Horizontal Gene Transfer Accelerates Genome Innovation and Evolution

Ravi Jain,†[1] Maria C. Rivera,*†‡[1] Jonathan E. Moore,† and James A. Lake*†‡§*

*Molecular, Cell, and Developmental Biology, †Molecular Biology Institute, ‡Institute of Geophysics and Planetary Physics Astrobiology, and §Human Genetics, University of California, Los Angeles

Horizontal gene transfer (HGT) spreads genetic diversity by moving genes across species boundaries. By rapidly introducing newly evolved genes into existing genomes, HGT circumvents the slow step of ab initio gene creation and accelerates genome innovation. However, HGT can only affect organisms that readily exchange genes (exchange communities). In order to define exchange communities and understand the internal and external environmental factors that regulate HGT, we analyzed approximately 20,000 genes contained in eight free-living prokaryotic genomes. These analyses indicate that HGT occurs among organisms that share similar factors. The most significant are genome size, genome G/C composition, carbon utilization, and oxygen tolerance.

Introduction

Horizontal gene transfer (HGT) refers to the acquisition of foreign genes by organisms. It occurs extensively among prokaryotes, especially in response to a changing environment and provides organisms with access to genes in addition to those that can be inherited (Spratt et al. 1992; Hilario and Gogarten 1993; Syvanen 1994; Rivera et al. 1998, Doolittle 1999; Martin 1999, Ochman, Lawrence, and Groisman 2000). By allowing the transfer of genes between unrelated species, HGT increases genetic diversity. By introducing new genes into existing genomes, and thereby circumventing the slow step of ab initio gene creation, HGT accelerates genome innovation and evolution, although not necessarily gene evolution (Jain, Rivera, and Lake 1999). We will refer to a collection of organisms that can share genes by HGT, but need not be in physical proximity, as an exchange community.

The influence of environmental factors on HGT, and hence on the structure of exchange communities, has been uncertain and confusing. For example, it is well known that horizontal gene transfer can be influenced by factors such as temperature and pH (Lorenz and Wackernagel 1994; Williams et al. 1996; Davison 1999), yet in vitro and in vivo experiments have shown that some proteins that function at high temperatures are fully capable of replacing their mesophilic orthologs (Piper et al. 1996; Thomas and Cavicchioli 2000). Proximity can have an overriding effect on HGT, since if a DNA is missing from a particular habitat, it will be impossible for residents of that habitat to acquire the DNA even if HGT is otherwise acceptable. Nevertheless, organisms living very close to each other do not always exchange genes by HGT. Regulatory regions can be strongly influenced by internal factors such as G/C composition, yet some pathogens have obtained the ability to infect a host through acquisition of large stretches of DNA called pathogenicity islands that possess G/C compositions markedly different from the remainder of the genome (Groisman and Ochman 1996; Salama and Falkow 1999). Thus, the influence of environmental and internal parameters on HGT has been uncertain. In this paper, we ask whether and how geographic, environmental, and internal parameters have influenced genetic exchange by HGT. We find HGT is not random, but depends critically upon internal and environmental factors. We also identify and quantify those parameters that affect HGT and thereby delineate exchange community boundaries.

[1] These authors contributed equally to the work.

Key words: horizontal gene transfer, lateral gene transfer, exchange community, genome evolution.

E-mail: lake@mbi.ucla.edu.

Mol. Biol. Evol. 20(10):1598–1602. 2003
DOI: 10.1093/molbev/msg154
Molecular Biology and Evolution, Vol. 20, No. 10,

Materials and Methods

Bootstrap Calculations and Phylogenetic Reconstructions

Bootstrap calculations were performed to estimate both the slope of the zero associativity lines and to estimate the variance of the experimentally determined slopes in the associativity plots. To calculate the slope of the zero associativity lines, the HT distance from the reference tree to each of the 10,395 possible eight taxon trees was calculated utilizing the cut and regraph distances of Allen and Steel (2001) and the EF-Tu/EF-1a (caterpillar) tree as reference. The Allen and Steel distances correspond to the minimum number of HTs required to convert an observed tree into the reference tree. Of these trees, one was zero HT steps, 90 were one HT steps, 1,978 were two HT steps, 7,486 were three HT steps, and 840 were four HT steps from the reference tree.

