Experimental Phylogenetics: Generation of a Known Phylogeny

DAVID M. HILLIS, JAMES J. BULL, MARY E. WHITE, MARTY R. BADGETT, IAN J. MOLINEUX

Although methods of phylogenetic estimation are used routinely in comparative biology, direct tests of these methods are hampered by the lack of known phylogenies. Here a system based on serial propagation of bacteriophage T7 in the presence of a mutagen was used to create the first completely known phylogeny. Restriction-site maps of the terminal lineages were used to infer the evolutionary history of the experimental lines for comparison to the known history and actual ancestors. The five methods used to reconstruct branching pattern all predicted the correct topology but varied in their predictions of branch lengths; one method also predicts ancestral restriction maps and was found to be greater than 98 percent accurate.

The development over the past four decades of explicit methods for phylogenetic inference (1) has permitted biologists to reconstruct the broad outlines of evolutionary history and to interpret comparative biological studies within an evolutionary framework (2). However, evolutionary history usually cannot be observed directly, at least over the course of relevant magnitudes of change, so that assessment of phylogenetic methods has relied on numerical simulations. Although simulations have provided considerable insight into the effectiveness of various

REFERENCES AND NOTES

11. Table I: A total of samples studied before high-T NMR runs, including bakers and those quenched from high-T NMR runs (0.3 to 0.7 g) and samples that were the same, and peak positions differed by less than 0.5 ppm.
12. 23Na. For 28Si high-T spectra, data from 1000 to 2000 pulses, requiring about 5 min to collect, were averaged. For 27Al and 29Si, 100 to 400 pulses were used, requiring less than 1 min. Signal to noise ratios for 23Na, 29Si, and 27Al were typically at least 100:1, 50:1, and 10:1, respectively. For these nuclei, samples of 1 M NaCl, 1 M Al(NO3)3, and pure SiO2 were run at ambient T as frequency references. For 27Al, 90° pulse times were measured at high T, and comparison with results for crystal line Al2O3 showed that liquid spectra collected above 1200°C were fully averaged over all transitions. Under these conditions, peaks for 23Na and 27Al were accurately fitted with Lorentzians, and dispersion versus absorption (DISPA) plots were circular. Both indicate that exchange dynamics were rapid enough to eliminate dynamic quadrupolar line shifts [J. G. Verweij, J. Skibsted, H. Bildsoe, N. C. Nielsen, J. Magn. Reson. 85, 173 (1990); and A. G. Marshall (J. Magn. Reson. 43, 443 (1981)]. MAS data were collected with a commercial high-speed probe. Spinning rates of about 6 kHz for 29Si (Al2O3 rotors) and up to 12 kHz for 27Al (ZrO2 rotors) were used in a probe from Dory Scientific. Acquisition delays were chosen to eliminate any effects of differential relaxation rates on peak shapes. We subtracted a significant amount of background from the probe by collecting data with an empty rotor under identical conditions to those for glass samples. For 27Al and the glass composition for 29Si, MAS peak centroids were determined that included the first set of spinning side bands. For 27Al, the centroid of the 1/2 to 1/2 transition spinning side bands was determined from side-band peaks and used in conjunction with the main peak to estimate isotopic chemical shifts in the glasses [H. J. Jacobsen, J. Skibsted, H. Bildsoe, N. C. Nielsen, J. Magn. Reson. 85, 173 (1990); E. Lippma, A. Samosen, M. Mägi, J. Am. Chem. Soc. 108, 1730 (1986)]. The relatively low intensity of the satellite side bands introduced an uncertainty of about 0.3 ppm in this estimation. We compared data for Gd-doped and undoped samples of the same lithium aluminosilicate composition (Li2AlSO4 and Li2AlSO3). 29Si and 27Al NMR lineshapes in both liquid and glass samples were the same, and peak positions differed by less than 0.5 ppm.
14. Stoichiometry requires about one nonbridging O per tetrahedral cation in most of the compositions studied. This does not allow for time-averaged local symmetry to be increased to cubic (reducing the electric field gradient to 0) simply by local motion of the cations around their mean positions.
20. This assumption might not be the case if there was extensive clustering of Al polyhedra. There is not clear evidence that this occurs in these compositions. In tecosilicate glass compositions, NMR studies have suggested that Al and Si distributions are random (see [14]).
25. For a Si site with four O neighbors, the number of O atoms bonded to other network-forming cations (Si2P39, and so forth) is indicated by n in Qn.
27. Supported under National Science Foundation graets EAR 85-53024 and ESR 89-05188. We thank B. Coté, J. P. Coutures, D. Massiot, P. F. McMillan, and F. Taulelle for useful discussions; P. Fiske for help with sample preparation; and two anonymous reviewers for helpful comments.
phylogenetic algorithms, they are limited by an incomplete knowledge of biology: all models incorporate untested assumptions about evolutionary processes.

