

Commentary

Back to the past: a new take on the timing of flowering plant diversification

‘In the case of plants, an adequate fossil record does not exist . . .’

(Boulter *et al.*, 1972)

‘Molecular clocks are like Santa Claus: everyone wants to believe in them, but no one really does.’

(H. Brad Shaffer, pers. comm.)

The second most interesting thing about Magallón *et al.*'s new analysis of the timing of angiosperm diversification in this issue of *New Phytologist* (pp. 437–453) is their compendium of molecular clock estimates for the age of angiosperms (their fig. 1). Since 1972, when Ramshaw and colleagues (in this journal) used some of the first amino acid sequence data from plants to estimate an age of angiosperms at 300–400 million years ago (Ma), molecular-based estimates have ranged wildly. Also since then many things have happened: genomics, phylogenetic biology, statistically sophisticated inference of divergence times, and a much better understanding of the early angiosperm fossil record, to name the most relevant. Why then are estimates of the ‘crown group’ age of angiosperms (see Fig. 1) typically 50–150 million years older than the earliest angiosperm fossils from the Early Cretaceous (*c.* 135 Ma); and why have the error bars on estimated ages not gotten much narrower?

‘Is a 5-million-year confidence interval on the crown age of angiosperms too good to be true?’

Mentioning possible explanations will delay getting to the most interesting aspect of Magallón *et al.*'s new paper, but it seems necessary. The Ramshaw *et al.* (1972) study exemplifies one essential problem – the assumption of a ‘molecular clock’; that is, a constant rate of molecular evolution extending from their fossil calibration point, at the split between birds and mammals, back through early eukaryotic history, then returning forward to flowering plants (that is something like three billion years of clock-like evolution). One thing the decades of molecular phylogenetic reconstruction since 1972 have made clear is that constant rates of evolution across a tree are rare in real DNA sequences at any time depth, much less across most eukaryotes. We found some of the strongest evidence for this in the angiosperm case where it could be seen in two-fold differences in ages estimated by sampling

fast-evolving ‘grocery store plants’ (rice, pea), which put the angiosperms back in the Permian, versus more slowly evolving long lived perennial plants (Sanderson & Doyle, 2001). This is not that surprising in hindsight. As Gillespie (1991, p. 140) put it, after spending a career modeling molecular evolution, ‘The molecular clock – in its strict incarnation as a constant rate across all lineages – cannot be derived from a mechanistic model of evolution . . . the clock does not exist.’ Undaunted by this, a number of workers developed ‘relaxed clock’ methods aiming to infer divergence times from less clocklike sequence data (full disclosure: my primitive nonparametric and semi-parametric methods were among them: Sanderson, 2002). Sadly, it is hard to see the persistence of very different estimates of angiosperm age cataloged in Magallón *et al.*'s paper as a triumph for these methods, despite their increasing sophistication in the last decade.

Apologias are emerging. Some blame the fossils (Rannala & Yang, 2007; Dos Reis & Yang, 2013), pointing to an intrinsic lower bound on the error of molecular dating because of errors in the dating or the phylogenetic placement of fossils used to calibrate the analysis. The flip side of this has received less notice, but it may be even more problematic. Even if fossil calibrations are known precisely, there is another lower bound on error of age estimates attributable to idiosyncratic departures of rates from a molecular clock (Britton, 2005). The escape from that problem is to hope that rates of different genes vary across the tree in uncorrelated ways, so that in some sense there exists a genome-scale ‘über-clock’ with individual patterns of rate variation canceling out (Britton, 2005; Rannala & Yang, 2007). Unfortunately, evidence from recent broad comparative genomics work raises the concern that genomes act more conspiratorially than that. In a survey of 1000+ genes from transcriptomes across the angiosperm clade Caryophyllales (spinach, pinks, cacti, etc., Yang *et al.*, 2015), there was a consistent difference in rate across genes depending on whether the plants were woody or herbaceous.

