

The future of molecular evolution

Antony M. Dean

There are just two questions to be asked in evolution: how are things related, and what makes them differ? Lamarck was the first biologist—he invented the word—to address both. In his *Philosophie Zoologique* (1809) he suggested that the relationships among species are better described by branching trees than by a simple ladder, that new species arise gradually by descent with modification and that they adapt to changing environments through the inheritance of acquired characteristics. Much that Lamarck imagined has since been superseded. Following Wallace and Darwin, we now envision that species belong to a single highly branched tree and that natural selection is the mechanism of adaptation. Nonetheless, to Lamarck we owe the insight that pattern is produced by process and that both need mechanistic explanation.

Questions of pattern, process and mechanism pervade the modern discipline of molecular evolution. The field was established when Zuckerkandl & Pauling (1965) noted that haemoglobins evolve at a roughly constant rate. Their “molecular evolutionary clock” forever changed our view of evolutionary history. Not only were seemingly intractable relationships resolved—for example, whales are allies of the hippopotamus—but also the eubacterial origins of eukaryotic organelles were firmly established and a new domain of life was discovered: the Archaea.

Yet, different genes sometimes produce different trees. Golding & Gupta (1995) resolved two-dozen conflicting protein trees by suggesting that Eukarya arose following massive horizontal gene transfer between Bacteria and Archaea. Whole genome sequencing has since revealed so many conflicts that horizontal gene transfer seems characteristic of prokaryote evolution. In higher animals—where horizontal transfer is sufficiently rare that the tree metaphor remains robust—rapid and inexpensive whole genome sequencing promises to provide a wealth of data for population

studies. The patterns of migration, admixture and divergence of species will be soon addressed in unprecedented detail.

Sequence analyses are also used to infer processes. A constant molecular clock originally buttressed the neutral theory of molecular evolution (Kimura, 1985). The clock has since proven erratic, while the neutral theory now serves as a null hypothesis for statistical tests of ‘selection’. In truth, most tests are also sensitive to demographic changes. The promise of ultra-high throughput sequencing to provide genome-wide data should help dissect selection, which targets particular genes, from demography, which affects all the genes in a genome, although weak selection and ancient adaptations will remain undetected.

In the functional synthesis (Dean & Thornton, 2007), molecular biology provides the experimental means to test evolutionary inferences decisively. For example, site-directed mutagenesis can be used to introduce putatively selected mutations into reconstructed ancestral sequences, the gene products are then expressed and purified and their functional properties determined *in vitro*. In microbial species, homologous recombination is used routinely to replace wild-type with engineered genes, enabling organismal phenotypes and fitnesses to be determined *in vivo*. The vision of Zuckerkandl & Pauling (1965) that by “furnishing probable structures of ancestral proteins, chemical paleogenetics will in the future lead to deductions concerning molecular functions as they were presumably carried out in the distant evolutionary past” is now a reality.

If experimental tests of evolutionary inferences open windows on past mechanisms, directed evolution focuses on the mechanisms without attempting historical reconstruction. Today’s ‘fast-forward’ molecular breeding experiments use mutagenic PCR to generate vast libraries of variation and high throughput screens to identify rare novel mutants (Romero & Arnold, 2009; Khersonsky & Tawfik,

2010). Among numerous topics explored are: the role of intragenic recombination in furthering adaptation, the number and location of mutations in protein structures, the necessity—or lack thereof—of broadening substrate specificity before a new function is acquired, the evolution of robustness, and the alleged trade-off between stability and catalytic efficiency. Few, however, have approached the detail found in those classic studies of evolved β -galactosidase (Hall, 2003) that revealed how the free-energy profile of an enzyme-catalysed reaction evolved. Even further removed from natural systems are catalytic RNAs that, by combining phenotype and genotype within the same molecule, allow evolution to proceed in a lifeless series of chemical reactions. Recently, two RNA enzymes that catalyse each other’s synthesis were shown to undergo self-sustained exponential amplification (Lincoln & Joyce, 2009). Competition for limiting tetranucleotide resources favours mutants with higher relative fitness—faster replication—demonstrating that adaptive evolution can occur in a chemically defined abiotic genetic system.

Lamarck was the first to attempt a coherent explanation of biological patterns in terms of processes and mechanisms. That his legacy can still be discerned in the vibrant field of molecular evolution would no doubt please him as much as it does us in promising extraordinary advances in our understanding of the mechanistic basis of molecular adaptation.

REFERENCES

- Dean AM, Thornton JW (2007) *Nat Rev Genet* **8**: 675–688
- Golding GB, Gupta RS (1995) *Mol Biol Evol* **12**: 1–6
- Hall BG (2003) *Genetica* **118**: 143–156
- Khersonsky O, Tawfik DS (2010) *Annu Rev Biochem* [Epub 17 Mar 2010]
- Kimura M (1985) *The Neutral Theory of Molecular Evolution*. Cambridge University Press
- Lincoln TA, Joyce GF (2009) *Science* **323**: 1229–1232
- Romero PA, Arnold FH (2009) *Nat Rev Mol Cell Biol* **12**: 866–876
- Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins*, Bryson V, Vogel HJ (eds), pp 97–166. Academic Press

Antony M. Dean is a professor in the Department of Ecology, Evolution and Behaviour, University of Minnesota, St Paul, MN, USA.
E-mail: deanx024@umn.edu

EMBO reports (2010) **11**, 409. doi:10.1038/embor.2010.78