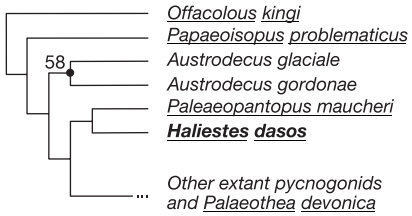

Figure 2 Strict consensus (consistency index = 0.38, retention index = 0.66) of three most-parsimonious trees of length 203 obtained from analysis of data matrix (see Supplementary Table). Bootstrap values, where over 50, are shown. Underlined species are fossils. Most extant species and *Palaeothea devonica* have been removed for clarity (see Supplementary Figure for full tree).

discovered among the diverse arachnomorph arthropods of the Cambrian period to undermine the significance of the chelicerae in this regard. Placement of the pycnogonids as sister to all other euarthropods[8] receives no new support from this discovery.

The phylogenetic position of *Haliestes* within the pycnogonids was analysed by adding it to a modified version (see Supplementary Table) of the most extensive published character matrix for pycnogonids, which included species from all extant families together with *Palaeoisopus*[20]. *Palaeopantopus*, *Palaeothea* and the chelicerate *Offacolus*[12] were also added, the last as a non-pycnogonid outgroup. Both *Haliestes* and *Palaeothea* are each based on a single specimen of unknown gender, and were coded on the arbitrary assumption that they are not sexually dimorphic (see Supplementary Methods).

Analysis of this data set using unweighted characters (Fig. 2; Supplementary Figure) suggests that *Haliestes* and *Palaeopantopus* may form a clade within the crown group, with *Palaeoisopus* lying in the stem group. The relative positions of *Palaeoisopus*, *Austrodecus* and *Palaeopantopus* are identical in unfigured implied weights analyses of this revised data set (following the methodology of ref. 20, Fig. 2; see also Supplementary Methods), although here *Haliestes* is in the stem, opposed to the crown group. The basal position of *Austrodecus* was not recovered before the addition of fossil taxa, but concurs with a recent study combining morphological and molecular data[21]. The crown-group position for *Palaeopantopus* differs from previous interpretations[4,17]. *Palaeothea* falls in a relatively derived position in both analyses, although 24 of the 39 characters for this taxon are coded as missing data. The analysis of Fig. 2 is poorly supported and the phylogenetics of the sea spiders remains to be resolved in detail[20,21], hence these results should be considered preliminary—nonetheless, our analyses suggest that *Haliestes* belongs near or in the pycnogonid crown group, which may thus have originated by the Silurian period.

Haliestes inhabited the outer shelf/upper slope of the subtropically positioned Anglo–Welsh Basin, and its mode of life and functional morphology were probably like those of extant pycnogonids (Supplementary Note 2). □

Methods

Morphology was reconstructed digitally following serial grinding at 20-μm intervals[12]. See Supplementary Methods for details of cladistic analyses.

Received 11 June; accepted 12 August 2004; doi:10.1038/nature02928.

