**Methods used for analysis of plant evolution**

Estimation of divergence times of clades

Phylogeny and absolute divergence times of clades were estimated using a Bayesian approach as implemented in software BEAST v. 1.8 (Drummond et al., 2012). The analyses were set up as described below, using software BEAUti v. 1.8 (Drummond et al., 2012). Two identical analyses were run, each for 100 million generations, with parameters logged every 10 000 generations. In addition a third analysis was run, identical to the other two but with sampling from priors only. Values sampled for different parameters and convergence of runs were examined and assessed in Tracer v. 1.5 (Rambaut and Drummond, 2007). Results were summarized using LogCombiner v. 1.8 and Tree Annotator v. 1.8 (Drummond et al., 2012). Graphic output was produced in FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

Substitution models. The data set was partitioned into three partitions corresponding to the three protein-coding genes used (*rbcL, atpB, rps4*). Substitution models were selected for each gene region according to the corrected Akaike information criterion and the Bayesian information criterion as estimated using the Perl script MrAIC v. 1.4 (Nylander, 2004) in in combination with PHYML v. 3.0 (Guindon et al., 2010). The general time reversible model (GTR), with gamma distributed rates (four categories and shape parameter estimated from the data) and a proportion of invariable sites, was estimated as best-fitting model for all three genes. The data partitions were unlinked and base frequencies were estimated separately for each data partition.

Clock models. The assumption of a strict molecular clock (i.e., a constant substitution rate) was clearly violated for the present data set, as shown by preliminary analyses using a relaxed clock with uncorrelated rates drawn from a lognormal distribution. While the prior settings of these runs put considerable weight on constant substitution rates, the marginal densities for the clock parameter ucld.stdev sampled during the analyses differed clearly from those sampled from the prior alone, and do not have any significant density at zero. A strict molecular clock model is thus inappropriate for our data. Initially we considered several options for how to relax the molecular clock and allow for among-lineage variation in substation rate. A random local clock model assumes the same rate within discrete regions of the tree (but allows these regions to diﬀer in rates), and recent methodological advances allows for integration over the uncertainty of number of rate shifts (number of local clocks) appropriate for the data at hand (Drummond and Suchard, 2010). However, implementation of random local clocks in a Bayesian framework are reported complicated (Drummond and Bouckaert, 2014), and our analytical outputs were unsatisfactory. The concept of auto-correlated clock rates was not considered for this data set of distantly related taxa; instead we selected uncorrelated clock rates, in which substitution rates of branches are allowed to vary by being independently drawn from an underlying known distribution. Exponential and lognormal distributions were tested for this purpose and yielded similar result but effective sample size (ESS) values for the clock parameters, the tree likelihood and the posterior were higher when using the exponential distribution. Thus, for our final analyses we employed a relaxed clock model, with uncorrelated substation rates drawn from an exponential distribution. The clock model was unlinked; i.e., the parameters of the clock model were estimated for each gene region separately. Exponential prior distribution was specified for the clock parameter (uced.mean: initial value 5, mean5, offset 0).

Tree priors. As tree prior for the present set of data, with samples from the large and ancient land plant clade, we selected a mechanistic birth-death model appropriate for data sets with incomplete sampling (Stadler, 2009). It models speciation rates as well as extinction rates. This tree prior was selected for all data partitions. Partitions were linked, and parameters of the tree prior were thus estimated for the entire data set. A bifurcating topology created from a parsimony analysis (see above) was used as staring tree.

Calibration to absolute time. The following clades were defined as monophyletic and prior age distributions were assigned with the aim to as best as possible describe the prior probabilities of age of clades based on fossil documentation of respective groups. All geological ages are taken from the most recent GSA Geological Time Scale (Gradstein et al., 2012).

*Land plants.* A lognormal prior distribution was specified for the root height, with an offset of 467 Ma (million years before present) and a mean in real space of 15 Ma (log stdev 0.5), based on the earliest evidence of land plants, “cryptospores” from the mid-Ordovician (Llanvirn) (Taylor et al., 2009). The clade to which these fossils belong must be older than the fossil, however, it appears likely that spores of land plants would have been found in older strata if present, and we therefore prefer a relatively narrow prior distribution. Our settings result in a prior distribution in which the median age of the land plant clade is 480 Ma, and 97.5% of the prior probability of the age of the clade is above (younger than) 502 Ma.