Direct tests of organismal phylogenetic histories are limited to a few studies of strains of laboratory animals (3) and plant cultivars (4). Even these cases have shortcomings: the organisms underwent little genetic differentiation, phylogenies were produced over the course of decades or centuries, and the histories are incompletely known. In contrast, viruses can be manipulated in the laboratory through thousands of generations per year, and mutation rates of viruses can easily be elevated through the use of mutagens, so that experimental studies of phylogenies with viruses should be feasible (5, 6). We report the creation of a known phylogeny of lineages derived from bacteriophage T7 and provide fine-scale restriction maps of the entire genomes of the experimental lineages, including the ancestors. Our purpose was to test the effectiveness of methods for inferring phylogenies and ancestral genetic characteristics by comparing the inferred evolutionary history against a known phylogeny and the true ancestors.

We chose to construct a symmetric phylogeny with equal distances among nodes (Fig. 1) (7). This topology (tree shape) has proven especially amenable to accurate reconstruction in theoretical studies (8–10) and may thus be regarded as a "null model" against which other topologies may be compared. We created a phylogeny of nine taxa (eight ingroup lineages and one outgroup lineage) on each node are labeled with letters A through F and W (the latter represents wild-type T7). The numbers represent the number of restriction-site differences scored between the phages at each node of the phylogeny.

**Fig. 1.** True phylogeny for the experimental lineages of bacteriophage T7. The ancestors at each node are labeled with letters A through F and W (the latter represents wild-type T7). The numbers represent the number of restriction-site differences scored between the phages at each node of the phylogeny.

bifurcating trees for this many taxa, so the likelihood of inferring the correct phylogeny by chance alone is minimal. We compared the actual phylogeny (Fig. 1) to estimated phylogenies from five reconstruction methods; estimates were based on restriction-site maps produced for 34 restriction endonucleases in all terminal lineages (Figs. 2 and 3). To avoid bias, the actual phylogeny was unknown to the person mapping the restriction sites. We also produced restriction maps for the ancestral phage at each of the nodes of the true phylogeny (Fig. 3). Three aspects of the inferred phylogeny were compared to the actual phylogeny: branching topology, branch lengths, and ancestral states. The five methods of phylogenetic inference evaluated were parsimony (12), the Fitch-Margoliash method (13), the Cavalli-Sforza method (14), neighbor-joining (15), and the unweighted pair-group method of arithmetic averages (UPGMA) (16).

All methods predicted the correct branching order of the known phylogeny, but no method predicted the actual branch lengths for every branch (17). To compare the five methods for their ability to predict branch lengths, the correlation between observed and predicted branch lengths was calculated for each method. These five correlations were significantly heterogeneous, with parsimony yielding the highest value and UPGMA yielding the lowest value (18). The UPGMA method is known to be sensitive to unequal rates of change (1), and the number of changes per branch was quite variable in the true phylogeny (although the number of changes per ingroup branch is not significantly heterogeneous from the expectation under a Poisson distribution; test from [19]).

