Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 35 (2005) 569–582
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Taxon sampling effects in molecular clock dating: An example from the African Restionaceae

H. Peter Linder*, Christopher R. Hardy, Frank Rutschmann

Institute for Systematic Botany, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland

Received 12 July 2004; revised 3 December 2004
Available online 21 January 2005

Abstract

Three commonly used molecular dating methods for correction of variable rates (non-parametric rate smoothing, penalized likelihood, and Bayesian rate correction) as well as the assumption of a global molecular clock were tested for sensitivity to taxon sampling. The test dataset of 6854 basepairs for 300 terminals includes a nearly complete sample of the *Restio*-clade of the African Restionaceae (272 of the 288 species), as well as 26 outgroup species. Of this, nested subsets of 35, 51, 80, 120, 150, and the full 300 species were used. Molecular dating experiments with these datasets showed that all methods are sensitive to undersampling, but that this effect is more severe in analyses that use more extreme rate smoothing. Additionally, the undersampling effect is positively related to distance from the calibration node. The combined effect of undersampling and distance from the calibration node resulted in up to threefold differences in the age estimation of nodes from the same dataset with the same calibration point. We suggest that the most suitable methods are penalized likelihood and Bayesian when a global clock assumption has been rejected, as these methods are more successful at finding optimal levels of smoothing to correct for rate heterogeneity, and are less sensitive to undersampling.

Keywords: Molecular dating; NPRS; Penalized likelihood; Bayesian dating; Sampling effects; Restionaceae; Lineage through time plots

1. Introduction

Dating the internal nodes of cladograms is useful for many evolutionary investigations, for example exploring plant–insect co-speciation (e.g. Percy et al., 2004), historical biogeographical analysis (e.g., Conti et al., 2002; Davis et al., 2002; Nagy et al., 2003; Vinnersten and Bremer, 2001), and relating speciation rate changes to palaeo-environmental changes (e.g., Kadereit et al., 2004; Linder, 2003). However, molecular dating is beset by a number of problems. For example, the pseudoprecision and errors that may result from the use of inadequate calibration points, and especially the use of derived calibration points which are not directly based on fossil evidence, have recently received attention (Graur and Martin, 2004; Hedges and Kumar, 2004; Lee, 1999; Shaul and Graur, 2002). Furthermore, the assumption of a global molecular clock has been shown to be invalid in many instances (Gaut, 1998). Various methods have been developed to accommodate rate variation: these include the removal of clades with deviant rates (Takezaki et al., 1995), excluding data-partitions that falsify the clock assumption (Kato et al., 2003), using several local clocks for rate-homogenous clades (i.e., the local clocks approach of Yoder and Yang, 2000), using non-parametric rate smoothing to constrain between internode rate variation (Sanderson, 1997), and searching for the optimal rates using Bayesian methods (Thorne et al., 1998) and penalized likelihood (Sanderson, 2002a). However, there seems to have been no investigation into the effects of sampling only a small proportion of the

* Corresponding author. Fax: +41 1 634 8403.
E-mail address: plinder@systbot.unizh.ch (H.P. Linder).

1055-7903/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2004.12.006

terminals (species) on the age estimates of the interior nodes. An understanding of how undersampling effects age estimates is important, as molecular phylogenetic investigations of clade ages are often based on sparse taxon samples.

Here, we investigate the sensitivity of various methods of obtaining molecular age estimates to incomplete taxon sampling in the "*Restio* clade" of African Restionaceae (Poales) which, with 288 species, is the largest clade of African Restionaceae. The African Restionaceae as a whole comprise 350 species of evergreen, rush-like plants that collectively dominate much of the fynbos vegetation of the species-rich Cape Floristic Region of Southern Africa (Linder, 1991, 2003; Taylor, 1978). Specifically, we evaluate effect on node age estimates of increasing or decreasing taxon sampling, and distance from the calibration node. Our data on the *Restio* clade are particularly suited this type of investigation because (1) taxon sampling is nearly complete (ca. 95%) and (2) phylogenetic relationships are well resolved and supported by over 6000 nucleotides of DNA sequence data.

2. Methods

2.1. Phylogeny estimation

Two hundred and seventy-two species (ca. 95%) of the 288 species of the "*Restio* clade" African Restionaceae were included in the current analysis. Additionally, both subspecies of *Restio dodii* and two accessions of the variable and widespread species *Ischyrolepis macer* were included, as they appear to represent two distinct chloroplast lineages and may be separate species. To allow the use of the basal dating node, we also included 24 species of the "*Willdenowia* clade" of African Restionaceae. As such, a total of 298 plants of African Restionaceae were sampled for this analysis. Of the 16 species of the "*Restio* clade" that were not included, three are possibly not taxonomically distinct (for detailed comment, see Linder, 2001), and the remainder could not be located in the field for the collection of extraction-quality plant material. Based on the phylogenetic studies of Briggs et al. (2000) and Linder et al. (2003), the tree was rooted to two terminals representing the ca. 150 species of Australian Restionaceae.

DNA sequences were generated from the plastid regions spanning the *trnL* intron and the *trnL–trnF* intergenic spacer (Taberlet et al., 1991), the complete gene encoding *rbcL* (Chase and Albert, 1998), the complete *atpB-rbcL* intergenic spacer (Chiang and Schaal, 2000; Cuénoud et al., 2000; Manen et al., 1994), and *matK* plus the flanking *trnK* intron (Hilu and Liang, 1997). Total DNA was isolated from silica gel-dried culms using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA). Sequences were generated using standard methods for PCR amplification and automated sequencing.

Raw sequence data files were analysed with the ABI Prism 377 Software Collection 2.1. Contigs were constructed in Sequencher and alignments were performed using the default alignment parameters in Clustal X (Thompson et al., 1997), followed by adjustment by eye. These sequences were assembled into a single matrix in WinClada (Nixon, 2002). The aligned matrix consisted of 6854 aligned bases, of which 1512 are parsimony informative. Additionally, indels were coded at the end of the matrix using Simple Indel Coding (Simmons and Ochoterena, 2000) as implemented in the program GapCoder (Young and Healy, 2001). The total matrix consists of 1782 parsimony–informative characters. All characters were weighted equally and treated as non-additive during tree searches. This data matrix has been deposited at www.treebase.org.

Parsimony searches were conducted using the parsimony ratchet (Nixon, 1999) as implemented from WinClada, running NONA vers. 1.6 (Goloboff, 1993) as a daughter process. Ten ratchet searches were conducted, each initiated with the generation of a Wagner tree, using a random taxon entry sequence, followed by TBR branch swapping with one tree retained and used as the starting point for 500 ratchet cycles. In the weighted/constrained half of each ratchet cycle, a randomly selected set of 10% of the characters were resampled, and a randomly selected set of 10% of the resolved clades were constrained. This analysis resulted in 885 equally most parsimonious cladograms ($L = 5415$, $CI = 0.44$, $RI = 0.84$; informative characters only). These were then pooled and swapped to obtain a total of 10,615 cladograms of length 5415. One of these cladograms was arbitrarily chosen for the subsequent investigation into the impact of taxon sampling on the estimation of absolute dates and divergence times.

2.2. Construction of smaller subset matrices and cladograms

Using our 300 taxon matrix (not including indels) and tree as fixed starting points, six smaller matrices and trees were constructed by deleting terminals in Mesquite 1.02 (Maddison and Maddison, 2003). These smaller datasets have 150, 120, 100, 80, 51, and 35 species/terminals, respectively. The list of species and sequences in each smaller set is a precise subset of the next larger set, and each employed the same relative alignment and tree topology as those obtained from the 300 taxon analysis. The only differences lie in the numbers of terminals and by the exclusion of extraneous gaps from the larger matrices that are no longer necessary in the smaller matrices. As such, each successively smaller matrix consists of 6623, 6547, 6480, 6399, 6248, and 6135 aligned bases. For the smallest (35 species) sampling, at least two

representatives of the basal lineages for each of the 32 clades depicted in Fig. 1 were chosen. Successively larger data matrices and cladograms simply added descendant species and, therefore, more distal nodes to these 32 nodes of interest (the "test nodes"). Thus, the proportion of the descendent species sampled differs enormously among the test nodes, as does the rate at which the sampling density increases (Table 1). This particular strategy was chosen because it results in a set of comparable test nodes for each sampling set (Simmons et al., 2004). Sampling basal lineages is also the method used by phylogeneticists to estimate the age of particular clades with incomplete sampling. Only test nodes 1–30 were used in the analysis. Node 32 is the constrained basal node, and node 31 is at the base of the *Willdenowia* clade, and as such is not part of the study group.

