Two New Fern Chloroplasts and Decelerated Evolution Linked to the Long Generation Time in Tree Ferns

Bojian Zhong[1,*], Richard Fong[1], Lesley J. Collins[2], Patricia A. McLenachan[1], and David Penny[1]

[1]Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

[2]Faculty of Health Sciences, Universal College of Learning, Palmerston North, New Zealand

*Corresponding author: E-mail: bjzhong@gmail.com.

Accepted: April 23, 2014

Data deposition: The two new chloroplast genomes (*Dicksonia* and *Tmesipteris*) have been deposited at GenBank under accessions KJ569698 and KJ569699, respectively.

Abstract

We report the chloroplast genomes of a tree fern (*Dicksonia squarrosa*) and a "fern ally" (*Tmesipteris elongata*), and show that the phylogeny of early land plants is basically as expected, and the estimates of divergence time are largely unaffected after removing the fastest evolving sites. The tree fern shows the major reduction in the rate of evolution, and there has been a major slowdown in the rate of mutation in both families of tree ferns. We suggest that this is related to a generation time effect; if there is a long time period between generations, then this is probably incompatible with a high mutation rate because otherwise nearly every propagule would probably have several lethal mutations. This effect will be especially strong in organisms that have large numbers of cell divisions between generations. This shows the necessity of going beyond phylogeny and integrating its study with other properties of organisms.

Key words: *Tmesipteris*, *Dicksonia*, ferns and fern allies, chloroplast genomes, generation time effect, mutation rates.

Introduction

We address three main types of questions in this study: The phylogeny of early land plants, the decelerated evolutionary rates of tree ferns, and the possible biological reasons for the observed differences in mutation rates. *Tmesipteris* (or "hanging fork fern") only grows in New Caledonia, New Zealand, and parts of eastern Australia, and so it is difficult for some researchers to obtain it for sequencing. *Tmesipteris* and *Psilotum* are both interesting plants in that *Psilotum* superficially resembles certain extinct early vascular plants, such as the rhyniophytes and the trimerophyte genus *Psilophyton* (Bierhorst 1977). The unusual features of *Psilotum* that suggest an affinity with early vascular plants include dichotomously branching sporophytes, aerial stems arising from horizontal rhizomes, a simple vascular cylinder, homosporous and terminal eusporangia, and a lack of roots. However, recent studies have tended to place the Psilotales (which includes *Tmesipteris*) closer to ferns (Pryer et al. 2001; Qiu et al. 2006, 2007). This left *Psilotum* as a "long branch" in the phylogenetic tree, and these are well known to be problematic. *Tmesipteris* was to help test the possibility that the more widespread *Psilotum* was misplaced because of "long branch attraction" artifact (Hendy and Penny 1989). *Dicksonia* was chosen to test whether it would also show the slowdown in rates (Korall et al. 2010) that was known for the chloroplast genes of the other main family of tree ferns (Cyatheaceae that includes the *Alsophila* genus). Removing the fastest evolving sites has helped the phylogenetic reconstruction (see Zhong et al. 2011, 2014) but it is not yet known how much effect, if any, it has on divergence time estimates. We further examined the question of removing the fastest evolving sites in order to help evaluate whether their removal had any major effect on the estimated times of divergence.

Next we turn to the question of evolutionary rates. Early in molecular evolution studies, researchers were surprised at the relatively equal rate of molecular evolution, for example, between vertebrates, fungi, and plants: There did not appear to be the expected correlation between mutation rates in diversified and nondiversified lineages (Kimura and Ohta 1974). This observation, together with the much higher than

expected genetic diversity within populations, led to the development of the neutral theory of molecular evolution (Kimura and Ohta 1974) where mutations that were neutral tended to outnumber those that were advantageous. This did lead to the concept of a "molecular clock," with a relatively constant rate of DNA evolution in different eukaryote groups.

However, there has been recently considerable interest in the actual variation in rates, and the observation of lineage-specific rate heterogeneity has been well characterized within fungi (Lumbsch et al. 2008), mammals (Goldie et al. 2011), seed plants (Smith and Donoghue 2008; Xiang et al. 2008; Bromham et al. 2013), and ferns (Soltis et al. 2002; Schneider et al. 2004; Korall et al. 2010; Rothfels et al. 2012). We still lack a good biological understanding of factors that might affect this observed variation in rates. There are at least three general types of explanation that might affect rate, the first is a general increase (or decrease) in mutation rate; the second is a change in the number of sites "free to vary" (i.e., a change in selection pressures); and the third is variations in the mechanisms that might, for example, lead to double-stranded breaks and subsequent repair. This last aspect of the heterogeneity has probably made it difficult for fully resolving the placental mammal phylogeny (Romiguier et al. 2013) because the location where genetic recombination (and increasing the number of double-stranded breaks, which are more error prone during their correction) appears to keep changing within Placentalia. It is important (essential) to understand the biological principles for the observed variation in rates of molecular evolution in different groups (e.g., Lanfear et al. 2014). We should be able to make *predictions* about what we expect.