The experimental system also enabled us to determine ancestral states directly. Of the methods tested, only parsimony makes predictions about ancestral character states (parsimony may be used to optimize states onto phylogenies inferred by other methods, but, for these data, all methods estimated the same branching pattern). In comparing inferred ancestral states to the actual ancestral states, three outcomes are possible: the ancestral states may be (i) correctly inferred, (ii) incorrectly inferred, or (iii) ambiguous (when more than one character optimization is possible). For 202 variable sites assayed in each of seven ancestors, parsimony correctly inferred 1369 states (97.3%), incorrectly inferred 18 states (1.3%), and was ambiguous about 20 states (1.4%). Seven states (of four sites) could not be observed in some ancestors because they fell under deletion mutations (Fig. 3). If the 91 wild-type states that were invariant in all lineages are included, the inferred restriction maps are an average of 98.6% identical to the actual maps (with either delayed or accelerated

7. The phage was grown in 1 ml cultures of Escherichia coli strain W3110 in the presence of the mutagen N-methyl-N-nitro-N-nitrosoguanidine (20 μg/ml). After lysis, a new culture was infected with 10 μl of the lysate. After every five serial lysates, phage were plated on agar and a single plaque was randomly selected for further propagation. The distance between each node in the ingroup phylogeny (linesages through P and their ancestors) was equal to 40 serial lysates; each of the consequences of two to three bursts (generations) of phage. The outgroup lineage (R) was derived from wild-type T7 by a similar protocol (6). To guard against contamination of the restriction digest, we regrew the lineages from the last contamination-free ancestor.


17. For the parsimony reconstruction, bootstrap confidence intervals were 100% for all nodes except A (92%; refer to Fig. 1 for this node), based on 1000 replications; the consistency index for informative character optimization used) (1). Three of the incorrectly inferred states involved sites that were found only in the ancestral lineages, hence would not have been detected as variable characters if ancestors were unavailable (as is typically the case in phylogenetic studies).

Fig. 3. Variable restriction enzyme cleavage sites and deletions of the terminal lineages (J through R) and ancestral nodes (A through F) compared to the wild-type T7 genome (W). Wild-type sites for these enzymes that are not listed are found among all the lineages. Sites with positions shown without decimals are wild-type sites that have been lost in some of the lineages; their positions were determined from the complete sequence of T7 DNA (11). The remaining site locations were mapped as shown in Fig. 2; these positions are indicated in kilobases. Sites marked with an asterisk are also MboI sites; they are listed only once unless multiple changes produced differences between MboI and the larger recognition sequence in some of the taxa (for example, the MboI and BamHI sites at position 5.2). Sites in ancestors that were inferred incorrectly by parsimony analysis of the terminal lineages are circled; those for which inference was ambiguous are surrounded by squares. A letter “D” indicates that the site occurs in a region that has been deleted in the respective lineage; these sites were coded as unknown in the analyses.

character optimization used) (1). Three of the incorrectly inferred states involved sites that were found only in the ancestral lineages, hence would not have been detected as variable characters if ancestors were unavailable (as is typically the case in phylogenetic studies).

The results of this study directly support the legitimacy of methods for phylogenetic estimation, not only with regard to reconstructing branching relationships, but also branch lengths and ancestral genotypes. Perhaps more importantly, they point the direction to a field of research in which methods of reconstruction can be tested against various known phylogenies of real organisms differing in topological and other evolutionary characteristics in the same fashion that tests have been conducted with simulated, theoretical phylogenies. Experimental phylogenetics is not a substitute for numerical studies, nor is it likely that laboratory phylogenies will ever display the full complexity of phylogenies produced over long-term evolution, but such studies will fill an important void in the science of phylogenetic reconstruction.

REFERENCES AND NOTES


Chemical Signals from Host Plant and Sexual Behavior in a Moth

ASHOK K. RAINA, TIMOTHY G. KINGAN, AUTAR K. MATTOO

In the phytophagous corn earworm, Helicoverpa (Heliothis) zea, females delay their reproductive behaviors until they find a suitable host on which to deposit their eggs. Perception of volatile chemical signals from corn silk triggers the production of sex pheromone followed by its release, which leads to mating. Several natural corn silk volatiles, including the plant hormone ethylene, induced pheromone production in H. zea females. Because H. zea larvae feed on the fruiting parts of a wide variety of hosts, ethylene, which is associated with fruit ripening, could act as a common cue.

