



*Molecular genetics in animal ecology*. Edited by P. A. Racey,  
P. J. Bacon, J. F. Dallas and S. B. Piertney

## Post-glacial re-colonization of European biota

GODFREY M. HEWITT

*Biological Sciences, University of East Anglia, Norwich NR4 7TJ*

Population structure is the result of both present processes and past history. Molecular markers are proving of great value in describing the former, and it is important to similarly determine the latter in order to understand their respective contributions. The study of palaeo-climates has also advanced significantly, and in particular that of the Pleistocene ice ages, which modified species ranges considerably. The last ice age and rapid post-glacial colonization of Europe is summarized. Possible population genetic consequences of expansion northward from southern refugia, and those of remaining in these mountainous regions are discussed. A series of recent case studies are detailed where DNA sequence information has been used to describe species genetic variation and subdivision across Europe. These include a grasshopper, the hedgehog, oak trees, the common beech, the black alder, the brown bear, newts, shrews, water vole, silver fir and house mice. These molecular data confirm southern peninsulas of Europe as major ice age refugia, and in most cases demonstrate that genetically distinct taxa emerged from them. They can thus define genomic differences and so greatly augment previous fossil data. The refugial genomes contributed differently in various species to the re-colonization of Europe, with three broad patterns described as paradigms—‘grasshopper’, ‘hedgehog’ and ‘bear’. These different expansion patterns produced clusters of hybrid zones where they made contact, and it is argued that many species genomes may be further cryptically subdivided. A reduction in diversity from southern to northern Europe in the extent of allelic variation and species subdivision is seen; this is attributed to rapid expansion northward and the varied topography of southern refugia allowing populations to diverge through several ice ages. The differences in DNA sequence indicate that some species have been diverging in refugial regions for a few ice ages at most, whilst distinct lineages in other species suggest much more ancient separation.

© 1999 The Linnean Society of London

ADDITIONAL KEY WORDS:—DNA sequence – range change – population structure – refugia – genetic divergence – biodiversity – phylogeography – hybrid zones – palaeoclimate – speciation.

### CONTENTS

Introduction . . . . .	88
Ice ages . . . . .	88
Post-glacial advance . . . . .	89
Genetic consequences of climatic range changes . . . . .	91
Suitable DNA markers . . . . .	93
Case studies . . . . .	93
General features . . . . .	104
Conclusions . . . . .	107
Acknowledgements . . . . .	109
References . . . . .	109

\* Email: [g.hewitt@uea.ac.uk](mailto:g.hewitt@uea.ac.uk)

## INTRODUCTION

Population structure is the distribution of genotypes in space and time and is the result of both present processes and past history (Hewitt & Butlin, 1997). Molecular markers are proving of great value in describing present processes such as drift, dispersal, mating and selection, with estimates of population size, gene flow, mate choice, reproductive success and relative genotype fitness. It is important to similarly determine the effects of historical and more distant past events in moving and moulding the population and species genomes that we presently study. For example, an extensive population may be genetically homogenous because of even environmental selection, or because of a past range expansion. The question has been nicely put: "Is it ancient or modern history that we can read in our genes?" (Nichols & Beaumont, 1996).

Along with molecular genetics, the study of palaeo-climates has made great strides in the last 20 years and recent discoveries need to be incorporated into our thinking about the causes of the genetic structure of populations and species. The data come from various physical and biological sources, including carbon and oxygen isotope levels, CO<sub>2</sub>, magnetic and mineral signatures, animal and vegetable remains. Together, they are providing an increasingly coherent picture of global climatic changes and their causes.

## ICE AGES

Perhaps the most dominant palaeoclimatic features are the ice ages, which became increasingly severe through the Pleistocene. The climate has been cooling for some 60 Myr with the Antarctic ice sheet forming some 35 Ma. It is from around 2.4 Ma that the Arctic ice cap has grown, producing progressively larger ice sheets across Eurasia and North America (Webb & Bartlein, 1992). These ice ages have about a 100 kyr periodicity with relatively short warm interglacials, as at present. The Croll–Milankovitch theory proposed that the pacemaker of these cycles was the orbital eccentricity of the earth around the sun which caused major changes in insolation (Hays, Imbrie & Shackleton, 1976). Axial tilt (41 kyr) and precession (23 kyr) cycles interact with the main 100 kyr eccentricity cycle to produce a complex of climatic oscillations with varying effects through time and across the globe.