The heterogeneous pattern of among-lineage rate variation has presented a significant challenge for accurately estimating divergence times. Various substitution models that relax the assumption of the strict molecular clock have been developed to account for rate heterogeneity between lineages in molecular phylogenetics (e.g., Thorne et al. 1998; Sanderson 2002; Drummond et al. 2006). It has been reported that the fast-evolving sites are one important source of systematic errors in molecular phylogenetics, and phylogenetic inference can be improved by removal of the most variable sites regardless of the mechanism of mutation (e.g., Goremykin et al. 2010; Zhong et al. 2011, 2014; Parks et al. 2012). However, it remains unknown whether the fastest evolving sites affect the accuracy of divergence time estimation, even using the relaxed clock models. To test the impact of divergence time estimation based on different sites with different evolutionary rates, and to investigate the relation between generation for a range of genome sizes and mutation rate, we designed an empirical study using the chloroplast genomes of land plants, which include two newly sequenced species (a tree fern and a fern ally) to give a total of 28 chloroplast genomes.

Results and Discussion

The two chloroplast genomes (*Dicksonia* and *Tmesipteris*) have been submitted to GenBank, and have accession numbers KJ569698 and KJ569699, respectively. The Herbarium numbers are MPN: 47797 for the *Dicksonia* sample and MPN: 47838 for the *Tmesipteris*. For this study, we had 34,386 aligned sites, and identified 3,250 rapidly evolving sites using the observed variability (OV)-sorting method (fig. 1). Thus, the reduced OV-sorted data are 31,136 aligned sites.

The first step was to reconstruct the phylogeny of early land plants. Zhong et al. (2011) and Goremykin et al. (2013) have reported that the OV-sorting method is effective in identifying the fastest evolving sites, and phylogenetic inference is significantly improved after their removal. The maximum-likelihood (ML) analyses with RAxML based on original and OV-sorted data produced both well-supported phylogenetic trees (figs. 2 and 3). All major groups (e.g., seed plants, monilophytes, and lycophytes) are monophyletic with high bootstrap support (BP), and the "fern ally" *Tmesipteris elongata* is the sister group to *Psilotum nudum*, and they are basal to ferns. The tree fern clade (i.e., *Dicksonia squarrosa* and *Alsophila spinulosa*) is strongly supported as monophyletic (BP = 100), and there is a major rate deceleration occurred along both tree ferns (notably shorter branches within the tree fern clade). Thus, the slowdown in rates does occur in both tree ferns. This slowdown is in marked contrast to the other ferns where there is rate acceleration, especially among the more advanced (derived) ferns.

The only difference between the two trees was the position of lycophytes. For the original data, lycophytes as sister to seed plants were weakly supported (BP = 64%; fig. 2). In contrast, after removing fastest evolving sites, the phylogenetic tree supported lycophytes close to (seed plant + monilophytes) (fig. 3) which was congruent with previous studies (Pryer et al. 2001; Qiu et al. 2006, 2007; Rai and Graham 2010; Zhong et al. 2013). This confirms that these fastest evolving sites may mislead the phylogenetic inference of position of lycophytes, so we used the phylogenetic tree based on OV-sorted data for divergence time estimation.

To evaluate the impact of divergence time estimation of the fastest evolving sites, we estimated the divergence times using the original and OV-sorted data with eight fossil records. We found that age estimates from most nodes did not vary substantially between original and OV-sorted data (table 1). For instance, relaxed molecular clock analyses using original and OV-sorted data yielded the similar mean estimates as 136.6 and 150.1 Myr before present (Ma) for crown angiosperms (node 1 in table 1 and fig. 3), and the estimated age of crown Tracheophyta (node 20) is 445.7 and 428.9 Ma, respectively. However, for some deeper nodes (e.g., nodes 25, 26, and 27), the mean ages and confidence intervals reduced considerably with OV-sorted data compared with the original data. This