2.3. NPRS, PL, and clock analyses

As preparation for the clock assumption (CL), nonparametric rate smoothing (NPRS, Sanderson, 1997), and penalized likelihood (PL, Sanderson, 2002a,b) approaches to dating, molecular branch lengths were estimated for each of the seven nested matrices and cladograms. We used the implementation of Modeltest (Posada and Crandall, 1998) in Hy-Phy ver. 0.99 beta for Windows (Kosakovsky Pond et al., 2004) to select a statistically adequate model from a set of 56 possible models of sequence evolution. Using the selected models (Table 2), likelihood ratio tests (Felsenstein, 1981) were performed in Hy-Phy to test for a significant departure from the hypothesis of a global molecular clock. In each case, the clock was rejected (Table 2). Branch lengths were estimated in PAUP* 4.0 (Swofford, 2002) using the appropriate model without a clock assumption.

These branches were made ultrametric using NPRS, as implemented in TreeEdit for Macintosh (Rambaut and Charleston, 2004), and penalized likelihood with r8s (ver. 1.6 for Linux, Sanderson, 2002a). For the latter, an optimal rate-smoothing parameter value was selected with the prerequisite cross-validation procedure (Sanderson, 2002a) for all except the largest two matrices (i.e., 150 and 300 terminals). Attempts to find an optimal smoothing parameter for the 150 and 300 taxon sets failed, possibly due to computational limitations with these large datasets, or possibly because of the presence of zero-length terminal branches. For the 120 taxon and fewer datasets, smoothing parameter values ranging from $10^{-3.5}$ to $10^{7.5}$ (in increments of $10^{0.5}$) were tested and the resulting values reported in Table 2. To compare the four methods, the optimal regression of the taxon sampling and average node age values for the 35–120 sample PL method were used to

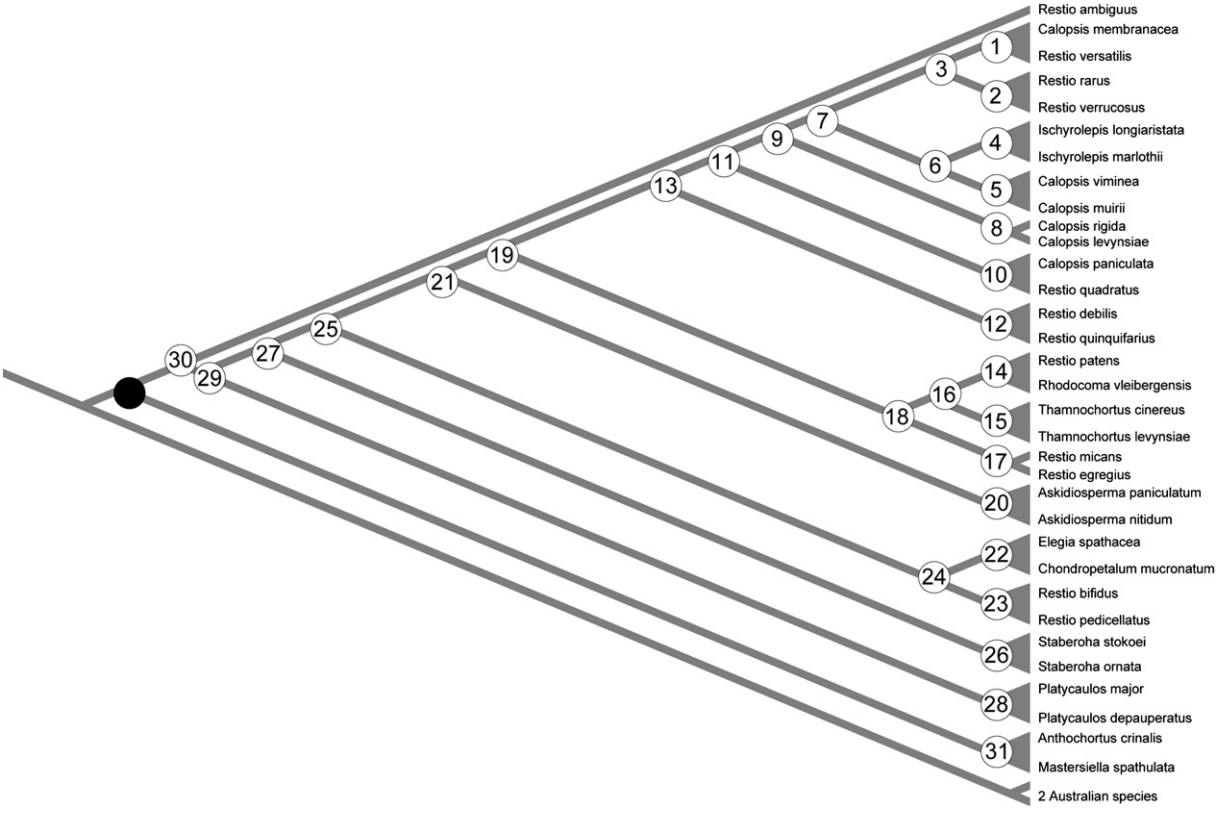

Fig. 1. Cladogram used in this study. Species listed are those used in the 35 taxon analysis. The black circle is the calibration node, the numbered open circles are the test nodes.

Table 1
Percentage of species sampled at each selected node for each sampling run

Nodes	Total species	Taxon						
		35	51	80	100	120	150	300
1	4	50	50	50	50	75	75	100
2	39	5	13	31	33	36	41	100
3	47	9	15	30	34	38	45	100
4	50	4	10	18	24	30	44	100
5	11	18	18	27	36	36	55	100
6	61	7	11	20	26	31	49	100
7	108	7	14	24	30	34	47	100
8	2	100	100	100	100	100	100	100
9	110	9	15	25	31	35	48	100
10	3	67	100	100	100	100	100	100
11	113	11	18	27	33	37	50	100
12	4	50	100	100	100	100	100	100
13	117	12	21	30	35	39	51	100
14	31	6	10	23	35	42	45	100
15	30	7	10	20	33	43	53	100
16	61	7	10	21	34	43	49	100
17	2	100	100	100	100	100	100	100
18	63	10	13	24	37	44	51	100
19	180	11	17	28	36	41	51	100
20	12	17	33	33	33	42	50	100
21	192	11	18	28	35	41	51	100
22	49	4	10	16	22	37	51	100
23	12	17	17	25	33	33	33	100
24	61	7	11	18	25	34	48	100
25	253	10	17	26	33	40	50	100
26	9	22	22	33	33	44	44	100
27	262	11	17	26	33	40	50	100
28	11	18	18	27	36	36	45	100
29	273	11	17	26	33	40	50	100
30	274	11	17	26	33	40	50	100
31	24	8	8	25	29	29	46	100

The second column gives the total number of species subtended by each selected node, the subsequent columns the percent of these species sampled.

predict the values for the 150 and 300 taxon samples. These predicted values were not included in any statistical testing.

Despite rejecting the clock, branch lengths were made ultrametric also under the assumption of a molecular clock, for comparative purposes, using the appropriate model in PAUP*.

We calibrated the trees against node 32 (connecting the *Restio* and *Willdenowia* clades). We used a secondary date (49.8 Ma, with a range of 42.7–55.9 Ma), obtained from the analysis of Linder et al. (2003). In this analysis the adjacent node (connecting the African and Australian Restionaceae) was dated from an African pollen deposit from the earliest Tertiary (Linder et al., 2003; Scholtz, 1985).