Females of most species of moths use sex pheromones to attract their mates. Production of the pheromone in a number of moth species is controlled by a peptide hormone, the pheromone biosynthesis-activating neuropeptide (PBAN) (1). Release of PBAN in laboratory-reared females of the corn earworm, Helicoverpa (Heliothis) zea, is regulated primarily by photoperiod (2). However, in nature, photo-periodic control appears to be superseded by signals from the host plant (2). For instance, female progeny of H. zea collected from a corn field did not produce pheromone during scotophase until they encountered corn, one of the hosts for this insect (3). The presence of silk from an ear of corn was sufficient to elicit pheromone production followed by “calling” behavior or pheromone release (3). Moreover, physical contact between the females and corn silk was not required, which indicated the involvement of a volatile factor (or factors) produced by the host plant.

Several years ago, trans-2-hexenal, a volatile component of oak leaves, was implicated in the induction of sex pheromone production in the polyphemus moth, Antheraea polyphemus (4, 5). However, this phenomenon could not be confirmed in a subsequent study (6). Hence, the signal (or signals) and the mechanism involved in the process have remained a matter of conjecture. We show the efficacy of several chemical constituents of corn silk volatiles, including the plant hormone ethylene, in inducing pheromone production in H. zea females. Ethylene, widely used by plants in fruit ripening (7), may act as a common cue because H. zea larvae feed on the fruiting parts of a wide variety of hosts. Thus, the corn earworm female can recognize these plant chemicals and exploit this ability to coordinate its reproductive behavior with the availability of food for the offspring.

Larvae of H. zea were collected from corn fields (8) and reared through one generation on artificial diet for many generations, which do not require host plant for the production of pheromone (2).

Because most species belonging to the Heliothis-Helicoverpa group feed on fruiting parts of various host plants and because the gaseous plant hormone ethylene is commonly produced by flowering plants, we tested the possibility that ethylene may act as a cue for pheromone production. First, using gas chromatography, we checked whether ethylene was a constitutive of the volatiles produced by corn silk. On a fresh weight basis, a gram of corn silk produced 2.07 ± 0.47 ng (n = 6) and 2.53 ± 0.67 ng of ethylene per hour over a 24-hour period in two separate tests. We then determined the effect of exogenous ethylene on pheromone production in H. zea females. Ethylene at various concentrations was introduced into containers holding individual H. zea females in their second photophase. Pheromone was extracted during the following scotophase.

Table 1. Effect of host plant on sex pheromone production by wild F1 females of H. zea.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>n</th>
<th>Z-11-hexadecenal (ng/female ± SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>10</td>
<td>2.1 ± 0.3</td>
</tr>
<tr>
<td>Corn silk-1</td>
<td>10</td>
<td>54.1 ± 6.4</td>
</tr>
<tr>
<td>Corn silk-2</td>
<td>5</td>
<td>11.1 ± 3.2</td>
</tr>
</tbody>
</table>

Fig. 1. Pheromone production in H. zea wild F1 females as a function of ethylene concentration (mean ± SEM, n = 7).

Expressions from the actual estimates was found to be significantly greater than that of the randomized data (P = 0.017), with the correlation for UPGMA being significantly less than the randomized value minimum (P < 0.011). The heterogeneity of correlations among the set of four methods with UPGMA removed is no longer significantly large, but the correlation for parsimony is larger than the maximum of randomized values at P = 0.06. By these criteria, UPGMA appears to be significantly worse than the other methods, and there is some evidence that parsimony is superior.

20. Supported by the NSF (to D.M.H.) and the NIH (to I.J.M.). We thank F. W. Studier, J. Tessman, and T. Kunkel for advice on mutagenesis in phage.

8 August 1991; accepted 22 November 1991