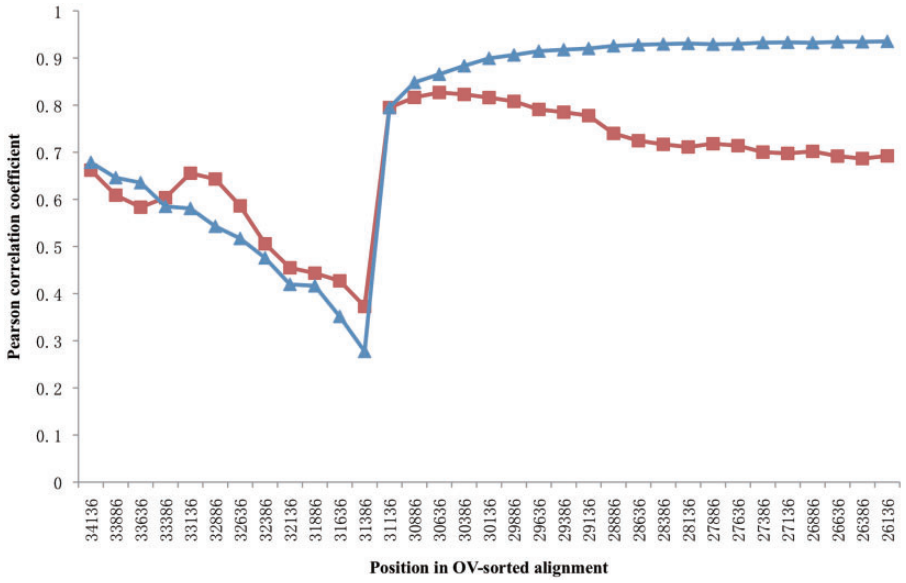

Fig. 1.—Pearson correlation results. The blue line indicates the Pearson correlation coefficient (*r*) of the ML distance calculated from "A" (more conserved) and "B" (less conserved) partitions. The red line indicates the *r* value of uncorrected *p* distances and ML distances for B partitions. The *r* values begin to increase significantly at 31,136 sites remaining and this is taken to indicate that the assumed model of nucleotide evolution is beginning to fit the data well.

does require more investigation to clarify such variation because estimates of divergence time are an important aspect of molecular evolution. Consequently, these results are encouraging in that they appear to give more realistic estimates for the deeper divergences.

In figure 4, results with numbers of mutations between generations show the expected number of mutations based on both the mutation rate and the generation time. At the longer generation times, there are predicted to be many more mutations in the offspring, and many of these are potentially lethal. The average "generation time" for tree ferns does not appear to be accurately known (nor for many organisms) and so an estimate of about 100 years for tree ferns being actively reproductive was used, based on results in Ash (1987), Shepherd and Cook (1988), and Large and Braggins (2004). In some cases an estimate of 200 years was available, but we limited it to, on average, 100 years generation time. We arbitrarily count half the genes, in that we allow some "lethal" mutations to have occurred in leaf tissue, and any such cells will simply die at that point, and not affect the ongoing activity of the leaf. So we ignore those lethal mutations. Basically, it is important to consider the effect to the next generation, many genes will only be expressed earlier in development or in root tissue—so they will not have been selected against during stem and leaf growth. In practice, there will also be an effect from cells being diploid, but that is not expected to alter the basic result in the longer term. There is clearly an effect of DNA replication error in both meiosis and mitosis. There will only be one meiosis per generation, and it will be subject to double-stranded breaks during recombination and repair. However, there will be many mitoses between generations. As we point out, the number of cell divisions per generation is a critical factor, and the absolute time may also be important for double-stranded breaks and their repair.

The phylogenetic aspects covered here appear largely to be as expected for the phylogeny of early land plants. Monilophytes are a monophyletic (natural) group, and are closest to seed plants. Lycophytes are the sister group of a clade comprising seed plants and monilophytes. Moreover, the fastest evolving sites do not appear to make many major changes to the estimated times of divergence, but the more recent divergences for the deeper nodes based on OV-sorted data (fig. 3) do warrant further testing.

The other aspect is that it appears that larger organisms (with more cell divisions and longer generation times) tend to have lower mutation rates, possibly in order to limit the number of mutations between a parent and its offspring. Lehtonen J and Lanfear R (in preparation) have reached a similar conclusion. If there are very large numbers of mutations between parents and offspring, then almost certainly some of these mutations would be detrimental, and this appears to place a limit on the mutation rate of organisms with a long generation time. It appears that the slow-down of mutation rates affects both the nuclear and chloroplast genes (Rothfels and Schuettpelz 2013). There need not be a close correlation between mutation rate and generation time—all that the present results imply that a high mutation rate is incompatible if combined with a long generation time—this

