Regardless of the mechanism controlling reef-water pH, our results suggest that corals at Flinders Reef have experienced a relatively wide range in pH (~0.3 pH units) over the past ~300 years. As a result, these corals have also experienced equivalent changes in the aragonite saturation state (Ω_{arag}), one of the main physicochemical controllers of coral calcification. Changes in Ω_{arag} have been derived from the Flinders pH record (Fig. 2D), with Ω_{arag} varying from ~3 to 4.5, assuming constant alkalinity (*10*, *24*). This encompasses the lower and upper limits of Ω_{arag} within which corals can survive (*37*). Despite such marked changes, skeletal extension and calcification rates for the Flinders Reef coral (Fig. 2E) fall within the normal range for *Porites* (*38*) and are not correlated with Ω_{arag} or pH. Therefore, the *Porites* coral at Flinders Reef seems well adapted to relatively large fluctuations in seawater pH and Ω_{arag}.

The interdecadal cycle in seawater pH observed at Flinders Reef has relevance for predicting its response to future ocean acidification, given that it will either enhance or moderate the local effects of the projected long-term decrease in pH (*3*, *4*). For example, the next rise in the ~50-year cycle of reef-water pH should counteract the lowering of pH values at Flinders Reef until ~2035 A.D. Conversely, the subsequent fall in the reef-water pH cycle will lead to an abrupt shift toward low pH reef water. The extent to which corals and other calcifying reef organisms can adapt to such rapid decreases in pH is largely unknown.

Our findings suggest that the effects of progressive acidification of the oceans are likely to differ between coral reefs because reef-water P_{CO_2} and consequent changes in seawater pH will rarely be in equilibrium with the atmosphere. Although the relatively large variations in seawater pH at Flinders Reef suggest that coral reefs may be resilient to the shorter term effects of ocean acidification, in the coming decades many reefs are likely to experience reduced pH that is unprecedented relative to "natural" levels. Additional paleo-pH records are required from a range of coral reef ecosystems to improve our understanding of the physical and biological controls on reef-water pH, and the long-term impacts of future ocean acidification.

