

1. Phylogeny reconstruction
2. The models of the distribution of branch-lengths
3. Statistical modelling
4. Branch-lengths
5. Estimation of statistical parameters
6. Determining ranges of priors, and results using maximum likelihood.
7. Long branches and results with short branches removed.
8. Papers from which datasets were taken

### 1. Phylogeny reconstruction

We used a phylogenetic mixture-model<sup>1,2</sup> to infer a posterior sample of 750 phylogenetic trees for each of the 101 gene-sequence datasets. Datasets varied from 9-229 taxa (mean = 42.9), with forty-seven sampling species from within a genus, forty-one within a family and thirteen within an order. See the below for a list of papers from which our datasets were taken. We ran a number of independent Markov chains to apparent convergence (at least 2,500,000 iterations) before sampling trees at widely spaced intervals (10,000 iterations) to ensure independence among successive trees in the sample. Trees were rooted for subsequent analyses, using the outgroup taxa reported in the source article. Outgroups were removed before analyzing the trees.

All the phylogenies included in this study were shown to be free from a well known artifact of phylogeny reconstruction called the node-density effect, using the methodology outlined in<sup>3,4</sup>, any samples of trees which showed evidence of the artifact were not included.

### 2. The models of the distribution of branch-lengths

**Exponential:** *Constant rate of net-speciation* – In a phylogenetic context, the probability distribution of branch lengths,  $x$ , under a model in which species have a fixed net-rate of speciating (speciation-extinction) is:

$$p(x) = \frac{1}{\beta} e^{-x/\beta}$$

Where  $\beta$  is the mean or net-speciation rate.

**Variable-Rates:** *Variable rates of net-speciation* – if the speciation rate,  $r$ , varies from species to species, different branches of a phylogenetic tree would be drawn from different exponential distributions. If we assume that speciation rates are distributed according to a gamma probability density we can find the probability density of branch lengths from integrating the exponential over all possible speciation rates,  $r$ , yielding<sup>5,6</sup>:

$$p(x) = \frac{\alpha\beta}{(1 + \beta x)^{1+\alpha}}$$

Here  $\alpha$  and  $\beta$  are the shape and scale parameters, respectively, of the gamma probability density, with mean equal to  $\alpha\beta$  and variance given by  $\alpha\beta^2$ .

**Weibull:** *net-speciation rate dependent on time from ancestral species* – if net-speciation rate decreases as divergence from the ancestral species increases,  $\alpha < 1$ . If the net speciation rate increases as divergence increases then  $\alpha > 1$ . When  $\alpha = 1$  the Weibull reduces to the exponential:

$$p(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha}$$

The parameter  $\beta$  is the scale of the distribution.

**Normal:** if the net-speciation rate is caused by the cumulative effects of many factors that act additively, we would expect the distribution of branch lengths to be approximately normal:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)/\sigma]^2/2}$$

Where  $\mu$  is the mean divergence until speciation, and  $\sigma^2$  is the variance.

**Log-normal:** if net-speciation rate is determined by the cumulative effects of many multiplicative (rather than additive) factors, we would expect the branch lengths to be log-normally distributed:

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-[\ln x - \mu]^2/2\sigma^2}$$

where  $\mu$  and  $\sigma^2$  are the mean and variance of the logarithm of the branch lengths.

### 3. Statistical modelling

We fitted the statistical models in two ways. In one, we implemented a reversible-jump<sup>7</sup> Markov chain that moved among the five statistical models estimating their parameters while simultaneously jumping among trees in the posterior sample (Methods and Supplementary Information). In the second method, we fitted each model separately in its own Markov chain that estimated the parameters of the statistical model while moving among trees, recording the harmonic mean of the likelihoods (based on samples from chains that were run for  $1.875 \times 10^9$  iterations) as an estimate of each model's marginal density<sup>7</sup>. In both approaches we used Bayesian prior distributions chosen to favour the four two-parameter models (see below)