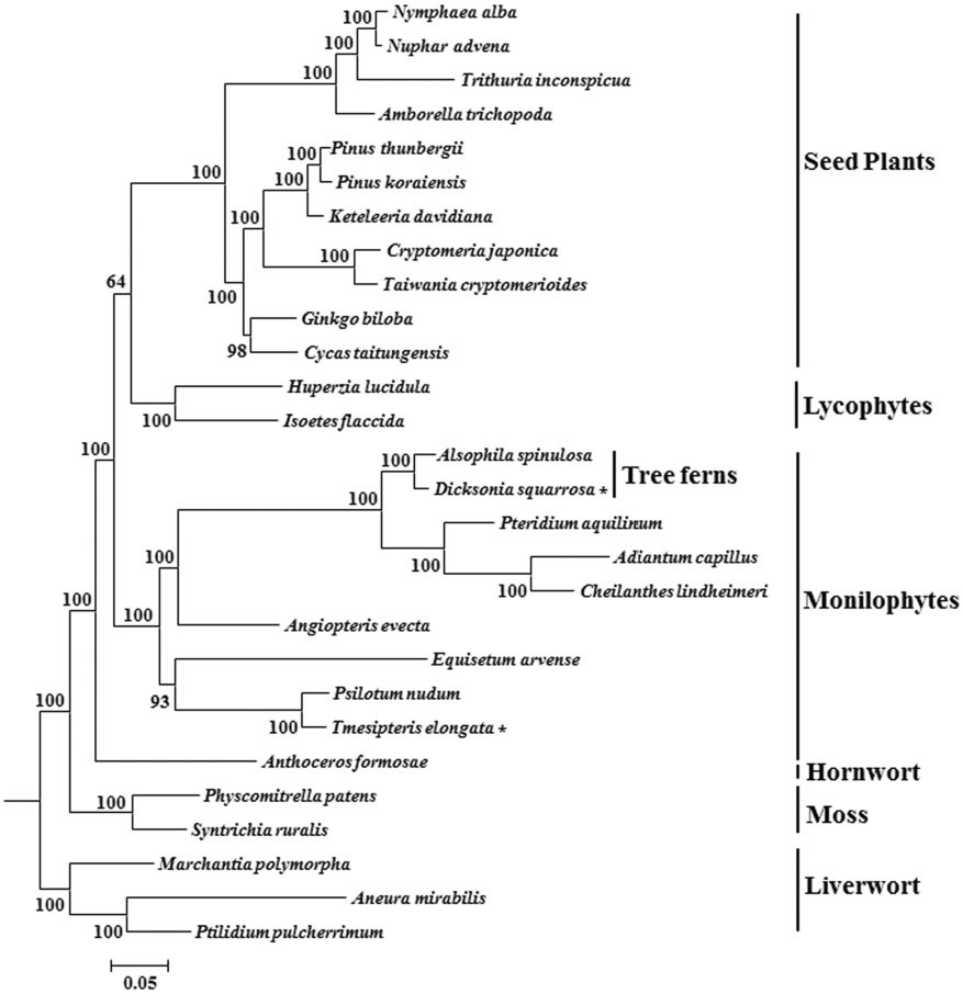

Fig. 2.—ML tree of land plants based on the original data (34,386 sites). Bootstrap support values are indicated along the branches. The two newly sequenced genomes are indicated as *. In this tree, the lycophytes are adjacent to the seed plants with weak bootstrap support (BP = 64%).

would eliminate only one combination of possible rates and generation times. It is also unclear whether the lower mutation rate means that the rate of DNA copying is slower, and that there is consequently more time for checking the new strand against the old strand—how does a lower mutation rate really occur? There are many questions that need to be followed up.

There has been considerable effort into testing some of the reasons behind variation in rates between different lineages. Lanfear et al. (2013) pointed out that for many trees, about a fifth of the rate variation can be explained by slower rates of mutation in taller trees. It may well be that the generation time effect is at least a partial explanation for why there appears, on a geological time scale, to be continued turnover of large organisms. However, it is always going to be important to take into account the number of cell divisions between generations. For most vertebrates (including humans), there are special germ-line cells set aside that may have fewer cell divisions than many of the somatic cells (see Kong et al. 2012).

Furthermore, Lynch and Abegg (2010) reported that larger populations can more quickly acquire combinations of mutations that might lead to useful longer-term innovations, than can smaller populations (such as might be found with larger individuals). Consequently, smaller organisms (with larger population sizes) may be better able to continue evolving, particularly gaining complex new features. It is important to determine the number of mutations between generations for a range of genome sizes, mutation rates, and population sizes, and the apparent slowdown in mutation rates among the tree ferns could be an important test of this hypothesis. We do need further tests on whether very long generation times are associated with lower mutation rates in other organisms—we predict that a high mutation rate and a long generation time are incompatible (see also Thomas et al. 2010). It is also important to understand the mechanisms involved in the lower

