In search of molecular darwinism

Paul M. Sharp

Attempts to detect adaptive evolution at the molecular level have met with little success. Studies of a digestive enzyme in primates, involving the reconstruction of DNA sequences that have long been extinct, show a way forward.

THE molecular genetics revolution has transformed evolutionary biology, but with rather more success in some areas than in others. About 70 years ago, in an essay predicting the future of biology, J. B. S. Haldane wrote[1] that "our ideal is to establish a family tree of plants and animals", in which he included being able to state when in the past the common ancestors of extant species lived. Haldane suggested that this could be achieved by chemical methods, and he was right, even if his timescale was not: Haldane predicted that it would take thousands of years, whereas phylogenetics has already been the outstanding success story of molecular evolution. Haldane foresaw not only the charting of the evolutionary past, but also the prospect of finding evidence for the central mechanism of darwinian evolution, namely adaptive change. He looked forward to being able to ask "What inheritable variations ... show any sign of being ... of advantage to their possessor?". In this respect, molecular data have, as yet, thrown up disappointingly few examples.

One apparent success concerns the enzyme lysozyme in primates. Colobine monkeys are leaf-eaters with a fermentative foregut where lysozyme is used in bacteriolysis in much the same way as in ruminants. Some years ago, it was shown that in one colobine, the hanuman langur (Fig. 1), lysozyme has amino-acid differences from other primates that parallel changes seen in the cow[2]. That the same amino-acid changes had occurred in two independent lineages provided firm hints that they were provoked by parallel selective pressures.

Messier and Stewart (page 151 of this issue[3]) have now examined the evolution of primate lysozyme in much more detail

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

E. & D. Hosking/FLPA

FIG. 1 Foregut in ferment — the hanuman (or common) langur (*Presbytis entellus*), one of the colobine Old World monkeys that have a foregut adapted to the digestion of leaves.

— they have pinned down the proposed adaptive changes to a short period during the early emergence of colobine monkeys (see the authors' Fig. 1 on page 152), and provide unexpected evidence for another period of adaptive evolution in the ancestors of hominoids (Fig. 2, page 153). Their approach has been to contrast the rates of non-synonymous and synonymous substitutions (K_A/K_S ratios; see box) among the lysozyme genes from a wide range of primate species. Comparisons of K_A and K_S have often been used elsewhere, but for reasons that are (with hindsight) fairly obvious, have identified few cases of adaptive molecular evolution.

Adaptive changes will be very hard to find using K_A/K_S ratios, particularly if we only look for ratios greater than one (see box). Genes encoding functional proteins are always expected to be subject to purifying selection, weeding out most amino-acid changes because of their detrimental effects, and thus lowering the value of K_A to substantially less than K_S. Consequently, K_A/K_S ratios may often be less than one even if some adaptive substitutions have occurred. For example, when 363 equivalent gene sequences were compared between mouse and rat[4], only one showed a ratio greater than one (Fig. 2, overleaf); yet it would be most surprising if this were the only one of these genes that had undergone adaptive changes during the divergence of the two species.

One way around this problem involves knowing where to look for adaptive evolution. Lysozyme in colobine monkeys was always likely to be a good candidate but, even there, comparisons of the entire gene sequence between extant species reveal adaptive change with low efficiency[3]. Alternatively, if we knew beforehand that specific sites within a protein were candidates for adaptive changes, we could concentrate on these and exclude many positions subject to purify-

Coding DNA substitutions

Many molecular evolutionary analyses, particularly those attempting to detect adaptive evolution, rely on distinguishing between synonymous and non-synonymous differences in DNA sequences (see ref. 8). A synonymous (or 'silent') substitution is one that, owing to the degeneracy of the genetic code, makes no change to the protein sequence encoded; a non-synonymous substitution results in an amino-acid replacement. The extent of each type of change can be estimated as K_A and K_S, respectively the numbers of non-synonymous and synonymous substitutions per site. For any comparisons other than those between very closely related sequences, these values have to be subjected to statistical methods to attempt to correct for multiple hits; that is, successive substitutions superimposed at the same site.

The approach that has been widely applied in the search for adaptive evolution of proteins is to contrast these non-synonymous and synonymous substitution rates, using the ratio K_A/K_S. Because synonymous changes seem (at least in many genes, and particularly those of mammals) to be largely neutral, and because population genetics theory indicates that neutral substitutions should occur at the neutral mutation rate, this K_A/K_S ratio is in effect examining whether non-synonymous substitutions have occurred faster or more slowly than this neutral rate. Most non-synonymous mutations are likely to be deleterious and therefore unlikely to become substitutions by spreading to fixation in the population. Thus, much of the natural selection at non-synonymous sites is expected to be negative or purifying, leading to a reduction in K_A below the neutral rate (and so to K_A/K_S ratios less than one). But if a non-synonymous mutation is advantageous, its process of fixation can happen much faster than the neutral rate. So a K_A/K_S ratio greater than one is taken as indicating that positively selected changes have inflated the numbers of non-synonymous substitutions. P.M.S.