1. King, P. E. *Pycnogonids* (Hutchinson, London, 1973).
2. Arnaud, F. & Bamber, R. N. The biology of Pycnogonida. *Adv. Mar. Biol.* **24,** 1–96 (1987).
3. Child, C. A. The marine fauna of New Zealand: Pycnogonida (sea spiders). *Natl Inst. Water Atmos. Res. Biodiv. Memoir* **109,** 1–71 (1998).
4. Waloszek, D. & Dunlop, J. A. A larval sea spider (Arthropoda: Pycnogonida) from the Upper Cambrian 'Orsten' of Sweden, and the phylogenetic position of pycnogonids. *Palaeontology* **45,** 421–446 (2002).
5. Vilpoux, K. & Waloszek, D. Larval development and morphogenesis of the sea spider *Pycnogonum litorale* (Ström, 1762) and the tagmosis of the body of Pantopoda. *Arthrop. Struct. Dev.* **32,** 349–383 (2003).
6. Boxshall, G. A. The evolution of arthropod limbs. *Biol. Rev.* **79,** 253–300 (2004).
7. Dunlop, J. A. & Arango, C. P. Pycnogonid affinities: a review. *J. Zool. Syst. Evol. Res.* (in the press).
8. Giribet, G., Edgecombe, G. D. & Wheeler, W. C. Arthropod phylogeny based on eight molecular loci and morphology. *Nature* **413,** 157–161 (2001).
9. Briggs, D. E. G., Siveter, David J. & Siveter, Derek J. Soft-bodied fossils from a Silurian volcaniclastic deposit. *Nature* **382,** 248–250 (1996).
10. Sutton, M. D., Briggs, D. E. G., Siveter, David J. & Siveter, Derek J. An exceptionally preserved vermiform mollusc from the Silurian of England. *Nature* **410,** 461–463 (2001).
11. Sutton, M. D., Briggs, D. E. G., Siveter, David J. & Siveter, Derek J. A three-dimensionally preserved fossil polychaete worm from the Silurian of Herefordshire, England. *Proc. R. Soc. Lond. B* **268,** 2355–2363 (2001).
12. Sutton, M. D., Briggs, D. E. G., Siveter, David J., Siveter, Derek J. & Orr, P. J. The arthropod *Offacolus kingi* (Chelicerata) from the Silurian of Herefordshire, England: computer based morphological reconstructions and phylogenetic affinities. *Proc. R. Soc. Lond. B* **269,** 1195–1203 (2002).
13. Siveter, David J., Sutton, M. D., Briggs, D. E. G. & Siveter, Derek J. An ostracode crustacean with soft parts from the Lower Silurian. *Science* **302,** 1749–1751 (2003).
14. Briggs, D. E. G., Sutton, M. D., Siveter, David J. & Siveter, Derek J. A new phyllocarid (Crustacea: Malacostraca) from the Silurian Fossil-Lagerstätte of Herefordshire, UK. *Proc. R. Soc. Lond. B* **271,** 131–138 (2004).
15. Orr, P. J., Briggs, D. E. G., Siveter, David J. & Siveter, Derek J. Three-dimensional preservation of a non-biomineralized arthropod in concretions in Silurian volcaniclastic rocks from Herefordshire, England. *J. Geol. Soc. Lond.* **157,** 173–186 (2000).
16. Miyazaki, K. J. On the shape of the foregut lumen in sea spiders (Arthropoda: Pycnogonida). *Mar. Biol. Ass. UK* **82,** 1037–1038 (2002).
17. Bergström, J., Stürmer, W. & Winter, G. *Palaeoisopus, Palaeopantopus* and *Palaeothea*, pycnogonid arthropods from the Lower Devonian Hunsrück Slate, West Germany. *Paläont. Zeit.* **54,** 7–54 (1980).
18. Bartels, C., Briggs, D. E. G. & Brassel, G. *The Fossils of the Hunsrück Slate* (Cambridge Univ. Press, 1998).
19. Wilby, P. R., Briggs, D. E. G. & Riou, B. Mineralization of soft-bodied invertebrates in a Jurassic metalliferous deposit. *Geology* **24,** 847–850 (1996).
20. Arango, C. P. Morphological phylogenetics of the sea spiders (Arthropoda: Pycnogonida). *Org. Divers. Evol.* **2,** 107–125 (2002).
21. Arango, C. P. Molecular approach to the phylogenetics of sea spiders (Arthropoda: Pycnogonida) using partial sequences of nuclear ribosomal DNA. *Mol. Phylogenet. Evol.* **28,** 588–600 (2003).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements This work was supported by the Leverhulme Trust, NERC and English Nature. K. Saunders is thanked for technical assistance; C. Arango, R. Bamber and D. Waloszek for discussion; G. Boxshall, J. Dunlop and D. Waloszek for pre-prints of their papers; S. de Grave for providing Recent material; and R. Fenn, T. Hall and J. Sinclair for general assistance.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.J.S. (derek.siveter@earth.ox.ac.uk).

Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous

Bryan Kolaczkowski[1] **& Joseph W. Thornton**[2]

[1]*Department of Computer and Information Science and* [2]*Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, Oregon 97403, USA*

All inferences in comparative biology depend on accurate estimates of evolutionary relationships. Recent phylogenetic analyses have turned away from maximum parsimony towards the probabilistic techniques of maximum likelihood and bayesian Markov chain Monte Carlo (BMCMC). These probabilistic techniques represent a parametric approach to statistical phylogenetics, because their criterion for evaluating a topology—the probability of the data, given the tree—is calculated with reference to an explicit evolutionary model from which the data are assumed to be identically distributed. Maximum parsimony can be considered nonparametric, because trees are evaluated on the basis of a general metric—the minimum number of character state changes required to generate the data on a given tree—without assuming a specific distribution[1]. The shift to parametric methods was spurred, in large part, by studies showing that although both approaches perform well most of the time[2], maximum parsimony is strongly biased towards recovering an

incorrect tree under certain combinations of branch lengths, whereas maximum likelihood is not[3–6]. All these evaluations simulated sequences by a largely homogeneous evolutionary process in which data are identically distributed. There is ample evidence, however, that real-world gene sequences evolve heterogeneously and are not identically distributed[7–16]. Here we show that maximum likelihood and BMCMC can become strongly biased and statistically inconsistent when the rates at which sequence sites evolve change non-identically over time. Maximum parsimony performs substantially better than current parametric methods over a wide range of conditions tested, including moderate heterogeneity and phylogenetic problems not normally considered difficult.

Functional constraints on sites in a gene sequence often change through time, causing shifts in site-specific evolutionary rates, a phenomenon called heterotachy (meaning 'different speeds')[7–16]. When an identically distributed evolutionary framework is imposed on sequences that evolve heterogeneously, parameter estimates are compromises over sites and lineages and are therefore incorrect for many or all sites. Likelihood-based techniques are guaranteed to recover the true phylogeny only when the correct model is used, and nonparametric statistical methods are often applied when the assumptions of parametric techniques are violated. On the other hand, parametric methods, including maximum likelihood, are generally more powerful than nonparametric techniques and can be robust to certain violations[17,18]. We used an experimental approach to evaluate the phylogenetic accuracy of parametric and nonparametric methods under a simple form of heterotachy. We simulated replicate DNA sequence alignments with two symmetrical rate partitions along a four-taxon tree; each partition represents a phylogenetically challenging problem—two clades, each consisting of a long branch (length p) and a short branch (length q)—but the sites with accelerated rates differ between partitions (Fig. 1a). To reveal the specific impact of heterogeneity, we compared phylogenetic accuracy (the fraction of replicates from which the true tree was recovered) on heterogeneous data with accuracy on control sequences simulated under corresponding evolutionary conditions without heterogeneity (see Methods).

