprising given that these methods have become highly complex, and methodological discourse involves highly specific discussions that are too dense for many readers (Yang and Rannala 2006; Rannala and Yang 2007; dos Reis and Yang 2013; Zhu et al. 2015; Barba-Montoya et al. 2017). Often, attempts by methodologists to communicate to a wider audience involve the declaration of “best practice” approaches that users of time-calibrated phylogenies are expected to follow (Parham et al. 2012). By contrast, a more widespread engagement with fundamental methodological issues may provide a stronger basis for users of time-calibrated phylogenies to evaluate the extent to which their time-calibrated phylogeny enables them to reduce uncertainty about a given aspect of macroevolutionary history.

CONCLUSION

It is probably impossible for any field to develop in a way which moves smoothly to a state of less uncertainty, as set out in Figure 1a. False premises and methodological conservatism are likely to lead to inferences that are directly contradicted by subsequent findings (Fig. 1b). However, with divergence time estimation, we are now in the position where we know that uncertainty stems directly from uncertainty about r and t. Further, it is now clear that the effects of this uncertainty are not ameliorated by complex analyses that pool multiple different sources of evidence, and which make complex but unreasonable assumptions that do not reflect existing knowledge (Carruthers and Scotland 2020). Because of this, it is both possible and necessary for future methodological developments in divergence time estimation to adequately account for uncertainty, and to provide a basis for moving closer to the state represented in Figure 1a. These methodological developments may lead to less precise divergence time estimates because they are better reflecting the true extent of uncertainty. Alternatively, it may also be the case that at some point in the future new methods produce more precise estimates as well as better reflecting uncertainty, if for example, better models are developed that account for the relationship between life history traits and *r* (Lartillot and Poujol 2011; Lanfear et al. 2013; Ho 2014; Berv and Field 2018). Regardless, better reflecting the true extent of uncertainty will result in estimates that provide a robust foundation for subsequent research to build upon, and will not be directly contradicted by subsequent findings.

ACKNOWLEDGEMENTS

We thank Jacob Berv and one anonymous reviewer for their thorough and insightful reviews of this manuscript.

REFERENCES

Angelis K., Alvarez-Carretero S., dos Reis M., Yang Z. 2018. An evaluation of different partition strategies for Bayesian estimation of species divergence times. *Syst. Biol.* 67:61-77.

Barba-Montoya J., dos Reis M., Yang Z. 2017. Comparison of different strategies for using fossil calibrations to generate the time prior in Bayesian molecular clock dating. *Mol. Phyolgenetics. Evol.* 114:386-400.

Baele G., Lemey P., Bedford T., Rambaut A., Suchard M. A., Alekseyenko A.V. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* 29:2157–2167.

Baele G., Li W. L. S., Drummond A. J., Suchard M. A., Lemey P. 2013. Accurate model selection of relaxed Molecular clocks in bayesian Phylogenetics. *Mol. Biol. Evol.* 30:239–243.

Berger B. A., Kriebel R., Spalink D., Sytsma K. J. 2016. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Mol. Phylogenetics Evol*. 95:116-136.

Berv J. S. and Field D. J. 2018. Genomic signature of an avian Lilliput effect across the K-Pg extinction. *Syst Biol.* 67:1-13.

Betts H. C., Puttick M. N., Clark J. W., Williams T. A., Donoghue P. C. J., Pisani D. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* 2:1556-1562.

Britten RJ. 1984. Rates of DNA sequence evolution differ between taxonomic groups. *Science*. 39:1393–1398.

Britton T. 2005. Estimating divergence times in phylogenetic trees without a molecular clock. *Syst. Biol*. 54:500–507.

Bromham L. 2006. Molecular dates for the Cambrian explosion: is the light at the end of the tunnel an oncoming train? *Palaeontol Electron.* 9:2004–2006.

Bromham L, Duchene S., Hua X., Ritchie A. M., Duchene D. A., Ho S. Y. W. 2018. Bayesian molecular dating: opening up the black box. *Biol. Rev*. 93:1165-1191.

Brown J. M. 2014. Predictive approaches to assessing the fit of evolutionary models. *Syst. Biol.* 63:289–292.

Brown J. M. 2014. Detection of implausible phylogenetic inferences using posterior predictive assessment of model fit. *Syst. Biol.* 63:334–348.

Brown J. W and Smith S. A. 2018. The past sure is tense: on interpreting phylogenetic divergence time estimates. *Syst. Biol*. 67:340-353.

Cardillo M, Weston P. H., Reynolds Z. K. M., Olde P. M., Mast A. R., Lemmon E. M., Lemmon A. R., Bromham L. 2017. The phylogeny and biogeography of *Hakea* (Proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution.* 71:1928-1943.

Carruthers T, Sanderson MJ, Scotland RW. 2019. The implications of lineage-specific rates for divergence time estimation. *Syst Biol. In press*.

