

GEOGRAPHICAL VARIATION IN THE RATE OF EVOLUTION: EFFECT OF AVAILABLE ENERGY OR FLUCTUATING ENVIRONMENT?

SAMRAAT S. PAWAR

Section of Integrative Biology, University of Texas, Austin, Texas 78712

E-mail: samraat@mail.utexas.edu

Abstract.—In a recent paper, Wright et al. (2003) argue for the hypothesis that greater biologically available energy elevates the rate of molecular evolution. However, their results are also consistent with alternative hypotheses that invoke either environmentally driven variation in effective population sizes, or natural selection, or both. The available energy gradient cited by Wright et al. is linearly correlated with temperature fluctuations, and the observed rate heterogeneity could be a consequence of this environmental variability. The distribution of phylogenetic branch lengths alone is insufficient to distinguish between the hypotheses, and complementary approaches are suggested.

Key words.—Biologically available energy, effective population size, fluctuating environment, natural selection, rate of evolution.

Received February 23, 2004. Accepted October 1, 2004.

Whether certain climatic regimes allow higher rates of biological evolution and speciation relative to others is an intriguing question. In the last few years, efforts have been made to understand the most fundamental aspect of the problem by testing whether rates of molecular evolution are correlated with geographic gradients, with equivocal results (Bleiweiss 1998; Bromham and Cardillo 2003). Recently, Wright et al. (2003) showed that net substitution rate appears to be higher in species from localities with greater mean annual temperatures (MAT), based on branch lengths from phylogenetic analyses of ribosomal DNA (rDNA) in the plant subgenus *Mearnsia*. They interpreted this correlation as support for the hypothesis that biological evolution is accelerated in regions with greater biologically available energy due to an increase in mutation rates and/or shortening of generation times (Stehli et al. 1969; Rohde 1992).

Of the many factors that can contribute to nucleotide substitution rate heterogeneity, biologically available energy could be a tenable explanation for the pattern found by Wright et al. (2003). However, these results are also consistent with other hypotheses of environment-driven rate heterogeneity. These fall under two main categories: changes in effective population size (N_e) or in natural selection due to environmental fluctuations. Although previous authors have mentioned that environmental stability might be a factor determining evolutionary rate and taxonomic diversity, they have not suggested specific mechanisms by which this could happen (Stehli et al. 1969; Currie 1991; Rohde 1992; but see Martin and Mckay 2004). I argue that environmental variability can affect substitution rate by imposing fluctuations in N_e and selection and therefore generate geographical gradients in substitution rate similar to those predicted by the biologically available energy hypothesis.

Figure 1 shows that a linear correlation exists between MAT and temperature fluctuations at two timescales (across months and years) along the localities sampled by Wright et al. This is a general trend; MAT decreases, and temperature variability increases with latitude at different timescales (Fig. 2; also see Jansson and Dynesius 2002; Vasseur and Yodzis 2004). Assuming that such environmental variation affects

N_e and selection regimen, it is important to consider whether the phylogenetic branch length pattern shown by Wright et al. (2003) is unique to an elevated mutation rate scenario. I will discuss this in the light of studies on the effects of selection and fluctuating N_e on genealogies.

Consider a sample of homologous DNA sequences taken from a diploid population of N_e individuals. Under the neutral Wright-Fisher model (Hudson 1990), the probability of two randomly chosen sequences in the current generation coming from a single ancestral sequence (i.e., coalescing) in the previous generation is $1/(2N_e)$, and that of coalescing $t + 1$ generations ago is $(1/(2N_e))(1 - 1/(2N_e))^t$, which can be approximated by an exponential distribution with mean $2N_e$ (Hudson 1990). Assume that the number of mutations in a sequence over t generations is a Poisson variable with expectation $t\mu$, where μ is the mutation rate per generation. Rescaling time according to the coalescent, that is, $T = t/(2N_e)$, the number of mutations becomes $T\theta/2$, where $\theta = 4N_e\mu$, the expected number of nucleotide differences between the two sequences (Hudson 1990). Thus, under neutral mutation-drift balance, an elevated μ and/or N_e inflates θ , increasing genealogical tree length. If μ or N_e are unequal across lineages, genealogical branch lengths will vary accordingly across these lineages. In essence, Wright et al. (2003) invoke the scenario of uneven μ (or generation time) to explain the rate heterogeneity inferred from their phylogenetic tree.