Under conditions of strong heterotachy ($p = 0.75$ substitutions per site, $q = 0.05$), the accuracy of both maximum likelihood and BMCMC is dramatically reduced compared with homogeneous controls (Fig. 1b). Both methods have zero accuracy when the internal branch length $r < 0.22$, and they reach 100% accuracy only when $r > 0.34$. Maximum parsimony is superior to the parametric methods when $0.15 < r < 0.35$, and it is never inferior. For each method, we used nonlinear regression to estimate the internal branch length at which 50% accuracy is achieved (BL_{50}) and found that maximum parsimony can reliably recover the true topology at significantly shorter internal branch lengths ($BL_{50} = 0.22$) than the two likelihood-based methods ($BL_{50} = 0.28$, $P < 0.001$). Maximum parsimony's performance is worse than that of the parametric methods on single-partition data (due to the well-known long branch attraction bias[5]), but it is not additionally hampered by evolutionary heterogeneity ($P = 0.76$). Maximum parsimony retains its performance advantage over maximum likelihood and BMCMC on heterotachous data when strong support is required to accept a tree as resolved (bootstrap or posterior probability >95%, Supplementary Fig. S1). These results indicate that heterotachy substantially reduces the accuracy of maximum likelihood and BMCMC on phylogenetic problems

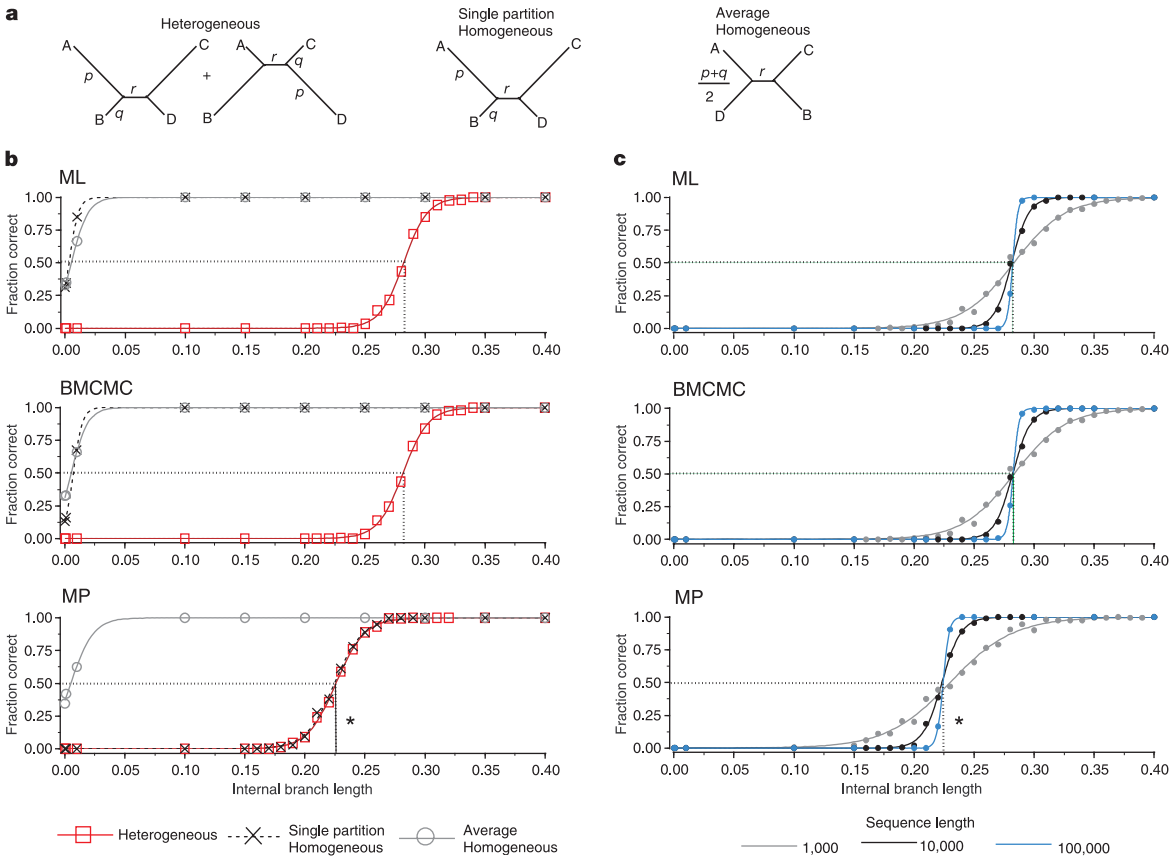

Figure 1 Likelihood-based methods are less accurate than maximum parsimony (MP) under heterogeneous conditions. **a**, Trees on which heterogeneous and control sequences were simulated. **b**, Heterotachy reduces the accuracy of likelihood methods. Accuracy is plotted against internal branch length for sequences with and without strong heterotachy. Dotted lines, BL_{50} for each method (asterisk: maximum parsimony < maximum likelihood (ML) and BMCMC, $P < 0.001$). **c**, Likelihood methods are inconsistent below the BL_{50} under strong heterotachy, recovering the incorrect tree with increasing frequency as the amount of data increases.

that are not difficult enough to impair maximum parsimony.