**4. Branch lengths.** We analyze branch-lengths in units of expected number of substitutions in preference to scaling them to 'time', for three reasons. One is that estimates of branch lengths in phylogenies are inaccurate when the phylogenetic tree is inferred using poorly-specified or under-specified models of sequence evolution. The mixture-model approach we used to infer trees has been shown to improve the estimation of branch lengths over standard single-rate-matrix models, and over partitioning of the data<sup>8,9</sup>, but is not implemented in packages that can be used to infer trees scaled to time. Another is that no existing model of sequence evolution accounts for the rapid bursts of evolution that are frequently associated with speciation<sup>3,10</sup>. If the substitutions these bursts account for are scaled to time at some uniform rate along with the remaining substitutions in a branch, the estimated times will be wrong (branch too long). The third reason derives from the recent finding that rates of sequence evolution often appear to have been far more rapid in recent time intervals than in the past, the so-called 'j-curve' of rates versus time<sup>11</sup>. There is not at present any general correction for this effect. The form of any correction that is proposed will have to derive from empirical estimates for given genes and taxa, and making use of several calibration points over the period of recent evolution. This is because methods for scaling substitution-based branch lengths to time can only act on the differences among branches in their length, and so will not have information about generalized or taxa-wide effects on rates of change that are related to the absolute time in the past that events occurred.

Branches on a phylogeny will be lengthened by the loss of taxa owing to extinction or

because of missed samples. Where such events have occurred, the observed branches will be sums of shorter segments. It can be shown (as kindly pointed out to us by Mike Steel) that if the underlying true segments are exponentially distributed, then the observed lengths will under a reasonable set of assumptions also follow an exponential. Consider that extinctions and missed samples are distributed geometrically, with  $q$  denoting the probability that a given node along an internal branch has been missed (either from extinction or sampling) and  $p = 1 - q$ . Then the probability that an observed internal branch is actually the sum of  $j$  internal branches is  $pq^{j-1}$ , for  $j = 1, 2, \dots$ . Let  $N$  denote the random number of true segments that contribute to each observed internal branch. By assumption,  $N$  is geometric with mean  $1/p$ . If we assume that the true segment lengths, denoted  $X_i$ , are independent and identically distributed as exponential random variables with mean  $\lambda$ , the observed internal branch lengths, denoted  $Y$ , can be represented as

$$Y = \sum_{i=1}^N X_i$$

Note that each  $X_i$  has moment generating function  $M_X(t) = E(e^{tX}) = 1/(1 - \lambda t)$ . Hence, using a standard conditional probability calculation, we see that the moment generating function of  $Y$  is simply

$$M_Y(t) = E(e^{tY}) = E[E(e^{tY} | N)] = E\left[\left(\frac{1}{1 - \lambda t}\right)^N\right] = \frac{1}{1 - \left(\frac{\lambda}{p}\right)t}.$$

This shows that the observed branch lengths are also exponentially distributed, with mean equal to the product of the average segment length and the mean number of segments.

**5. Estimation of statistical parameters.** We constructed Markov chains to estimate the parameters of the statistical distributions (Table 1) and to compare their posterior likelihoods. We did this two ways. In one, we implemented a Bayesian reversible-jump<sup>7</sup> procedure that moves among the five statistical models estimating their parameters while simultaneously jumping among trees in the posterior sample. At each iteration of the Markov chain we either propose new values of the statistical parameters of the current model, propose a different model (along with proposing new parameters for it), or propose a jump to a new tree in the

Bayesian posterior sample of trees previously inferred for that dataset. In the usual way, the chain moves to a newly proposed model (or set of parameters or tree) if the proposed state improves on the current state of the chain; otherwise, the proposed state is accepted with probability proportional to the ratio of its likelihood to that of the previous model in the chain.

When the reversible-jump model is allowed to run for many iterations, the proportion of time the chain spends in each model measures that model's relative probability of describing those data. Formally, this proportion is an estimate of the percentage of the marginal probability of the data that that model accounts for, where the marginal probability is the probability of the data integrated over all possible values of the various models' parameters and trees in the sample.