2.4. Bayesian dating

The Bayes dating method (Thorne and Kishino, 2002; Thorne et al., 1998) uses a probabilistic model to describe the change in evolutionary rate over time and uses the Markov chain Monte Carlo (MCMC) procedure to derive the posterior distribution of rates and time. It allows multiple calibration windows and provides direct credibility intervals for estimated divergence times and substitution rates. The procedure we followed is divided into three different steps and programs, and is described in more detail in a step-by-step manual available at http://www.plant.ch/software.html. It was performed on a 3 Ghz Pentium IV machine running Windows XP. In a first step, we estimated the model parameters for the the F84 + G model (Felsenstein, 1993; Kishino and Hasegawa, 1989), the most complex model of nucleotide substitution implemented in the software below so far. By using the program *Baseml*, which is part of the PAML package (Yang, 1997), we estimated base frequencies, transition/transversion rate κ, and the α shape parameter (describing the rate heterogeneity among sites under a discrete gamma model; five categories of rates). Then, by using these parameters, we estimated the maximum likelihood of the branch lengths of the rooted evolutionary tree together with a

Table 2
Results of Modeltest, clock tests, and the penalized likelihood (PL) cross-validation procedure for each of the seven nested taxon samplings

	Taxon						
	35	51	80	100	120	150	300
Modeltest results	GTR + G + I	GTR + G + I	GTR + G + I	GTR + G + I	GTR + G + I	GTR + G + I	GTR + I
LR test: clock	Rejected ($\chi^2 = 106.5$, df = 33, $p < 0.01$)	Rejected ($\chi^2 = 145.0$, df = 49, $p < 0.01$)	Rejected ($\chi^2 = 234.0$, df = 78, $p < 0.01$)	Rejected ($\chi^2 = 308.7$, df = 98, $p < 0.01$)	Rejected ($\chi^2 = 329.6$, df = 118, $p < 0.01$)	Rejected ($\chi^2 = 4830.3$, df = 148, $p < 0.01$)	Rejected ($\chi^2 = 7297.2$, df = 298, $p < 0.01$)
PL smoothing value (log 10)	6.5	5.0 (4.5–6.0 were equally optimal)	4.0	5.0	1.5	0.0, 3.5, 6.5[a]	0.0, 3.5, 6.5[a]

[a] Attempts to determine optimal smoothing value failed; therefore, three zero or positive values were chosen (consistent with the range of values determined for the smaller datasets) to bracket the range of probable values.

variance–covariance matrix of the branch length estimates by using the program *Estbranches* (Thorne et al., 1998). The maximum likelihood scores obtained in *Baseml* and *Estbranches* were then compared to check if both approaches were able to optimize the likelihood. The third program we used, *Multidivtime* (Kishino et al., 2001; Thorne and Kishino, 2002), approximates the posterior distributions of substitution rates and divergence times by using a multivariate normal distribution of estimated branch lengths (provided here by *Estbranches*) and running a Markov chain Monte Carlo procedure. Two constraints for the age of node 32 were set: a lower constraint of 42.7 Ma, and a higher one of 55.9 Ma, representing the extreme values obtained for this node by Linder et al. (2003). The other settings for the prior distributions were: 50 for both *rttm* (mean of the prior distribution for the time separating the ingroup root from the present) and *rttmsd* (the prior's standard deviation), 0.004 for both *rtrate* (mean of the prior distribution for the rate of molecular evolution at the ingroup root node, calculated by taking the mean distance between the ingroup root and the ingroup tips obtained from estbranches) and *rtratesd* (the prior's standard distribution). *Brownmean* (the mean of the prior distribution for the Brownian motion parameter ν, which determines the permitted rate change between ancestral and descendant nodes) was initially left at the default value of 0.4. Later, we changed that value to 0.02 and repeated the analysis, following the manual's recommendation that *rttm* multiplied with ν should be about 1. As this did not affect the divergence time estimates significantly, we report here only the results from the first analysis. *Brownsd*, the prior's standard deviation was chosen to be 0.4. For the parameter *bigtime*, a number that should be set higher than the time units between the tips and the root in the user's wildest imagination, we've chosen a value of 100.

We ran the Markov chain for at least 10^4 cycles and collected one sample every 100 cycles, after an unsampled burnin of 10^4 cycles. We performed each analysis at least twice by using different initial conditions to assure convergence of the Markov chain, although it is not possible to say with certainty that a finite sample from an MCMC algorithm is representative of an underlying stationary distribution (Cowles and Carlin, 1996).

2.5. Statistical testing and lineage through time plots

The hypothesis that the test node ages obtained are related to the number of taxa sampled was tested using the Wilcoxon paired sample test, which compares the number of instances in which the larger sample finds an older date compared to the number of times it finds a younger date. The hypotheses that changes in age estimation are related to the distance from the calibration point, and to the degree of sampling of the subclade subtended by each test node, were evaluated using linear regressions. This allowed us to statistically test both the extent to which the variation in the data was accounted for by the regression line, and also whether the slope of the regression line deviates significantly from horizontal. All statistical tests were conducted using SPSS. For each analysis a lineage through time (LTT) plot (Nee et al., 1992) was constructed. The rate constancy of the radiation in Restionaceae was tested using the constant rates test of Pybus and Harvey (2000), as implemented in Gammastatistic v1.0 (Griebeler, 2004).

3. Results

3.1. Undersampling

All four methods find more or less the same ages for the 30 test nodes when only 35 taxa are sampled. However, when more taxa are sampled, the age estimates of the test nodes diverge rapidly (Fig. 2, Table 3). The proportion of species sampled (thus the proportion of nodes distal to the test nodes) clearly has a major impact on the ages estimated for the test nodes (Fig. 3). For all four methods the mean estimated ages of the nodes are significantly less with a sparser taxon sampling than when all taxa are included in the final calculation. Thus, not including all taxa in the sample results in a "younger" estimation of the test node ages. Furthermore, for all four methods the degree of age underestimation increases logarithmically with the proportion of undersampling (Fig. 3).

However, taxon undersampling has very different effects in the four methods. The CL analysis and PL are only slightly affected: for the 35 taxon samples the average ages using CL are 91%, and using PL 88%, of those obtained with the 300 taxon sample. The regression line explains only 73% of the variation in these data for CL, and 72% in the case of PL. Whilst the more severe undersampling in both CL and PL resulted in significant age change, in both there is no significant change in the age estimates between the 150 and 300 taxon samples for CL, suggesting that at 50% sampling an asymptote had been reached, at least for CL (the values for the 150 and 300 taxon samples were inferred for PL, and so the asymptote cannot be calculated).

For Bayesian and NPRS analyses the effects are more dramatic, and no age asymptote is reached. Thus, any change in sampling resulted in significantly different age estimations. For NPRS the 35 taxon sample the average age estimate is only 56% of the estimate with the 300 taxon dataset, and for the fitted regression line $r^2 = 0.9804$. For Bayesian analysis the equivalent values are 72% and $r^2 = 0.9819$.

The regular decrease in the age estimations with decreasing sample size is reflected in the very good fit of the data to the logarithmic regression, and indicates that these patterns are not random.

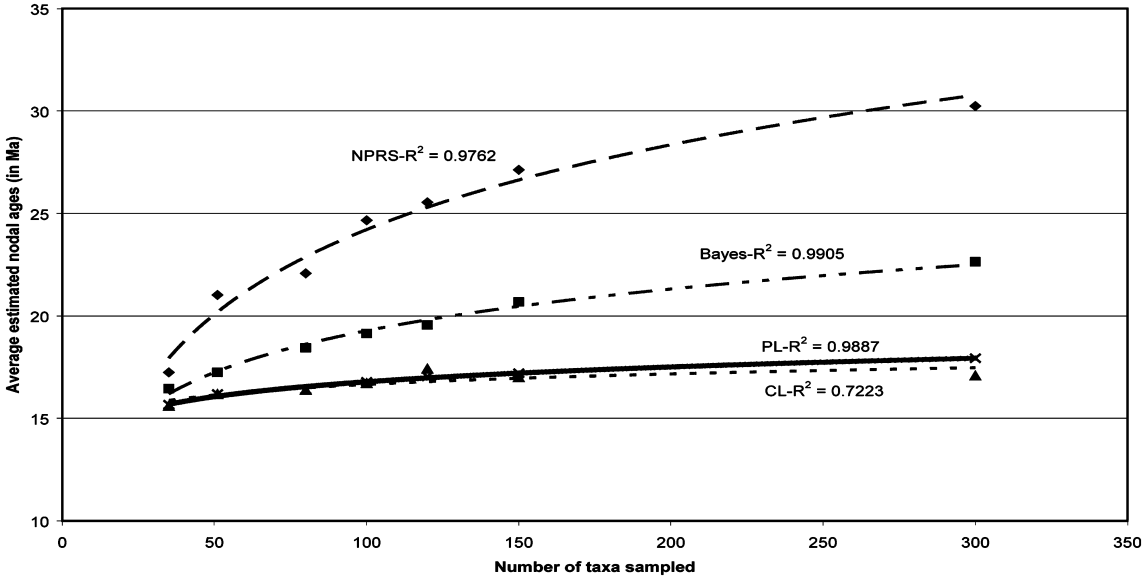

Fig. 2. A comparison of the actual ages returned by the four different molecular dating methods for the seven nested taxon samples. The *x* axis gives the number of taxa sampled, the *y* axis the average age of the 30 test nodes. Diamonds, NPRS; squares, Bayesian; crosses, PL; and triangles, CL.