The last full glacial cycle from the Eemian interglacial (135 ka) to the present is the best understood, and in particular the last warming from full glacial conditions some 18 000 BP through to the current warm interglacial climate. Marine sediments, beetle exoskeletons and pollen cores provide particularly valuable biological data (e.g. Coope, 1977, 1994; Huntley & Birks, 1983; Beaulieu & Reille, 1992; Webb & Bartlein, 1992; Poneel, 1997), while recent analysis of deep ice cores in Greenland (GRIP) have provided detailed evidence of striking climatic changes through this period (Dansgaard *et al.*, 1993). These long ice cores provided evidence of dramatic switches in temperature on the ice sheet through the ice age and the Eemian interglacial. Average temperatures would seem to have changed by 10–12% over 5–10 years and lasted for periods of 70–5000 years. Such massive, often rapid, changes in climate will modify species distributions and should be reflected in terrestrial biological records with sufficient detail to resolve them. Beetle and pollen data have given indications of sharp changes in frequency and distribution of species for some time

(e.g. Atkinson, Briffa & Coope, 1987; Beaulieu & Reille, 1992) and current research is confirming a close correlation between vegetation changes and oscillations in climate over the last 135 kyr (e.g. Whitlock & Bartlein, 1997; Guiot, 1997).

#### POST-GLACIAL ADVANCE

Apart from the highlands of Scandinavia (which was completely glaciated during the last ice age), the major mountain ranges of Europe are in the south and in general run east–west, i.e. Cantabrians, Pyrenees, Alps, Transylvanians and Caucasus. These all had extensive ice caps during the ice age. To their south, also in an east–west orientation, are the Mediterranean and Black Seas, with the fairly mountainous peninsulas of Iberia, Italy, Greece and Turkey. Between the main ice sheet and southern mountain blocks was a plain of permafrost, tundra and cold steppe, which extended eastwards across Russia to the Urals (Fig. 1). This particular geography is expected to influence greatly the movement of species in response to climate changes. Both the seas and the mountains are formidable barriers to most organisms today, and the ice-covered mountains would have been more so during the last ice age. From plant and animal remains it is clear that most organisms presently distributed across Europe were in refugia in the south at the height of glaciation 18 000 BP, many in the peninsulas of Iberia, Italy and the Balkans, and some possibly near the Caucasus and Caspian Sea.

From about 16 000 BP, after the last Heinrich event, the climate warmed, the ice retreated, and species expanded their ranges out of the refugia northwards. The pollen data are particularly informative on this great expansion. It is clear that species responded individually to this warming with each tracking their particular set of environmental conditions. This created mixtures of species different from today, which were transitory so that communities were not stable (Huntley, 1990). Around 13 000 BP the pollen maps of Europe (Huntley & Birks, 1983) show that plant and tree species spread much more quickly up the east of Europe between the Caspian Sea and White Sea (35°E) than in the central and western parts. Whilst not as dramatic, there is evidence of early spread of pine, oak, elm and alder up the western Atlantic fringe to Brittany, Ireland and Scotland, perhaps transported by water currents or animals. Beetle remains show that species with a present day Mediterranean distribution had reached Britain by this time. The climate in Britain may well have been warmer than now (Atkinson, Briffa & Coope, 1987). Around 11 000 BP this rapid northern advance was sharply reversed for 1000 years in the Younger Dryas period. The Atlantic Polar Front moved from Britain to Iberia, the ice readvanced in places, tundra spread down through France, the birch trees died out in Northern Europe, the mediterranean beetles vanished from Britain, and in Southern Europe the pine and oak retreated again. The Greenland ice cores show this severe event, and since it is the most recent and accessible example of the major climatic oscillations revealed by GRIP, it is particularly useful as a model of what has driven the distribution of plants and animals through the last ice age, and probably previous ones.

The Younger Dryas cold spell came to an end around 10 000 BP, the polar front shifted north again, the climate warmed and vegetation advanced rapidly over Europe. By 6000 BP the vegetation pattern broadly resembled that of today,