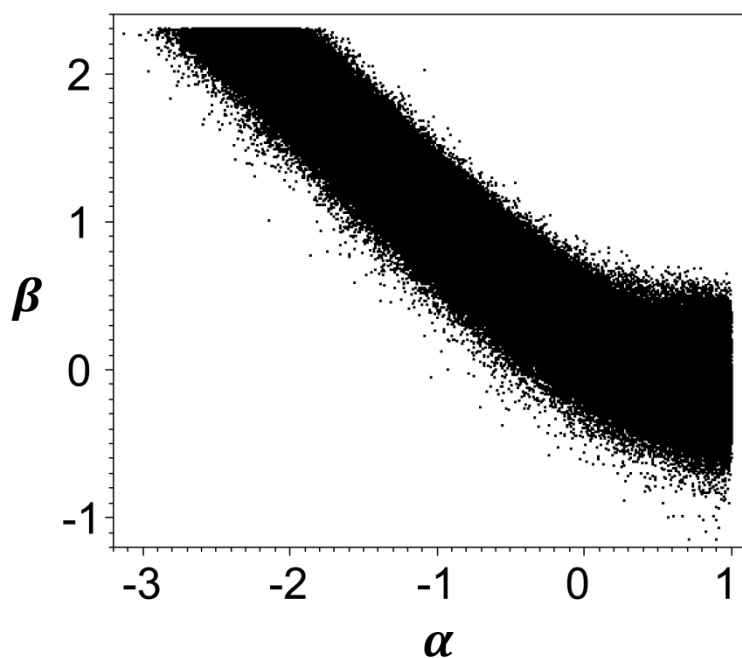
We simulated Weibull distributed data to test the sensitivity of our reversible-jump procedure to departures from the exponential distribution. The Weibull distribution provides a stringent test because it becomes identical to the exponential when its shape parameter  $\alpha$  takes the value 1.0. Simulations with an average tree size for our datasets ( $n=40$  taxa) showed that the reversible-jump procedure rejected the exponential model for values of the Weibull shape parameter less than 0.8 or greater than 1.2. For values of  $\alpha$  in the range  $1\pm 0.2$  the Weibull and exponential are nearly indistinguishable. For the other distributions, the reversible-jump algorithm favoured the exponential even when the distributions were manipulated to have mean/variance ratios that conform to the exponential.

In the second method, we fitted each model separately in its own Markov chain that estimated the parameters of the statistical model while moving among trees, and keeping track of the harmonic mean of the likelihoods. Because harmonic means can be unstable we ran the chains a minimum of  $1.875 \times 10^9$  iterations. The harmonic mean estimates a model's marginal density<sup>7</sup> integrated over trees and parameters of the model, and we chose the model with the best marginal density for each dataset.

**6. Determining ranges, priors and results using maximum likelihood.** When a new model is proposed in a Markov chain the acceptance probability rule must keep track of the product of the model's likelihood and its prior probability. The acceptance rule then compares these two products as above to determine whether to stay in the current model or move to the proposed model. Four of the models have two parameters and so the prior probabilities of these models is given by the product of the two separate priors for their parameters. The exponential has just one prior.

It is necessary to specify the distribution from which the parameters of each model are drawn, in order to calculate their prior probabilities. The choice of priors influences how often a model is visited, because the priors act to 'penalise' a model for having an extra parameter. The wider or less certain the prior is, the greater is the penalty. To give the more complex models every chance, we chose lenient (narrow) priors as follows. We used flat or uninformative priors (uniform distribution) for all our proposal mechanisms. To establish the range of these uniform priors we conducted a series of analyses with each model using priors on a 0-100 interval (with the exception of the lognormal  $\mu$  parameter for which we used a prior on a -100 to 100 interval) and then inspected the posterior distributions of the models' parameters. These posterior distributions were well behaved for the exponential, Weibull, log-normal and normal models. For these models, then, we chose our priors to be equal to these posteriors, with most on a uniform scale of width three.

The variable-rates model is an exception. This model is difficult to estimate because its two parameters have a tendency to trade-off producing a line of approximately equal likelihood over a wide range of the parameters (Figure S1, axes on log-scales). Thus a high value of the  $\beta$  parameter and a low value of  $\alpha$  can produce a likelihood that is equivalent to that from a low value of  $\beta$  and a high value of  $\alpha$ . The trade-off can be accommodated by restricting the model to explore just one of the two equivalent combinations. Based on a number of trials we discovered that by searching for  $\alpha$  in the range 0-10 and  $\beta$  in the range 0-200 we could describe the data with no overall change to the likelihood. Even this prior range, however, is wide and its cost of roughly 5.3 log-units for the  $\beta$  parameter would greatly disadvantage the variable rates model against the exponential. The trade-off curve however provides a way to remove this burden. Given that the value of  $\beta$  is linked to the value of  $\alpha$ , we treated the variable rates model as a one-parameter model, charging only for the  $\alpha$  parameter using a uniform 0-10 prior.