3.2. Distance from the calibration point

For NPRS and the Bayesian analysis there is a significant (at $p < 0.01$), positive, linear relationship between the degree of underestimation of the age of a node and its distance (time) from the calibration point, for the 35 taxon sample (Fig. 4). The slope of the regression is somewhat steeper for NPRS than for the Bayesian analysis, and also explains more of the variation ($r^2 = 0.9087$ compared to $r^2 = 0.7534$). In both these cases the slope deviates significantly from the horizontal. Thus, the more distant a node is from the calibration point, the more sensitive its age estimation is to the effects of undersampling. And conversely, the closer it is to its calibration point, the less sensitive it is to taxon undersampling. For the NPRS 35-taxon analysis, the most distant test node from the calibration point is dated to only 37.7% of the age indicated by the 300 taxon sample (7 instead of 18.7 Ma for node 5). These effects are much less severe in the Bayesian analysis, where this node is dated to 6.6 instead of 10.4 Ma (an underestimation of 63%). Conversely, the most proximal node in the NPRS 35 taxon analysis is estimated to be 79% of the value of the 300 taxon sample (32.96 instead of 41.66 Ma).

The CL analysis and PL are less sensitive to this distance effect, and in their cases the regression explains very little of the variation ($r^2 = 0.1469$ and $r^2 = 0.0344$, respectively). Interestingly, both these analyses show a much wider scatter, consistent with the assumption of a more clock-like molecular variation. For both these analyses neither the variation explained, nor the deviation of the regression slope from horizontal, is significant.

3.3. Impact of undersampling of individual clades

The sensitivity of the various methods to variation in the sampling density of the clades subtended by each evaluated node cannot be rigorously tested from our data, since most of these clades were rather poorly sampled in the 35 taxon sample. For most clades less than 30% of the species were included, one clade includes 50% of the species, one 63%, and one all species. Nonetheless, it appears as if there is no relationship between the sampling density of the individual clades, and the age estimation of their subtending nodes (Fig. 5) for any of the four methods used. Neither the variance in the data, nor the deviation of slope of the regression line from horizontal, is significant, suggesting that the subclade sampling has no impact on the results. Thus, only the average sampling density on the whole tree under investigation has an impact on the results, not the sampling density of the individual subclades.

3.4. Age estimates by the four methods

The four methods result in rather different age estimates (Table 3, Fig. 6) for the 300 taxon data set. As expected, the extremes are formed by CL and NPRS. The NPRS analysis returns results that are between 10 and 15 million years older (thus up to double) the age estimates of the CL analysis. This suggests that all the nodes have been made older, and this could only be achieved by interpreting the basal branch of the whole tree, between the calibration point and the first speciation events, as being much shorter than the unsmoothed data indicates. Bayesian analysis and PL

Table 3
Mean node ages in millions of years obtained for the four different molecular dating methods with the seven different sampling strategies

Nodes	Methods	Taxon						
		35	51	80	100	120	150	300
1	CL	7.56	7.72	7.98	8.50	9.23	8.56	8.81
	PL	7.65	7.81	7.96	8.60	8.36	8.43	8.80
	Bayes	6.53	7.04	7.99	9.22	8.65	10.77	13.38
	NPRS	7.82	10.26	11.56	13.79	14.13	14.94	19.20
2	CL	10.48	10.61	10.77	10.91	11.09	10.68	11.10
	PL	10.52	10.58	10.69	10.92	10.14	10.23	10.67
	Bayes	9.47	10.04	11.52	12.01	12.19	13.26	15.90
	NPRS	10.40	13.26	15.08	17.96	18.58	20.28	24.81
3	CL	12.12	12.58	12.96	13.15	13.71	13.30	13.61
	PL	12.16	12.52	12.86	13.17	12.70	12.81	13.36
	Bayes	11.09	11.92	13.52	14.06	14.45	15.70	18.20
	NPRS	12.12	15.85	17.62	20.36	21.11	22.98	27.00
4	CL	12.18	13.45	13.34	13.64	14.07	13.76	14.14
	PL	12.40	13.47	13.44	13.56	12.87	12.98	13.54
	Bayes	10.79	11.55	12.81	13.09	13.46	14.80	16.76
	NPRS	11.97	15.49	16.89	18.15	19.31	21.56	25.53
5	CL	7.94	8.27	9.86	10.03	10.31	9.99	9.81
	PL	7.90	8.22	9.74	9.89	9.25	9.33	9.73
	Bayes	6.65	7.12	9.06	9.34	9.64	10.55	13.54
	NPRS	7.00	8.65	11.62	12.01	12.94	14.79	18.72
6	CL	12.77	13.45	13.59	13.64	14.07	13.76	14.14
	PL	12.74	13.47	13.66	13.56	12.87	12.98	13.54
	Bayes	11.47	12.35	13.33	13.72	14.13	15.51	18.61
	NPRS	12.36	15.49	17.16	18.15	19.31	21.56	25.53
7	CL	14.45	15.10	15.28	15.48	16.04	15.58	15.90
	PL	14.47	15.04	15.34	15.49	14.92	15.05	15.70
	Bayes	13.44	14.36	15.59	16.11	16.54	17.85	20.37
	NPRS	14.48	18.84	19.94	22.63	23.45	25.36	28.97
8	CL	9.00	9.18	9.33	9.36	9.71	9.43	13.45
	PL	9.16	9.38	9.49	9.53	9.31	9.39	9.80
	Bayes	8.94	9.48	10.26	10.67	10.96	11.80	13.33
	NPRS	9.76	12.73	13.51	15.56	16.34	18.06	21.83
9	CL	15.12	15.64	15.76	15.93	16.54	16.16	16.62
	PL	15.20	15.53	15.80	15.93	15.58	15.72	16.39
	Bayes	14.45	15.27	16.52	17.01	17.52	18.88	21.58
	NPRS	15.60	19.95	20.87	23.60	24.50	26.38	29.91
10	CL	8.98	9.18	9.25	9.27	9.61	9.29	9.28
	PL	9.00	9.24	9.36	9.40	9.27	9.35	9.75
	Bayes	9.21	9.79	10.66	11.02	11.37	12.29	14.04
	NPRS	9.76	12.88	13.77	15.98	16.83	18.67	22.62
11	CL	15.63	16.09	16.21	16.40	17.05	16.66	17.05
	PL	15.78	16.06	16.30	16.41	16.13	16.27	16.97
	Bayes	15.38	16.14	17.42	17.96	18.53	19.93	22.76
	NPRS	16.41	20.81	21.70	24.51	25.42	27.26	30.70
12	CL	14.20	13.61	14.06	13.90	14.60	14.13	14.05
	PL	14.20	13.86	14.00	14.08	14.38	14.51	15.13
	Bayes	14.50	14.92	16.11	16.72	17.16	18.34	20.16
	NPRS	15.62	19.54	20.43	23.20	24.15	25.95	29.35
13	CL	16.66	17.13	17.64	17.53	18.38	17.96	18.12
	PL	16.80	17.18	17.50	17.61	17.69	17.85	18.61
	Bayes	16.83	17.77	19.13	19.78	20.36	21.77	24.46
	NPRS	17.96	22.65	23.53	26.42	27.34	29.07	32.26
14	CL	10.92	11.23	12.39	13.30	14.25	13.81	12.71
	PL	10.92	11.54	12.82	13.50	14.22	14.35	14.96

(*continued on next page*)

Table 3 (*continued*)