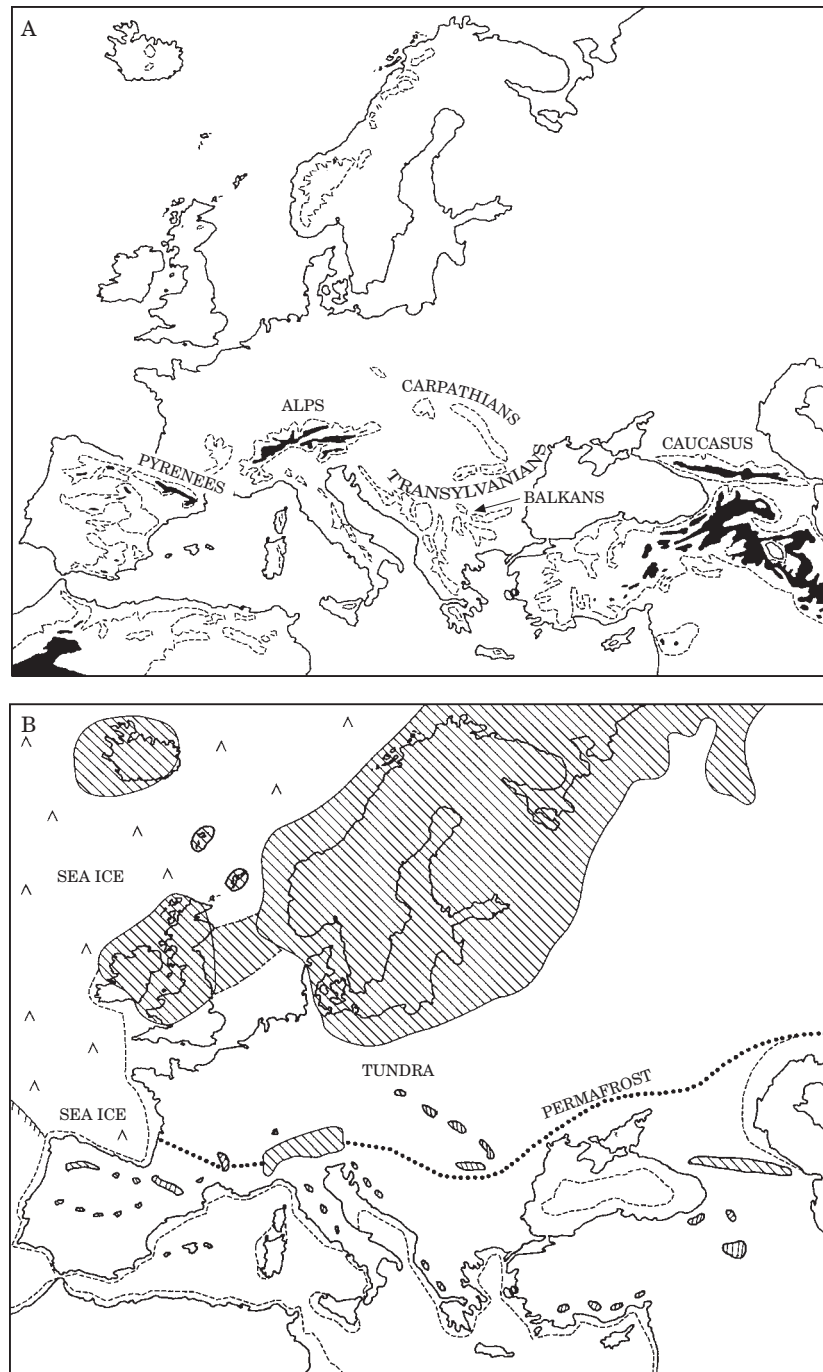


Figure 1. A, physical geography of Europe, showing the predominance of mountains in the south running east-west. Black regions=over 2000 m, dashed line=over 1000 m altitude. B, ice cover (hatched) and extent of permafrost at the end of last ice age 18 000 BP.

but some vegetational changes continue until present times. Huntley (1990) has emphasized these changing vegetation mixtures for which there is no modern analogue. In the far north the Scandinavian ice sheet remained only on the highlands by 8000 BP; similarly the glacial blocks on southern mountains like the Pyrenees and Alps had shrunk and formed lesser barriers to dispersal.

Such sudden warming allowed rapid advances across Europe, where not impeded by ice, mountains or seas. Radiodated pollen distributions provide estimates of the rate of spread for a number of species in this period (Huntley & Birks, 1983; Bennett, 1986), with most species managing 50–500 m/year. Across the European plains pine and hazel apparently reached 1500 m/year, and alder 2000 m/year for a while. Beetle remains show that they track climate changes more closely, and they advanced very rapidly where food was available, Mediterranean species reaching Britain by 13 000 BP (Coope, 1990). Flightless grasshoppers like *Chorthippus parallelus* reached England before the rising sea level cut the English Channel and Irish Sea, which would have required a dispersal rate of some 300 m/year to expand from a Southern European refuge (Hewitt, 1990). Such rapid advance should not surprise us, since in historical times invading species have shown similar rates of spread, e.g. cheat grass in western North America (Mack, 1981) and the collared dove from Turkey across Europe (Hengeveld, 1989). Since plant species were expanding at different rates, the animals dependent on them were differently limited. For example, insects feeding on pioneer grasses and herbs may spread more quickly than those dependent on trees, and those with broad nutritional tolerances would spread wider than more host specific ones.