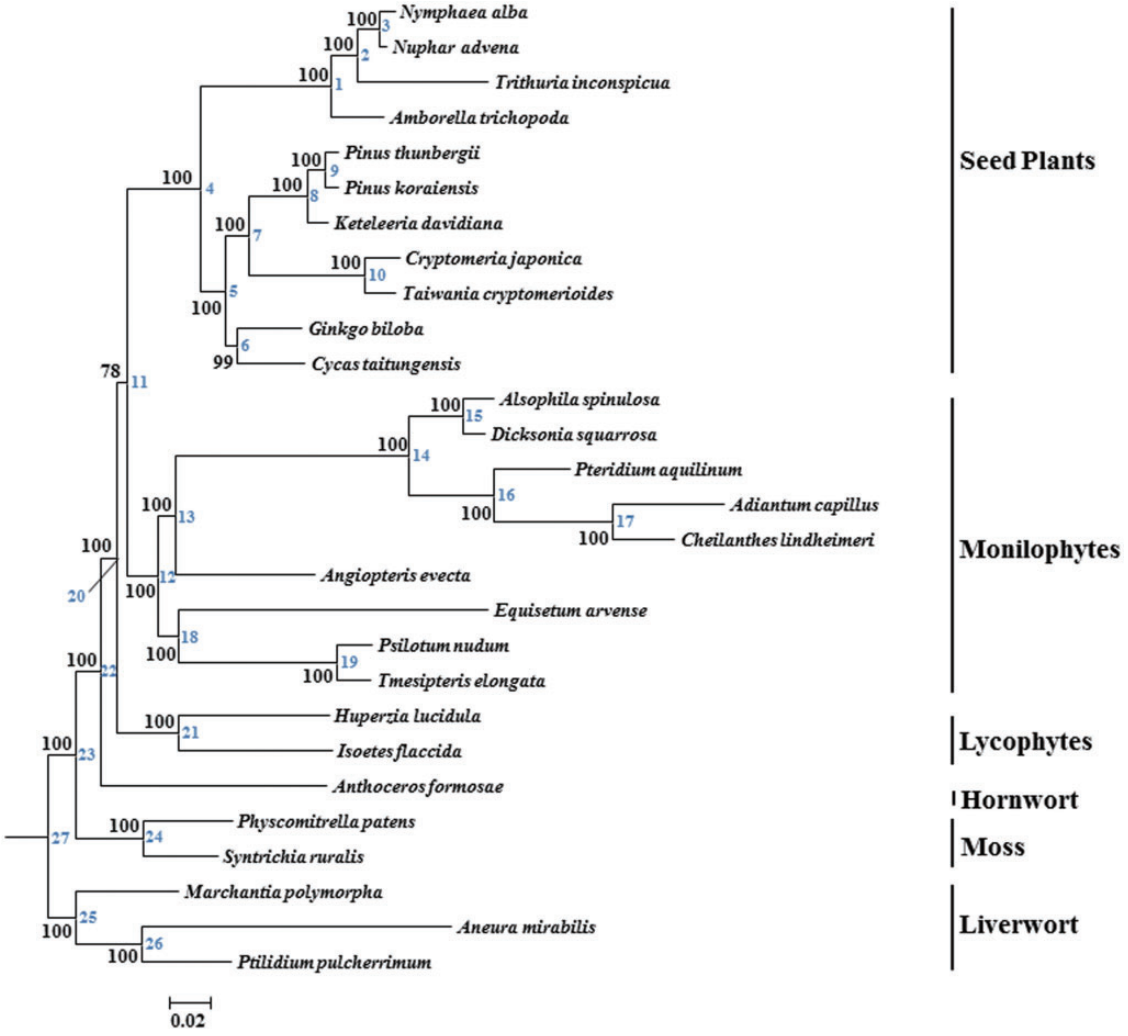

Fig. 3.—ML tree of land plants based on the OV-sorted matrix (31,136 sites). Bootstrap support values are indicated along the branches and node numbers are marked as blue. This ML tree is the same as figure 1 except that the lycophytes are now adjacent to (seed plant + monilophytes) with 100% bootstrap support.

mutation rates (as observed in tree ferns). Is higher accuracy (lower mutation rates) associated with slower copying of DNA (and therefore more time for checking of potential errors during copying)? Or is it some intrinsic mechanism that is independent of the rate of DNA copying? All aspects, the phylogeny of early land plants, the use of slowly evolving sites for time estimates, and the effect of life cycle and generation time, certainly warrant continued study.

Materials and Methods

DNA Sequencing and Data Assembly

The tree fern *D. squarrosa* and the "fern ally" *T. elongata* were collected and sourced from Palmerston North, New Zealand. The *Dicksonia* sample was a cultivated plant from Palmerston North, and the *Tmesipteris* sample was growing on the trunks of tree ferns at the beginning of the Sledge track in the Kahuterawa valley, inland from Palmerston North; its location was 40.47 (south), 175.60 (east). Total genomic DNA (~50 ng) from each sample was extracted using the Qiagen Plant DNeasy kit according to the manufacturer's protocols, and then sequenced using Illumina GAIIx sequencing platform with 100-bp paired-end reads. The short reads were filtered with the error probability <0.05, and were then assembled using Velvet (Zerbino and Birney 2008). The contigs were further assembled using Geneious software version 5.6 (www.geneious.com, last accessed May 12, 2014). Protein-coding genes were annotated using DOGMA (Wyman et al. 2004) with manual correction. Each protein-coding gene from 28 taxa was aligned using MUSCLE (Edgar 2004), and trimmed to exclude poorly aligned positions using Gblocks (Castresana 2000)

Table 1
Estimated Times of Divergence Using Original and Reduced OV-Sorted Matrices

