The first steps in adaptive evolution

James J Bull & Sarah P Otto

The first empirical test of an evolutionary theory provides support for a mutational landscape model underlying the process of adaptation. The study shows that it is possible to predict at least the first step in an adaptive walk and also shows the importance of incorporating mutation bias in the fitness effects of mutations.

Adaptation is everywhere. Sometimes it gets in our way, as with drug-resistant microbes, pesticide-resistant insects and cancer. Sometimes it does us good, as in the domestication of plants and animals and industry's use of directed evolution to create useful molecules. Nevertheless, although the products of adaptation are well known, the mechanism and quantitative nature of the adaptive process remain poorly understood. Early attempts to describe the adaptive process on geometrical grounds[1] did not lead much beyond a rudimentary understanding of whether small-effect or large-effect mutations contribute most to adaptation[2]. Real progress towards understanding the adaptive process came from explicit models of genetic sequences[3–6]. These models predicted the fitness effects of mutations that arise and fix within a population as it adapts to its environment. On page 441 of this issue, Rokyta *et al.*[7] now provide an empirical test of this theory, finding that the first steps that adaptation takes are consistent with the theoretical predictions of the adaptive process.

Testing evolution

Theories that make real, *a priori* predictions about adaptation have been gaining momentum, and the report from Rokyta *et al.* in this issue provides the empirical support needed to launch these theories into the spotlight. This paper presents a quantitative experimental test of a theory of adaptation initially developed by John Gillespie in the 1980s (refs. 3–5) and recently extended by Allen Orr[6]. Gillespie realized that in a population slightly displaced from its closest fitness optimum, there will be but a handful of mutations that improve fitness compared with an overwhelming number that reduce fitness[3–5]. Therefore, beneficial mutations that lead to adaptation should represent the most-fit tail of the distribution of all possible mutations.

Gillespie's model made use of a common statistical theory called extreme-value theory, which indicates that samples drawn from the tail of a distribution have properties that do not depend on the exact nature of the distribution. Applied to evolution, extreme-value theory predicts an ordered progression of fitness effects among the handful of beneficial alleles: the best allele should be substantially fitter than the next-best allele, and fitness differences between pairs of next-most beneficial alleles should decline so that most of the beneficial alleles have small effects. Using this insight, Orr[6] derived predictions about the distribution of fitness effects of the mutations that arise and fix during an adaptive walk. Orr's model allowed for testable predictions about the course of adaptive evolution. In particular, he derived the probability that a mutated allele with a given fitness rank would be fixed at the next adaptive step.

The first adaptive step

Rokyta *et al.*[7] now supply the needed empirical test of this theory. The authors used a

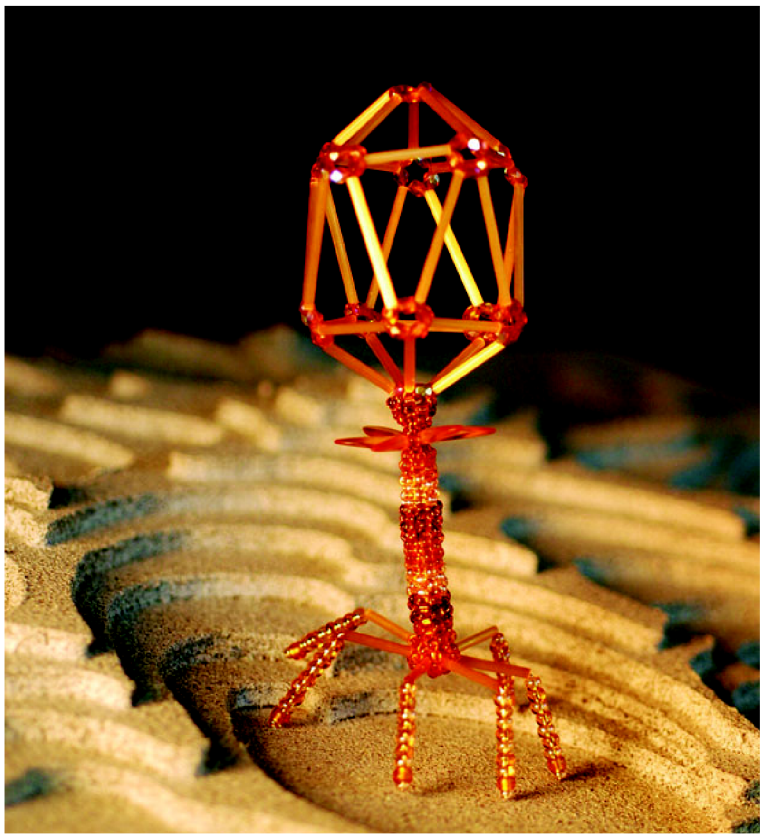

Figure 1 A model bacteriophage takes its first adaptive step on a fitness landscape. An adaptive walk along a mutational landscape, reflecting all possible mutations deriving from the initial sequence, can represent the evolution of a virus. Now, a virus's first steps in an adaptive walk have been defined to within a likely mutational landscape. Photo by J. Palmersheim; phage sculpture by H. Wichman; landscape by A. Johnston.

James J. Bull is in the Section of Integrative Biology and Institute for Cellular and Molecular Biology, University of Texas, Austin, Texas 78712-0253, USA. Sarah P. Otto is in the Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada.
e-mail: bull@bull.biosci.utexas.edu

single-stranded DNA virus to determine whether the beneficial mutation fixed at the first substitution in an adaptive walk agreed with that predicted by Orr's theory (**Fig. 1**). The empirical test was not trivial, requiring considerable replication as well as detailed information on the identity and fitness of substituted alleles. To accomplish this, Rokyta *et al.* focused on the first step in adaptation, allowing for greater replication and predictive ability.