Selection plays no part in the neutral coalescent model under the combined assumption of neutrality and the over-riding effect of drift (Kaplan et al. 1988). However, there are now a number of studies on the effect of selection on genetic variation and genealogies (see Barton and Etheridge 2004, and references therein). Depending upon the level of and fluctuations in genetic variation maintained by the modeled selection regime, these studies predict shortening or lengthening of genealogies relative to the expectation of neutral coalescent theory. In general, selection regimes that reduce neutral genetic variation (such as selective sweeps and purifying selection) shorten branch lengths, analogous to the effect of reduction in N_e . For example, for a diploid, equal-

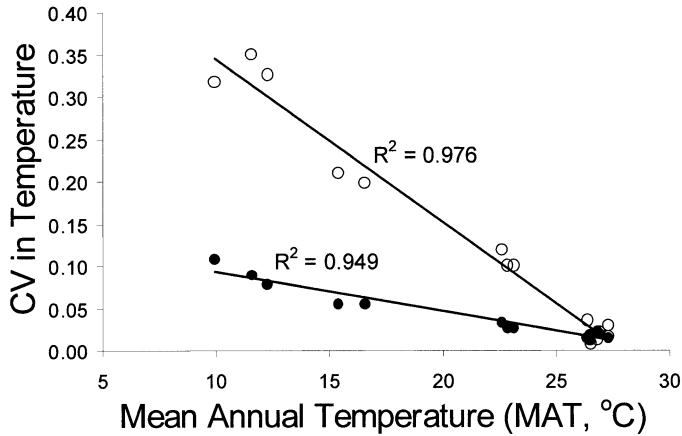


FIG. 1. The relationship between mean annual temperature (MAT) and fluctuations in temperature, measured as coefficient of variation (CV) in mean temperature across months (open circles) and years (closed circles) for the areas covered by Wright et al. (2003). Both decrease linearly with increasing MAT, that is, sites with higher MAT also have a less variable temperature regime across months as well as years. All three variables are themselves a function of latitude; sites on the right extreme of the graph are from equatorial regions. Coefficients of variation were calculated from across-month (data from same month averaged across years) and yearly (across-month data averaged for every year) temperature data for representative sites from sampling localities for the genus *Mearnsia* given by Wright et al. (2003, from their Fig. 1; data from http://bonnet19.cs.qc.edu:7778/pls/rschdata/rd_start.main).

effects mutation model, Charlesworth et al. (1993) predict that for a mean selection coefficient of a new mutation σ , and dominance coefficient h , the expected length of each branch of the genealogy will be the neutral branch length reduced by a factor of $f_o = e^{-\mu/(2h\sigma)}$ provided that $\mu/(2h\sigma)$ is sufficiently large (i.e., $f_o \ll 1$). Balancing selection, however, can either shorten or lengthen genealogies, depending upon the association of neutral sequences with the region under selection and the persistence time of a balanced polymorphism (Charlesworth et al. 2003; Barton and Etheridge 2004).

Figure 3 shows some possible genealogies under the scenarios outlined above. Clearly, if environmental changes cause variation in N_e and/or selection, branch lengths and inferred net substitution rates will be affected, and similar tree structure could be observed with elevated mutation rate, fluctuating N_e , or periodic selection. For example, fluctuations in N_e driven by environmental factors will tend to reduce genetic diversity and shorten genealogies (Fig. 3c; Charlesworth et al. 2003).

A crucial question here is whether temperature fluctuations along Wright et al.'s geographic gradient actually affect N_e or selection regimens. Temporal environmental variations can affect populations by creating fluctuations in population size (Wichmann et al. 2003) and direction/strength of selection (Meyers and Bull 2002). Of these effects, the most obvious causal link between temperature fluctuation and net substitution rate is the effect on N_e , which is subject to more periodic reductions in higher latitudes relative to the tropics (Vasseur and Yodzis 2004). Alternatively, if the evolution of rDNA regions used by Wright et al. (or linked loci) is in any way influenced by temperature (or environmental covariates), the gradient of fluctuations could impose selection

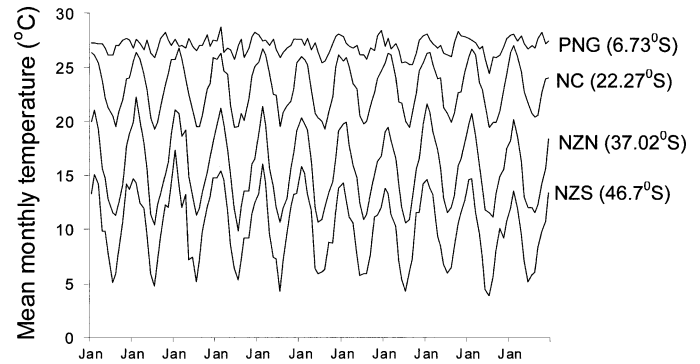


FIG. 2. Variation in temperature across months over 11 years for four sites spanning Wright et al.'s (2003) geographical gradient. NZS, New Zealand South; NZN, New Zealand North; NC, New Caledonia; PNG, Papua New Guinea. The latitude of each site is in parentheses. All sites are from a similar altitude (0–20 m above sea level). Note that the fluctuations increase away from the equator. Also, Vasseur and Yodzis (2004) showed that even after removing the effect of seasonality, higher latitudes tend to have more random environmental noise.

regimes analogous to those assumed by the above-mentioned models. For example, shifts in adaptive optima driven by temperature fluctuations could lead to bouts of genetic variation—reducing selective sweeps.