While species were expanding northwards, southern populations would die out as the southern edge of the species tolerance range moved north also. The present day distributions of many birds and butterflies across Europe and Asia occur in bands running E–W with northern and southern limits (Harrison, 1982; Higgins & Hargreaves, 1983). The ranges of species generally would move north and south with the climatic oscillations. On the other hand, as conditions warmed refugial populations could also climb up one of the many mountains in the south when one was nearby. When the conditions deteriorated again such populations could descend and repopulate the lower refugial areas, while those populations in the north of the range largely went extinct (Hewitt, 1993a). Indeed the great variation in topography, climate and habitat in the south of Europe provides much more opportunity for a species to find a nearby suitable habitat through the climatic cycles. This has consequences for dispersal, genetic variation and retention of diversity over time.

#### GENETIC CONSEQUENCES OF CLIMATIC RANGE CHANGES

Several earlier authors have implicated post-glacial expansion in structuring genome distribution (ref. Hewitt, 1989). The rapid northward expansion across European plains would be expected to have very different genetic consequences from the slower altitudinal shifts in the mountainous southern parts (Hewitt, 1993a, 1996). As the climate warmed rapidly, populations at the northern limits of the refugial range would expand into the relatively large areas of suitable territory. Of course most species are more or less dependent on others and would necessarily follow their resource provider, but since the primary species should disperse quickly,

this should be generally true. Such dispersal at the leading edge would likely be by long distance dispersants that set up colonies far ahead of the main population. These pioneers could expand rapidly to fill the area before significant numbers of other dispersants arrived, and so their genes would dominate the new population genome. This long distance dispersal would be repeated many times over long expansion distances. The series of founder events that this represents leads to loss of alleles and to homozygosity. This process has been modelled to compare leptokurtic (with more long distance migrants), stepping stone and normal dispersal, along with growth rates and carrying capacity (Ibrahim, Nichols & Hewitt, 1996). As leptokurtic expansion proceeds increasingly large patches of homozygosity are produced, in comparison to the fragmented patchwork of high and low frequency areas produced in the other forms of dispersal. These large patches of homozygosity persist and can grow larger with time. This tendency to homozygosity may be increased by smaller climatic oscillations in this general advance, since if a reversal eliminates most of the recently established leading populations, the surviving ones will retain their homozygous nature and expand leptokurtically when conditions improve. Repeated oscillations can create more bottlenecks.

These considerations predict that rapid continued expansion would produce large areas of reduced genetic diversity in Northern Europe, and any region subject to the same form of colonization. A number of studies now show greater homozygosity in northern expansion areas (Hewitt, 1996). Slower expansion would have different consequences, with much more genetic diversity maintained. One would expect population shifts in Southern Europe, particularly in the regions of mountains and refugia, to have been more of this type, involving relatively slow ascents and descents of mountains. The resemblance to these two extreme forms of expansion, 'pioneer and phalanx' (Nichols & Hewitt, 1996), will clearly depend on the sharpness of the climatic change, the latitude and the topography of the region, and of course the dispersal and reproductive capabilities of the organism. One can predict that if the colonization of Northern Europe begins from several southern refugia with different genomes, then when the spread is rapid one of these may cover much of the continent with the others remaining in the south. If the colonization is slower then several of the refugial genomes may be involved in the spread north (Hewitt, 1996).

Another property of leading edge expansion from the north of a species refugial range is that once a pioneer area has been colonized, it is much more difficult for a migrant from behind the front to contribute to the population and influence its genome. This is simply a matter of density, and logistic rather than exponential growth. (Hewitt, 1993a). If there are differences between the front line residents' genome and that of the migrants from behind which are negatively heterotic and cause any hybrid unfitnes, then this barrier to mixing is stronger: it is in effect a simple hybrid zone (Hewitt, 1993a). Consequently populations, genomes and subspecies which are behind the northern front line in the centre and south of the refugium range will not be able to advance readily. They must survive where they are, or climb mountains. This property, along with many mountains and several climatic oscillations could generate a packing of genomes and subspecies in southern ice age refugia. It has been emphasized that the E-W mountains of Europe could act as impediments to dispersal, which would tend to isolate the populations in the southern Mediterranean peninsulas of Iberia, Italy and Greece. Turkey would also be largely isolated by the Black Sea and Mediterranean Sea. The pollen and fossil record clearly identify these regions as the major ice age refugia for the recolonization