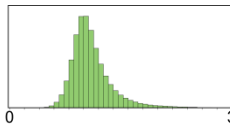
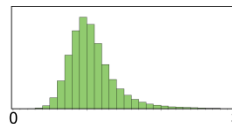
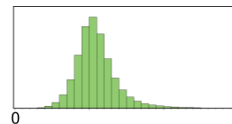
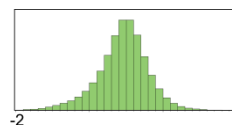
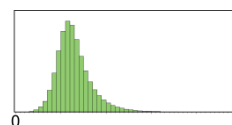
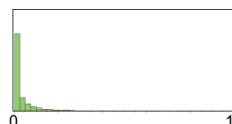
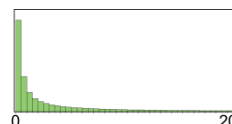
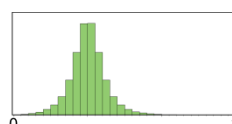
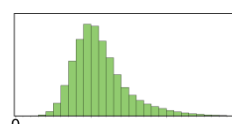
**Figure S1 | Scatterplot of the parameters of the variable-rates model, sampled from a Markov chain at convergence.** The two parameters trade-off such that high values of  $\alpha$  and low values of  $\beta$  return similar likelihoods to the opposite configuration.

The ranges of the priors for the various models are shown in Table S1 along with the corresponding posterior distributions of the model parameters. All priors were drawn from uniform distributions but yielded posteriors that moved away from the priors. Our procedures for choosing the priors ensured that they were as narrow as possible for the two-parameter models, without constraining the parameter space the model could explore, and translates into the two-parameter models having less prior weight to overcome than if we had set priors from the usual position of relative ignorance. The average prior cost we assessed the two-parameter models translates to having to overcome a ‘debt’ of about 1.1 log-units. That is, to perform better than the exponential the two-parameter model would need to improve the log-likelihood by this amount. This is less than a typical likelihood ratio test, which uses a criterion of 1.92 log-units per parameter, or an Akaike Information Criterion test which penalises models 2 log-units per parameter.

Table S2 compares our results using these priors to those obtained using maximum likelihood to estimate the parameters of the models, and then choosing among models using the likelihood ratio statistic and requiring the conventional amount of an additional 1.96 log-units for each extra parameter in a model. The differences are small but in the expected direction of the maximum likelihood analyses finding more support for the exponential model.

It has not escaped our notice that our results have implications for Bayesian MCMC phylogenetic reconstruction methods. Many MCMC phylogenetic inference programs use an exponential prior on branch lengths – our findings support this choice.

**Table S1** | Shows the models, their parameters and the prior ranges from which the parameters for each model were drawn (all priors were based on uniform distributions). The histograms show the observed posterior distribution of parameters summed over all of the datasets.

Model	Parameters	Prior ranges (above, uniform distributions) and (below) observed posterior distributions	
Exponential $p(x) = \frac{1}{\beta} e^{-x/\beta}$	$\beta$ =scale	$\beta = 0$ to 3 	
Weibull $p(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha}$	$\alpha$ =Shape, $\beta$ =scale, when $\alpha=1$ Weibull = Exponential	$\alpha = 0$ to 3 	$\beta = 0$ to 3 
Log-normal $p(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\ln x - \mu)^2 / 2\sigma^2}$	$\mu$ =mean of $\ln x$ , $\sigma^2$ =variance	$\mu = -2$ to 1 	$\sigma^2 = 0$ to 5 
Variable Rates $p(x) = \frac{\alpha\beta}{(1 + \beta x)^{1+\alpha}}$	$\alpha$ =Shape, $\beta$ =scale (of the gamma distribution of rates)	$\alpha = 0$ to 10 	$\beta = 0$ to 200* 
Normal $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)/\sigma]^2 / 2}$	$\mu$ =mean, $\sigma^2$ variance	$\mu = 0$ to 3 	$\sigma^2 = 0$ to 3 

\*but see text.



**Table S2** | Percentage of datasets for which each model provided the best description of the branch-length distribution

	Exponential	Weibull	Variable Rates	Lognormal	Normal
Reversible-Jump	78	8	6	8	0
Harmonic mean	79	6	14	1	0
Maximum Likelihood	81	4	3	12	0

**7. Long branches and results with short branches removed.** Underestimation of long branches could bias against long-tailed distributions such as the log-normal, Weibull or variable-rates, but we think our phylogenetic methodology makes this unlikely and Figure 2d (in the main text) reveals no shortage of long branches

The exponential and variable-rates models expect very short branches but the lognormal does not. If very short branches are poorly estimated in the phylogenetic inference, this could bias results in favour of the exponential. To ensure that we did not have a bias towards short branches we used uniform priors on branch lengths at the phylogenetic inference step (it makes no difference to our results if we use the conventional exponential prior on branch lengths). Before fitting the statistical models, we then removed all branches with an expectation of having less than or equal to 0.5 nucleotide substitutions per branch. Table 3S shows the results with these branches removed – they do not qualitatively differ from the complete analysis and only the exponential model achieves more than the null expectation of fitting 20% of the datasets.