References and Notes

1. J. R. Petit *et al.*, *Nature* **399**, 429 (1999).
2. C. L. Sabine *et al.*, *Science* **305**, 367 (2004).
3. D. A. Wolf-Gladrow, U. Riebesell, S. Burkhardt, J. Bijma, *Tellus* **51B**, 461 (1999).
4. K. Caldeira, M. E. Wickett, *Nature* **425**, 365 (2003).
5. A. Ridgwell, R. E. Zeebe, *Earth Planet. Sci. Lett.* **234**, 299 (2005).
6. J. C. Zachos *et al.*, *Science* **308**, 1611 (2005).
7. R. A. Feely *et al.*, *Science* **305**, 362 (2004).
8. U. Riebesell *et al.*, *Nature* **407**, 364 (2000).
9. J. P. Gattuso, D. Allemand, M. Frankignoulle, *Am. Zool.* **39**, 160 (1999).
10. J. A. Kleypas *et al.*, *Science* **284**, 118 (1999).
11. F. Marubini, M. J. Atkinson, *Mar. Ecol. Prog. Ser.* **188**, 117 (1999).
12. J. Dore, R. Lukas, D. Sadler, D. M. Karl, *Nature* **424**, 754 (2003).
13. Data are reported with the delta notation relative to the National Bureau of Standards (NBS) 951 boric acid standard defined as $\delta^{11}B$ (‰) = (Rs/Rstd – 1) × 1000, where Rs = $^{11}B/^{10}B$ of the sample and Rstd = $^{11}B/^{10}B$ of NBS 951. External precision of $\delta^{11}B$ values are estimated at ±0.2‰ ($2\sigma_{mean}$ level) (*24*).
14. A. Vengosh, Y. Kolodny, A. Starinsky, A. R. Chivas, M. T. McCulloch, *Geochim. Cosmochim. Acta* **55**, 2901 (1991).
15. N. G. Hemming, G. N. Hanson, *Geochim. Cosmochim. Acta* **56**, 537 (1992).
16. A. J. Spivack, C.-F. You, J. Smith, *Nature* **363**, 149 (1993).
17. A. Sanyal, N. G. Hemming, G. N. Hanson, W. S. Broecker, *Nature* **373**, 234 (1995).
18. P. N. Pearson, M. R. Palmer, *Science* **284**, 1824 (1999).
19. M. R. Palmer, P. N. Pearson, *Science* **300**, 480 (2003).
20. J. Gaillardet, C. Jean Allègre, *Earth Planet. Sci. Lett.* **136**, 665 (1995).
21. N. G. Hemming, T. P. Guilderson, R. G. Fairbanks, *Global Biogeochem. Cycles* **12**, 581 (1998).
22. B. Hönisch *et al.*, *Geochim. Cosmochim. Acta* **68**, 3675 (2004).
23. S. Reynaud, N. G. Hemming, A. Juillet-Leclerc, J. P. Gattuso, *Coral Reefs* **23**, 539 (2004).
24. Materials and methods are available as supporting material on *Science* Online.
25. F. Böhm *et al.*, *Geochem. Geophys. Geosyst.* **3**, 10.1029/2001GC000264 (2002).
26. S. J. Fallon, T. P. Guilderson, K. Caldeira, *Geophys. Res. Lett.* **30**, 10.1029/2003GL018049 (2003).
27. S. Power, T. Casey, C. K. Folland, A. Colman, V. Mehta, *Clim. Dyn.* **15**, 319 (1999).
28. N. J. Mantua, S. R. Hare, *J. Oceanogr.* **58**, 35 (2002).
29. S. Minobe, *Geophys. Res. Lett.* **24**, 683 (1997).
30. M. J. Salinger, J. Renwick, A. B. Mullan, *Int. J. Climatol.* **21**, 1705 (2001).
31. A. Gershunov, T. P. Barnett, *Bull. Am. Meteorol. Soc.* **79**, 2715 (1998).
32. E. Calvo *et al.*, in preparation.
33. J. C. Andrews, M. J. Furnas, *Cont. Shelf Res.* **6**, 491 (1986).
34. M. A. Cane, *Science* **222**, 1189 (1983).
35. A. Suzuki, H. Kawahata, *Tellus B* **55**, 428 (2003).
36. B. A. Taft, W. S. Kessler, *J. Geophys. Res.* **96**, 12599 (1991).
37. J. A. Kleypas, J. W. McManus, L. A. B. Menez, *Am. Zool.* **39**, 146 (1999).
38. J. M. Lough, D. J. Barnes, *J. Exp. Mar. Biol. Ecol.* **211**, 29 (1997).
39. PDO index data (available at ftp://ftp.atmos.washington.edu/mantua/pnw_impacts/INDICES/PDO.latest).
40. Australian Government, Bureau of Meteorology (www.bom.gov.au).

41. We thank L. Kinsley for assistance with Triton mass spectrometry at Research School of Earth Sciences, S. Power and the UK Met. Office for providing the IPO data and W. Müller, E. Hendy, G. Foster, C. Coath, B. Windel, B. Hönisch, A. Suzuki, R. Key, T. McConnaughey, M. Church, J. Dore, and L. Pena for stimulating discussions and help throughout this research. Australian Postdoctoral Fellowships awarded to C.P. and E.C. (Australian Research Council grants DP0342702 and DP0450682) are gratefully acknowledged. C.P. and E.C. also acknowledge funding from the Spanish Ministry of Education and Science during the early stages of this work. The coral drilling by Peter Isdale and Bruce Parker, Australian Institute of Marine Science, was funded by an Australian National Greenhouse Advisory Committee grant to J. Chappell. The data will be made available on the NOAA Paleoclimatology Web site (www.ngdc.noaa.gov/paleo/data.html).

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5744/2204/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 and S2
References

18 April 2005; accepted 11 August 2005
10.1126/science.1113692

Phylogenetic MCMC Algorithms Are Misleading on Mixtures of Trees

Elchanan Mossel[1] and Eric Vigoda[2]

Markov chain Monte Carlo (MCMC) algorithms play a critical role in the Bayesian approach to phylogenetic inference. We present a theoretical analysis of the rate of convergence of many of the widely used Markov chains. For *N* characters generated from a uniform mixture of two trees, we prove that the Markov chains take an exponentially long (in *N*) number of iterations to converge to the posterior distribution. Nevertheless, the likelihood plots for sample runs of the Markov chains deceivingly suggest that the chains converge rapidly to a unique tree. Our results rely on novel mathematical understanding of the log-likelihood function on the space of phylogenetic trees. The practical implications of our work are that Bayesian MCMC methods can be misleading when the data are generated from a mixture of trees. Thus, in cases of data containing potentially conflicting phylogenetic signals, phylogenetic reconstruction should be performed separately on each signal.