Nodes	Methods	Taxon						
		35	51	80	100	120	150	300
	Bayes	11.66	12.60	14.19	15.43	15.99	16.86	18.53
	NPRS	12.64	15.96	17.66	21.21	22.70	24.15	24.14
15	CL	14.16	14.84	15.21	15.11	15.84	15.58	15.17
	PL	14.29	14.90	15.48	15.62	16.07	16.21	16.91
	Bayes	14.55	15.34	16.35	17.21	17.63	18.55	19.73
	NPRS	16.04	20.01	20.46	23.99	25.43	26.95	29.80
16	CL	15.71	16.36	16.77	17.17	17.90	17.56	16.83
	PL	15.81	16.51	17.05	17.31	18.07	18.23	19.01
	Bayes	16.36	17.31	18.38	19.31	19.79	20.83	22.16
	NPRS	17.85	22.29	22.78	26.15	27.38	28.88	31.58
17	CL	13.57	14.33	14.43	14.35	15.16	15.95	14.58
	PL	13.68	14.11	14.47	14.78	15.40	15.54	16.20
	Bayes	14.58	15.43	16.48	17.09	17.50	18.43	19.11
	NPRS	15.80	19.62	20.39	23.09	24.10	25.46	28.35
18	CL	18.58	19.22	19.68	19.94	21.29	20.61	19.97
	PL	18.75	19.35	19.80	20.18	20.96	21.15	22.05
	Bayes	19.35	20.35	21.70	22.49	23.06	24.25	25.67
	NPRS	21.00	25.55	26.46	29.43	30.39	31.86	34.51
19	CL	19.35	19.99	20.42	20.62	21.77	21.23	21.03
	PL	19.48	20.03	20.43	20.80	21.55	21.74	22.68
	Bayes	20.29	21.31	22.73	23.50	24.08	25.35	27.04
	NPRS	21.73	26.28	27.25	30.15	31.05	32.51	35.17
20	CL	15.59	15.83	16.05	16.17	16.95	16.30	16.16
	PL	15.52	15.83	16.02	16.31	17.05	17.20	17.94
	Bayes	17.02	17.95	19.08	19.78	20.38	21.55	23.09
	NPRS	18.10	22.14	22.88	25.66	26.65	28.26	32.03
21	CL	20.24	20.91	21.15	21.60	22.51	21.98	21.89
	PL	20.20	20.95	21.02	21.62	22.23	22.43	23.39
	Bayes	21.47	22.36	23.82	24.69	25.19	26.48	28.47
	NPRS	22.84	27.25	28.10	30.98	31.82	33.22	35.86
22	CL	12.59	14.62	14.61	15.52	16.12	15.48	15.21
	PL	12.74	14.81	14.85	15.68	15.26	15.39	16.06
	Bayes	13.63	14.96	16.10	16.92	17.55	18.32	22.19
	NPRS	14.72	17.74	18.34	22.07	22.49	24.31	27.53
23	CL	13.04	13.56	13.59	14.49	15.02	14.57	14.93
	PL	13.20	13.97	13.63	14.59	15.57	15.71	16.38
	Bayes	15.43	16.41	17.75	18.42	18.85	19.96	18.29
	NPRS	16.44	20.18	21.40	24.72	25.46	27.21	30.74
24	CL	17.68	19.02	19.37	20.27	21.07	20.84	20.80
	PL	17.86	19.11	19.45	20.29	20.72	20.90	21.80
	Bayes	19.59	20.70	22.29	23.27	23.71	25.21	27.01
	NPRS	20.93	25.24	26.33	29.45	30.19	31.86	34.56
25	CL	20.94	21.88	22.00	22.61	23.56	23.03	23.15
	PL	20.98	21.77	21.93	22.58	23.25	23.46	24.46
	Bayes	22.50	23.54	25.03	25.97	26.49	27.78	30.07
	NPRS	23.70	28.28	29.44	31.98	32.80	34.13	36.69
26	CL	18.11	18.75	18.32	18.76	19.93	19.37	19.88
	PL	18.20	18.64	18.40	18.68	19.98	20.16	21.02
	Bayes	18.97	19.83	21.23	22.06	22.88	24.00	26.16
	NPRS	20.10	24.37	25.39	28.24	29.23	30.66	33.87
27	CL	21.25	22.15	22.23	22.90	23.79	23.20	23.36
	PL	21.28	22.03	22.14	22.78	23.47	23.68	24.70
	Bayes	23.14	24.17	25.68	26.61	27.14	28.44	30.97
	NPRS	23.99	28.55	29.44	32.22	33.04	34.35	36.69

Table 3 (*continued*)

Nodes	Methods	Taxon						
		35	51	80	100	120	150	300
28	CL	22.89	23.32	21.86	22.26	22.91	22.36	22.50
	PL	22.78	22.79	21.58	22.24	23.75	23.96	24.99
	Bayes	26.07	26.69	27.25	27.95	28.17	29.09	30.33
	NPRS	26.91	30.61	30.55	32.65	33.37	34.56	37.31
29	CL	27.10	27.77	27.71	28.39	29.10	28.48	28.24
	PL	26.90	27.28	27.36	28.01	29.15	29.41	30.67
	Bayes	29.99	30.74	31.89	32.86	33.10	34.06	35.49
	NPRS	30.37	34.07	34.83	36.93	37.54	38.51	40.31
30	CL	29.35	29.72	29.80	30.26	31.07	30.47	30.12
	PL	29.44	29.67	29.74	30.23	31.31	31.59	32.95
	Bayes	32.45	33.08	34.18	35.04	35.33	36.21	38.36
	NPRS	32.96	36.16	36.92	38.68	39.24	40.06	41.66

The PL values for 150 and 300 taxon were predicted from the 35 to 120 taxon samples by optimal regression (see section2). CL, assuming a global, constant clock; PL, penalized likelihood; Bayes, Bayesian; NPRS, non-parametric rate smoothing.

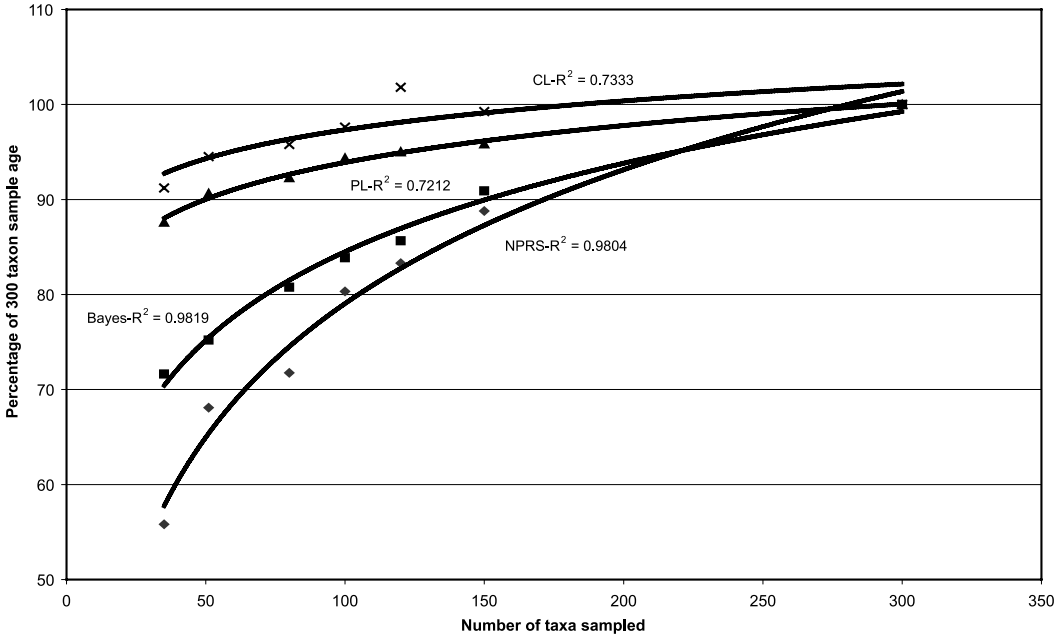

Fig. 3. Effect of the species sampling density on the average test node age estimated, for the four different molecular dating methods. The *x* axis is the taxon sample size. The *y* axis indicates, for each method, the average proportion of the 300 taxon sample age obtained for each species sample size.

return intermediate ages for the nodes. However, for the nodes further from the calibration point, Bayesian ages approach those of PL and CL, while nodes closer to the calibration point are more intermediate. The most extreme disparity is found in the middle section of the tree, for example node 10, which CL dates as 9.28, Bayesian as 14.04 and NPRS as 22.62 Ma, thus more than twofold differences.

The LTT plots for the four methods are remarkably different (Fig. 7). In all analyses the nodes are shifted towards the base of the tree, indicating that with time the speciation rate slowed down. For NPRS the shift is highly significant ($p < 0.01$), for CL and PL weakly significant ($p < 0.05$), and for Bayesian it is not significant.

4. Discussion

The differences in the node ages reported by the different methods, and for different sampling intensities, are remarkably large. On average, the highest age estimate for each node is 2.09 times larger than the smallest estimate, this factor ranges from a minimum of 1.41 to a maximum of 2.94. This indicates a potentially substantial source of error for dating studies. This error is determined by both the overall sampling density and the distance from the calibration node.