Prune and regraft distances were calculated as follows (Moore and Lake, unpublished data). Initially, all possible prunings of the reference tree were regrafted to a different location on the tree. These new trees were thus one horizontal transfer step (HT = 1) from the reference tree. Then all HT = 1 trees were used to calculate trees that were two steps from the reference (i.e., those new prune and regraft trees that were neither HT = 0 nor HT = 1 trees). The process was iterated until all distances were determined.

Parsimony scores were calculated for each tree according to the two-step method of Williams and Fitch (1989). A brief description of their method follows. In their first step, the range of possible parsimony values is progressively selected for each node. In the second step, the most parsimonious value (or values) for each node is calculated, working from a common node to the tips. It is assumed that to get from character i to $i + n$, one passes through n states. To determine the set of values that could fit a node, one proceeds by sweeping the tree from the branches furthest from the root towards the root (the root

was placed at the center of the four internal branches, although the position of the root does not affect the calculations).

The zero associativity slopes were calculated by sampling from the set of all possible trees, whereas the slope and the variance of the experimental associativities were calculated from the set of observed trees. For both types of calculations the number of trees sampled (with replacement) for each horizontal transfer value were the same as the number of experimentally observed trees: HT = 0 (one tree), HT = 1 (nine trees), HT = 2 (60 trees), HT = 3 (65 trees), and HT = 4 (five trees). The mean of each zero associativity slope was determined as the mean averaged over 10,000 bootstrap replicates, and the distributions of experimentally determined associativity slopes were determined from 10,000 bootstraps for each environmental parameter. The bootstrap results were observed to be normally distributed, and standard deviations were calculated and used for *Z*-score analyses.

Diverse strategies were employed to reduce, or to estimate, errors potentially introduced into the analyses. For example, because G/C composition was one of the factors being analyzed for its effect on HGT, it was vital that our analyses be independent of G/C compositional biases. Hence we measured HGT using phylogenetic trees, rather than use methods that monitor G/C composition, to remove this potential source of bias. Also trees were calculated using paralinear (LogDet) distances to remove potential G/C biases, since these distances are demonstrably independent of G/C composition (Lake 1994). Phylogenetic reconstructions were performed using star alignments, since these alignments are much less sensitive to biases associated with pairwise alignments (Lake 1991), and distances were compensated for site-to-site variation using pattern filtering as previously described (Lake 1998; Rivera et al. 1998; Jain, Rivera, and Lake 1999). Because both the associativity (– parsimony) score and HGT distances were based on phylogenetic trees, the slopes of the parsimony/HGT plots were not significantly affected by small, intragene horizontal transfers that may have been undetected in this study. To assess the effects of using different reference trees, results were calculated using the three most probable trees as the reference tree. Together, these three trees represented 94% of all bootstrap replicates. We observed no statistically significant differences in our results when any of these three trees were used as the reference.

Results and Discussion

Eight complete genomes spanning the diversity of prokaryotic life (Bult et al. 1996; Kaneko et al. 1996; Blattner et al. 1997; Klenk et al. 1997; Kunst et al. 1997; Smith et al. 1997; Deckert et al. 1998; Kawarabayasi et al. 2001) and consisting of approximately 20,000 open reading frames were analyzed for the impact of the following factors on HGT: optimum, minimum, and maximum growth temperature; oxygen (anaerobic, microaerophillic, and aerobic); pH; salinity (marine, fresh, or either); $\log_{10}$pressure at site of isolation, G/C content, carbon utilization (heterotroph, autotroph, or either), and genome size (Stetter et al. 1990; Stetter 1996). The taxa and their corresponding values for each of these variables are given in the online Supplementary Material.

To test whether HGT preferentially transfers genes among organisms living in similar or different environments, we introduce the concept of associative horizontal gene transfer. The effect of each environmental parameter on HGT is described by its associativity value. For example, if HGT occurs preferentially among organisms living at similar temperatures, illustrated by adjacent shallow marine environments in figure 1*a*, then the temperature associativity is positive. If HGT occurs preferentially from organisms living at high temperatures into organisms living at lower temperatures, or vice versa, then the temperature associativity is negative, as illustrated by glacial and volcanic environments in figure 1*a*.