Additionally, it must be noted that Figures 1 and 2 are a simplified representation of the actual nature of environmental variation along this latitudinal gradient, which has fluctuations at a wider range of temporal scales (from within month to between years and longer), and randomness (cf. Jansson and Dynesius 2002; Vasseur and Yodzis 2004). Such variation spans the generation times of at least a subset of *Mearnsia* species, which have a variety of life-history strategies (cf. Wright et al. 2003). This is important because the above-mentioned impacts of environmental fluctuation on population and evolutionary dynamics depend on its amplitude, randomness, and frequency in relation to the organism's generation time. For example, under an infinite allele model, autocorrelated fluctuating selection of a frequency greater than or equal to the organism's generation time is expected to maintain periodic cycling of genetic variation (Gillespie 1991). Fluctuations shorter than an organism's generation time, however, are less likely to drive population-level changes in genetic variation, favoring the evolution of phenotypic plasticity (Meyers and Bull 2002).

On the whole, it appears that Wright et al.'s results could be generated by alternative mechanisms, and distribution of branch lengths alone is insufficient to differentiate between them. Resolving this problem is not easy, as no single method is likely to be able to differentiate between these convergent effects of different environmental factors. I suggest a few approaches that could be used to separate the effects of some of these factors on evolutionary rate.

1. Determine geographic gradients in genetic variation

A direct approach would be to explore environmental correlates of genetic and phenotypic variation in a single species' populations distributed along geographic gradients. This will

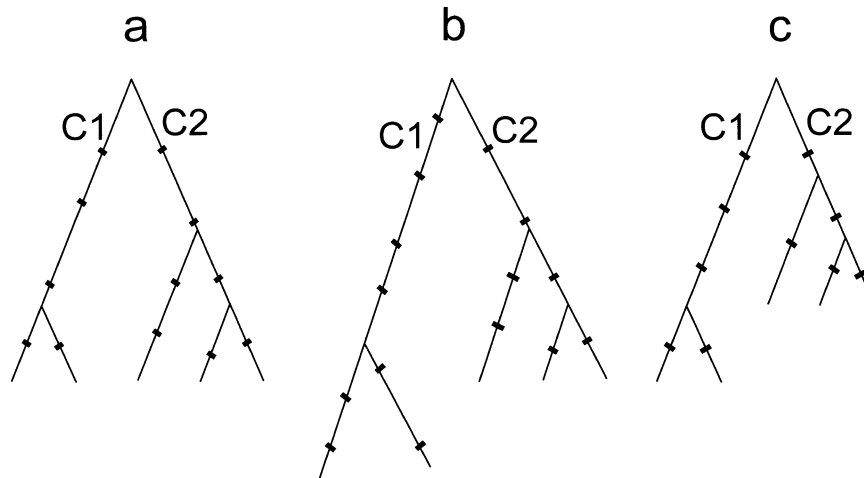


FIG. 3. Possible genealogies with (a) neutral evolution without rate heterogeneity, where the number of mutations is distributed proportionally throughout the tree; (b) rate heterogeneity between clades C1 and C2 due to elevated mutation rates in C1; and (c) effect of variation-reducing mechanisms such as reduced N_e , purifying selection, or selective sweeps in C2, which results in a shortening of the branches. The tick marks on the branches represent mutations. Note that the relative branch lengths in (b) and (c) are similar.

allow direct inferences about whether and how environmental differences create gradients in substitution rate.

2. Control for covariation between MAT and temperature fluctuations

Substitution rate heterogeneity could be tested along geographical gradients where the MAT and temperature variability do not covary. For example, altitudinal gradients generally have a low covariation between MAT and temperature variability. Bleiweiss (1998) found a positive relationship between available energy and substitution rate along an elevation gradient, but his results were weakened by failure to correct for phylogenetic relatedness.

