



## Unravelling post-glacial colonization through molecular techniques: new insight from estuarine invertebrates

The profusion of recent phylogeographical studies is linked to the increasing ease and ubiquity of molecular methods. Phylogeography itself is concerned primarily with assessing intraspecific variation between populations in order to deduce processes influencing the spatial distribution of genealogical lineages. In practice, this usually involves analysing molecular diversity by defining the haplotypes present in a population – variants of a specified gene, assigned according to the exact positions of nucleotide polymorphisms in the gene sequence. Such techniques have increasingly been applied to shed light on the historic routes of post-glacial colonization in European species. While the existence of southerly Balkan and Iberian refugia has been firmly established (Schmitt, 2007), evidence has also suggested the existence of cryptic northern refugia in many species (Stewart & Lister, 2001). However, efforts to verify such northern oases have largely focused on terrestrial plants and mammals; recently, new data have emerged that further support such refugia in marine and estuarine species (Remerie *et al.*, 2008).

The field of phylogeography has evolved substantially since its first inception. The first phylogeographical investigations used restriction fragment length polymorphism (RFLP) techniques (Avice *et al.*, 1979), where haplotypes were reconstructed by means of fragment sizes following the digestion of DNA molecules with restriction enzymes. RFLP was soon replaced by direct sequencing methods, which have the advantage of being much less time-consuming and, increasingly, cheaper as technology progresses. Microsatellites and amplified fragment length polymorphisms (AFLPs) have also been used to assess sequence diversity at short, hypervariable gene regions, but their usefulness has been disputed because of homoplasy and the vagueness of resulting genealogies (Hewitt, 2004).

Despite the dropping costs of sequencing, phylogeographical studies often require analysis of a large number of individuals

to ensure that all haplotypes are captured within a population and across a species' range. Sequencing hundreds of individuals can quickly become expensive; cheaper methods such as single-strand conformation polymorphism (SSCP) are often used alongside direct sequencing to detect haplotypes (Remerie *et al.*, 2008). In this method, physical migration across a polyacrylamide gel is directly related to the folded shapes of DNA molecules, and different migration patterns reflect base changes and thus denote different haplotypes. SSCP has been promoted as a cheap and accurate way of assessing genetic diversity – only a subset of individuals need be sequenced to obtain the corresponding nucleotide sequence for each haplotype (Sunnucks *et al.*, 2000).

Regardless of method, mitochondrial genes are usually the preferred markers for phylogeographical studies in animals; mitochondrial DNA is present in a high copy number in animal tissue, and many robust primers exist that allow for easy polymerase chain reaction (PCR) amplification. The focus on mitochondrial genes can be attributed to their supposed uniparental mode of inheritance, lack of recombination, and effectively neutral selection. Maternal inheritance means that the effective population size of mitochondrial DNA (mtDNA) is one-quarter that of nuclear genes; thus, a gene in any given individual had only a single ancestor in the previous generation. Genetic variation can therefore accrue over relatively short time scales, and the distribution of haplotypes is considered to reflect demographic rather than selective events (Beebe & Rowe, 2004). These assumptions do not hold true in all species, and the complexity of mitochondrial evolution is often underestimated and ignored in many studies (for a full review see Ballard & Whitlock, 2004). An array of statistical analyses can be used to test hypotheses of neutral selection, and these are recommended for mitochondrial sequence data.

While microsatellite markers can be used effectively for phylogeographical studies,

direct sequencing of a single nuclear gene does not often provide the resolution needed to assess recent or subtle genetic differentiation between populations. Nuclear loci are often employed to reconstruct the deep evolutionary history of species, and can offer powerful insight for phylogeographical studies, especially when used in conjunction with mitochondrial genes, or when data from multiple loci are combined. However, mitochondrial genes are still preferred in the vast majority of phylogenetic studies (Beebe & Rowe, 2004).

Such molecular methods have been used extensively in recent studies to elucidate post-glacial dispersal of species in Europe. The advance and retreat of ice sheets has been a defining feature of the Quaternary period, with the most recent Ice Age occurring in the late Pleistocene. Climatic oscillations occurred on a massive scale, including temperature changes of 7–15°C within a few decades, expansion of ice and permafrost at high latitudes, and compression of suitable habitats in temperate and tropical regions (Hewitt, 2004). The existing fossil record provides substantial evidence for range compression and local extinctions for many species during this time, with remnant pockets of refugia providing source populations for re-expansion when these cold periods subsided.

Genetic studies have aimed to pinpoint historic refugia and reconstruct range expansions by mapping gene flow and haplotype diversity. For studies centred on temperate Europe, the overwhelming majority focus on terrestrial species – particularly vertebrates, invertebrates and some plants (Schmitt, 2007). From these results, it appears that the distribution of mitochondrial haplotypes needs to be evaluated on a species-specific basis; survival of local populations can be unpredictable, and can be directly affected by ecological factors and the dispersal capabilities of a given species. Despite this, general patterns seem to indicate former zones of shared refugia in Iberia and the Balkans, and the presence of

HOLLY M. BIK

*Nematode Research Group, Zoology  
Department, The Natural History Museum,  
Cromwell Road, London SW7 5BD, UK.  
E-mail: h.bik@nhm.ac.uk*

hybrid-rich suture zones seated around major topographic barriers such as the Alps and Pyrenees where populations met and mixed following post-glacial expansions (Schmitt, 2007).