Rokyta *et al.* carried out 20 replicate single-step adaptations using a wild relative of ΦX174 grown in liquid culture, allowing each line to adapt independently to the same conditions. In each replicate, the first mutation to both arise and spread in the line was identified by whole-genome sequencing. The fitness effect of each mutation was measured as the growth rate of the virus, and the 20 fitness effects were ranked. Of the 20 first adaptive steps examined, all mutations were nonsynonymous, involving nine distinct amino acid changes, and all increased fitness (from 11% to 39%).

Comparing these experimental results with previous theory, Rokyta *et al.* found that Orr's model fit the observed fitness distribution of the mutations reasonably well. The predictions of Orr, however, are only expectations over all possible genomic starting points and over all possible adaptive walks. Rokyta *et al.* found a substantially improved fit to the data by incorporating mutation rate differences between their starting sequence and each of the nine observed amino acid changes. A slightly better fit was obtained using all the available data (including the mutation rate differences, fitness effects and population size dynamics). Thus, the authors found that models tailored to the specifics of the population could better describe the process of adaptation. It is notable that, without this additional knowledge of the starting and mutant sequences, Orr's predictions faired so well.

Rigorous biological tests of these models are presently limited to small genomes (viruses, plasmids and single molecules subjected to *in vitro* selection) and to computer models of fitness landscapes. Tests using bacteria, yeast and higher eukaryotes await the cost-feasibility of sequencing large genomes with multiple replicates of a single experiment. But even now, tests are feasible and of obvious relevance for predicting drug resistance evolution in viruses, including HIV and influenza, two viruses for which drugs have or could have a crucial role in treatment and for which we already know that evolution causes problems. In some cases, we might want the theory tailored to the individual genome, a combination of Orr's model and the modifications offered by Rokyta *et al.*

Next steps

In showing the relevance of existing adaptation theories to real experimental conditions, this work brings new excitement to adaptive walks. For any particular system, general properties about the course of adaptive walks can now be predicted based on only a few parameters. To borrow from Fisher, no practical biologist would have dared imagine that the details of adaptive walks might be largely independent of the biology, yet that is precisely what the current results suggest. This work shows that it is possible to predict the spectrum of possible first steps of an adaptive walk. The next steps will be to extrapolate this work to the full course of an adaptive walk. Such predictive power would be extremely valuable when anticipating the evolutionary response of pathogens to new antimicrobial drugs and when using directed evolution to create molecules with specific functions for industry.

1. Fisher, R.A. *The Genetical Theory of Natural Selection* (Oxford University Press, Oxford, UK, 1930).
2. Orr, H.A. *Nat. Rev. Genet.* **6**, 119–127 (2005).
3. Gillespie, J.H. *Theor. Popul. Biol.* **23**, 202–215 (1983).
4. Gillespie, J.H. *Evolution* **38**, 1116–1129 (1984).
5. Gillespie, J.H. *The Causes of Molecular Evolution* (Oxford University Press, Oxford, UK, 1991).
6. Orr, H.A. *Evolution* **56**, 1317–1330 (2002).
7. Rokyta, D.R., Joyce, P., Caudle, S.B. & Wichman, H.A. *Nat. Genet.* **37**, 441–444 (2005).

The X chromosome: not just her brother's keeper

Eric J Vallender, Nathaniel M Pearson & Bruce T Lahn

The X chromosome has traditionally been characterized as a conscientious sister to her derelict brother that is the Y. Beyond dutifully maintaining the family heritage, however, the X has developed its own unique identities. Now, the complete sequence of the human X allows us to appreciate its distinctiveness at an unprecedented resolution.

Since its discovery, the human X chromosome has been defined by its relationships to other chromosomes. Its hemizygosity in males and unusual patterns of inheritance immediately separate it from its autosomal cousins, and its length and gene content set it apart from its Y chromosome sibling. The two sex chromosomes in mammals descended from a pair of autosomes[1]. The Y underwent massive degeneration, losing size and gene content, whereas the X was maintained, retaining its size and most of its genes. But the X is much more than a faithful copy of its autosomal progenitor; it also evolved many distinctive features[2]. The most notable of these include X inactivation, the extensive flux (both accretion and loss) of sex-specific genes, and a deficit in polymorphism. Additionally, the X has a disproportionately large representation of genes involved in mendelian diseases, probably owing to the relative ease of identifying such genes when X-linked. As recently reported by Mark Ross and colleagues in *Nature*[3], the sequence of the human X brings these and other features of this chromosome into sharp focus.

Eric J. Vallender, Nathaniel M. Pearson and Bruce T. Lahn are in the Howard Hughes Medical Institute, Department of Human Genetics, University of Chicago, Chicago, Illinois 60637, USA. e-mail: blahn@bsd.uchicago.edu

Origin of the sex chromosomes

The mammalian sex chromosomes arose from autosomal progenitors ~300 million years ago. Before then, sex determination probably relied on environmental cues such as egg incubation temperature, as is the case in many extant reptiles. Ross *et al.*[3] confirmed that much of the long arm of the human X is homologous to the short arm of chicken chromosome 4, whereas most of the short arm of the human X matches a stretch of chicken chromosome 1. The bird sex chromosomes Z and W, on the other hand, show homology to human chromosome 9. These observations demonstrate the independent origins of genetic sex determination in mam-