**Table S3** | Percentage of datasets for which each model provided the best description of the branch-length distribution when very short branches were removed.

	Exponential	Weibull	Variable Rates	Lognormal	Normal
Reversible-Jump	72	9	0	19	0
Harmonic mean	75	11	7	7	0

## 8. Papers from which datasets were taken.

- U. Amason *et al.*, *Proc Natl Acad Sci U S A* 99, 8151 (2002).
- C. M. Austin, N. P. Murphy, *Zoologica Scripta* 34, 187 (2005).
- N. P. Barker, A. Vanderpoorten, C. M. Morton, J. P. Rourke, *Mol Phylogenet Evol* 33, 845 (2004).
- C. D. Bell, R. Patterson, L. Hamilton, A, *Systematic Botany* 24, 632 (1999).
- M. Beltran *et al.*, *Mol Biol Evol* 19, 2176 (2002).
- F. R. Blattner, *Mol Phylogenet Evol* 33, 289 (2004).
- C. Bleidorn, L. Vogt, T. Bartolomaeus, *Mol Phylogenet Evol* 34, 673 (2005).
- M. C. Brandley, A. Schmitz, T. W. Reeder, *Syst Biol* 54, 373 (2005).
- E. S. Bridge, A. W. Jones, A. J. Baker, *Mol Phylogenet Evol* 35, 459 (2005).
- L. Bromham, M. Woolfit, *Syst Biol* 53, 758 (2004).
- W. R. Buck, B. Goffinet, A. J. Shaw, *Mol Phylogenet Evol* 16, 180 (2000).
- T. R. Buckley, C. W. Cunningham, *Mol Biol Evol* 19, 394 (2002).
- J. Bunsawat, N. Elliott, K. Hertweck, E. Sproles, A. L. A., *Systematic Botany* 29, 959 (2004).
- A. Burk, M. Westerman, M. S. Springer, *Syst Biol* 47, 457 (1998).
- J. J. Clarkson *et al.*, *Mol Phylogenet Evol* 33, 75 (2004).
- W. L. Clement *et al.*, *American Journal of Botany* 91, 905 (2004).
- B. L. Cohen, S. Stark, A. B. Gawthrop, M. E. Burke, C. W. Thayer, *Proc Biol Sci* 265, 475 (1998).
- J. Crayan, R., *Systematic Entomology* 30, 563 (2004).
- B. J. Crespi, C. P. Sandoval, *Journal of Evolutionary Biology* 13, 249 (2000).
- R. H. Cruickshank, R. Thomas, H., *Evolution Int J Org Evolution* 53, 1796 (1999).
- R. L. Cunha, R. Castilho, L. Ruber, R. Zardoya, *Syst Biol* 54, 634 (2005).
- B. N. Danforth, *Proc Natl Acad Sci U S A* 99, 286 (2002).
- D. L. Des Marais, A. R. Smith, D. B. Britton, K. M. Pryer, *International Journal of Plant Sciences* 164, 737 (2003).
- M. Dorris, M. E. Viney, M. L. Blaxter, *Int J Parasitol* 32, 1507 (2002).
- S. Ekman, *Mycological Research* 105, 783 (2001).
- D. J. Emlen, J. Marangelo, B. Ball, C. W. Cunningham, *Evolution Int J Org Evolution* 59, 1060 (2005).
- T. Eriksson, M. Hibbs, S., A. D. Yoder, C. Delwiche, F., M. J. Donoghue, *international Journal of Plant Sciences* 164, 197 (2003).
- B. J. Evans, D. B. Kelley, R. C. Tinsley, D. J. Melnick, D. C. Cannatella, *Mol Phylogenet Evol* 33, 197 (2004).
- M. Fishbein, C. Hirsch-Jetter, D. E. Soltis, L. Hufford, *Syst Biol* 50, 817 (2001).
- P. K. Flook, S. Klee, C. H. Rowell, *Syst Biol* 48, 233 (1999).
- J. J. Flynn, J. A. Finarelli, S. Zehr, J. Hsu, M. A. Nedbal, *Syst Biol* 54, 317 (2005).
- M. Friedrich, D. Tautz, *Syst Biol* 46, 674 (1997).
- P. W. Fritsch, F. Almeda, S. S. Renner, A. B. Martins, C. B., *Am J Bot* 91, 1105 (2004).
- P. W. Fritsch, B. C. Cruz, F. Almeda, Y. Wang, S. Suhua Shi, *Systematic Botany* (IN PRESS).
- T. Galewski, J. F. Mauffrey, Y. L. Leite, J. L. Patton, E. J. Douzery, *Mol Phylogenet Evol* 34, 601 (2005).
- J. Glos, F. A. Glaw, M. Vences, *Copeia* 2, 252 (2005).
- A. Greslebin, K. Nakasone, K. M. Rajchenberg, *Mycologia* 96, 260 (2004).