of Europe, and also inflow from the east near the Caspian Sea. Consequently the populations of species surviving the ice ages in these isolated places will not exchange genes, and may well be subject to different selection; so they will diverge genetically. When two such diverged genomes expand from different refugia they will form hybrid zones where they make contact. The genomes of many species are divided across their range into sub-species, races and forms by such narrow zones. These were originally identified by classical taxonomic differences in morphology and behaviour, but the use of chromosomal markers revealed many more cryptic differences and new zones, and the application of allozyme and DNA methods is extending this (Hewitt, 1988, 1993a). The dynamics and locations of a number of these major hybrid zones indicates that they have remained broadly in the same place during the relatively stable Holocene since their post glacial formation.

#### SUITABLE DNA MARKERS

Whilst it has been recognized for some time that this hybrid zone subdivision is a product of secondary contact after expansion from refugia (Hewitt, 1975, 1989; Hewitt & Barton, 1981), the advent of modern DNA techniques for population studies provides the possibility of much greater genetic discrimination, with the tracing of lineages, routes of expansion and identification of relevant refugia. The Pleistocene with its ice ages occurred just 2.5 Mya and the last post glacial warming began only 15 000 BP; consequently, one needs fast evolving DNA sequences or loci to provide fine discrimination. MtDNA, for historical and technical reasons, is by far the most used method in animal studies, while cpDNA is widely used in plant phylogenetic studies. While nuclear sequences such as introns are being investigated, there are few studies directly comparing nuclear and organellar rates of divergence in the same organisms. MtDNA is still generally the fastest (1.4–2.6% Myr) and different regions vary in rate. Intron sequences are 5–10 times slower and cpDNA sequences are 10 times slower again (e.g. Slade, Moritz & Heidemann, 1994; Gaut *et al.*, 1996; Bohle *et al.*, 1996). Consequently, over the time period of interest new haplotypes that are established will differ by only a few base substitutions and occasional insertions/deletions. Thus much of the significant geographic sequence variation may comprise the sorting of more ancient divergence among sequences. Reciprocal monophyly of mtDNA (<4%) will be the most useful. Microsatellite loci have much higher mutation rates producing large numbers of alleles and variation among populations; however because they are simple sequence repeats there is not a clear genealogy and allele homoplasy is quite striking (Orti, Pearse & Avise, 1997). When used carefully they can be useful for phylogenetic purposes (Angers & Bernatchez, 1998). Due to hybridization and reassortment one or more markers may not be concordant in defining the boundaries of the genome; this is clearly seen in some mtDNA studies (Hewitt, 1993a, b, 1996). Clearly the combination of information from several markers is desirable, but is expensive (Hillis, Moritz & Mable, 1996).

#### CASE STUDIES

Several studies using molecular markers here recently appeared, allowing us to begin to address these issues and examine the expectations raised by considering

the effects of ice ages on the population structure of species. They are most valuable when a suitably variable and discriminatory sequence is analysed in a large number of samples taken from across the European range. There are a few detailed cases using DNA markers, and several more with fewer samples (Hewitt, 1996; Taberlet *et al.*, 1998).