	Mean Estimates (Ma)		95% Credibility Intervals		
Node	Full Matrix (34,386 sites)	OV-Sorted Matrix (31,136 sites)	Full Matrix	OV-Sorted Matrix	Fossil Calibrations (Ma)
1	136.6	150.1	67.0–208.7	75.6–242.7	
2	85.8	97.7	38.4–140.1	42.8–176.6	
3	22.9	28.2	4.1–47.7	3.8–60.6	
4	315.8	317.5	306.2–333.2	306.2–339.4	>306.2[a]
5	225.0	224.1	168.5–287.8	165.4–285.3	
6	187.0	163.3	108.8–262.2	75.9–257.7	
7	160.9	161.4	147.0–187.8	147.0–187.3	>147.0[b]
8	57.3	62.6	22.7–99.7	21.6–108.6	
9	23.8	26.9	5.3–47.7	5.7–54.9	
10	55.1	58.4	12.2–98.8	16.5–106.6	
11	413.1	404.8	388.2–447.9	388.2–429.2	>388.2[c]
12	366.2	368.5	354.0–388.2	354.0–390.7	>354.0[d]
13	327.8	336.7	280.8–365.4	291.5–378.8	
14	221.9	228.8	179.8–264.1	187.5–270.1	
15	165.7	168.0	159.0–180.6	159.0–185.3	>159.0[e]
16	144.8	154.3	91.1–201.1	93.1–217.0	
17	73.1	76.5	36.0–116.5	34.5–122.0	
18	296.1	296.2	203.6–364.9	189.9–370.3	
19	69.1	72.3	18.5–147.7	14.7–147.1	
20	445.7	428.9	403.3–492.9	400.1–463.5	
21	387.7	386.3	377.4–406.9	377.4–403.0	>377.4[f]
22	483.4	454.4	423.0–553.2	413.6–501.5	
23	534.7	487.6	450.9–629.1	435.6–550.8	>420.4[g]
24	190.3	178.3	51.3–353.1	37.5–364.7	
25	677.3	375.8	277.4–1030.3	172.5–569.7	
26	468.5	228.3	130.4–758.6	99.1–397.8	
27	775.5	529.8	502.4–1042.0	449.0–629.5	449–1042[h]

Note—Node numbers are shown in figure 3.
[a]References: Davydov et al. (2004) and Heckel (2008).
[b]Spalleti et al. (1982).
[c]House and Gradstein (2004).
[d]Bateman (1991), Galtier and Phillips (1996), and Bek and Psenicka (2001).
[e]Lantz et al. (1999) and Skog (2001).
[f]Grierson and Bonamo (1979).
[g]Edwards and Feehan (1980) and Zalasiewicz et al. (2009).
[h]Cooper and Sadler (2004) and Turnbull et al. (1996).

with default settings. These alignments were concatenated to generate a matrix of 34,386 sites.

Phylogenetic Inference and Divergence Time Estimation

The OV-sorting method (Goremykin et al. 2010) was used to rank the original concatenated alignment from the most to the least variable sites. The most variable sites were then successively removed from the original matrix, in increments of 250 sites. The stopping point for site removal was determined as the point at which the two correlations showed significant improvement (see Goremykin et al. [2010, 2013] for details of the method). ML phylogeny on each dataset was conducted using RAxML (Stamatakis 2006) with the GTRGAMMA model.

The divergence times were estimated using the Bayesian software BEAST version 1.7.2 (Drummond and Rambaut 2007). The optimal substitution model was selected using ModelTest (Posada and Crandall 1998). Rate heterogeneity among lineages was modeled using an uncorrelated relaxed-clock (UCLN) model (Drummond et al. 2006). Samples from the posterior distribution were drawn every 2,000 steps over 10^8 steps of a single chain, with the first 10% of samples discarded as burn-in. Four independent MCMC chains were run. Convergence was checked based on time-series plots of the likelihood scores using Tracer program (Drummond and Rambaut 2007). Eight fossil-based calibrations were utilized for molecular dating analyses. The root age was set at 449–1,042 Ma (Turnbull et al. 1996; Cooper and Sadler 2004). The other internal fossil calibrations were representatives of

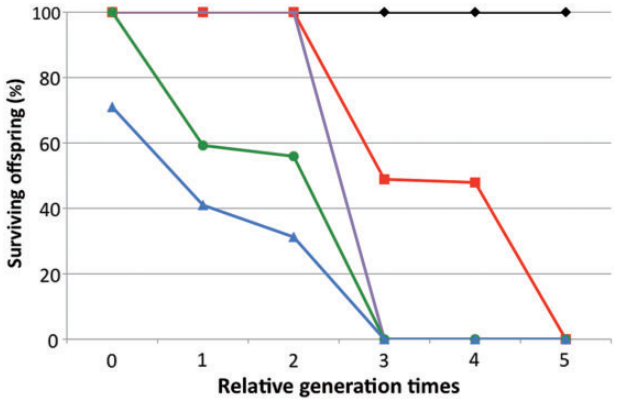

Fig. 4.—Estimates of the percentage of surviving offspring (*y* axis) at lowered mutation rates and increasing relative generation times (*x* axis). The mutation rate increases when going from the black to red, purple, green, and blue lines. There are three overlapping lines at the top left, and three overlapping lines at the bottom right. As the generation time increases, particularly with higher mutation rates, the chance of successful offspring (with no lethal mutations) strongly declines. A high mutation rate and a long generation time are incompatible.

the oldest-known clades to provide minimum age constraints (see table 1).