*Vascular plants.* The two oldest fossils that can be assigned to either daughter lineage of the ancestor of crown group vascular plants are *Asteroxylon mackiei* (Kidston and Lang, 1920) and *Eophyllophyton bellum* (Hao, 1988). The former is an early lycopsid (Kenrick and Crane, 1997) from deposits of Rhynie chert radiometrically dated to early Devonian (Pragian) (408-411 Ma) by Rice et al. (1994), and the latter a stem-relative of Euphyllophytes (Kenrick and Crane, 1997) from the Posongchong Formation, Yunnan, China, also dated to early Devonian (Pragian) by Hao (1989). However, to accommodate for the controversial but possible (see Hueber, 1992; Kenrick and Crane, 1997) presence of lycopsids already in the late Silurian as represented by *Baragwanathia longifolia* (Lang and Cookson, 1935), possibly from Ludlow (427-423 Ma), Garratt and Rickards, 1987; Garratt et al., 1984), a lognormal prior with an offset of 408 Ma, a mean in real space of 20 (log stdev of 0.75) was specified for crown group vascular plants. These settings result in a prior age distribution of the clade with a median of 423 Ma and 97.5% probability cutoff of 473.7 Ma.

*Euphyllophytes.* A lognormally distributed age prior with an offset of 394 Ma and a mean in real space of 10 (log stdev of 0.75) was specified for crown group Euphyllophytes, based on the oldest known descendant of this node (Kenrick and Crane, 1997), the fossil *Pertica varia* (Granoff et al., 1976) from the early Devonian Battery Point Formation, Canada, from the mid to late Emsian (Hoffman and Tomescu, 2013), corresponding to about 402-394 Ma.

*Heterosporpus lycopods.* A lognormally distributed prior with an offset of 331 Ma, a mean in real space of 20 and log(stdev) of 0.75 was specified for the heterosporous lycopod crown group, based on the oldest fossils that have been safely assigned to either daughter lineage of the *Selaginella-Isoetes* ancestor: the isoetalean taxa *Paralycopodites* (Bateman and DiMichele, 1994), and the selaginellalean *Selaginellites resimus* (Korall et al., 1999; Rowe, 1988) from the early Visean (347-331 Ma). There are in addition older taxa that are often discussed as possible isoetalean taxa, such as *Protostigmaria, Cyclostigma, Lepidosigillaria* and *Clevelandodendron* (Bateman and DiMichele, 1994; Chitaley and Pigg, 1996; Kenrick and Crane, 1997; Korall et al., 1999). Although the phylogenetic position of these taxa have not been established with confidence (see also Stewart and Rothwell, 1993, p. 127), their probable rhizomorphic habit and/or abscising leaves indicate that they may be members of the *Selaginella-Isoetes* crown group, which would push back the split between *Isoetes* and *Selaginella* to the Late Devonian (Korall et al., 1999), i.e., 383-359 Ma. Moreover, heterosporous lycopsid strobili are reported from as far back in time as the Middle Devonian (Schweitzer and Cai, 1987), i.e., 393-383 Ma. The prior distribution was set to accommodate for these indications.

*Seed plants*. A relatively broadly distributed lognormal prior was specified for seed plants, with an offset of 312 Ma, a mean in real space of 20, and log(stdev) of 0.5, based on the oldest members of crown group seed plants, the Cordaites (Doyle, 2008; Hilton and Bateman, 2006). Cordaites are unequivocally documented from the mid-Pennsylvanian (312 Ma), but were probably present already during the late Mississippian (Taylor et al., 2009), i.e., 323-331 Ma.

*Angiosperms.* Unequivocal angiosperm remains in the form of isolated plant organs and pollen with reticulate-collumellate pollen wall are not reported prior to the Valanginian-Hauterivian (Friis et al., 2011), i.e., 131-139 Ma. The age of the clade is likely somewhat older and a lognormally distributed age prior was specified, with an offset of 135 Ma, a mean in real space of 15 Ma, and log(stdev) of 0.75.

*Eudicots.* Rarely occurring triaperturate pollen from the late Barremian to early Aptian represent the earliest evidence of the eudicots (Friis et al., 2011). This pollen type is distinct and would most probably have been found in older strata if present. Thus, we specify a narrowly distributed age prior of eudicots, an exponential prior with an offset of 126 Ma with a mean of 1.