#### *The grasshopper*

*Chorthippus parallelus*, the meadow grasshopper occurs all over Europe into Siberia (Reynolds, 1980); it is found on the high Sierra Nevada in Spain and in southern Finland. Three subspecies are recognized: in Iberia *C.p. erythropus*, in Greece *C.p. tenuis*, and over the rest of Europe *C.p. parallelus*. A narrow well studied hybrid zone along the Pyrenees between the Spanish and French subspecies has been analysed for differences in morphology, chromosomes and behaviour (Butlin & Hewitt, 1985; Hewitt, 1993b). Only two out of 40 allozyme loci tested showed subspecific differences, so mtDNA, rDNA and ncDNA sequences were investigated (Cooper & Hewitt, 1993; Cooper, Ibrahim & Hewitt, 1995; Szymura, Lunt & Hewitt, 1996; Lunt, Ibrahim & Hewitt, 1998). An anonymous nuclear fragment of 393 bp has proved the most revealing so far. This sequence was chosen because it showed significant variation among geographic regions in a pilot study. Some 350 individuals were sequenced for 88 populations across Europe identifying 72 haplotypes. Several parsimony and distance methods of analysis were tried, and the analysis of distance measured by Kst and presented as a Fitch tree was the most instructive (Fig. 2). This incorporates both haplotype frequency and sequence divergence into the measure of population similarity; whereas, for example, Brook's Parsimony concerns only presence/absence of a haplotype in a population and haplotype distance or parsimony trees do not consider population frequency. This analysis shows that the genome of *C. parallelus* is divided into at least five major geographic regions—Iberia, Italy, Greece, Turkey and the rest of Northern Europe including western Russia. The Spanish and French haplotypes show an average sequence divergence of 2.5% and no cases of haplotype introgression across the hybrid zone. Interestingly, the divergence between the Italian genome and that north of the Alps was not indicated by any previous taxonomic work, and suggested the existence of another hybrid zone. This has now been demonstrated with the discovery of hybrid dysfunction and chromosomal differences (Flanagan *et al.*, 1999).

The four southern Mediterranean regions—Spain, Italy, Greece and Turkey—all contained a high proportion of unique haplotypes clearly demonstrating their distinct genomes and possible subspecific status. In contrast to these southern regions there is less haplotype diversity across Northern Europe and little differentiation from the populations in the Balkans. These results address directly a number of the postulated effects of ice age range changes on population genomes. (1) They strongly argue that Northern Europe is populated by an expansion from a Balkan refugium in the last post glacial warming, and that this expansion was rapid; the sequence similarity and lower haplotype diversity in these northern areas supports this. (2) Spain and Italy had their own refugia in the south which repopulated these peninsulas, but were prevented from expanding further north by the Pyrenees and Alps, and importantly the presence of the Balkan/North European genome with which hybrid zones formed. The rapid post glacial expansion from the east is noteworthy, and is reflected in some pollen records. (3) The distinct populations in Greece were denied



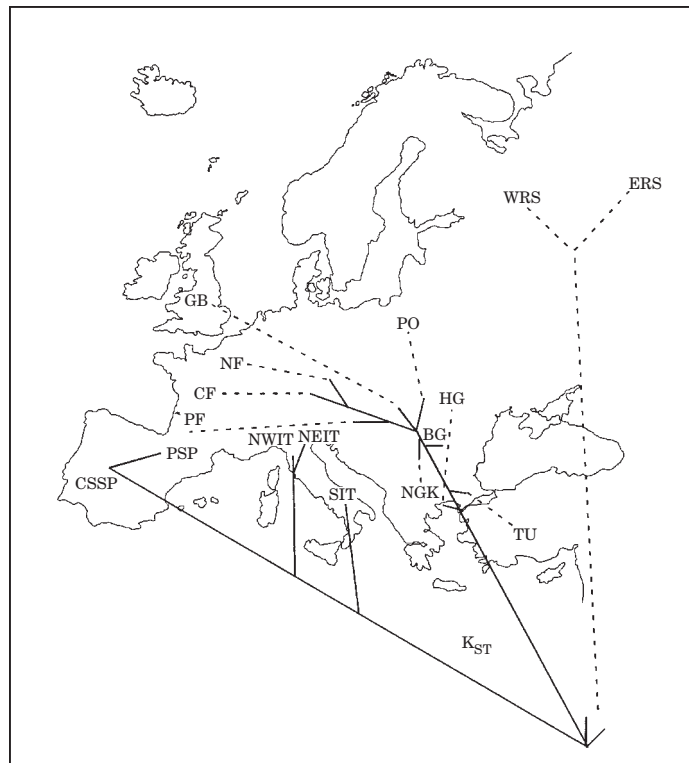


Figure 2. A Fitch tree of  $K_{ST}$  genetic distance among populations of *Chorthippus parallelus* from different areas of Europe. Note the similarity among Balkan and Northern European regions. (CSSP = Central & Southern Spain, PSP = Pyrenean Spain, NWIT = NW Italy, NEIT = NE Italy, SIT = S Italy, WRS = W Russia, ERS = E Russia, TU = Turkey, NGK = N Greece, BG = Bulgaria, HG = Hungary, PO = Poland, GB = Great Britain, NF = N France, CF = Central France, PF = Pyrenean France) (See Cooper *et al.*, 1995).