Such large differences have been reported before. For example, Klak et al. (2004) reported a twofold difference in the age estimation of the start of the radiation time of the

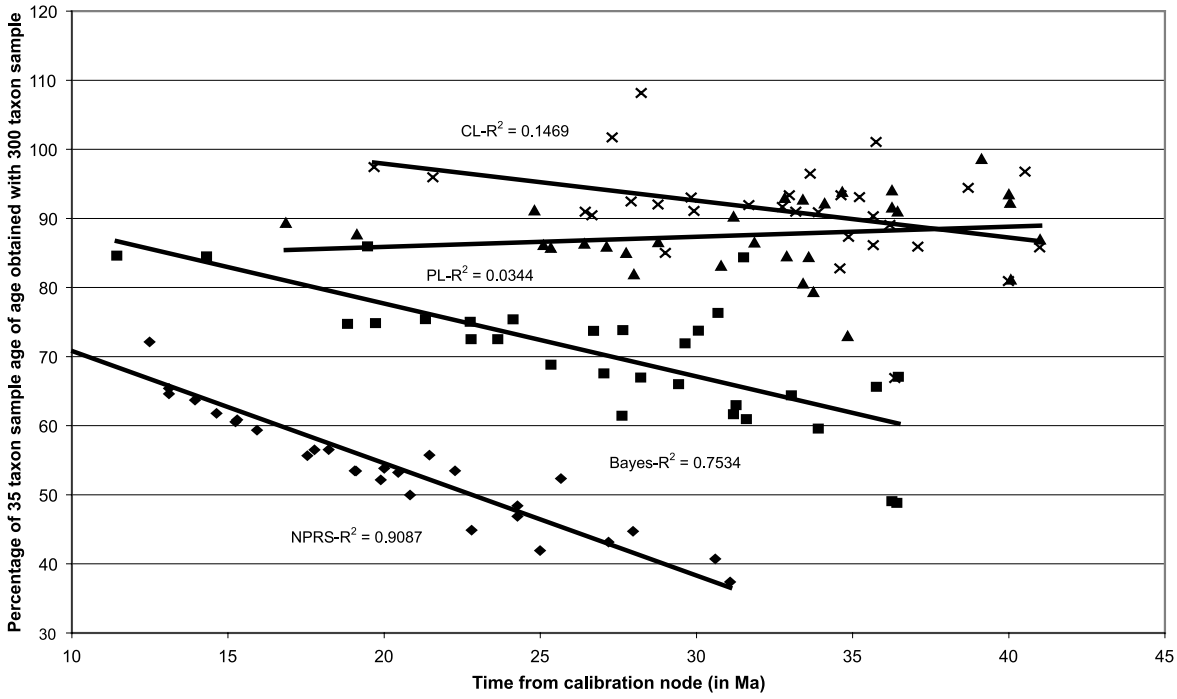

Fig. 4. Proportion of node age underestimation relative to distance from the calibration point. The *x* axis indicates the time between the test node and the calibration node, calculated with each method based on the 300 taxon sample. The *y* axis represents the proportion of the age obtained with the 35 taxon sample of the age obtained with the 300 taxon sample, for each method of analysis. For both Bayes and NPRS the slope deviates significantly (at $p < 0.01$) from 0, and the r^2 is significant at the same p value.

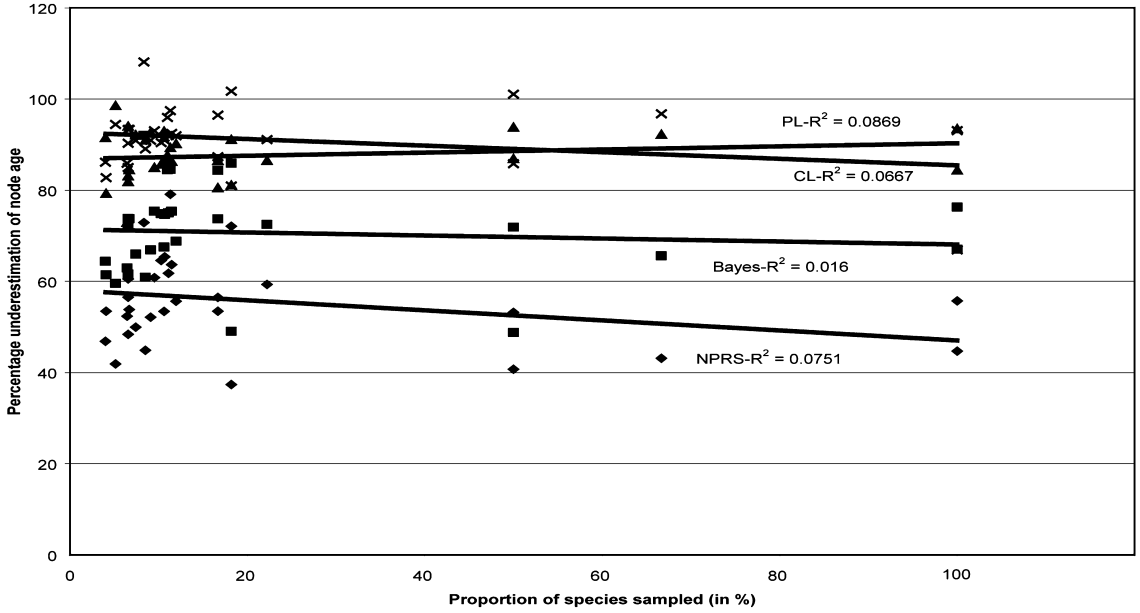

Fig. 5. Differential effects of sampling on each test node individually in the 35 taxon analysis, for the four different molecular dating methods used. The *x* axis indicates the proportion of the species sampled above each node. The *y* axis represents the percentage of age undersampling for each node.

African Rushioideae for two different genes; however, one gene was sampled for twice as many species as the other gene. Our results point to the possibility that this discrepancy may not due to differences in the molecular evolution of the two genes, but to differences in taxon sampling.

4.1. Undersampling

Undersampling has a severe impact on the results in rate-smoothed analyses, and with increasing undersampling the impact rapidly becomes more extreme.

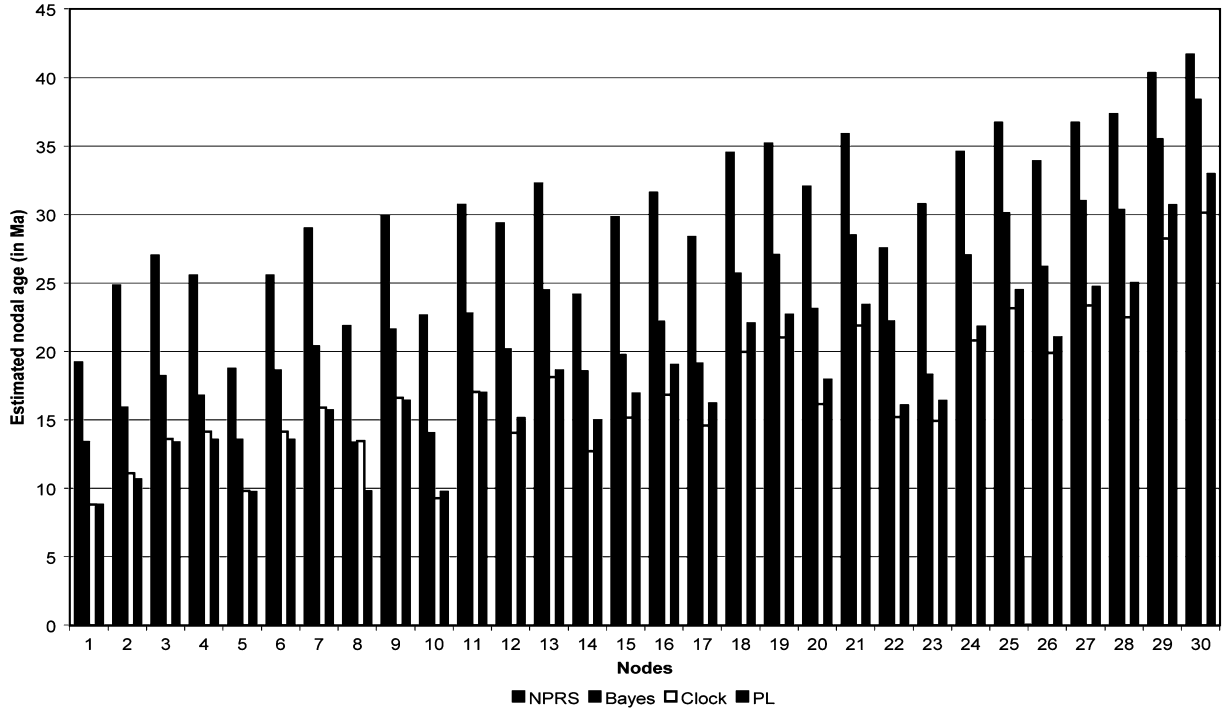

Fig. 6. Comparison of the average nodal ages estimated by the four different molecular dating methods for the 300 taxon sample. The nodes are numbered as in Fig. 1.