Bayesian inference is one of the most popular methods in phylogeny reconstruction (*1*). Many widely used software packages, such as MrBayes (*2*), BAMBE (*3*), and PAML (*4*), rely on Markov chain Monte Carlo (MCMC) methods. These algorithms are often known as BMCMC. Part of the appeal of BMCMC is that they are supposed to be more robust and faster than standard maximum likelihood approaches. Our results show that these appealing features are overly optimistic in some settings.

The basis of the MCMC algorithms is a Markov chain whose stationary distribution is the desired posterior distribution. Reliable phylogenetic estimates depend on the Markov chain converging to the posterior distribution

before any phylogenetic inferences are made. Typically it is elementary to establish that the Markov chains eventually converge to the posterior distribution. However, convergence after an infinite number of iterations is not of practical use. The chains need to converge quickly to the posterior distribution in order to be considered useful. Unfortunately, it is notoriously difficult to rigorously analyze the convergence rate of the Markov chains used for phylogeny. In practice, heuristics [such as multiple starting states and convergence of log-likelihood plots (*5*)] are commonly used to determine when the Markov chains have converged to the posterior distribution.

The major difficulty in analyzing these Markov chains is our poor understanding of the geometry of the space of tree topologies weighted by the likelihood function (this geometric space is often referred to as the tree space). Work has been done on the likelihood function for fixed trees on three and four leaves (*6*, *7*) [abstract properties of tree space on any number of leaves also have been analyzed (*8*–*11*)].

In our work, we consider N characters generated from a uniform mixture of two trees on $n \geq 5$ leaves and show that BMCMC takes an exponential number of iterations to converge. Our proofs also yield detailed information on the geometry of the likelihood function on the tree space for five or more leaves. It has its basis in combinatorial analysis techniques that are further discussed in (*12*).

The bases of our results are the following two trees of five taxa: T_1 is given by ((12),3),(45) in the standard Newick format, whereas T_2 is given by ((14),3),(25); see the yellow trees in Fig. 1 for an illustration. Our results apply to cases where the data are generated from a uniform mixture of two trees, $(T_1^*,\vec{L}_1^*)$ and $(T_2^*,\vec{L}_2^*)$, on $n \geq 5$ leaves, where for some subset S of leaves, the subtree of $(T_1^*,\vec{L}_1^*)$ on S is $(T_1,\vec{l}_1)$, and the subtree of $(T_2^*,\vec{L}_2^*)$ on S is $(T_2,\vec{l}_2)$. Moreover, the branch lengths on subtrees $(T_1,\vec{l}_1)$ and $(T_2,\vec{L}_2^*)$ must lie in the following zone: The branch lengths of terminal branches (i.e., edges incident to the leaves) are between a and a^2; the branch lengths of internal branches are between a and $2a$; and the number a satisfies $0 < a < b$, where b is some small positive constant. The assumptions on the branch lengths of T_1 and T_2 are essential in the details of the proof.

For each of the trees $(T_1^*,\vec{L}_1^*)$ and $(T_2^*,\vec{L}_2^*)$, the character data are generated by using any of the standard mutation models, such as the Cavender-Farris-Neyman (CFN) model, the Jukes-Cantor model, and Kimura's two parameter model [see (*13*) for an introduction to these models]. Moreover, our results hold for almost any prior distribution on branch lengths used in BMCMC including those discussed in (*14*, *15*); see (*12*) for more details.

Our results are valid for two families of BMCMC. In the first family, the MCMC performs a random walk on the discrete set of tree topologies. The transition probabilities are determined by the Metropolis rule (*16*) using the Bayesian probability of tree topology given the data (*14*). In the second family, we consider MCMC performing a random walk on the continuous space of tree topologies and branch lengths (*3*, *17*). For both families the moves that change the tree topology may be nearest-neighbor interchanges (NNI), subtree pruning and regrafting (SPR), or tree bisection and reconnection (TBR) moves; see (*14*) for an introduction to these transitions.

In order to measure convergence of the Markov chain, we use the notion of mixing time, T_{mix}, which is standard in probability theory (*18*). The mixing time is, for the worst initial state T_0, the first time that the total variation distance between the distribution of T_t (i.e., the chain at time t) and the stationary distribution is at most 1/4. (The constant 1/4 is somewhat arbitrary and simply needs to be <1/2.) The above definition of mixing time implies that for any $\varepsilon > 0$, after $\leq T_{\text{mix}}$ iterations, the Markov chain is variation distance $\leq \varepsilon$ from the stationary distribution.