Under the heterotachous conditions studied, maximum likelihood and BMCMC are statistically inconsistent, converging on the wrong answer as the amount of data grows. For internal branch lengths below the BL_{50}, accuracy declines to zero as sequence length increases, indicating that parametric methods are statistically inconsistent in this region of parameter space; the BL_{50} therefore represents an 'inconsistency point' (Fig. 1c and Supplementary Fig. S2). This inconsistency is due to a directional bias: maximum likelihood and BMCMC specifically infer the erroneous tree ((AC),(BD)) with high support when the internal branch is shorter than the BL_{50}, including length zero (Supplementary Fig. S3). This is the same tree towards which maximum parsimony is biased on single-partition data, but heterotachy causes likelihood-based methods to infer the incorrect tree over a wider range of parameter values and with stronger apparent support.

Heterotachy reduces the performance of parametric methods across a broad range of evolutionary conditions. Whenever the short terminal branch length $q < 0.3$, maximum parsimony significantly outperforms both likelihood-based methods. Even fairly weak heterotachy—a ratio of branch lengths among partitions as low as 0.5:0.2—is sufficient to produce a significant performance disparity between the likelihood-based methods and maximum parsimony (Fig. 2a). The more intense the heterotachy, the greater the performance difference. Furthermore, maximum likelihood's accuracy can be reduced to zero when only a small fraction of sites deviate in rate from the rest of the sequence. Fewer heterotachous sites are required to impair performance as heterotachy grows more intense or the phylogenetic problem becomes more difficult (Fig. 2b).

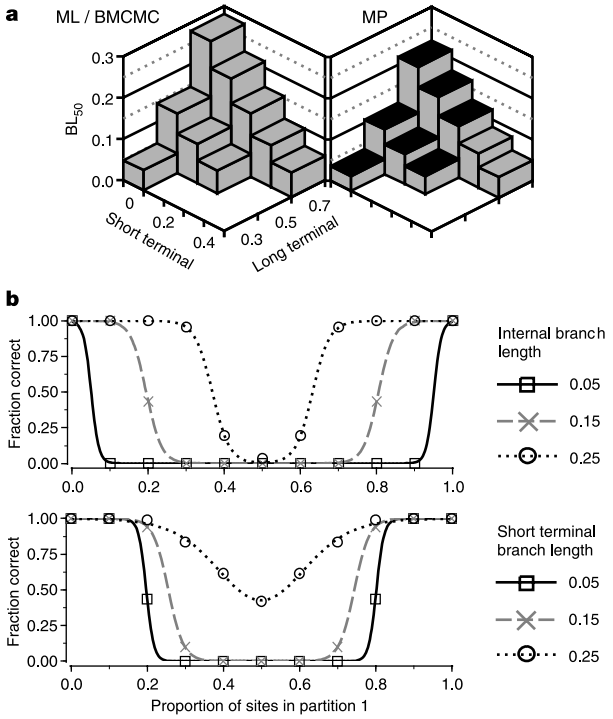

Figure 2 Parsimony outperforms likelihood over a wide range of heterotachous conditions. **a**, Maximum parsimony is more accurate than likelihood-based methods on data with weaker heterotachy. Bars show the BL_{50} for combinations of long and short terminal branch lengths in heterotachous data sets (black: maximum parsimony < maximum likelihood and BMCMC, $P < 0.001$). The BL_{50}s for maximum likelihood and BMCMC are equivalent for all conditions ($P > 0.91$). **b**, Maximum likelihood accuracy is impaired when only a small fraction of sites are heterotachous. Accuracy is plotted against the fraction of heterotachous sites as the phylogenetic problem becomes more difficult (upper panel: $p = 0.75$, $q = 0.05$) and heterotachy more intense (lower panel: $p = 0.75$, $r = 0.15$).

We used several existing likelihood models that account for among-site or -lineage rate variation by applying identically distributed models of heterogeneity, including gamma, invariant sites and covarion models, but none improve the performance of maximum likelihood or BMCMC (Fig. 3a). Using amino acid instead of nucleotide sequences substantially increases the accuracy of maximum parsimony ($BL_{50} = 0.08$) because convergence is less likely with 20 than with 4 possible states. In contrast, maximum likelihood and BMCMC improve to a much smaller extent ($BL_{50} = 0.18$). As a result, maximum parsimony's performance advantage on heterotachous protein sequences is even greater than on DNA (Fig. 3b). Denser taxon sampling to break up long branches[19] improves the accuracy of all methods by about equal proportions (Fig. 3b).