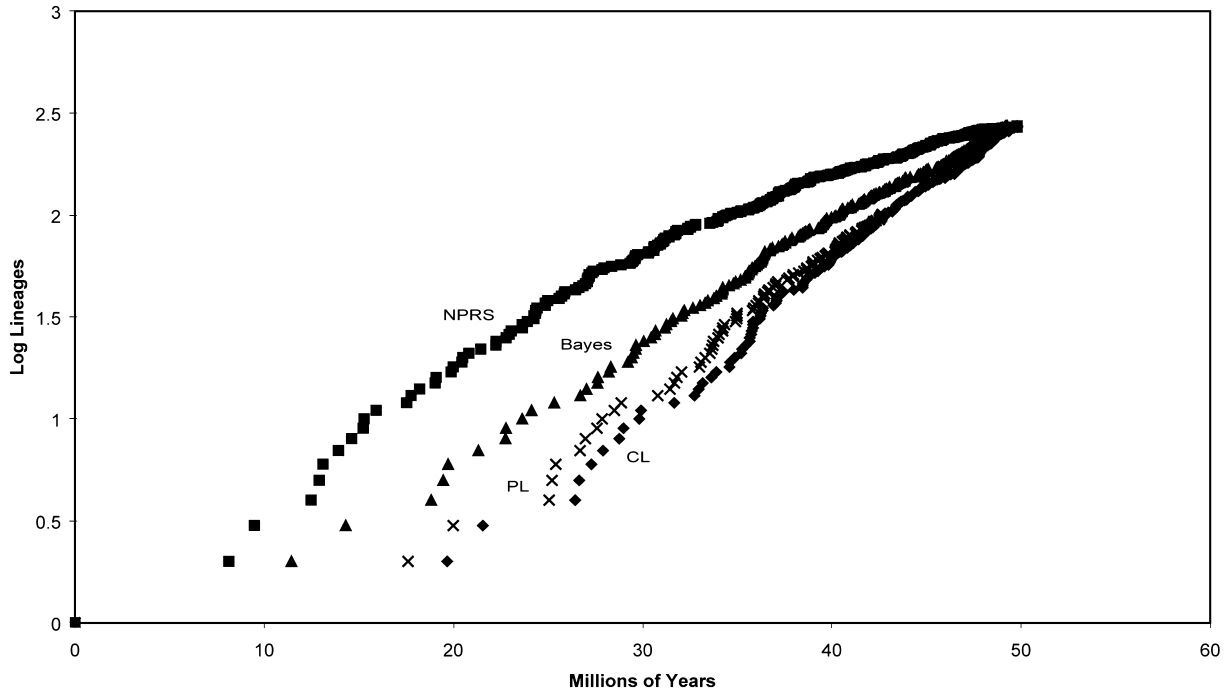

Fig. 7. For the 300 taxon sample, using four different molecular dating methods. Squares, NPRS; triangles, Bayesian; crosses, PL; and diamonds, CL. Note that the PL values for the 300 taxon sample were estimated.

With NPRS, a taxon sampling of 10% can result in age estimates that are half of the correct value (assuming that sampling all species gives the correct value). The fitted curve to the undersampling effect indicates that it is logarithmic, which means that increasing the undersampling might increasingly rapidly exacerbate the age underestimation. Many recent studies included less than 10% of their species, suggesting that they could be prone to the undersampling error.

Although sampling significantly effects analyses that have limited (PL) no (CL) rate smoothing, this effect in our studies was less than 15%. Experiments with PL, involving changing the smoothing parameter, showed that decreasing the smoothing function λ comes with the cost of increased sensitivity to sampling effects (data not shown). It therefore does not automatically follow that the use of PL would eliminate sampling effects. In addition, we do not know what happens when the sampling is less than 10%, and would caution against using these results to suggest that in all circumstances undersampling can be accommodated by using CL or PL.

It is most likely that the effect is not due to simple undersampling. In our experiments we added only nodes that were distal to the test nodes (relative to the calibration node), and it is possible that if nodes were added both proximally and distally to the test nodes, there might be no undersampling effect. Adding only proximal nodes might result in the test nodes being shifted down the tree (further into the past). This shifting of the nodes is manifested as changing age estimates. However, it is difficult to avoid biased sampling. Unbiased sampling is only possible if we know the relative time positions of all the nodes, and can select them to keep the proportions of proximal and distal nodes equal. Unfortunately, we do not know their relative positions without first including all of them in a dating analysis. Random sampling does not help, since randomly selected species bias towards retrieving the deeper nodes (Pybus and Harvey, 2000). Furthermore, most investigators are interested in establishing the age of a preset number of nodes, deep inside the tree (e.g., the starting age of the radiation of Rushioideae (Klak et al., 2004), Angiosperms (Sanderson and Doyle, 2001; Wikström et al., 2001), *Phylica* (Richardson et al., 2001)) and these nodes can only be retrieved if the two basalmost descendents of the node are included in the analysis. This forces an unbalanced sampling.

Furthermore, the vagaries of extinction and speciation are not likely to have left a temporally evenly spaced set of surviving taxa. Clusters of short branch lengths in a phylogeny have been reported in very divergent groups, such as commelinid monocots (Duvall et al., 1993) and sponge-dwelling snapping shrimps (Morrison et al., 2004). Consequently, we cannot establish what a full species sample constitutes, and so the extent of undersampling cannot be determined. Thus, only methods that would minimize the undersampling effect can be considered to be robust. Since all methods we tried showed some undersampling effect, it might be useful to search for an asymptote, as was shown by the CL and to a lesser extent the PL methods.

4.2. Distance from the calibration node

There is a remarkably linear relationship between the degree of under estimation of the test node ages, and the distances from the calibration point for NPRS and the Bayesian analysis. Although a weak trend is visible, there is no significant linear relationship for the CL and PL analyses.

This argues, at least in analyses using NPRS or Bayesian analysis, that the calibration nodes should be situated within the study group, as has been suggested by Shaul and Graur (2002). Arguments for multiple calibration points are usually to protect against errors in single calibration points (Lee, 1999), and a second strong argument is that dated nodes are so placed in the proximity of fixed nodes, thus reducing the error that might accumulate over longer time spans.

4.3. Impact of undersampling of individual nodes

Somewhat surprisingly, we show that the age estimate for a test node is not affected by how complete the sampling is for the clade subtended by the node. Instead, the average level of sampling for the whole data set is of importance. Thus, the sampling effect cannot be avoided by sampling one clade exhaustively, and leaving a number of place-holders for the rest of the study group. Clearly the effects of sampling density are spread more or less evenly across the ingroup. This is advantageous in that groups that are species poor, possibly due to extinction, are not intrinsically impossible to date, but it does argue against strategies of sampling a group of interest in detail, while the related groups are undersampled.

4.4. Age and rate of diversification of Restionaceae

As the sampling becomes better, the age estimates for the African Restionaceae diverge. Thus, more and better data do not result in a convergence to a single possibly correct answer. This indicates that not only taxon sampling, but also the choice of algorithm, is important.

The clock assumption (CL) is fairly robust to undersampling and to distance from the calibration node, and should therefore be used whenever possible. However, if the clock assumption is rejected, as it was for our Restionaceae data, CL cannot be used. On the LTT plot, CL results in a "wobbly" line (Fig. 7), which could either be interpreted as a variable net diversification rate, or as violations of the clock. Since the latter has been demonstrated, we can ignore these results.

NPRS returned remarkably divergent results from the other methods with respect to the estimated age of the nodes. Its great sensitivity to both sampling effects and the distance from the calibration node suggest that the danger of oversmoothed results is real. More remarkable is the effect of NPRS on the LTT plot. Not only does the radiation start earlier than predicted by the other methods, but the initial phases of the radiation are interpreted to be much more rapid than by the other methods, so that a constant rates test shows a significant

change in the net diversification rate. These results may be due to "oversmoothing" (Sanderson, 2002a).

Bayesian and PL were the most resilient to undersampling with our Restionaceae data. Both methods are computationally very intensive, and our 300 taxon Bayesian analysis required more than four weeks of computation time. Possibly PL is the best, since it is not sensitive to the distance from calibration, but we were not able to complete the cross-validation analyses to obtain the optimal smoothing values for the PL analyses of 150 and 300 taxa. The only difference between the two methods is the estimation of the start date of the radiation. Thus, when a global clock assumption is rejected, we recommend that either PL or Bayesian analyses should be used.

4.5. Beyond the Restionaceae

It is difficult to generalize from our results, since they are based on the results of a single study. Generality can be achieved by using simulated data sets, but then we don't know how well simulations will mimic the real situation. We have not attempted to evaluate our results with simulated data sets, largely because of the enormous computing effort that would be needed to analyse a sufficiently large set of replicates. However, our results clearly demonstrate that caution is required when using rate-smoothing methods, and that an understanding of the potential effects of sampling and calibration position on age estimates is a pre-requisite to any study.

Acknowledgments

We are grateful to Barbara Briggs for sequence data of Australian Restionaceae, which were used as outgroups in this study, and to Philip Moline for sequence data for the *Elegia* clade of the African Restionaceae. The study was financially supported by the Swiss Science Fund. Matt Lavin kindly shared his expertise with PL, and the work benefitted from numerous discussions with Elena Conti, Chloé Galley, Philip Moline and Timo van der Niet. We would also like to thank an anonymous reviewer for constructive comments.