Although some phylogeographical investigations have focused on marine species, estuarine species remain understudied despite their unique environment. The considerable geographical distance separating estuaries, combined with physiological boundaries (such as salinity ranges), are both critical factors for the survival and dispersal of fauna. Estuaries have been proposed as speciation hotspots, with some animal taxa showing genetic differentiation between adjacent estuaries and even within a single system (Bilton *et al.*, 2002). Low-dispersing taxa are thought to harbour strong genetic signals of population fragmentation and subsequent range expansion following Pleistocene glaciations. In contrast, any such genetic signal in highly dispersing taxa has probably been eradicated as a result of frequent gene flow between populations.

Recent investigation of the low-dispersing estuarine mysid *Neomysis integer* has presented solid evidence for the existence of an additional northern glacial refugia for estuarine species during the last Ice Age, supplementary to the established Iberian and Balkan refugia in the south (Remerie *et al.*, 2008). Remerie *et al.* (2008) found no haplotypes to be common to both glaciated and unglaciated regions, and no northerly decrease in haplotype diversity was uncovered. This retention of haplotype diversity often reflects dispersal ability, and *N. integer* populations are thought to have followed a slow, gradual post-glacial expansion. Many terrestrial and marine species exhibit a loss of genetic diversity at northern latitudes, due to successive bottlenecks at the leading edge as populations rapidly expanded from southern refugia (Schmitt, 2007). The haplotype pattern of *N. integer* populations appears to cluster geographically; several unique haplotypes appear only north of the English Channel, and the Channel itself exhibits a high genetic diversity with several unique but common haplotypes. Mysids collected from an Irish estuary formed a

distinct group with private haplotypes; when correlated with evidence from other marine (Remerie *et al.*, 2008) and terrestrial species (Stewart & Lister, 2001), the data present substantial evidence for Pleistocene refugia on an ice-free Irish coast, and may also suggest an additional refugium in a palaeoriver system within the English Channel. Further investigation is needed to determine whether such refugial populations may have existed in other estuarine taxa.

Despite the increasing wealth of information about post-glacial expansion, questions still remain. Many will require more sophisticated methodology and analysis in order to be answered accurately. For example, molecular data from one gene can be especially confusing if past events were especially complex and layered, as in Alpine habitats where the evolutionary history of species may be particularly intricate (Hewitt, 2004). In order to fully resolve the picture for certain species, analysis of multiple genes will be necessary. As sequencing technology continues to become cheaper and more widely accessible, it is likely that future work will move towards whole-genome comparisons.

Perhaps the trickiest problems currently arise when attempting to assign dates of lineage divergence using molecular clocks. A pertinent question is whether such divergences came as a direct result of climate changes, or if lineages had begun to split long before. Phylogeographical studies require estimated timings of ancestral splits, but some authors counsel caution because of dates derived from the current outdated and uncalibrated clock methodology (Remerie *et al.*, 2008). Current estimates seem to indicate speciation in conjunction with Pleistocene glaciations; however, further work is needed to calibrate clocks with taxon-specific rates of sequence divergence, and then assess whether published date estimates can be trusted. If future clocks are proven accurate, it would offer immense potential for deducing the exact timings of speciation events – a method that could revolutionise our understanding of evolutionary and biogeographical processes.

## REFERENCES

- Avise, J.C., Giblin-Davidson, G., Laerm, J., Patton, J.C. & Lansman, R.A. (1979) Mitochondrial DNA clones and matrilineal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Sciences USA*, **76**, 6694–6698.
- Ballard, J.W. & Whitlock, M.C. (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.
- Beebe, T.J.C. & Rowe, G. (2004) *An introduction to molecular ecology*. Oxford University Press, Oxford.
- Bilton, D.T., Paula, J. & Bishop, J.D.D. (2002) Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuarine, Coastal, and Shelf Science*, **55**, 937–952.
- Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.
- Remerie, T., Vierstraete, A., Weekers, P.H.H., Vanfleteren, J.R. & Vanreusel, A. (2008) Phylogeography of an estuarine mysid, *Neomysis integer* (Crustacea, Mysida) along the northeast Atlantic coasts. *Journal of Biogeography*, doi: 10.1111/j.1365-2699.2008.01970.x.
- Schmitt, T. (2007) Molecular biogeography of Europe: Pleistocene cycles and post-glacial trends. *Frontiers in Zoology*, **4**, 11.
- Stewart, J.R. & Lister, A.M. (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608–613.
- Sunnucks, P., Wilson, A.C.C., Beheregaray, L.B., Zenger, K., French, J. & Taylor, A.C. (2000) SSCP is not so difficult: the application and utility of single-stranded conformational polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology*, **9**, 1699–1710.

Editor: Alistair Crame