expansion north by the Balkan incumbents, whose northern populations formed the leading edge of expansion. There may well be hybrid zones in this region, and also near the Bosphorus with the Turkish genome. These deduced refugia and European expansion routes are shown in Figure 4. (4) It is worth noting that some divergent haplotypes were recorded from the more easterly Russian samples; it may be that another refugia existed near the Caspian, but further samples are needed. (5) The extent of sequence divergence among these distinct geographic genomes indicates that they have been diverging for about 0.5 Myr, and this is supported by mtDNA divergence of <1% on average for several genes between Spanish and French genomes. Whilst exercising due caution about the accuracy of molecular clocks, such a difference argues that the populations in Spain, Italy, Greece and the Balkans arrived some 4–6 ice ages ago and have been resident since. If there were any previous populations, they were exterminated by perhaps the particularly severe glaciation around 450 ka, as evidenced by Red Sea levels (Rohling *et al.*, 1998). The colonization came possibly from Turkey, with its high haplotype diversity, genetic relatedness, geographic proximity and more southerly refugia.

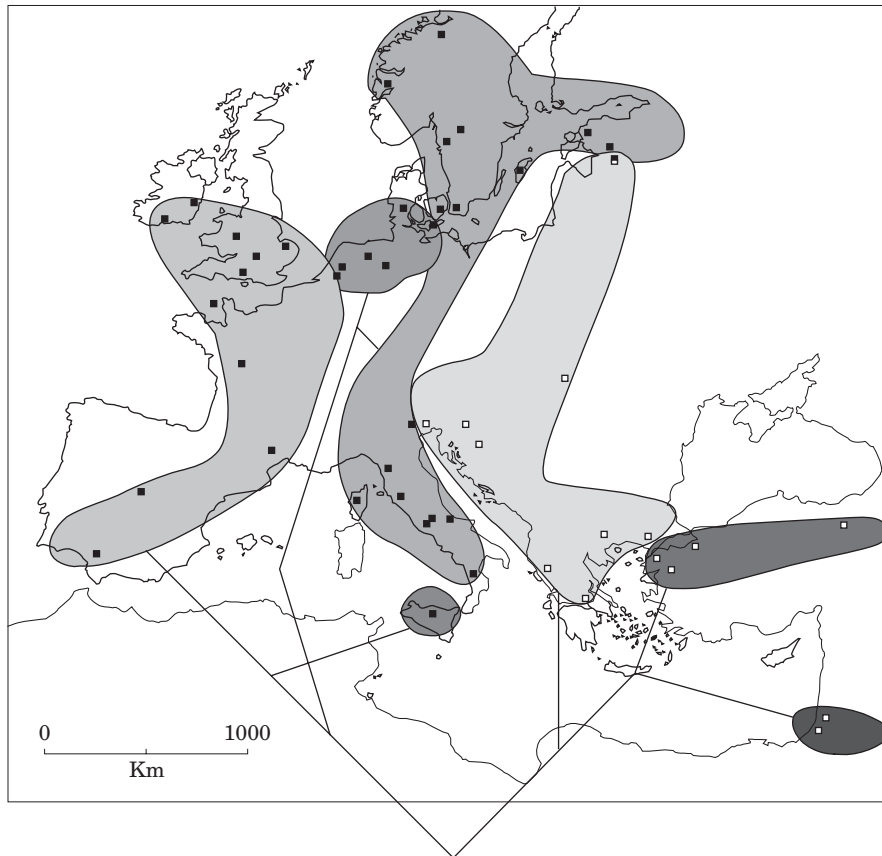


Figure 3. A phylogeography of mtDNA haplotypes for *Erinaceus* hedgehogs based on a NJ tree. Note the deep west/east division of *E. europaeus* and *E. concolor* with more recent dichotomies in each.

#### *The hedgehog*

European hedgehogs occur from Finland to Israel and from Spain to beyond the Caspian. They are currently recognized as two parapatric species—*Erinaceus europaeus* in the west and *E. concolor* in the east—with possible hybridization in their contact zone from the Baltic to the Adriatic Sea. There are indications of further geographic morphological differences, but Corbet (1988) has called for a molecular based investigation to solve this confused taxonomy. Recently an excellent allozyme study (Filipucci & Simson, 1996) has shown that Turkish and Balkan *E. concolor* are rather distinct, as are Iberian and Italian *E. europaeus*. In order to better resolve their phylogeography a 486 bp fragment of the 5' end of the *cyt b* gene mtDNA was sequenced from 56 animals across much of the range (Santucci, Emerson & Hewitt, 1998). Recent collections have extended these data to 154 individuals with more samples from the north of Europe (Santucci *et al.*, in prep.). Some 33 haplotypes have been described, and parsimony and distance trees for these agree in topology (Fig. 3). The genomes of *E. concolor* and *E. europaeus* are quite distinct; within these two clades, the Iberian haplotypes are distinct from those of Italy and Germany, and the Balkan haplotypes are distinct from those of Turkey and Israel. The Iberian

and Italian haplotypes appear to continue northward, as do the Balkan ones, so that much of Europe is divided into three major genome strips, with a fourth genome in Turkey and Israel.