The accuracy of likelihood-based methods declines because they erroneously impose homogeneous branch lengths across sites. On heterotachous data with internal branch lengths below the inconsistency point, maximum likelihood overestimates the length of the internal branch and infers the lengths of the long and short terminals as approximately the average over the two partitions (Supplementary Fig. S4). To test whether these errors are responsible for phylogenetic bias, we compared the standard homogeneous maximum likelihood model (ML_{homo}) with an a priori partitioned model in which the branch lengths for each site are constrained to their true values (ML_{true}). As Fig. 4a shows, ML_{true} has much better performance ($P < 0.001$). Models that set only the internal (ML_{term}) or the internal and long terminal branches (ML_{short}) to their true lengths did not improve performance. Correcting the short terminal (ML_{long}), however, yields a substantial improvement in phylogenetic accuracy. Erroneous optimization of the short terminal length using 'compromise' branch lengths is therefore the primary cause of heterotachy-induced phylogenetic error in maximum likelihood (Fig. 4a).

Maximum likelihood's bias is caused by misinterpretation of specific character state patterns. We analysed the contribution each character state pattern makes to the likelihood of the true and erroneous trees and compared net support for the true tree using ML_{homo} to that using the heterogeneous model ML_{true} (Fig. 4b). Patterns that provide the most support for the correct tree under ML_{true} (*xxyy* and *xxyz*) only weakly support the true tree when ML_{homo} is used; this occurs because ML_{homo} overestimates the probability that these patterns are due to convergence on branches whose lengths are overestimated. In contrast, the convergent patterns *xyxy* and *xyxz* support the wrong tree using either method. As a result, the likelihood of the incorrect tree becomes greater than that of the true tree when ML_{homo} is used on heterotachous data. Under the same conditions, maximum parsimony recovers the true

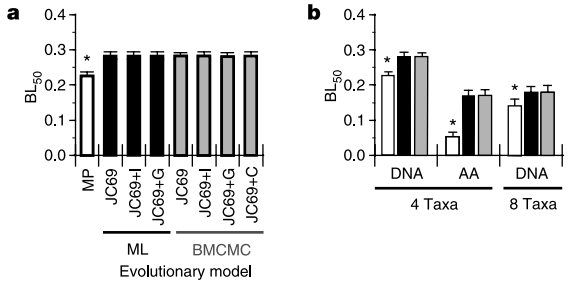

Figure 3 Maximum parsimony is more accurate than likelihood methods when techniques to improve phylogenetic performance are used. **a**, Accuracy of likelihood-based methods on heterotachous data does not improve when evolutionary models that incorporate among-site rate variation (+G, gamma distribution; +I, invariant sites) or covarion heterotachy (+C) are used. BL_{50}s are shown under strong heterotachy; bars indicate 99% confidence intervals. Asterisks show lower BL_{50} values ($P < 0.001$). **b**, Maximum parsimony (white) outperforms maximum likelihood (black) and BMCMC (grey) on amino-acid sequences and 8-taxon data sets with strong heterotachy.

tree because the frequency of *xxyy* is greater than that of *xyxy*; the patterns *xxyz* and *xyxz*, which taken together mislead ML_{homo}, are not informative in a nonparametric context.

The bias of parametric methods arises due to heterogeneity in the data and the resulting violation of the identical distribution assumption, as predicted theoretically[20]. We implemented a novel likelihood method using a mixed model ($BMCMC_{hetero}$) that incorporates heterotachy by including two branch length sets for each topology. For each sequence site, the likelihood is calculated for each branch length set, weighted by the posterior probability of the site being in that set and then summed to yield the total likelihood. This model, which corresponds to the true evolutionary conditions but assuming an identical data distribution, performs dramatically better than both maximum parsimony and the standard maximum likelihood or BMCMC algorithms ($BL_{50} = 0.045$, $P < 0.001$) on heterotachous data (Fig. 4c). It did not perform as well, however, as a non-identically distributed method ($BMCMC_{true}$) that uses the true evolutionary model with a priori sorting of sites into their true partitions. Furthermore, $BMCMC_{hetero}$ remains statistically inconsistent, converging on the wrong tree as sequence length increases at internal branch lengths $r < BL_{50}$, (Supplementary Fig. S5). $BMCMC_{true}$ is consistent under all conditions examined. These results indicate that violating the identical distribution assumption can cause inconsistency, even when the 'true' evolutionary model is used.

