

alter their expression and/or translation [11]. As Adams *et al.* [1] have shown us, we have much to learn about the genetic processes that occur upon polyploid formation. And what we will learn promises to reshape our understanding of genomic evolution.

#### References

- 1 Adams, K.L. *et al.* (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4649–4654
- 2 Keightley, P.D. and Lynch, M. (2003) Toward a realistic model of mutations affecting fitness. *Evolution* 57, 683–685
- 3 Otto, S.P. and Whitton, J. (2000) Polyploid incidence and evolution. *Annu. Rev. Genet.* 34, 401–437
- 4 Takahata, N. and Maruyama, T. (1979) Polymorphism and loss of duplicate gene expression: a theoretical study with application to tetraploid fish. *Proc. Natl. Acad. Sci. U. S. A.* 76, 4521–4525
- 5 Force, A. *et al.* (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545
- 6 Lynch, M. and Force, A. (2000) The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154, 459–473
- 7 Hughes, A.L. (1994) The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond. Ser. B* 256, 119–124
- 8 Osborn, T.C. *et al.* (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* 19, 141–147
- 9 Ferris, S.D. and Whitt, G.S. (1979) Evolution of the differential regulation of duplicate genes after polyploidization. *J. Mol. Evol.* 12, 267–317
- 10 Hughes, M.K. and Hughes, A.L. (1993) Evolution of duplicate genes in a tetraploid animal, *Xenopus laevis*. *Mol. Biol. Evol.* 10, 1360–1369
- 11 Dunlap, J.C. and Wu, C-T., eds (2002) *Homology Effects*, Academic Press

0169-5347/\$ - see front matter © 2003 Elsevier Ltd. All rights reserved.  
doi:10.1016/S0169-5347(03)00213-1

## The end of the beginning for neutral theory

Sean Nee and Graham Stone

Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh, UK EH9 3JT

**The point of ecology is to understand the distribution and abundance of species. Neutral models of species' abundances, which suppose that these fluctuate entirely at random, are currently enjoying considerable attention. To date, much of the credibility of such models has been based on their ability to mimic observed abundance distributions. A new paper by Brian McGill shows that, at least for some data, neutral theory performs worse than does an older, simpler null model. However, this is not the end for neutral theory: rather, it is the 'end of the beginning'.**

In the 1960s, gel electrophoresis revealed an unexpected amount of molecular variation in natural populations, which, it was thought, was too high to be explained by the action of natural selection. The neutral allele theory of molecular evolution [1] was a proposed explanation, in which selectively neutral variants (alleles) arise by mutation and then fluctuate at random in abundance. Tropical forests have posed a similar problem for ecologists: how can we understand their enormous tree diversity in the traditional terms of the 'one species/one niche' requirement for coexistence? After all, surely all trees want basically the same thing, a patch of ground, some water and some light?

The unified neutral theory of biodiversity and biogeography [2] proposes the same explanation: species are, ecologically, all the same, originate at random and their abundances simply fluctuate randomly over time. Originally proposed a quarter of a century ago [3,4], this idea is currently enjoying considerable interest for numerous reasons, not least of which is that it might be a very pure

null model with utility far beyond understanding tropical forest diversity [2,5]. And it appears to pass the minimum requirements for serious consideration: with suitable parameter tuning, the stochastic theory can generate patterns that appear similar to observed patterns of relative abundances of species, such as the tree species studied in the famous 50-ha plot on Barro Colorado Island in Panama [2].

It is not the only candidate for a simple null model, of course ('null' in the sense of not containing any explicit consideration of the usual biological ingredients that ecologists consider to be important [6,7]). Another simple model is this: suppose that numerous factors influence the size of any particular population and they do so in a multiplicative fashion. So, for example, an epidemic decimates the population, a good season doubles the per capita birth rate, a fire halves the amount of suitable habitat for the species and so on. This model predicts a lognormal distribution of species abundances; that is, the logarithms of their abundances should be normally distributed [8].

#### Best-fit models

So, an obvious question arises: which model fits data better? This is the question addressed by Brian McGill in a recent *Nature* paper [9]. There are actually two interpretations of this question. First, which model is more readily fitted to data? There is no contest: the lognormal is trivially easy to fit from the mean and variance of the logged species' abundances. Fitting the neutral distribution is hard and requires extensive simulation: the procedure requires five pages and a flow chart for its description [2] and, even then, it is

Corresponding author: Sean Nee (sean.nee@ed.ac.uk).

incomplete [9]. The difficulty arises from the fact that the theory imagines a large ‘metacommunity’ in which both speciation and random drift in abundances occur, connected by migration to a local community (corresponding to what is actually studied) in which only drift occurs.