Possibly in reaction to these troubles with clocks, Magallón *et al.* have taken a now quite unusual tack. They use the fossil record to estimate the crown age of the angiosperms – then use molecular data to date the major radiations *within* it. To me, this is the most interesting aspect of their paper, partly because it so flies in the face of the past three decades of molecular approaches, and also because it challenges a fairly entrenched view that the origin of angiosperms is too shrouded in mystery to provide a useful calibration point itself. However, this ‘abominable mystery’ that so troubled Darwin (Friedman, 2009) was due in large part to a paucity of fossils at the earliest phases of angiosperm diversification, a situation that has greatly improved since the 1960s (Doyle, 2012). Fossil evidence from pollen, flowers and leaves clearly implies a rapid (but not instantaneous) temporal sequence of increasing trait diversity and complexity, much more in line with a rapid diversification from one common ancestor in

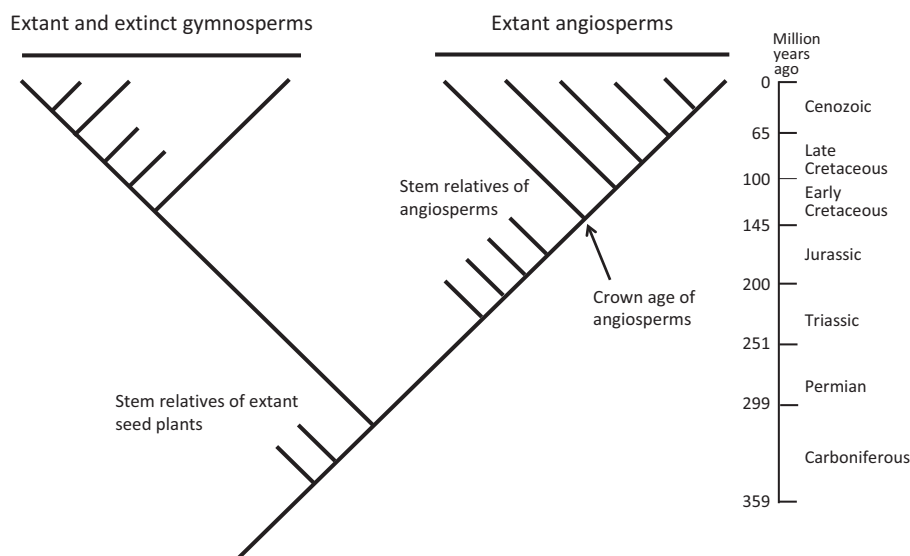


Fig. 1 Schematic phylogenetic tree of extant and extinct seed plants showing position of crown group node of angiosperms, stem relatives, and geologic timescale discussed in the text (highly modified from Doyle, 2012).

the Early Cretaceous or Late Jurassic than from many independent lines stemming from a much older common ancestor. (Interestingly, Ramshaw *et al.* (1972) took this latter view to such an extreme that they argued angiosperms might even be polyphyletic, something we know today to be wrong.)

Using statistics invented by Marshall (2008), Magallón *et al.* estimate the true age of the most recent common ancestor of living angiosperms to be within a brief 99% confidence interval from 136 to 141 Ma, the younger date corresponding to the oldest angiosperm fossils, and the older date estimated from Marshall's method. The surprisingly narrow interval follows from the authors' application of 137 fossil calibration points within the angiosperm tree. Under Marshall's model of uniform random fossilization, this exceptionally large number makes the probability of undiscovered angiosperm fossils existing that are much older than 136 Ma very small – essentially because the large number of angiosperm calibrations implies the discovery rate is high, which is incompatible with a long pre-Cretaceous interval with no fossils. Armed with this result, they then estimate timing of origin of clades within flowering plants. Using taxonomic families as a proxy of morphological and functional distinctiveness, they go on to conclude that a large fraction of angiosperm trait diversity originated in the Late Cretaceous, tapering off by the early Tertiary, although the exceptional species richness that is a trademark of angiosperms may have come much later.

Is a five-million-year confidence interval on the crown age of angiosperms too good to be true? Marshall (2008) cautioned that if preservation rates of fossils in the era before the oldest fossil in a clade were significantly reduced – the assumption of uniform sampling violated – the true confidence interval could extend much deeper into the past. This helps to reconcile the clocks versus rocks controversy over the age of primates (Tavaré *et al.*, 2002), for example. However, there are many fossil occurrences of other seed plant groups, and other plants in general in the Jurassic, and in the absence of a quantitative assessment of preservation rates, this does not seem to provide a resounding reason to worry that Magallón *et al.*'s confidence intervals are far off.