**References**

Bateman RM, DiMichele WA. Heterospory: the most iterative key innovation in the evolutionary history of the plant kingdom. Biol Rev. 1994;69:345-417.

Chitaley S, Pigg KB. *Clevelandodendron ohioensis*, gen. et sp. nov., a slender upright lycopsid from the Late Devonian Cleveland Shale of Ohio. Am J Bot. 1996;83:781-9.

Doyle JA. Integrating molecular phylogenetic and paleobotanical evidence on origin of the flower. Int J Plant Sci. 2008;169:816-43.

Drummond AJ, Bouckaert RR. Bayesian evolutionary analysis with BEAST 2. Cambridge University Press; 2014.

Drummond AJ, Suchard MA, Dong X-P, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29:1969-73.

Drummond AJ, Suchard MA. Bayesian random local clocks, or one rate to rule them all. BMC Biol. 2010;8:114.

Friis EM, Crane PR, Pedersen KR. Early flowers and angiosperm evolution. New York, NY: Cambridge University Press, 2011.

Garratt MJ, Rickards RB. Pridoli (Silurian) graptolites in association with *Baragwanathia* (Lycophytina). Bull Geol Soc Denmark. 1987;35;135-9.

Garratt MJ, Tims JD, Rickards RB, Chambers TC, Douglas JG. The appearance of *Baragwanathia* (Lycophytina) in the Silurian. Bot J Linn Soc. 1984;89:355-8.

Gradstein FM, Ogg JG, Schmitz M, Ogg G. The Geologic Time Scale 2012. Oxford, UK: Elsevier; 2012.

Granoff JA, Gensel PG, Andrews HN. A new species of *Pertica* from the Devonian of eastern Canada. Palaeontographica B. 1976;155:119-28.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate Maximum-Likelihood Phylogenies: Assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307-21.

Hao S-G. A new Lower Devonian genus from Yunnan, with notes on the origin of leaves. Act Bot Sinica. 1988;30:441-48.

Hao S-G. A new zosterophyll from from the Lower Devonian (Siegenian) of Yunnan, China. Act Bot Sinica. 1989;31:954-61.

Hilton J, Bateman RM. Pteridosperms are the backbone of seed-plant phylogeny. J Torrey Bot Soc. 2006;133:119-68.

Hoffman LA, Tomescu AMF. An early origin of secondary growth: *Franhueberia gerriennei* gen. et sp. nov. from the Lower Devonian of Gaspé (Quebec, Canada). Am J Bot. 2013;100:754-63.

Hueber FM. Thoughts on the early lycopsids and zosterophylls. Ann Mo Bot Gard. 1992;79:474-99.

Kenrick P, Crane PR. The origin and early diversification of land plants: a cladistic study. Washington: Smithsonian Institution Press; 1997.

Lang WH, Cookson IC. On a flora, including vascular land plants, associated with *Monograptus*, in rocks of Silurian age, from Victoria, Australia. Philos Trans R Soc Lond B Biol Sci. 1935;224:421-49.

Korall P, Kenrick P, Therrien J. Phylogeny of Selaginellaceae: evaluation of generic/subgeneric relationships based on *rbcL* gene sequences. Int J Plant Sci. 1999;160:585-94.

Nylander JAA. MrAIC.pl. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University. 2004.

Rice CM, Ashcroft WA, Batten DJ, Boyce AJ, Caulfield JBD, Fallick AE, Hole MJ, Jones E, Pearson MJ, Rogers G, Saxton JM, Stuart FM, Trewin NH, Turner G. A Devonian auriferous hot spring system, Rhynie, Scotland. J Geol Soc London. 1994;152:229-50.

Rowe NP. A herbaceous lycophyte from the Lower Carboniferous Drybrook Sandstone of the Forest of Dean, Gloucestershire. Palaeontology. 1988;31:69-83.

Schweitzer HJ, Cai C. Beiträge zur Mitteldevon-Flora Südchinas. Palaeontographica. 1987;207B:1-109.

Stadler T. On incomplete sampling under birth-death models and connections to the sampling-based coalescent. J Theor Biol. 2009;261:58-66.

Stewart WN, Rothwell GW. Paleobotany and the evolution of plants. Cambridge, UK: Cambridge University Press; 1993.

Taylor TN, Taylor EL, Krings M. Paleobotany - The biology and evolution of fossil plants. New York, NY: Academic Press; 2009.