We can now state our main theoretical result: There exists a constant $c > 0$ such that in the setting described above, given N characters, $\vec{D} = (D_1,\ldots,D_N)$, generated from the mixtures of $(T_1^*,\vec{L}_1^*)$ and $(T_2^*,\vec{L}_2^*)$ on $n \geq 5$ taxa, with probability at least $1 - \exp(-cN)$ over the data generated, the mixing time of MCMC algorithms with NNI, SPR, or TBR transitions is at least $\exp(cN)$.

The formal proof of this statement appears in (*12*). We follow with a heuristic argument. The algorithmic computations below were performed for both the binary CFN model (*19*–*21*) and the Jukes-Cantor model. In both of these models, the branch length $a = \alpha t$ where α is the rate from state i to j, for $i \neq j$, and t is time.

The convergence properties of a Markov chain requires a detailed understanding of the weighted geometry of the state space. One aspect of the geometry is depicted in Fig. 1. In this figure, two trees are joined by an edge if they are connected by a single NNI transition. One can see that our two generating trees T_1 and T_2 have maximum distance.

The second aspect of the geometry are the posterior probabilities of tree topologies. The posterior probability of tree topology T is denoted by $w(T)$. It is natural to expect that for long sequences, $w(T)$ is essentially determined by the branch lengths that maximize the expected log-likelihood. In other words,

$$\frac{\log w(T)}{N} \sim J(T) = \max_{\vec{l}} E[\log \Pr(D|T,\vec{l})]$$

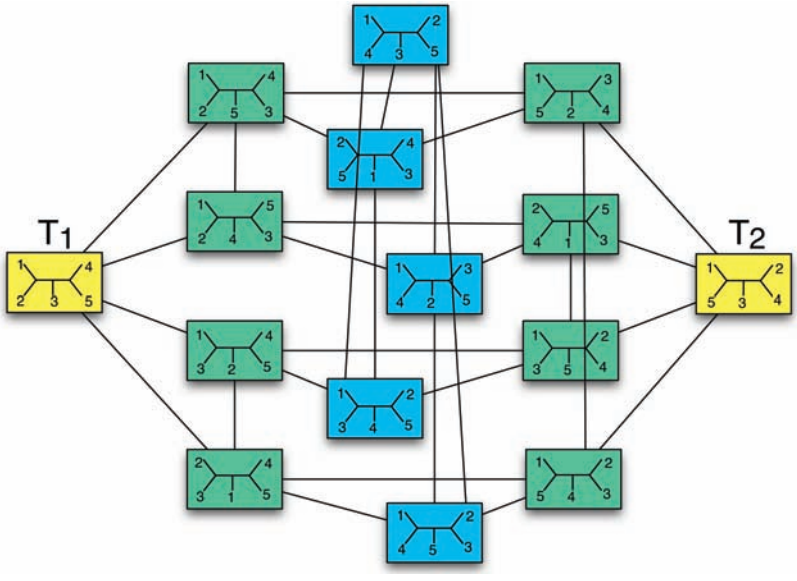

Fig. 1. The space of phylogenetic trees on five taxa connected by NNI transitions. For the mixture distribution on T_1 and T_2 in the CFN model with internal branch lengths $a = 0.1$ and terminal branch lengths $a^2 = 0.01$, the yellow trees, with optimal branch lengths, had expected log likelihood $J(T) \approx -1.887$, the green trees had $J(T) \approx -1.934$, and the blue trees had $J(T) \approx -1.986$. For the Jukes-Cantor model with $a = 0.1$, the yellow trees had $J(T) \approx -4.293$, the green trees had $J(T) \approx -4.338$, and the blue trees had $J(T) \approx -4.544$. Note, for SPR and TBR transitions, each yellow tree is connected to 12 other trees but not to each other. Thus, to travel between the two yellow trees by NNI, SPR, or TBR transitions, we need to traverse through a valley (i.e., trees with lower log-likelihood).

[1]Department of Statistics, University of California at Berkeley, Berkeley, CA 94720, USA. E-mail: mossel@stat.berkeley.edu [2]College of Computing, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: vigoda@cc.gatech.edu

where the expectation is over the probability distribution μ generating the data.