That said, what *is* missing from the fossil record is something much more particular: a diverse grade of seed plants closely related to angiosperms, having some but not all of angiosperms' key innovations, well-preserved in the Jurassic fossil record. Based on recent phylogenetic analyses that include both fossils and living forms, there are a few candidates for these 'stem relatives' (Fig. 1) such as *Caytonia*, Bennettitales, and others, but their phylogenetic position is still controversial largely because they all exhibit quite a few unique traits not shared with angiosperms (Doyle, 2012). Could additional stem relatives, or even early true angiosperms, have been hiding out in relatively restricted regions in the Jurassic in environments with poor preservation potential and at relatively low species diversity, just waiting for their chance, an idea championed recently by Feild *et al.* (2004; see also Doyle, 2012)? Maybe, but this kind of 'crypto-depauperon' hypothesis (apologies for the new jargon: see Donoghue & Sanderson, 2015) has a serious weakness: the fate of most low-species-diversity clades in the history of life is, well, extinction, similar to the fate of most new mutations in populations (Strathmann & Slatkin, 1983).

Postulating such crypto-depauperons may not really be necessary. After all, the origin of angiosperms is hardly the only abominable mystery in plant evolution. A number of much more recent clades of angiosperms could be interpreted as mini-abominations: that is, they have poor fossil records at their base, novel innovations with unclear transitional forms among related taxa, and, were it not for recent molecular phylogenies, uncertain phylogenetic relationships: think cacti or legumes, for example. Do all such clades require a complex scenario for their origin, or are there just some rapid diversification events preserved in the fossil record with little trace of their early history? Magallón *et al.*'s paper embraces an oddly remarkable but increasingly sensible idea: maybe the earliest Cretaceous angiosperm fossils *were* the first angiosperms, or near enough to them in time for it not to matter. If so, the new timescale for angiosperm diversification inferred by Magallón *et al.* has exciting implications, including many not discussed in the paper. It will be interesting to see, for example, whether new studies of molecular

evolutionary rates in angiosperms will take these results to heart, since the impact of these new dates on reconstructions of the tempo of evolution will likely be dramatic.

Michael J. Sanderson

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA
(tel +1 520 626 6848;
email sanderm@email.arizona.edu)

References

- Boulter D, Ramshaw JAM, Thompson EW, Richardson M, Brown RH. 1972. A phylogeny of higher-plants based on amino-acid sequences of cytochrome-c and its biological implications. *Proceedings of the Royal Society Series B: Biological Sciences* 181: 441–455.
- Britton T. 2005. Estimating divergence times in phylogenetic trees without a molecular clock. *Systematic Biology* 54: 500–507.
- Dos Reis M, Yang ZH. 2013. The unbearable uncertainty of Bayesian divergence time estimation. *Journal of Systematics and Evolution* 51: 30–43.
- Donoghue MJ, Sanderson MJ. 2015. Confluence, synnovation, and depauperons in plant diversification. *New Phytologist* 270: 260–274.
- Doyle JA. 2012. Molecular and fossil evidence on the origin of angiosperms. *Annual Review of Earth and Planetary Sciences* 40: 301–326.
- Feild TS, Arens NC, Doyle JA, Dawson TE, Donoghue MJ. 2004. Dark and disturbed: a new image of early angiosperm ecology. *Paleobiology* 30: 82–107.
- Friedman WE. 2009. The meaning of Darwin's "abominable mystery". *American Journal of Botany* 96: 5–21.
- Gillespie JH. 1991. *The causes of molecular evolution*. New York, NY, USA: Oxford University Press.
- Magallón S, Gómez-Acevedo S, Sánchez-Reyes LL, Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* 207: 437–453.
- Marshall CR. 2008. A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *American Naturalist* 171: 726–742.
- Ramshaw JAM, Richardson DL, Meatyard BT, Brown RH, Richardson M, Thompson EW, Boulter D. 1972. Time of origin of the flowering plants determined by using amino-acid sequence data of cytochrome C. *New Phytologist* 71: 773–779.
- Rannala B, Yang ZH. 2007. Inferring speciation times under an episodic molecular clock. *Systematic Biology* 56: 453–466.
- Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19: 101–109.
- Sanderson MJ, Doyle JA. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from rbcL and 18S rDNA data. *American Journal of Botany* 88: 1499–1516.
- Strathmann RR, Slatkin M. 1983. The improbability of animal phyla with few species. *Paleobiology* 9: 97–106.
- Tavaré S, Marshall CR, Will O, Soligo C, Martin RD. 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature* 416: 726–729.
- Yang Y, Moore MJ, Brockington SF, Soltis DE, Wong GK-S, Carpenter EJ, Zhang Y, Chen L, Yan Z, Xie Y *et al.* 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Molecular Biology and Evolution*. doi: 10.1093/molbev/msv081.

Key words: angiosperm diversification, crown age of angiosperms, Early Cretaceous, evolution, fossil record, molecular clock.



About New Phytologist

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**