- K. Halaných, M. R. Feldman, A. R. Vrijenhoek, C., *The Biological Bulletin* 201, 65 (2001).
- K. M. Halaných, T. J. Robinson, *Mol Phylogenet Evol* 7, 294 (1997).
- J. C. Hall, H. H. Iltis, K. J. Sysma, 29, 654 (2004).
- C. Lopez-Vaamonde, H. C. Godfray, J. M. Cook, *Evolution Int J Org Evolution* 57, 1804 (2003).
- C. Lopez-Vaamonde, J. Y. Rasplus, G. D. Weiblen, J. M. Cook, *Mol Phylogenet Evol* 21, 55 (2001).
- H. T. Lumbsch, R. Lindemuth, *Mycological Research* 105, 901 (2001).
- F. Lutzoni, M. Pagel, V. Reeb, *Nature* 411, 937 (2001).
- K. Maekawa, M. Kon, K. Araya, *Entomological Science* 8, 121 (2005).
- R. Mallarino, E. Bermingham, K. R. Willmott, A. Whinnett, C. D. Jiggins, *Mol Phylogenet Evol* 34, 625 (2005).
- G. Manganelli, N. Salomone, F. Giusti, *Biological Journal of the Linnean Society* 85, 501 (2005).
- R. J. Mason-Gamer, *Syst Biol* 53, 25 (2004).
- P. B. Matheny, *Mol Phylogenet Evol* 35, 1 (2005).
- C. A. Matthee, B. J. van Vuuren, D. Bell, T. J. Robinson, *Syst Biol* 53, 433 (2004).
- L. Matthias, B. Robert, B. Dominik, O. Franz, *Mycologia* 96, 614 (2004).
- S. L. Miller, B. Buyck, *Mycological Research* 106, 259 (2002).
- Y. Minegishi *et al.*, *Mol Phylogenet Evol* 34, 134 (2005).
- A. Mitchell, C. Mitter, J. C. Regier, *Syst Biol* 49, 202 (2000).
- R. L. Mueller, J. R. Macey, M. Jaekel, D. B. Wake, J. L. Boore, *Proc Natl Acad Sci U S A* 101, 13820 (2004).
- N. P. Murphy, C. M. Austin, *Zoologica Scripta* 34, 187 (2005).
- W. J. Murphy *et al.*, *Science* 294, 2348 (2001).
- C. J. O'Kelly, W. K. Bellows, B. Wyszor, *Journal of Phycology* 40, 209 (2004).
- C. J. O'Kelly, B. Wyszor, W. K. Bellows, *Journal of Phycology* 40, 789 (2004).
- R. D. Page, P. L. Lee, S. A. Becher, R. Griffiths, D. H. Clayton, *Mol Phylogenet Evol* 9, 276 (1998).
- K. M. Pryer *et al.*, *Nature* 409, 618 (2001).
- S. S. Renner, L. B. Zhang, *Syst Biol* 53, 422 (2004).
- C. C. Ribas, R. Gaban-Lima, C. Y. Miyaki, J. Cracraft, *Journal of Biogeography* 32, 1409 (2005).
- M. J. Riesing, L. Kruckenhauser, A. Gamauf, E. Haring, *Mol Phylogenet Evol* 27, 328 (2003).
- R. P. Roberts, L. Urbatsch, E., *Systematic Botany* 29, 199 (2004).
- K. Roelants, F. Bossuyt, *Syst Biol* 54, 111 (2005).
- A. Rokas *et al.*, *Mol Phylogenet Evol* 26, 36 (2003).
- G. Rouhan *et al.*, *Mol Phylogenet Evol* 33, 745 (2004).
- G. W. Rouse, S. K. Goffredi, R. C. Vrijenhoek, *Science* 305, 668 (2004).
- I. Ruiz-Trillo, M. Riutort, D. T. Littlewood, E. A. Herniou, J. Baguna, *Science* 283, 1919 (1999).
- J. A. Sánchez, S. D. Cairns, *Zool. Med. Leiden* 78, 265 (2004).
- J. A. Sanchez, C. S. McFadden, S. France, C., H. R. Lasker, *Marine Biology* 142, 975 (2003).
- G. M. Schneeweiss, A. Colwell, J. M. Park, C. G. Jang, T. F. Stuessy, *Mol Phylogenet Evol* 30, 465 (2004).
- H. Schneider *et al.*, *Proc Biol Sci* 272, 455 (2005).
- H. Schneider *et al.*, *Systematic Botany* 29, 260 (2004).
- L. M. Schultheis, M. J. Donoghue, *Systematic Botany* 29, 77 (2004).
- S. Shoup, L. A. Lewis, *Journal of Phycology* 39, 789 (2003).
- B. B. Simpson, J. A. Tate, A. Weeks, *Systematic Botany* 29, 933 (2004).
- C. H. Slamovits, J. A. Cook, E. P. Lessa, M. S. Rossi, *Mol Biol Evol* 18, 1708 (2001).
- R. Small, L., *Systematic Botany* 29, 385 (2004).