Carruthers T. and Scotland R. W. 2020. Insights from empirical analyses and simulations on using multiple fossil calibrations with relaxed clocks to estimate divergence times. *Mol Biol Evol*. 37:1508-1529.

Condamine F. L., Nagalingum N. S., Marshall C. R., Morlon H. 2015. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evol. Biol.* 15:65.

Crisp M. D., Trewick S. A., Cook L. G. 2011. Hypothesis testing in biogeography. *Trends Ecol. Evol.* 26:66-72.

Drummond A. J., Ho S. Y. W., Phillips M. J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLOS Biol.* 4:e88.

Drummond A. J. and Suchard M. A. 2011. Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 8:114.

Fischhoff B. and Davis L. A. 2014. Communicating scientific uncertainty. *Proc. Natl. Acad. Sci. USA.* 111:13664-13671.

Fitch W. M. 1976. Molecular evolutionary clocks. Pp. 160-178 *in* F. J. Ayala, ed. Molecular evolution. Sinauer, Sunderland, Mass.

Folk R. A., Stubbs R. L., Mort M. E., Cellinese N., Allen J., Soltis P. S., Soltis D. E., Guralnick R. P. 2019. Rates of niche and phenotype evolution lag behind diversification in a temperate radiation. *Proc. Natl. Acad. Sci. USA.* 116:10874-10882.

Foster C. S. P. and Ho S. Y. W. 2017. Strategies for partitioning clock models in phylogenomic dating: application to the angiosperm evolutionary timescale. *Genome Biol. Evol.* 9:2752-2763.

Gillespie J. H. 1989. Lineage effects and the index of dispersion of molecular evolution. *Mol Biol Evol.* 6:636-647.

Gillespie J. H. 1991. The causes of molecular evolution. Oxford: Oxford University Press.

Ho S.Y.W. 2014. The changing face of the molecular evolutionary clock. *Trends Ecol Evol.* 29:496-503.

Hohmann N., Wolf E. M., Lysak M. A., Koch M. A. 2015. A time-calibrated road map of Brassicaceae species evolutionary history. *Plant Cell.* 27:2770-2784.

Hoorn C., Wesselingh P., ter Steege H., Bermudez M. A., Mora A., Sevink J., Sanmartín I., Sanchez-Meseguer A., Anderson C. L., Figueiredo J. P., Jaramillo C, Riff D., Negri F. R., Hooghiemstra H., Lundberg J., Stadler T., Särkinen T., Antonelli A. 2010. Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. *Science.* 330:927-931.

Huelsenbeck J. P., Larget B., Alfaro M. E. 2004. Bayesian phylogenetic model selection using reversible jump Markov Chain Monte Carlo. *Mol. Biol. Evol.* 21:1123.

Hughes CE. and Eastwood R. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proc. Natl. Acad. Sci. USA.* 103:10334-10339.

Hughes C. E., Pennington R. T., Antonelli A. 2013. Neotropical plant evolution: assembling the big picture. *Bot. J. Linn. Soc.* 171:1-18.

Jarvis E. D. et al. 2014. Whole genome analyses resolve early branches in the tree of life of modern birds. *Science.* 346:1320-1331.

Kishino H, Thorne JL, Bruno WJ. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Mol Biol Evol.* 18:352–361.

Lagomarsino L. P., Condamine F. L., Antonelli A., Mulch A., Davis C. C. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New. Phytol.* 210:1430-1442.

Lanfear R, Ho SYW, Davies TJ, Moles AT, Aarssen L, Swenson NG, Warman L, Zanne AE, Allen AP. 2013. Taller plants have lower rates of molecular evolution. *Nat Commun.* 4:1879.

Langley CH and Fitch WM. 1974. An examination of the constancy of the rate of molecular evolution. *J Mol Evol.* 3:161-177.

Lartillot N and Poujol R. 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol. Biol. Evol.* 28:729-744.

Lavin M., Schrire B. P., Lewis G., Pennington R. T., Delgado-Salinas A., Thulin M., Hughes C. E., Matos A. B., Wojciechowski M. F. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Philos. Trans. Royal. Soc.* 359:1509-1522.

Li W. H. and Tanimura M. 1987. The molecular clock runs more slowly in man than in apes and monkeys. *Nature.* 326:93-96.

Louca S and Pennell M. W. 2020. Extant timetrees are consistent with a myriad of diversification histories. *Nature*. 580:502-505.

Magallón S., Gomez-Acevedo S., Sanchez-Reyes L. L., Hernandez-Hernandez T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytol.* 207:437-453.

Magallón S. and Sanderson M. J. 2005. Angiosperm divergence times: the effect of genes, codon positions, and time constraints. *Evolution*. 59:1653-1670.