Estimation of Number of Mutations

For the rates of evolution, we used a Python script to estimate the number of mutations that were expected to occur for a set number of genes and for a given mutation rate. In order to make the calculation in a reasonable amount of time, the mutation rate and numbers of genes were scaled to keep the same proportion. In practice, the genomes started with 1,000 genes, each 1,000 nt long with an error rate in copying DNA of about 10^{-9} per errors per nucleotide—this is a realistic rate for eukaryotes (Drake 1999). The number of mutations was recorded, and a gene was considered lethal if there were more than 10 mutations in it. Alternatively, if two mutations occurred at the same amino acid site (a double hit) this was also taken as a lethal mutation—and led to a "dead (nonfunctional) gene." The mutation rate was the same for all genes. In general, photosynthetic organisms seem to have a higher number of genes often having 30,000 genes (Raven et al. 2013). In practice, many genes will be functioning in the photosynthetic leaf tissue, so any lethal mutation may only lead to a cell that takes no further function in the leaf. Therefore, we assume that less than half the genes may only function in tissues such as embryos or roots, and any lethal mutation here will affect the next generation.

Acknowledgments

The authors thank Patrick Brownsey and Leon Perrie, both from Te Papa Museum, Wellington, for information about generation times of tree ferns, and thank Vadim Goremykin for assistance of OV-sorting method, and Simon Ho of BEAST analyses. This work was supported (in part) by the NZ Marsden Fund. They also thank two reviewers and the editor for providing useful comments.

Literature Cited

Ash J. 1987. Demography of *Cyathea hornei* (Cyatheacee), a tropical fern from Fiji. Aust J Bot. 35:331–341.

Bateman RM. 1991. Palaeobiology and phylogenetic implications of anatomically preserved *Archaeocalamites* from the Dinantian of Oxroad Bay and Loch Humphrey Bun, Scotland. Palaeontographica Abt B Paläephytol. 223:1–59.

Bek J, Psenicka J. 2001. *Senftenbergia plumose* (Artis) emend. and its spores from the Carboniferous of the Kladno and Pilsen Basins, Bohemian Massif, and some related and synonymous taxa. Rev Palaeobot Palynol. 116:213–232.

Bierhorst DW. 1977. The systematic position of *Psilotum* and *Tmesipteris*. Brittonia 29:3–13.

Bromham L, Cowman PF, Lanfear R. 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. BMC Evol Biol. 13:126.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 17: 540–552.

Cooper RA, Sadler PM. 2004. The Ordovician period. In: Gradstein FM, Ogg J, Smith A, editors. A geologic time scale 2004. Cambridge: Cambridge University Press. p. 165–187.

Davydov V, Wardlaw BR, Gradstein FM. 2004. The Carboniferous period. In: Gradstein FM, Ogg JG, Smith AG, editors. A geologic time scale 2004. Cambridge: Cambridge University Press. p. 222–248.

Drake JW. 1999. The distribution of rates of spontaneous mutation over viruses, prokaryotes and eukaryotes. Ann NY Acad Sci. 870: 100–107.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 7:214.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32: 1792–1797.

Edwards D, Feehan J. 1980. Records of Cooksonia-type sporangia from late Wenlock strata in Ireland. Nature 287:41–42.

Galtier J, Phillips TL. 1996. Structure and evolutionary significance of Palaeozoic ferns. In: Camus JM, Gibby M, Johns RJ, editors. Pteridology in perspective. Kew (United Kingdom): Royal Botanic Gardens. p. 417–434.

Goldie X, Lanfear R, Bromham L. 2011. Diversification and the rate of molecular evolution: no evidence of a link in mammals. BMC Evol Biol. 11:286.

Goremykin VV, et al. 2013. The evolutionary root of flowering plants. Syst Biol. 62:51–62.

Goremykin VV, Nikoforova SV, Bininda-Emonds OPP. 2010. Automated removal of noisy data in phylogenomic analyses. J Mol Evol. 71: 319–331.

Grierson JD, Bonamo PM. 1979. Leclercqia complexa: earliest ligulate lycopod (Middle Devonian). Am J Bot. 66:474–476.

Heckel PH. 2008. Carboniferous period. In: Ogg JG, Ogg G, Gradstein FM, editors. The concise geologic time scale. New York: Cambridge University Press. p. 73–83.

Hendy MD, Penny D. 1989. A framework for the quantitative study of evolutionary trees. Syst Zool. 38:297–309.