- K. Hansen, K. F. Lobuglio, D. H. Pfister, *Mol Phylogenet Evol* 36, 1 (2005).
- T. A. Hedderson, D. J. Murray, C. J. Cox, T. L. Nowell, *Systematic Botany* 29, 29 (2004).
- D. M. Helfgott, R. J. Mason-Gamer, *Systematic Botany* 29, 850 (2004).
- M. D. Herron, T. A. Castoe, C. L. Parkinson, *Mol Phylogenet Evol* 31, 1015 (2004).
- D. S. Hibbett, M. Binder, *The Biological Bulletin* 201, 319 (2001).
- D. M. Hillis, T. P. Wilcox, *Mol Phylogenet Evol* 34, 299 (2005).
- S. M. Ickert-Bond, M. Wojciechowski, F., *Systematic Botany* 29, 834 (2004).
- P. Inderbitzin, S. Landvik, M. A. Abdel-Wahab, M. L. Berbee, *Am J Bot* 88, 52 (2001).
- A. Ingram, L., J. Doyle, J., *Systematic Botany* 29, 545 (2004).
- M. Jaarola *et al.*, *Mol Phylogenet Evol* 33, 647 (2004).
- M. Janda, D. Folkova, J. Zrzavy, *Mol Phylogenet Evol* 33, 595 (2004).
- P. Järvinen *et al.*, *American Journal of Botany* 91, 1834 (2004).
- D. G. Kelch, B. G. Baldwin, *Mol Ecol* 12, 141 (2003).
- M. Kennedy, R. D. Gray, H. G. Spencer, *Mol Phylogenet Evol* 17, 345 (2004).
- M. Kennedy, H. G. Spencer, *Mol Phylogenet Evol* 31, 31 (2004).
- K. Kiontke *et al.*, *Proc Natl Acad Sci U S A* 101, 9003 (2004).
- K. M. Kjer, *Syst Biol* 53, 506 (2004).
- S. O. Klanten, L. Herwerden Lv, J. H. Choat, D. Blair, *Mol Phylogenet Evol* 32, 221 (2004).
- E. Larsson, K. H. Larsson, *Mycologia* 95, 1037 (2003).
- K. H. Larsson, E. Larsson, U. Koljalg, *Mycological Research* 108, 983 (2004).
- S. Lavery, T. Y. Chan, Y. K. Tam, K. H. Chu, *Mol Phylogenet Evol* 31, 39 (2004).
- D. H. Les, M. L. Moody, S. L. L. Jacobs, *Systematic Botany* 30, 503 (2005).
- L. A. Lewis, B. D. Mishler, R. Vilgalys, *Mol Phylogenet Evol* 7, 377 (1997).
- R. Leys, S. J. Cooper, M. P. Schwarz, *Mol Phylogenet Evol* 17, 407 (2000).
- S. Liede-Schumann, A. Rapini, D. J. Goyder, M. W. Chase, *Systematic Botany* 30, 185 (2005).
- D. A. Lijtmaer, N. M. Sharpe, P. L. Tubaro, S. C. Loughheed, *Mol Phylogenet Evol* 33, 562 (2004).
- C. P. Lin, B. N. Danforth, T. K. Wood, *Syst Biol* 53, 400 (2004).
- M. D. Sorenson, C. N. Balakrishnan, R. B. Payne, *Syst Biol* 53, 140 (2004).
- T. Sota, Y. Takami, G. B. Monteith, B. P. Moore, *Mol Phylogenet Evol* 36, 391 (2005).
- P. Q. Spinks, H. Bradley Shaffer, J. B. Iverson, W. P. McCord, *Mol Phylogenet Evol* 32, 164 (2004).
- M. Staats, P. van Baarlen, J. A. van Kan, *Mol Biol Evol* 22, 333 (2005).
- M. J. Stanhope *et al.*, *Proc Natl Acad Sci U S A* 95, 9967 (1998).
- S. Steppan, R. Adkins, J. Anderson, *Syst Biol* 53, 533 (2004).
- S. J. Steppan, B. L. Storz, R. S. Hoffmann, *Mol Phylogenet Evol* 30, 703 (2004).
- R. F. Sullivan, J. F. White, *Applied and Environmental Microbiology* 66, 425 (2000).
- F.-J. Sun, S. R. Downie, R. L. Hartman, *Systematic Botany* 22, 419 (2004).
- C. H. Tagliaro *et al.*, *Genetics and Molecular Biology* 28, 46 (2005).
- A. Tartar, D. G. Boucias, J. J. Becnel, B. J. Adams, *Int J Syst Evol Microbiol* 53, 1719 (2003).
- M. C. Tebbitt *et al.*, *American Journal of Botany* 91, 905 (2004).
- M. Thollessen, J. L. Norenburg, *Proc Biol Sci* 270, 407 (2003).
- M. Thulin, M. Lavin, R. Pasquet, A. Delgado-Salinas, *Systematic Botany* 29, 903 (2004).
- T. Townsend, A. Larson, E. Louis, J. R. Macey, *Syst Biol* 53, 735 (2004).
- M. Vences, F. A. Glaw, *Herpetological Journal* 15, 37 (2005).
- M. Vences, F. A. Glaw, *African Journal of Herpetology* 54, 77 (2005).
- A. Vogler, A. Cardoso, T. Barraclough, *Syst Biol* 54, 4 (2005).
- S. Wagstaff, J. J. Wege, *American Journal of Botany* 89, 965 (2002).
- A. Wang, M. Yang, J. Liu, *Ann Bot (Lond)* 96, 489 (2005).
- Y. Wang *et al.*, *American Journal of Botany* 91, 1901 (2004).
- S. Watanabe, A. Himizu, L. A. Lewis, G. L. Floyd, P. A. Fuerst, *Journal of Phycology* 36, 596 (2000).
- R. K. Wayne *et al.*, *Systematic Biology* 46, 622 (1997).
- T. P. Wilcox, D. J. Zwickl, T. A. Heath, D. M. Hillis, *Mol Phylogenet Evol* 25, 361 (2002).
- R. C. Winkworth, M. J. Donoghue, *Mol Phylogenet Evol* 33, 109 (2004).
- H. Xiao *et al.*, *Mol Phylogenet Evol* 36, 67 (2005).
- A. D. Yoder, Z. Yang, *Mol Ecol* 13, 757 (2004).