Positive associative HGT produces gene trees in which organisms from similar temperatures are clustered together, and negative associative HGT juxtaposes high temperature organisms with low temperature organisms (fig. 1*b*). The resulting clustering of environmental factors on gene trees can be assessed by parsimony, which provides a numerical measure of the extent of the grouping of organisms having similar parameters. For example, positive associativity results in low parsimony scores, since most temperature changes encountered in going from one branch in the tree to the next will be small (fig. 1*b*). In contrast, high parsimony scores are produced by negative associativity, since many temperature changes will be encountered due to the juxtaposition of high temperature organisms with low temperature organisms (fig. 1*b*).

The relationship between environmental factors and HGT was assessed by correlating the clustering of factor values on each gene tree, (Williams and Fitch 1989; Allen and Steel 2001). This correlation provides a measure of whether a given factor is positively or negatively associative. The zero associativity line corresponds to the random horizontal transfer of genes independent of the donors' and recipients' environmental parameters and separates the regions of positive and negative associativity (fig. 2*a*). For optimum growth temperature, for example, zero associativity would produce an assortment of trees sometimes containing groups of like temperature organisms and sometimes containing vastly different temperature organisms juxtaposed. When correlated with horizontal transfer, the parsimony scores under the zero associativity model will be, on average, less than those for negative associativity (fig. 2*a* [top region]) but greater than those for positive associativity (fig. 2*a* [bottom region]).

Parsimony scores computed for each set of orthologs are plotted against the number of horizontal transfers and illustrated for maximum growth temperature in figure 2*b*. In this example, the regression line for the data and for the zero associativity model are both shown. Since the slope of the regression over the experimental data in figure 2*b* is less than the zero associativity slope, these data imply positive associativity, indicating that HGT occurs preferentially among organisms with similar maximum growth temperatures. All parameters examined were positively associative, although the strength, and hence statistical significance, of the relationship varied with the parameter.

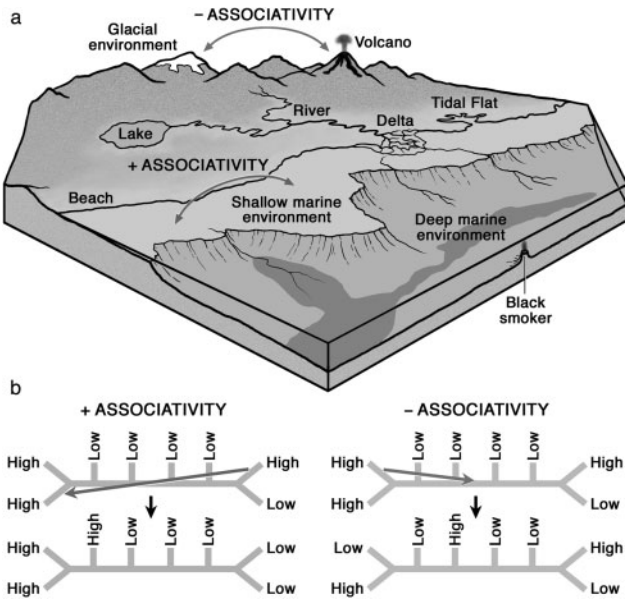

FIG. 1.—Illustrations of positive associative HGT and negative associative HGT. (*a*) Positive associativity refers to the preferential movement of DNA among similar environments, as shown here between two shallow marine environments. Negative associativity describes the preferential exchange of DNA among disparate environments, as illustrated here between a glacial environment and a volcanic one. (*b*) A hypothetical example illustrating how a two-state environmental parameter mapped onto an eight taxon tree changes in response to positive associativity (at the left) or to negative associativity (at the right). The directions of the arrows reflect the movement of a branch from one region of the tree to another and correspond to the movement of a gene in the opposite direction. Positive associativity preferentially results in organisms with similar parameters being clustered on the tree (low parsimony scores), whereas negative associativity preferentially results in organisms with different parameters being juxtaposed on the tree (high parsimony scores).