Morris J. L., Puttick M. N., Clark J. W., Edwards D., Kenrick P., Pressel S., Wellman C. H., Yang Z., Schneider H., Donoghue P. C. J. 2018. The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. USA.* 115:E2274-E2283.

Muñoz-Rodríguez P., Carruthers T., Wood J. R. I., Williams B. R. M., Weitemier K., Kronmiller B., Ellis D., Anglin N. L., Longway L., Harris S. A., Rausher M. D., Kelly S., Liston A., Scotland R. W. 2018. Reconciling conflicting phylogenies in the origin of sweet potato and dispersal to polynesia. *Curr Biol*. 28:1246-1256.

Near TJ, Meylan PA, Shaffer HB. 2005. Assesing concordance of fossil calibration points in molecular clock studies: an example using turtles. *Am Nat.* 165:137-146.

Near TJ and Sanderson MJ. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Phil Trans R Soc Lond B.* 359:1477-1483.

Parham J. F., Donoghue P. C. J., Bell C. J., Calway T. D., Head J. J., Holroyd P. A., Inoue J. G., Irmis R. B., Joyce W. G., Ksepka D. T., Patané J. S. L., Smith N. D., Tarver J. E., van Tuinen M., Yang Z., Angielczyk K. D., Greenwood J. M., Hipsley C. A., Jacobs L., Makovicky P. J., Müller J., Smith K. T., Theodor J. M., Warnock R. C. M., Benton M. J. 2012. Best practices for justifying fossil calibrations. *Syst. Biol.* 61:346-359.

Perrigo A., Hoorn C., Antonelli. 2020. Why mountains matter for biodiversity. *J. Biogeogr.* 47:315-325.

Rahbek C., Borregaard M. K., Antonelli A., Colwell R. K., Holt B. G., Nogues-Bravo D., Christian M. Ø. Rasmussen C. M. Ø., Richardson K., Rosing M. T., Whittaker R. J., Fjeldså J. 2019. Building mountain biodiversity: Geological and evolutionary processes. *Science.* 365:1114-1119.

Ramírez-Barahona S., Sauquet H., Magallón S. 2020. The delayed and geographically heterogeneous diversification of flowering plant families. *Nat Ecol Evol.* 4:1232-1238.

Rannala B. and Yang Z. 2007. Inferring speciation times under and episodic molecular clock. *Syst. Biol.* 56:453-466.

dos Reis M. and Yang Z. 2013. The unbearable uncertainty of Bayesian divergence time estimation. *J Syst Evol.* 51:30-43.

dos Reis M., Zhu T., Yang Z. 2014. The impact of the rate prior on Bayesian estimation of divergence times with multiple loci. *Syst. Biol.* 63:555–565.

Sanderson M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* 14:1218–1231.

Sanderson M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19:101–109.

Suchard M. A., Weiss R. E., Sinsheimer J. S. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* 18:1001–1013.

Takezaki N., Rzhetsky A., Nei M. 1995. Phylogenetic clock and linearized trees. Mol. Biol. Evol. 12:823-833.

Thorne J. L., Kishino H., Painter I. S. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol.* 15:1647–1657.

Wertheim J. O., Sanderson M. J., Worobey M., Bjork A. 2010. Relaxed molecular clocks, the bias-variance trade-off, and the quality of phylogenetic inference. *Syst. Biol.* 59: 1-8.

Wilf P., Carvalho M. R., Gandalfo M. A., Ruben Cuneo N. 2017. Eocene lantern fruits from Gondwanan Patagonia and the early origins of Solanaceae. *Science.* 355:71–75.
Wilf P. and Escapa I. H. 2015. Green Web or megabiased clock? Plant fossils from Gondwanan Patagonia speak on evolutionary radiations. *New Phytol.* 207:283–290.

Yang Z. and Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23:212–226.

Zhang S., Che L., Li Y., Liang D., Pang H., Slipinski A., Zhang P. 2018. Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nat. Commun.* 9:205.

Zhu T., dos Reis M., Yang Z. 2015. Characterization of the uncertainty of divergence time estimation under relaxed molecular clock models using multiple loci. *Syst. Biol.* 64:267–280.

CAPTIONS

Figure 1. The perceived change in uncertainty in recent studies compared to older studies. The grey shaded area indicates uncertainty about the parameter value, the grey horizontal dotted-line indicates the correct parameter value. In **a)** uncertainty is adequately expressed and reduces in more recent studies, such that statements become more precise and accurate. In **b)** there is a misleading perception of uncertainty such that statements are overly precise and inaccurate. Perceived uncertainty in older studies is directly contradicted by perceived uncertainty in more recent studies, and perceived uncertainty does not incorporate the correct parameter value.

Figure 2. A summary of the different types of data and assumptions that are required when inferring divergence times. n = number of substitutions, r = substitution rate, t = time.