3. Search for evidence of geographical gradients in selection regime

(a) Using one of many available techniques (Kreitman 2000; Ford 2002), test for geographical variation in selection at protein coding loci. One possible approach would be to examine nonsynonymous-to-synonymous substitution rate ratio ($\omega = dN/dS$) at protein coding loci, with ω values <1 , $=1$, and >1 indicating purifying selection, neutral evolution, and positive selection, respectively. Using phylogenetically independent contrasts, Bromham and Cardillo (2003) tested for heterogeneity in dN and dS between bird taxa from contrasting latitudes, but found no significant difference. (b) Compare substitution rate at a locus influencing temperature adaptation and an unlinked locus that does not, for species or populations across environments. Evolution at both types of loci should be proportionally greater in areas with higher MAT under the available energy as well as the fluctuating N_e hypotheses (because both affect the entire genome), whereas the rate should vary disproportionately under the geographic selection gradient hypothesis (because selection imposed by temperature variation should affect only the temperature-adaptation locus). (c) Use statistical tests based on

the shape and symmetry of trees for detecting effects of selection on phylogenetic trees (but see Maia et al. 2004).

The challenges of determining multiple, often correlated environmental effects on the rate of evolution are both theoretical and logistical. For instance, identifying quantitative trait loci under the direct influence of specific environmental factors will be a big step, but this is easier said than done. Additionally, factors such as elevated mutation rates, fluctuating N_e , and selection seldom operate independently, and it will be a challenge to extract testable predictions from theory. Wright et al. (2003) have made a timely contribution. However, considering the complexity of the problem, additional effort in terms of data collection and analysis will be needed to resolve this particular question about evolutionary rate heterogeneity.

ACKNOWLEDGMENTS

I am grateful to A. C. Birand, S. Biswas, J. Brown, C. Estrada, S. Hedtke, G. B. Pauly, and S. Sarkar for discussions and comments. I would also like to thank an anonymous reviewer for helpful insights and suggestions that greatly improved this manuscript.

LITERATURE CITED

- Barton, N. H., and A. M. Etheridge. 2004. The effect of selection on genealogies. *Genetics* 166:1115–1131.
- Bleiweiss, R. 1998. Slow rate of molecular evolution in high-elevation hummingbirds. *Proc. Natl. Acad. Sci. USA* 95:612–616.
- Bromham, L., and M. Cardillo. 2003. Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. *J. Evol. Biol.* 16:200–207.
- Charlesworth, B., D. Charlesworth, and N. H. Barton. 2003. The effects of genetic and geographic structure on neutral variation. *Annu. Rev. Ecol. Syst.* 34:99–125.
- Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303.
- Currie, D. J. 1991. Energy and large-scale patterns of animal species and plant species richness. *Am. Nat.* 137:27–49.

- Ford, M. J. 2002. Applications of selective neutrality tests to molecular ecology. *Mol. Ecol.* 11:1245–1262.
- Gillespie, J. H. 1991. The causes of molecular evolution. Oxford series in ecology and evolution. Vol. 2. Oxford Univ. Press, Oxford, U.K.
- Hudson, R. R. 1990. Gene genealogies and the coalescent process. *Oxf. Surv. Evol. Biol.* 7:1–44.
- Jansson, R. and M. Dynesius. 2002. The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annu. Rev. Ecol. Syst.* 33:741–777.
- Kaplan, N. L., T. Darden, and R. R. Hudson. 1988. The coalescent process in models with selection. *Genetics* 120:819–829.
- Kreitman, M. 2000. Methods to detect selection in populations with applications to the human. *Annu. Rev. Genomics Hum. Genet.* 1:539–559.
- Maia, L. P., A. Colato, and J. F. Fontanari. 2004. Effect of selection on the topology of genealogical trees. *J. Theor. Biol.* 226:315–320.
- Martin, P. R., and J. K. Mckay. 2004. Latitudinal variation in genetic divergence of populations and the potential for future speciation. *Evolution* 58:938–945.
- Meyers, L. A., and J. J. Bull. 2002. Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol. Evol.* 17:551–557.
- Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* 65:514–527.
- Stehli, F. G., R. G. Douglas, and N. D. Newell. 1969. Generation and maintenance of gradients in taxonomic diversity. *Science* 164:947–949.
- Vasseur, D. A., and P. Yodzis. 2004. The color of environmental noise. *Ecology* 85:1146–1152.
- Wichmann, M. C., K. Johst, K. A. Moloney, C. Wissel, and F. Jeltsch. 2003. Extinction risk in periodically fluctuating environments. *Ecol. Modell.* 167:221–231.
- Wright, S. D., R. D. Gray, and R. C. Gardner. 2003. Energy and the rate of evolution: inferences from plant rDNA substitution rates in the western Pacific. *Evolution*. 57:2893–2898.

Corresponding Editor: G. May