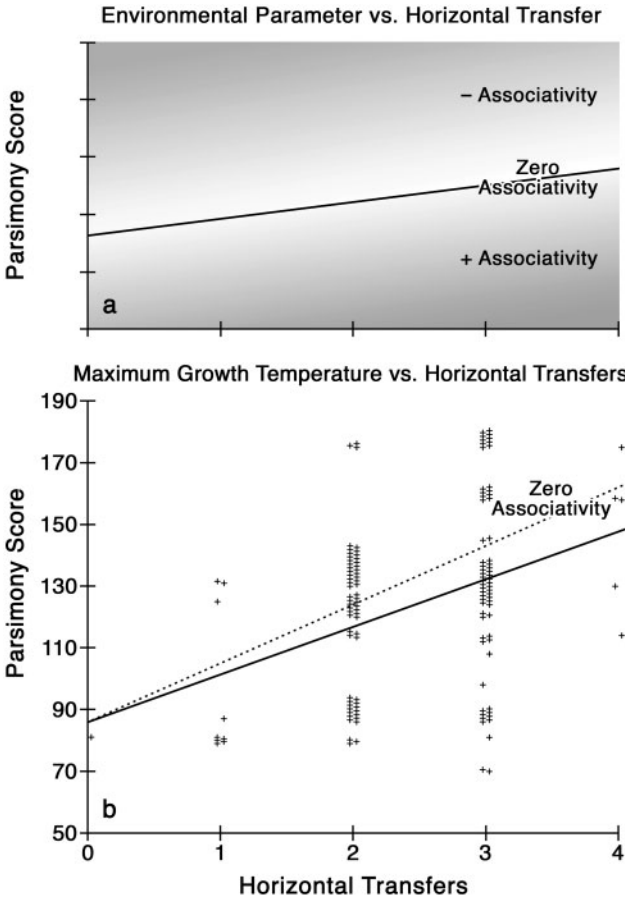

FIG. 2.—Illustrations of how associativities are measured. (*a*) For a hypothetical environmental parameter, the zero associativity regression between parsimony score and number of horizontal transfers is shown as a black line. When an observed regression line has a slope greater than that of the zero associativity line, this corresponds to negative associativity, since the parsimony scores are increasing faster than they would if horizontal gene transfers were independent of the environmental parameter. When an observed regression line has a slope less than that of the zero associativity line, this corresponds to positive associativity, since the parsimony scores are increasing more slowly than they would if horizontal gene transfers were independent of the environmental parameter. (*b*) A linear regression between parsimony score and horizontal transfers for the maximum growth temperature data (solid line) and for the zero associativity model (dotted line). Since the regression line for the data has a smaller slope than the zero associativity model, maximum growth temperature restricts HGT, resulting in lower parsimony scores for the observed data than the zero associativity model data. Alternating data points have been offset for illustration purposes on the X-axis by 0.025 horizontal transfers and on the Y-axis by 0.6 parsimony scores.

Bootstrap calculations were used to estimate the statistical significance of each of the associativities. Bootstraps indicated that the slopes for each of the 10 parameters were normally distributed and suggested a *Z*-score analysis of slopes. For each parameter, the experimentally determined mean slopes and variances from the bootstraps (10,000 replicates each) were transformed to $Z = 0$, and $\sigma^2 = 1$, respectively, as shown by the bell curve in figure 3. The zero associativity slopes for each of the 10 parameters were also transformed using the same *Z*-score transformations and are indicated in figure 3. In all cases, the position of the zero associativity slope lies to the right of the distribution, indicating positive associativity for all factors. Negative *Z*-scores would have corresponded to negative associativity. The closer their *Z*-scores are to zero, the less influence each particular environmental parameter has on restricting HGT and defining exchange groups. For the mean of the 10 factors, the associativity is significantly positive ($P < 0.0005$, $Z = 3.55$, by the two-tailed test). Moreover, each of the individual parameters correspond to positive associativities, albeit to different degrees. The internal discriminants genome size, G/C content, and carbon utilization have the greatest positive associativities (likes exchanging with likes). These are followed by oxygen and all three temperature variables, which also have statistically significant positive restrictions on HGT. Salinity, pH, and $\log_{10}$pressure have only weak effects.