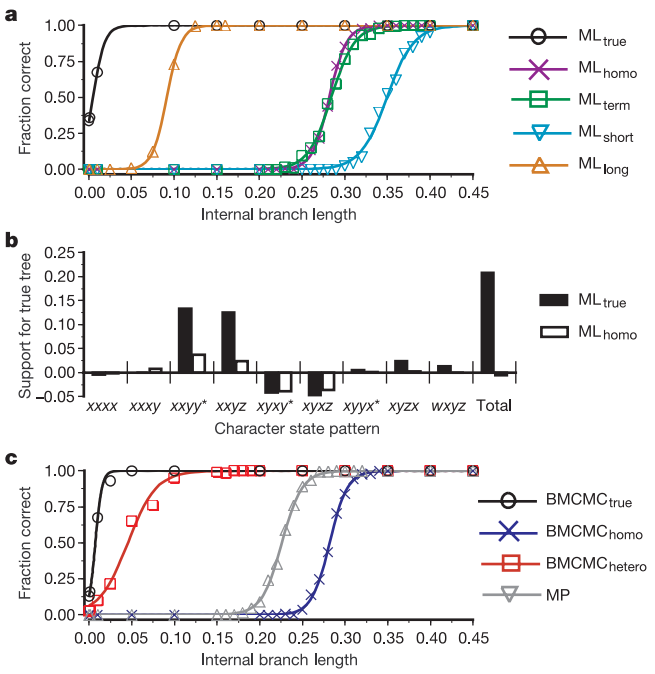

Figure 4 Poor maximum likelihood performance is due to assuming homogeneous branch lengths. **a**, Maximum likelihood error is caused primarily by overestimating short terminal branch lengths due to heterogeneity. Accuracy on strongly heterotachous sequences is shown as the internal branch length increases, using several likelihood models that constrain all (ML_{true}), some (ML_{term}, ML_{short}, ML_{long}) or no branches on the tree (ML_{homo}) to their true lengths for all sites. **b**, Support for the true tree by specific character state patterns is reduced due to strong heterogeneity when ML_{homo} is used. For each character state pattern and model, net support is shown as the ratio of the likelihood of the true topology to the likelihood of the incorrect ((AC),(BD)) tree, weighted by the frequency of the pattern. Asterisks indicate parsimony-informative patterns. **c**, Incorporating heterotachy improves the accuracy of parametric methods. Accuracy on strongly heterotachous data are shown for the homogeneous model ($BMCMC_{homo}$), a model that allows two independent branch length sets and correct a priori paritioning of sites ($BMCMC_{true}$), and a novel model with two branch length sets and likelihoods calculated on the basis of a posteriori weighting ($BMCMC_{hetero}$).

The form of heterotachy studied here is only one way that heterotachy can be distributed on a tree. Our additional work (not shown) indicates that several other forms of heterotachy can also impair the accuracy of parametric methods. The evolutionary model used in our simulations is a simplified one; the extent to which phylogenetic accuracy is impaired by the more complex evolutionary dynamics likely to affect real-world sequences is currently unknown. There are numerous sequence data sets from which parametric methods have failed to infer otherwise well-corroborated phylogenies[21–24], including one in which heterotachy has recently been implicated[13].

There are two ways to avert the negative effects of heterogeneity on parametric methods. One is to use maximum parsimony, which is not affected by heterotachy because it does not assume an identically distributed evolutionary process. The other is to develop more complex parametric models. Our results indicate that a new likelihood method using mixed branch length models may offer substantially improved accuracy on heterotachous sequences, but there are reasons for caution. The model that performed well in our tests matched the true evolutionary process, which we knew a priori. With real sequences, we do not know the true number of branch length partitions, so imposed models will usually use either too many or too few branch length parameters. For many sequences, the actual number of branch length categories may approach the number of sites; under these conditions, the true one-category-per-site likelihood model is formally equivalent to maximum parsimony[25]. Finally, the computational burden of mixed-model phylogenetic inference grows exponentially with the number of branch length sets. With current algorithms and computing power, incorporating heterotachy into a likelihood framework will often require sacrifices in the number of sequences analysed or the rigor with which tree and parameter space are searched, which may also reduce phylogenetic accuracy[1].

Our findings place those who infer and use phylogenetic trees in an uncertain position. Previous research has shown that parametric methods are superior or equal to nonparametric approaches when evolutionary heterogeneity is not present, but our work shows that maximum parsimony can substantially outperform current likelihood-based methods when it is. Worse still, heterotachy-induced bias leaves no obvious signature because the inferred trees have moderate branch lengths and strong support for erroneous nodes. With no reliable a posteriori diagnostic for heterotachy-induced phylogenetic error, how can we know which method to choose or, when trees from different methods conflict, which one to favour? The overall frequency and severity of the conditions that favour likelihood as compared with those that favour parsimony is not yet known for real-world sequences. At present, we recommend reporting nonparametric analyses along with parametric results and interpreting likelihood-based inferences with the same caution now applied to maximum parsimony trees. In the future, it is possible that new mixed-model techniques may improve likelihood's performance to the point that it is consistently superior to nonparametric methods. □

Methods

Simulations

We simulated sequences along a 4-taxon tree ((A,B),(C,D)) with two independent partitions that were concatenated into one heterogeneous alignment. In one partition, long terminal branches ($p \in (0.3, 0.75)$) lead to A and C, and short terminals ($q \in (0.001, 0.4)$) lead to B and D. In the other partition, terminal branches to B and D have length p, whereas A and C have length q. The internal branch length ($r \in (0.0, 0.5)$) is equal in both partitions. The two partitions were of equal size unless otherwise noted. Two-hundred replicate alignments of 1,000, 5,000, 10,000 and 100,000 characters were simulated under each set of conditions using the JC69 (DNA) or Poisson (protein) model. Average homogeneous control data were simulated using the same internal branch length as in the experimental condition and terminals with the mean length over the two partitions. Single-partition homogeneous controls were simulated using conditions for one of the experimental partitions (Fig. 1a). Sequences were also simulated on 8-taxon trees derived from 4-taxon trees by bisecting each terminal branch at the halfway point.