House MR, Gradstein FM. 2004. The Devonian period. In: Gradstein FM, Ogg JG, Smith AG, editors. A geologic time scale 2004. Cambridge: Cambridge University Press. p. 202–221.
Kimura M, Ohta T. 1974. On some principles governing molecular evolution. Proc Natl Acad Sci U S A. 71:2848–2852.
Kong A, et al. 2012. Rate of de novo mutations and the importance of father's age to disease risk. Nature 488:471–475.
Korall P, Schuettpelz E, Pryer KM. 2010. Abrupt deceleration of molecular evolution linked to the origin of arborescence in ferns. Evolution 64: 2786–2792.
Lanfear R, et al. 2013. Taller plants have lower rates of molecular evolution. Nat Commun. 4:1879.
Lanfear R, Kokko H, Eyre-Walker A. 2014. Population size and the rate of evolution. Trends Ecol Evol. 29:33–41.
Lantz TC, Rothwell GW, Stockey RA. 1999. *Conantiopteris schuchmanii*, gen. et sp. nov., and the role of fossils in resolving the phylogeny of Cyatheaceae *s.l.* J Plant Res. 112:361–381.
Large MF, Braggins JE. 2004. Tree ferns. Portland (OR): Timber Press.
Lumbsch H, Hipp AL, Divakar PK, Blanco O, Crespo A. 2008. Accelerated evolutionary rates in tropical and oceanic parmelioid lichens (Ascomycota). BMC Evol Biol. 8:257.
Lynch M, Abegg A. 2010. The rate of establishment of complex adaptations. Mol Biol Evol. 27:1404–1414.
Parks M, Cronn R, Liston A. 2012. Separating the wheat from the chaff: mitigating the effects of noise in a plastome phylogenomic data set from Pinus L. (Pinaceae). BMC Evol Biol. 12:100.
Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.
Pryer KM, et al. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409:618–622.
Qiu YL, et al. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. Proc Natl Acad Sci U S A. 103:15511–15516.
Qiu YL, et al. 2007. A non-flowering land plant phylogeny inferred from nucleotide sequences of seven chloroplast, mitochondrial and nuclear genes. Int J Plant Sci. 168:691–708.
Rai HS, Graham SW. 2010. Utility of a large, multigene plastid data set in inferring higher-order relationships in ferns and relatives (monilophytes). Am J Bot. 97:1444–1456.
Raven JA, Beardall J, Larkum AWD, Sanchez-Baracaldo P. 2013. Interactions of photosynthesis with genome size and function. Philos Trans R Soc Lond B Biol Sci. 368:20120264.
Romiguier J, Ranwez V, Delsuc F, Galtier N, Douzery E. 2013. Less is more in mammalian phylogenomics: AT-rich genes minimize tree conflicts and unravel the root of placental mammals. Mol Biol Evol. 30: 2134–2144.
Rothfels CJ, et al. 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. Syst Biol. 61:490–509.
Rothfels CJ, Schuettpelz E. 2013. Accelerated rate of molecular evolution for vittarioid ferns is strong and not driven by selection. Syst Biol. 63: 31–54.
Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol Biol Evol. 19: 101–109.
Schneider H, et al. 2004. Ferns diversified in the shadow of angiosperms. Nature 428:553–557.
Shepherd W, Cook WC. 1988. The Botanic Garden, Wellington: a New Zealand history, 1840–1987. Wellington (New Zealand): Millwood Press.
Skog JE. 2001. Biogeography of Mesozoic leptosporangiate ferns related to extant ferns. Brittonia 53:236–269.
Smith SA, Donoghue MJ. 2008. Rates of molecular evolution are linked to life history in flowering plants. Science 322:86–89.
Soltis PS, Soltis DE, Savolainen V, Crane PR, Barraclough TG. 2002. Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. Proc Natl Acad Sci U S A. 99:4430–4435.
Spalleti L, Iniguez Rodriguez AM, Mason M. 1982. Edades radimetricas de piroclastitas y volvanitas del Grupo Bahia Laura, Gran Bajo de San Julian, Santa Cruz. Rev Asoc Geol Argent. 37:483–485.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.
Thomas JA, Welch JJ, Lanfear R, Bromham L. 2010. A generation time effect on the rate of molecular evolution in invertebrates. Mol Biol Evol. 27:1173–1180.
Thorne JL, Kishino H, Painter IS. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol. 51:1647–1657.
Turnbull MJM, Whitehouse MJ, Moorbath S. 1996. New isotopic age determinations for the Torridonian, NW Scotland. J Geol Soc Lond. 153: 955–964.
Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20: 3252–3255.
Xiang Q-Y, et al. 2008. Rates of nucleotide substitution in Cornaceae (Cornales)—pattern of variation and underlying causal factors. Mol Phylogenet Evol. 49:327–342.
Zalasiewicz JA, et al. 2009. Graptolites in British stratigraphy. Geol Mag. 146:785–850.
Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18:821–829.
Zhong B, et al. 2011. Systematic error in seed plant phylogenomics. Genome Biol Evol. 3:1340−1348.
Zhong B, et al. 2014. Origin of land plants revisited using heterogeneous models and three new chloroplast genomes. Mol Biol Evol. 31: 177–183.
Zhong B, Liu L, Yan Z, Penny D. 2013. Origin of land plants using the multispecies coalescent model. Trends Plant Sci. 18:492–495.

Associate editor: Shu-Miaw Chaw