As our data show, internal and external environmental factors strongly influence which genetic material a prokaryote may acquire by HGT. The positive associativity of some of the parameters tested here is readily understood. For example, prokaryotes preferentially acquire genes from other prokaryotes when both live in the same oxygenic environments because there are many enzymes that do not function in the presence, or absence, of

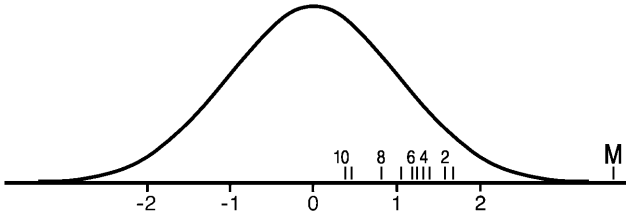

FIG. 3.—Z-score analysis for the observed associativity of each environmental parameter and for the mean of all parameters. The vertical lines mark the Z-score transformed zero associativity slopes for each of the environmental parameters and for the mean of all environmental parameters. The numbers, and letters, from right to left, correspond to the following environmental parameters: M, mean of all environmental parameters, $Z = 3.550$ ($P < 0.001$, two-sided); 1, genome size, $Z = 1.730$ ($P < 0.05$, one-sided); 2, G/C content, $Z = 1.599$ ($P < 0.1$, one-sided); 3, carbon, $Z = 1.406$ ($P < 0.1$, one-sided); 4, oxygen, $Z = 1.288$ ($P < 0.1$, one-sided); 5, maximum growth temperature, $Z = 1.221$ ($P < 0.1$, one-sided); 6, minimum growth temperature, $Z = 1.214$ ($P < 0.1$, one-sided); 7, optimum growth temperature, $Z = 1.112$ ($P < 0.2$, one-sided); 8, salinity, $Z = 0.832$; 9, pH, $Z = 0.428$; 10, $\log_{10}$pressure, $Z = 0.364$. The zero associativity slopes all lie on the positive side of the normal curve, corresponding to positive horizontal transfer associativity.

oxygen (Schneegurt et al. 1994; Ghirardi, Togasaki, and Seibert 1997).

Similarly, temperature differences can restrict HGT. In the case of a gene being transferred from a mesophilic environment to a thermophilic one, restriction may be due to the inactivation of mesophilic proteins at higher temperatures. In the reverse direction, a thermophilic protein may not function due to the need for higher temperatures for enzyme catalysis. Also, naked DNA is much more labile at higher temperatures (Kozyavkin et al. 1995), hampering the circulation of mesophilic DNA lacking thermal protective mechanisms within high temperature environments.

The mode of harnessing energy (carbon utilization) also has a restrictive influence on HGT. Heterotrophs prefer to exchange genes with each other, as do autotrophs, perhaps because the opportunity to utilize a novel carbon source may be advantageous.

The two factors with the strongest positive associative influences are the internal determinants G/C content and genome size. The strong positive associativity of these two factors may be due to their direct effects on the incorporation of new DNA into existing organisms, and possibly also due to a cumulative effect of other environmental factors on G/C composition and on genome size. For example, nucleotide preferences can restrict the ensembles of possible regulatory signals, which are recognized, thereby preferentially aiding the incorporation of genes with similar G/C ratios. Regarding genome size, free-living carbon heterotrophs generally have larger genomes than autotrophs (e.g., *Rhodopseudomonas palustris* versus *Synechocystis PCC6803*) (Kaneko et al. 1996). One reason for this may be that heterotrophs need to exist on diverse carbon substrates, requiring large numbers of ancillary enzymes and pathways that must be encoded. There is some evidence that an increase in genome size in heterotrophs carries with it an increase in metabolic diversity (Deckert et al. 1998).

We note that the three most strongly supported factors are independent of proximity. Genome size, G/C composition, and carbon utilization vary widely within microbial communities, and yet have strong associativities, indicating that exchange communities are not necessarily in physical proximity.