Phylogenetic analysis

Phylogenies were analysed using PAUP* v4.0b10 (ref. 26), PAML v3.14 (ref. 27) and MrBayes v3.0b4 (ref. 28). To determine the best-fit likelihood model for nucleotide data, hierarchical likelihood ratio tests were performed on 100 randomly selected alignments chosen from experimental data sets (using Modelltest v3.06 (ref. 29), $\alpha = 0.05$). The true JC69 model was strongly supported and was used for all maximum likelihood and BMCMC DNA analyses. The true Poisson model was used for protein analysis, and maximum parsimony used equal weights. We performed likelihood-based analyses with and without gamma and invariant sites models to determine their effect on accuracy. The covarion model implemented in MrBayes was also used. Topology searches were exhaustive for maximum parsimony and maximum likelihood. BMCMC analysis involved four chains (three heated) run well past stationarity.

To determine support, we used nonparametric bootstrapping (1,000 replicates) for maximum parsimony and maximum likelihood and posterior probability for BMCMC, with a support cutoff value of 95% to construct strongly supported consensus trees. (See Supplementary Information for details on phylogenetic methods.)

Accuracy

The accuracy of each method was calculated as the proportion of replicates for which the correct topology was uniquely recovered (ϕ). Nonlinear regression was performed using the logistic equation $\phi = 1/(1 + \exp((BL_{50} - r)H))$, in which BL_{50} is the estimated internal branch length that produces 50% correct recovery, and H estimates the steepness of the performance curve. The significance of differences among BL_{50}s was examined by a t-test.

Bias and error

The type I error rate for each method was determined by analysing data sets generated under strong heterotachy with zero-length internal branches and determining the fraction of replicates falsely resolved with 95% bootstrap or posterior probability support[30]. The presence of bias was determined by calculating the proportion of erroneous estimates consistent with each possible incorrect topology over all internal branch lengths. The intensity of bias was investigated by calculating the proportion of erroneous topology estimates consistent with each possible incorrect topology when a 95% support cutoff was imposed.

To determine the impact of homogeneous optimization of branch lengths on maximum likelihood error, we compared the standard maximum likelihood algorithm that estimates a single set of branch lengths (ML_{homo}) with several partitioned maximum likelihood models with constrained branch lengths. ML_{true} constrains all branch lengths for each site to the true values used to simulate data sets. ML_{term} constrains the internal branch lengths to the true value for each site, but terminal branches have the lengths homogeneously optimized under ML_{homo}. ML_{short} assumes the true internal and long terminal branches but uses the short terminal length from ML_{homo}. ML_{long} constrains the internal and short terminal branches to their true values and takes the long terminal branch length from ML_{homo}.

Support for the true topology by each character-state pattern was calculated from a 100,000-site data set constructed under strong heterotachy ($p = 0.75$, $q = 0.05$, $r = 0.254$). Net support for the true tree is defined as the likelihood ratio of the true tree to the incorrect tree for each pattern x, weighted by the frequency of x ($f(x)$) in the data set: $S_{((AB),(CD)),x} = \frac{P(x|((AB),(CD)))}{P(x|((AC),(BD)))} f(x)$.

To determine the performance impact of violating the identical distribution assumption when the true evolutionary model is used, we implemented a novel likelihood model ($BMCMC_{hetero}$) that incorporates heterotachy a posteriori by applying two sets of branch lengths to the data. For each sequence site x_i, the likelihood of tree t with branch length sets b_1 and b_2 is $L(t|x_i) = \sum_{j=1}^{2} [\rho_{i,j} \cdot P(x_i|t, b_j)]$, where $\rho_{i,j}$—the posterior probability that x_i is in branch length set b_j—is calculated from the data as $\rho_{i,j} = P(x_i|t, b_j)/\sum_{k=1}^{2} P(x_i|t, b_k)$. The overall posterior probability of each topology is calculated using BMCMC (see Supplementary Information). This new method was compared with a BMCMC analysis using the true heterogeneous model and correct a priori data partitioning ($BMCMC_{true}$).

Received 17 May; accepted 6 August 2004; doi:10.1038/nature02917.