References

Briggs, B.G., Marchant, A.D., Gilmore, S., Porter, C.L., 2000. A molecular phylogeny of Restionaceae and allies. In: Wilson, K.L., Morrison, D.A. (Eds.), Systematics and Evolution of Monocots, vol. 1 of Proceedings of the Second International Conference on the Comparative Biology of the Monocots. Sydney, September 1998. CSIRO, Melbourne, pp. 661–671.

Chase, M.W., Albert, V.A., 1998. A perspective on the contribution of plastid *rbcL* DNA sequences to angiosperm phylogenetics. In: Soltis, D.E., Soltis, P.S., Doyle, J.J. (Eds.), Molecular Systematics of Plants II: DNA Sequencing. Kluwer Academic Publishers, Boston, pp. 488–507.

Chiang, T., Schaal, B.A., 2000. Molecular evolution of the *atpB-rbcL* noncoding spacer of chloroplast DNA in the moss family Hylocomiaceae. Bot. Bull. Acad. Sin. 41, 85–92.

Conti, E., Eriksson, T., Schonenberger, J., Sytsma, K.J., Baum, D.A., 2002. Early tertiary out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. Evolution 56, 1931–1942.

Cowles, M.K., Carlin, B.P., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. J. Am. Stat. Assoc. 91, 883–904.

Cuénoud, P., Martinez, M.A.D., Loizeau, P.A., Spichiger, R., Andrews, S., Manen, J.F., 2000. Molecular phylogeny and biogeography of the genus *Ilex* L. (Aquifoliaceae). Ann. Bot. 85, 111–122.

Davis, C.C., Bell, C.D., Mathews, S., Donoghue, M.J., 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proc. Natl. Acad. Sci. USA 99, 6833–6837.

Duvall, M.R., Clegg, M.T., Chase, M.W., Clark, W.D., Kress, W.J., Hills, H.G., Equiarte, L.E., Smith, J.F., Gaut, B.S., Zimmer, E.A., Learn, G.H., 1993. Phylogenetic hypotheses for the monocotyledons constructed from *rbcL* sequence data. Ann. Mo. Bot. Gard. 80, 607–619.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 35, 292–303.

Felsenstein, J., 1993. Phylogeny Inference Package (PHYLIP). University of Washington, Seattle.

Gaut, B.S., 1998. Molecular clocks and nucleotide substitution rates in higher plants. Evol. Biol. 30, 93–120.

Goloboff, P.A., 1993. NONA. Published Privately, New York.

Graur, D., Martin, W., 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends Genet. 20, 80–86.

Griebeler, E.M., 2004. Gammastatistic V 1.0. Institut für Zoologie, Abteilung Ökologie, Johannes Gutenberg-Universität, Germany, Mainz. Available from: <http://www.oekologie.biologie.uni-mainz.de/people/evi/download/gammastatistic>.

Hedges, S.B., Kumar, S., 2004. Precision of molecular time estimates. Trends Genet. 20, 242–247.

Hilu, K.W., Liang, H.P., 1997. The *matK* gene: sequence variation and application in plant systematics. Am. J. Bot. 84, 830–839.

Kadereit, J.W., Griebeler, E.M., Comes, H.P., 2004. Quaternary diversification in European alpine plants: pattern and process. Philos. Trans. R Soc. Lond. B Biol. Sci. 359, 265–274.

Kato, Y., Aioi, K., Omori, Y., Takahata, N., Satta, Y., 2003. Phylogenetic analyses of *Zostera* species based on *rbcL* and *matK* nucleotide sequences: Implications for the origin and diversification of seagrasses in Japanese waters. Genes Genet. Syst. 78, 329–342.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170–179.

Kishino, H., Thorne, J.L., Bruno, W.J., 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18, 352–361.

Klak, C., Reeves, G., Hedderson, T.A., 2004. Unmatched tempo of evolution in Southern African semi-desert ice plants. Nature 427, 63–65.

Kosakovsky Pond, S.L., Muse, S.V., Frost, S.D.W., 2004. Hyphy package: hypothesis testing using phylogenies. Available from: <http://www.hyphy.org>.

Lee, M.S.Y., 1999. Molecular clock calibrations and metazoan divergence dates. J. Mol. Evol. 49, 385–391.

Linder, H.P., 1991. A review of the southern African Restionaceae. Contrib. Bolus Herb. 13, 209–264.

Linder, H.P., 2001. The African Restionaceae. Contributions from the Bolus Herbarium, Cape Town.

Linder, H.P., 2003. The radiation of the Cape flora, southern Africa. Biol. Rev. 78, 597–638.

Linder, H.P., Eldenäs, P., Briggs, B.G., 2003. Contrasting patterns of radiation in African and Australian Restionaceae. Evolution 57, 2688–2702.

Maddison, W.P., Maddison, D., 2003. Mesquite. Available from: <www.gnu.org>.

Manen, J., Natali, A., Ehrendorfer, F., 1994. Phylogeny of Rubiaceae–Rubieae inferred from the sequence of a cpDNA intergene region. Plant Syst. Evol. 190, 195–211.

Morrison, C.L., Rios, R., Duffy, J.E., 2004. Phylogenetic evidence for an ancient rapid radiation of Caribbean sponge-dwelling snapping shrimps (Synalpheus). Mol. Phylogenet. Evol. 30, 563–581.

Nagy, Z.T., Joger, U., Wink, M., Glaw, F., Vences, M., 2003. Multiple colonization of Madagascar and Socotra by colubrid snakes: evidence from nuclear and mitochondrial gene phylogenies. Proc. R. Soc. Lond. Ser. B Biol. Sci. 270, 2613–2621.

Nee, S., Mooers, A.O., Harvey, M., 1992. Tempo and mode of evolution revealed from molecular phylogenies. Proc. Natl. Acad. Sci. USA 89, 8322–8326.

Nixon, K.C., 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15, 407–414.

Nixon, K.C., 2002. WinClada. Published by the author, Ithaca, NY.

Percy, D.M., Page, R.D.M., Cronk, Q.C.B., 2004. Plant–insect interactions: double-dating associated insect and plant lineages reveals asynchronous radiations. Syst. Biol. 53, 120–127.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lond. Ser. B Biol. Sci. 267, 2267–2272.

Rambaut, A., Charleston, M., 2004. TreeEdit: an application for organising, viewing and manipulating sets of phylogenetic trees. Published privately, Oxford. Available from: <http://evolve.zoo.ox.ac.uk/software.html?id = treeedit>.

Richardson, J.E., Weitz, F.M., Fay, M.F., Cronk, Q.C.B., Linder, H.P., Reeves, G., Chase, M.W., 2001. Rapid and recent origin of species richness in the Cape flora of South Africa. Nature 412, 181–183.

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

Sanderson, M.J., 2002. r8s. Published privately, Davis. Available from: <http://ginger.ucdavis.edu/r8s/>.

Sanderson, M.J., Doyle, J.A., 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. Am. J. Bot. 88, 1499–1516.

Scholtz, A., 1985. The palynology of the upper lacustrine sediments of the Arnot Pipe, Banke, Namaqualand. Ann. S. Afr. Mus. 95, 1–109.

Shaul, S., Graur, D., 2002. Playing chicken (*Gallus gallus*): methodological inconsistencies of molecular divergence date estimates due to secondary calibration points. Gene 300, 59–61.

Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.

Simmons, M.P., Pickett, K.M., Miya, M., 2004. How meaningful are Bayesian support values?. Mol. Biol. Evol. 21, 188–199.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Gielly, L., Patou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.

Takezaki, N., Rzhetsky, A., Nei, M., 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12, 823–833.

Taylor, H.C., 1978. Capensis. In: Werger, M.J.A. (Ed.), Biogeography and Ecology of Southern Africa. Junk, The Hague, pp. 171–229.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Thorne, J.L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51, 689–702.

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.

Vinnersten, A., Bremer, K., 2001. Age and biogeography of major clades in Liliales. Am. J. Bot. 88, 1695–1703.

Wikström, N., Savolainen, V., Chase, M.W., 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. Lond. Ser. B Biol. Sci. 268, 2211–2220.

Yang, Z., 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. CABIOS or Computer Applications in the Biosciences 13, 555–556.

Yoder, A.D., Yang, Z.H., 2000. Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17, 1081–1090.

Young, N.D., Healy, J., 2001. GapCoder: a computer program for including indels in phylogenetic analysis. Published privately, Available from: <http://www-home.cr.duq.edu/~youngnd/GapCoder>.