At the other end of the spectrum of restrictions, the pressure at which prokaryotes live has the least positive influence on HGT. The optimum pH for growth also seems to have little effect on HGT, but this is most likely a result of using pH values of the growth medium, rather than the pH of the environment at the site of isolation of each species. Such environmental pH values often vary and are subject to environmental fluxes. Because of their variability, any possible effect a particular pH may have on HGT remains hidden. To test properly what role pH may play on HGT, genomes from organisms living in stable pH environments are needed. Lastly, the parameter salinity has a slight positive associativity. This might be explained either by enzymes having reduced activity in environments with suboptimal salt concentrations or by internal salt concentrations not being strongly correlated with external environments (Chrost 1991).

It is difficult to ascertain how much HGT has accelerated prokaryotic genome innovation, but the acceleration is significant. Assuming the number of novel genes originating per unit time is proportional to the number of cells producing them, then the average increase in innovation due to HGT can be calculated by dividing the number of cells/exchange group by the number of cells/species. The following numbers then are useful for calculation. It has been estimated that there are 10^9 prokaryotic species on Earth containing 10^{30} prokaryotic cells (Dykhuizen 1998; Whitman, Coleman, and Wiebe 1998), corresponding to about 10^{21} prokaryotic cells per species. The sizes of exchange communities are unknown, but some of the parameters characterizing them are not too different from those of some terrestrial ecosystems. The median prokaryotic population size of 12 diverse soil ecosystem types, as reviewed by Whitman, Coleman, and Wiebe (1998) is about 10^{28} prokaryotes, suggesting that an exchange community could contain some 10^{28} prokaryotes. If so, then the relative increase in innovation due to HGT would be $10^{28}/10^{21}$, or 10^7. Allowing three orders of magnitude for the inexactness of our estimate, then the increase in innovation afforded by HGT could be as small as 10^4 and as large as 10^{10}. Either would constitute huge HGT-dependent increases in innovation.

This study provides evidence for a significant effect of some components of the environment on HGT. Genes are preferentially exchanged among organisms sharing similar genome size, genome G/C composition, carbon utilization, and oxygen tolerance. Indeed HGT may be responsible for a remarkable increase in genome innovation that greatly exceeds anything that could have been accomplished by clonal evolution alone.

Acknowledgments

This work was supported by DOE, NSF, and NASA Astrobiology Institute grants to J.A.L. R.J. and J.E.M. were both supported by the IGERT Bioinformatics

Program, J.E.M by the NIH CMB training grant, and R.J. by the NIH genetics training grant. We thank L. Margaret Kowalczyk for help with the illustrations, Todd Gary for helpful comments, and two anonymous reviewers for the helpful advice.

Literature Cited

Allen, B. L., and M. Steel. 2001. Subtree transfer operations and their induced metrics on evolutionary trees. Ann. Combinatorics **5**:1–15.

Blattner, F. R., G. Plunkett, C. A. Bloch et al. (14 co-authors). 1997. The complete genome sequence of *Escherichia coli K-12*. Science **277**:1453–1462.

Bult, C. J., O. White, G. J. Olsen et al. (37 co-authors). 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science **273**:1058–1073.

Chrost, R. J. 1991. Microbial enzymes in aquatic environments. Springer-Verlag, New York.

Davison, J. 1999. Genetic exchange between bacteria in the environment. Plasmid **42**:73–91.

Deckert, G., P. V. Warren, T. Gaasterland et al. (12 co-authors). 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature **392**:353–358.

Doolittle, W. F. 1999. Lateral genomics. Trends Genet. **15**:M5–M8.

Dykhuizen, D. E. 1998. Santa rosalia revisited: Why are there so many species of bacteria? Antonie Leeuwenhoek **73**:25–33.

Ghirardi, M. L., R. K. Togasaki, and M. Seibert. 1997. Oxygen sensitivity of algal h-2-production. Appl. Biochem. Biotechnol. **67**:182–182.

Groisman, E. A., and H. Ochman. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87**:791–794.

Hilario, E., and J. P. Gogarten. 1993. Horizontal transfer of ATPase genes—the tree of life becomes a net of life. Biosystems **31**:111–119.

Jain, R., M. C. Rivera, and J. A. Lake. 1999. Horizontal gene transfer among genomes: the complexity hypothesis. Proc. Natl. Acad. Sci. USA **96**:3801–3806.

Kaneko, T., S. Sato, H. Kotani et al. (24 co-authors). 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis sp*. Strain pcc6803. Ii. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. **3**:109–136.