1. Sanderson, M. J. & Kim, J. Parametric phylogenetics? *Syst. Biol.* **49,** 817–829 (2000).
2. Hillis, D. M., Huelsenbeck, J. P. & Cunningham, C. W. Application and accuracy of molecular phylogenies. *Science* **264,** 671–677 (1994).
3. Felsenstein, J. Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.* **27,** 401–410 (1978).
4. Kuhner, M. K. & Felsenstein, J. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Mol. Biol. Evol.* **11,** 459–468 (1994).
5. Huelsenbeck, J. P. Systematic bias in phylogenetic analysis: is the Strepsiptera problem solved? *Syst. Biol.* **47,** 519–537 (1998).
6. Gaut, B. S. & Lewis, P. O. Success of maximum likelihood phylogeny inference in the four-taxon case. *Mol. Biol. Evol.* **12,** 152–162 (1995).
7. Huelsenbeck, J. P. Testing a covariotide model of DNA substitution. *Mol. Biol. Evol.* **19,** 698–707 (2002).
8. Miyamoto, M. M. & Fitch, W. M. Testing the covarion hypothesis of molecular evolution. *Mol. Biol. Evol.* **12,** 503–513 (1995).
9. Lopez, P., Casane, D. & Philippe, H. Heterotachy, an important process of protein evolution. *Mol. Biol. Evol.* **19,** 1–7 (2002).
10. Fitch, W. M. The molecular evolution of cytochrome c in eukaryotes. *J. Mol. Evol.* **8,** 13–40 (1976).
11. Pollock, D. D., Taylor, W. R. & Goldman, N. Coevolving protein residues: maximum likelihood identification and relationship to structure. *J. Mol. Biol.* **287,** 187–198 (1999).
12. Pupko, T. & Galtier, N. A covarion-based method for detecting molecular adaptation: application to the evolution of primate mitochondrial genomes. *Proc. R. Soc. Lond. B* **269,** 1313–1316 (2002).
13. Inagaki, Y., Susko, E., Fast, N. M. & Roger, A. J. Covarion shifts cause a long-branch attraction artifact that unites Microsporidia and Archaebacteria in EF-1{alpha} phylogenies. *Mol. Biol. Evol.* **21,** 1340–1349 (2004).
14. Lockhart, P. J. *et al.* A covariotide model explains apparent phylogenetic structure of oxygenic photosynthetic lineages. *Mol. Biol. Evol.* **15,** 1183–1188 (1998).
15. Misof, B. *et al.* An empirical analysis of mt 16S rRNA covarion-like evolution in insects: site-specific rate variation is clustered and frequently detected. *J. Mol. Evol.* **55,** 460–469 (2002).
16. Philippe, H. & Lopez, P. On the conservation of protein sequences in evolution. *Trends Biochem. Sci.* **26,** 414–416 (2001).
17. Donaldson, T. S. Robustness of the F-test to errors of both kinds and the correlation between the numerator and denominator of the F-ratio. *J. Am. Stat. Assoc.* **63,** 660–676 (1968).
18. Sullivan, J. & Swofford, D. L. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? *Syst. Biol.* **50,** 723–729 (2001).
19. Hillis, D. M. Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Syst. Biol.* **47,** 3–8 (1998).
20. Chang, J. T. Inconsistency of evolutionary tree topology reconstruction methods when substitution rates vary across characters. *Math. Biosci.* **134,** 189–215 (1996).
21. Rokas, A., Williams, B. L., King, N. & Carroll, S. B. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* **425,** 798–804 (2003).
22. Russo, C. A., Takezaki, N. & Nei, M. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. *Mol. Biol. Evol.* **13,** 525–536 (1996).
23. Delarbre, C., Gallut, C., Barriel, V., Janvier, P. & Gachelin, G. Complete mitochondrial DNA of the hagfish, *Eptatretus burgeri*: the comparative analysis of mitochondrial DNA sequences strongly supports the cyclostome monophyly. *Mol. Phylogenet. Evol.* **22,** 184–192 (2002).
24. Naylor, G. J. & Brown, W. M. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. *Syst. Biol.* **47,** 61–76 (1998).
25. Tuffley, C. & Steel, M. Links between maximum likelihood and maximum parsimony under a simple model of site substitution. *Bull. Math. Biol.* **59,** 581–607 (1997).
26. Swofford, D. L. *PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods, v.4.0b10* (Sinauer Associates, Sunderland, Massachusetts, 1998).
27. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13,** 555–556 (1997).
28. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17,** 754–755 (2001).
29. Posada, D. & Crandall, K. A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14,** 817–818 (1998).
30. Swofford, D. L. *et al.* Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Syst. Biol.* **50,** 525–539 (2001).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank J. Conery for advice, support and programming advice. P. Phillips, R. DeSalle and S. Proulx provided comments and discussion. We benefited from discussions of mixed model methods with D. Zwickl. B.K. was supported by an NSF IGERT training grant in Evolution, Development and Genomics to the University of Oregon.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.T. (joet@uoregon.edu).

Ecological constraints on diversification in a model adaptive radiation

Rees Kassen[1,2], Martin Llewellyn[1]* & Paul B. Rainey[1,3]

[1]*Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK*
[2]*Department of Biology & Centre for Advanced Research in Environmental Genomics, University of Ottawa, Ottawa K1N 6N5, Canada*
[3]*School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand*

* Present address: London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Taxonomic diversification commonly occurs through adaptive radiation, the rapid evolution of a single lineage into a range of genotypes or species each adapted to a different ecological niche[1,2]. Radiation size (measured as the number of new types) varies widely between phylogenetically distinct taxa[2–4] and between replicate radiations within a single taxon where the ecological opportunities available seem to be identical[5,6]. Here we