## References

1. Pagel, M. & Meade, A. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst Biol* **53**, 571-81 (2004).
2. Pagel, M. & Meade, A. Mixture models in phylogenetic inference. in *Mathematics of evolution and phylogeny* (ed. Gascuel, O.) 121-139 (Oxford University press, New York, 2005).
3. Pagel, M., Venditti, C. & Meade, A. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science (Wash.)* **314**, 119-21 (2006).
4. Venditti, C., Meade, A. & Pagel, M. Detecting the Node-Density Artifact in Phylogeny Reconstruction. *Syst. Biol.* **55**, 637-343 (2006).
5. Pagel, M., Meade, A. & Scott, D. Assembly rules for protein networks derived from phylogenetic-statistical analysis of whole genomes. *BMC Evol. Biol.* **7**, S16 (2007).
6. Pagel, M. The history, rate and pattern of world linguistic evolution. in *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form* (eds. Knight, C., Studdert-Kennedy, M. & Hurford, J.R.) 391-416 (Cambridge university press, Cambridge, 2000).
7. Green, P.J. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biomet* **82**, 711-732 (1995).
8. Pagel, M. & Meade, A. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst. Biol.* **53**, 571-81 (2004).
9. Venditti, C., Meade, A. & Pagel, M. Phylogenetic mixture models can reduce node-density artifacts. *Syst. Biol.* **57**, 286-93 (2008).
10. Webster, A.J., Payne, R.J. & Pagel, M. Molecular phylogenies link rates of evolution and speciation. *Science (Wash.)* **301**, 478 (2003).
11. Ho, S.Y.W., Phillips, M.J., Cooper, A. & Drummond, A.J. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* **22**, 1561-1568 (2005).