Kawarabayasi, Y., Y. Hino, H. Horikawa et al. (27 co-authors). 2001. Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. DNA Res. **8**:123–140.

Klenk, H. P., R. A. Clayton, J. F. Tomb et al. (48 co-authors). 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature **390**:364–370.

Kozyavkin, S. A., A. V. Pushkin, F. A. Eiserling, K. O. Stetter, J. A. Lake, and A. I. Slesarev. 1995. DNA enzymology above 100 degrees C: topoisomerase V unlinks circular DNA at 80–120 degrees C. J. Biol. Chem. **270**:13593–13595.

Kunst, F., N. Ogasawara, I. Moszer et al. (148 co-authors). 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature **390**:249–256.

Lake, J. A. 1991. The order of sequence alignment can bias the selection of tree topology. Mol. Biol. Evol. **8**:378–385.

———. 1994. Reconstructing evolutionary trees from DNA and protein sequences—paralinear distances. Proc. Natl. Acad. Sci. USA **91**:1455–1459.

———. 1998. Optimally recovering rate variation information from genomes and sequences: pattern filtering. Mol. Biol. Evol. **15**:1224–1231.

Lorenz, M. G., and W. Wackernagel. 1994. Bacterial gene-transfer by natural genetic-transformation in the environment. Microbiol. Rev. **58**:563–602.

Martin, W. 1999. Mosaic bacterial chromosomes: A challenge en route to a tree of genomes. Bio-Essays **21**:99–104.

Ochman, H., J. G. Lawrence, and E. A. Groisman. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405**:299–304.

Piper, P. W., C. Emson, C. E. Jones, D. A. Cowan, T. M. Fleming, and J. A. Littlechild. 1996. Complementation of a pgk deletion mutation in *Saccharomyces cerevisiae* with expression of the phosphoglycerate kinase gene from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Curr. Genet. **29**:594–596.

Rivera, M. C., R. Jain, J. E. Moore, and J. A. Lake. 1998. Genomic evidence for two functionally distinct gene classes. Proc. Natl. Acad. Sci. USA **95**:6239–6244.

Salama, N. R., and S. Falkow. 1999. Genomic clues for defining bacterial pathogenicity. Microbes Infect. **1**:615–619.

Schneegurt, M. A., D. M. Sherman, S. Nayar, and L. A. Sherman. 1994. Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium *Cyanothece* sp strain atcc-51142. J. Bacteriol. **176**:1586–1597.

Smith, D. R., L. A. Doucettestamm, C. Deloughery et al. (34 co-authors). 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum Delta H*: Functional analysis and comparative genomics. J. Bacteriol. **179**: 7135–7155.

Spratt, B. G., L. D. Bowler, Q. Y. Zhang, J. J. Zhou, and J. M. Smith. 1992. Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal *Neisseria* species. J. Mol. Evol. **34**: 115–125.

Stetter, K. O. 1996. Hyperthermophilic procaryotes. FEMS Microbiol. Rev. **18**:149–158.

Stetter, K. O., G. Fiala, G. Huber, R. Huber, and A. Segerer. 1990. Hyperthermophilic microorganisms. FEMS Microbiol. Rev. **75**:117–124.

Syvanen, M. 1994. Horizontal gene-transfer—evidence and possible consequences. Annu. Rev. Genet. **28**:237–261.

Thomas, T., and R. Cavicchioli. 2000. Effect of temperature on stability and activity of elongation factor 2 proteins from antarctic and thermophilic methanogens. J. Bacteriol. **182**:1328–1332.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: the unseen majority. Proc. Natl. Acad. Sci. USA **95**:6578–6583.

Williams, H. G., M. J. Day, J. C. Fry, and G. J. Stewart. 1996. Natural transformation in river epilithon. Appl. Environ. Microbiol. **62**:2994–2998.

Williams, P. L., and W. M. Fitch. 1989. Finding the minimal change in a given tree. Pp. 453–470 *in* B. Fernholm, K. Bremer, and H. Joernwall, eds. The hierarchy of life: molecules and morphology in phylogenetic analysis. Elsevier Science Publishers B. V., Amsterdam.

William Martin, Associate Editor

Accepted April 18, 2003