Figure 1 considers the expected log-likelihood $J(T)$ for data generated by taking independent samples from our mixture μ of the two trees T_1 and T_2 on five taxa, where all internal branches have length $a = 0.1$ and all terminal branches have length $a^2 = 0.01$. In Fig. 1, our generating trees have maximum expected log-likelihood $J(T)$, and the intermediate trees in this space have smaller expected log-likelihood. Thus, to traverse between the two maxima trees, we need to traverse a valley. Such a valley implies slow convergence, via mathematical techniques known as conductance and isoperimetric inequalities (*18*).

A similar picture holds for SPR and TBR transitions. On trees on five taxa TBR and SPR moves are identical. To traverse between the trees with maximum expected likelihood, T_1 and T_2, we need to pass through trees with smaller expected likelihood. This is the key property that implies slow convergence. For SPR and TBR transitions, tree T_1 and T_2 are connected to 12 other trees but not to each other.

In Fig. 2 we show how the maximum expected log-likelihood $J(T)$ varies with the NNI distance from the generating trees T_1 and T_2 for varying values of internal branch length a in the generating trees, where the terminal branch lengths are given by a^2.

The implications of mixture distributions to phylogeny has recently received considerable theoretical attention (*22*, *23*) and has clear practical importance. A simple example that often contains characters from multiple trees is molecular data consisting of DNA sequences for more than one gene. It is well known that phylogenetic trees can vary between genes [for example, see (*24*) for a discussion].

The numerical values of the constants a and c are not explicit in our results. However, simulations suggest that even moderate values such as $a = 0.1$ and $N = 1000$ have very slow convergence, and in fact starting at the tree T_1 or T_2 it will never visit any other tree topology. Moreover, in both cases, the likelihood plot suggests quick convergence. For these parameters, the behavior of the chain on data coming from mixtures or from data generated from a single tree is indistinguishable for as long as we run our experiments (millions of iterations). See fig. S1 (*12*) for log-likelihood plots illustrating these examples.

For small trees one can hope to overcome the slow convergence by using multiple starting states. However, mixtures coming from large trees may contain multiple species subsets where one tree has T_1 as an induced subtree and the other has T_2. If there are k such subsets, then about 15^k random starting points will be needed. Thus if there are 10 disagreement subsets, then 15^{10} random starting points will be needed in order to sample from the posterior distribution.

A popular MCMC program for phylogeny, MrBayes (*2*), uses Metropolis-coupled MCMC (*25*), which is designed to avoid bottlenecks in the state space. A key open question is to understand whether Metropolis-coupled MCMC is successful in avoiding bottlenecks created by mixtures. Resolving this question will require more delicate and detailed mathematical analysis. It is known that Metropolis-coupled MCMC converges exponentially slowly in some settings but avoid bottlenecks in others (*26*), but even in these simple cases a very detailed understanding of the state space is needed. If Metropolis-coupled MCMC successfully avoids the bottlenecks created by mixtures, then it may serve as a useful tool for identifying data generated from mixtures.

In our setting, BMCMC methods fail in a clearly demonstrable manner. We expect that there is a more general class of mixtures where BMCMC methods fail in more subtle ways. These subtle failures may occur for many real-world examples where the Markov chains quickly converge to some distribution other than the desired posterior distribution. Users of BMCMC methods should ideally avoid mixture distributions that are known to produce degenerate behavior in various phylogenetic settings (*27*, *28*). A good practice is to decompose the data into nonconflicting signals and perform phylogenetic reconstruction separately on each signal. Our work highlights important unresolved questions: how to verify homogeneity of genomic data and what phylogenetic methods can efficiently deal with mixtures.