This complexity ultimately endows the distribution to be fitted – given the name ‘zero-sum multinomial’ (ZSM) – with three tunable parameters compared with two for the lognormal. However, when goodness of fit to actual data is assessed using usual measures such as  $r^2$ , the lognormal is found to fit the data better, although the ZSM appears to do well [9] – the data being tree abundances from Barro Colorado Island and bird abundances from the North American Breeding Bird Surveys. With all the usual caveats (there are other data sets in the world, precisely how to fit the ZSM is still unclear, etc.), we suspect that the ZSM will not be widely used as a means of describing data.

### The future of the neutral theory

So, where stands the neutral theory of biodiversity? Certainly, what we could call the ‘early’ period is over, dominated by demonstrations that the theory can generate patterns that look like those in nature [2,5]. Although this is important to know (otherwise we would not consider the theory further), it is also well known that completely incompatible processes can generate the same distributions [10], so that the distributions themselves are uninformative. Gillespie [11] presented models of fluctuating selection that generate allele frequency distributions that are exactly the same as those predicted by the neutral theory of molecular evolution. This strongly suggests that the precisely analogous ecological ‘lottery’ models for coexistence [12] could also account for ‘neutral’ distributions.

The future of the neutral theory of biodiversity depends on what people can think of to actually do with it. The neutral theory of molecular evolution has had a profound impact on biology. Molecular phylogenetics [13] relies extensively on neutral models, which also inform the use of sequence data to reveal demographic histories of populations [14]. Finally, the models provide tests of the assumption of neutrality itself (e.g. [15]).

Although neutral models of sequence evolution and of community structure share similarities, there are, of course, enormous differences in both the underlying processes of change and the data available for testing them. First, all DNA-based genomes contain the same four basic nucleotides. In even the most complex models, the sequence changes possible at a given point in the genome are defined by a limited set of specific mutation processes (e.g. [16]). These different processes are computationally straightforward to incorporate into specific mutation models and these can, in turn, be tested using the gigabytes of sequence data available [17]. The division of genomes into loci whose evolution is effectively independent also enables us to

detect the impact of processes, such as selective sweeps, which affect only part of the genome of an individual (e.g. [18]). By contrast, neutral models of community composition involve a far larger number of alternative states (the number of species), linked by undefined processes of replacement. This makes specific models much harder to define, and there are currently fewer data available for testing them. Furthermore, in the unified neutral theory of biodiversity and biogeography, each community is equivalent to a single locus: there is no possibility here of the equivalent of tests across different loci.

However, it is entirely possible that trees have properties that are ideal for some studies. Rephrasing a result from population genetics (e.g. [19]), neutral theory predicts that, if we take a random sample of trees from the forest, the probability that any particular species is the oldest, in terms of when it arose by speciation, is equal to its frequency in the sample. Age can be assessed by where a species connects to the molecular phylogeny of the sampled species. This could be developed into a feasible test of the neutral theory for trees that is rather harder to implement for alleles.

### References

- Kimura, M. (1983) *The Neutral Allele Theory of Molecular Evolution*, Cambridge University Press
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press
- Hubbell, S.P. (1979) Tree dispersion, abundance and diversity in a tropical dry forest. *Science* 203, 1299–1309
- Caswell, H. (1976) Community structure: a neutral model analysis. *Ecol. Monogr.* 46, 327–354
- Bell, G. (2001) Neutral macroecology. *Science* 293, 2413–2417
- Gotelli, N.J. and Graves, G.R. (1996) *Null Models in Ecology*, Smithsonian Institution Press
- Tokeshi, M. (1999) *Species Coexistence*, Blackwell Science
- May, R.M. (1975) Patterns of species abundance and diversity. In *Ecology and Evolution of Communities* (Cody, M.L. and Diamond, J.M., eds), pp. 81–120, Belknap Press of Harvard University Press
- McGill, B.J. (2003) A test of the unified neutral theory. *Nature* 422, 881–885
- Nee, S. The unified phenomenological theory of biodiversity. In *Macroecology: Concepts and Consequences* (Gaston, K.J. and Blackburn, T.M., eds), Blackwell Publishing (in press)
- Gillespie, J.H. (1991) *The Causes of Molecular Evolution*, Oxford University Press
- Chesson, P.L. and Warner, R.R. (1981) Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.* 117, 923–943
- Hillis, D.M. *et al.* (1996) *Molecular Systematics*, Sinauer
- Emerson, B.C. (2001) Revealing the demographic histories of species using DNA sequences. *Trends Ecol. Evol.* 16, 707–716
- Hudson, R.R. *et al.* (1987) A test for neutral molecular evolution based on nucleotide data. *Genetics* 116, 153–159
- Yang, Z. *et al.* (1994) Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. *Mol. Biol. Evol.* 11, 316–324
- Posada, D. and Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818
- Rokas, A. *et al.* (2001) Understanding patterns of genetic diversity in the oak gallwasp *Biorhiza pallida*: demographic history or a *Wolbachia* selective sweep? *Heredity* 87, 294–305
- Tavare, S. (1984) Line-of-descent and genealogical processes and their applications in population genetics models. *Theor. Pop. Biol.* 26, 119–164