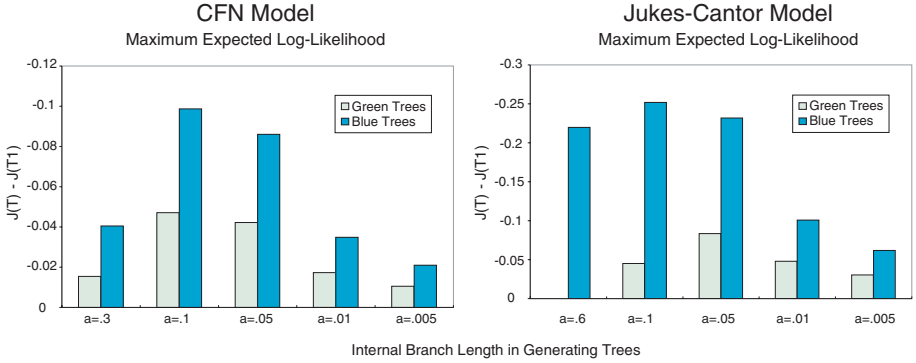

Fig. 2. The depth of the valley between the generating trees T_1 and T_2 with five taxa. The chart shows the maximum expected log-likelihood for different classes of trees for varying generating parameter a where the internal edges have length a and the terminal edges have length a^2. For the green trees and blue trees from Fig. 1, we plot the value of $J(T) - J(T_1)$. For the yellow trees (i.e., the generating trees T_1 and T_2), this value is 0.

References and Notes

1. J. P. Huelsenbeck, F. Ronquist, R. Nielsen, J. P. Bollback, *Science* **294**, 2310 (2001).
2. J. P. Huelsenbeck, F. Ronquist, *Bioinformatics* **17**, 754 (2001).
3. B. Larget, D. L. Simon, *Mol. Biol. Evol.* **16**, 750 (1999).
4. Z. Yang, B. Rannala, *Mol. Biol. Evol.* **14**, 717 (1997).
5. J. P. Huelsenbeck, B. Larget, R. E. Miller, F. Ronquist, *Syst. Biol.* **51**, 673 (2002).
6. Z. Yang, *Proc. R. Soc. London B. Biol. Sci.* **267**, 109 (2000).
7. B. Chor, M. D. Hendy, B. R. Holland, D. Penny, *Mol. Biol. Evol.* **17**, 1529 (2000).
8. P. W. Diaconis, S. P. Holmes, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14600 (1998).
9. L. J. Billera, S. P. Holmes, K. Vogtmann, *Adv. Appl. Math.* **27**, 733 (2001).
10. M. Develin, B. Sturmfels, *Doc. Math.* **9**, 1 (2004).
11. D. Speyer, B. Sturmfels, *Adv. Geom.* **4**, 389 (2004).
12. Details are available in the supporting online material on *Science* Online.
13. M. Nei, S. Kumar, *Molecular Evolution and Phylogenetics* (Oxford Univ. Press, New York, 2000).
14. J. Felsenstein, *Inferring Phylogenies* (Sinauer, New York, 2004).
15. B. Rannala, Z. Yang, *J. Mol. Evol.* **43**, 304 (1996).
16. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
17. S. Li, D. K. Pearl, H. Doss, *J. Am. Stat. Assoc.* **95**, 493 (2000).
18. A. Sinclair, *Algorithms for Random Generation and Counting: A Markov Chain Approach* (Birkhäuser, Boston, MA, 1993).
19. J. A. Cavender, *Math. Biosci.* **40**, 271 (1978).
20. J. S. Farris, *Syst. Zool.* **22**, 250 (1973).
21. J. Neyman, *Statistical Decision Theory and Related Topics* (Academic Press, New York, 1971), pp. 1–27.
22. B. Kolaczkowski, J. W. Thorton, *Nature* **431**, 980 (2004).
23. M. Steel, *Trends Genet.* **21**, 307 (2005).
24. D. Graur, W.-H. Li, *Fundamentals of Molecular Evolution* (Sinauer, Sunderland, MA, ed. 2, 1999).
25. C. J. Geyer, *Computer Science and Statistics: Proceedings of the 23rd Symposium Interface* (Interface Foundation, Fairfax Station, VA, 1991), pp. 156–163.
26. N. Bhatnagar, D. Randall, *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms* (New Orleans, LA, 2004), pp. 478–487.
27. J. T. Chang, *Math. Biosci.* **134**, 189 (1996).
28. M. A. Steel, L. A. Székely, M. D. Hendy, *J. Comp. Biol.* **1**, 153 (1994).

29. We thank S. Holmes, M. Steel, and S. Yi for useful comments. E.M. was supported by a Miller fellowship in computer science and statistics, by a Sloan fellowship in mathematics, and by NSF grant DMS-0504245. E.V. was supported by NSF grant CCF-0455666.

Supporting Online Material

www.sciencemag.org/cgi/content/full/309/5744/2207/DC1
Materials and Methods
SOM Text
Figs. S1 and S2

31 May 2005; accepted 23 August 2005
10.1126/science.1115493