This hedgehog phylogeography clearly indicates a rather different post glacial expansion from *C. parallelus* (Fig. 4). (1) Northern Europe has been colonized by three distinct genomes emanating from three ice age refugia in Spain, Italy and the Balkans. (2) The Pyrenees and the Alps do not appear to have been effective barriers to their northward expansions. (3) Turkey and the Near East had a distinct refugium possibly near Israel, but we have no data yet from Russia to examine a Caucasus/Caspian refugium hypothesis. (4) The depth of sequence divergence between the genomes is striking. The *E. concolor/europaeus* split is some 12% (using Tamura–Nei distances), while between the Spanish and Italian clades of *E. europaeus* it is 5.4%, and between the Balkan and Turkish *E. concolor* it is 6%. Part of the *cyt b* mtDNA sequence may evolve fairly quickly, but using the normal mtDNA rate such divergences indicate separation at about 6 Ma and 2.7–3 Ma respectively. (5) The current data also indicate a more recent divergence within the eastern clade of *E. europaeus* between the Italian and German populations of about 0.5 Myr, and one of similar magnitude in the eastern clade of *E. concolor* between Turkish and Israeli populations.

It is worth contemplating if these putative times of divergence have any particular significance for the evolution of the hedgehog genome. The earliest one (6 Ma) between *E. concolor* and *E. europaeus* was in the early Pliocene, which is when the first fossils closely related to *Erinaceus* are found across Europe (Butler, 1988) and also was a time of increased global glaciation. It is tempting to suggest that the divergence between the two species was instigated in western and eastern refugia by this Pliocene cold period. The time of the second divergences between east and west clades of both *E. concolor* and *E. europaeus* would coincide with the advance of the glaciers from the Arctic. Thus during each ice age hedgehogs have been restricted to refugia in Spain, Italy, Southern Balkans and Turkey/Near East, and diverged in them. The results of their northern expansions and hybridizations in warm interglacials would have been largely eliminated by the rapid climatic cooling at the start of each ice age. The third most recent discernable divergence at 0.5 Ma between the Italian and German clades coincides with the particularly severe glacial period noted previously for *Chorthippus* (Rohling *et al.*, 1998). It indicates that present day northern hedgehogs may have had a separate refugium in Italy or S. France during the last few ice ages allowing divergence and lineage sorting. Clearly much more sampling is required in critical locations to resolve these possibilities.

#### *The brown bear*

In Europe *Ursos arctos* had a broad distribution that has been reduced by man to a few isolated pockets in the Cantabrians, Pyrenees, Alps and Appenines, with larger populations in the Balkans, Carpathians, Scandinavia and Russia. Taberlet & Bouvet (1994) sequenced 269 bp of mtDNA Control Region in some 60 animals and showed 7.13% divergence (Kimura 2 parameter) between eastern and western mitochondrial lineages. The eastern lineage included Russian, Estonian, Finnish, Lap, Slovakian and Romanian samples, while the western lineage was further subdivided into two clades comprising the Balkan and Italian populations and the Pyrenean, Cantabrian and southern Scandinavian populations with some 2.7% divergence. From this the

*Chorthippus parallelus**Erinaceus spp**Ursus arctos**Alnus glutinosa**Quercus spp**Sorex araneus**Fagus sylvatica**Abies alba**Arvicola terrestris**Triturus cristatus**Mus musculus**Crocidura suaveolens*

Figure 4. Proposed post-glacial expansion routes deduced from molecular data and fossil evidence. Each case is described in the text. Three broad patterns emerge—the grasshopper, *Chorthippus parallelus*, the hedgehog, *Erinaceus europaeus* & *concolor*, and the bear, *Ursus arctos*. Species with similar patterns are shown below each of these.

authors postulate three major ice age refugia, with post-glacial colonization of Western Europe and southern Scandinavia from Iberia and Eastern Europe. The colonization of Eastern Europe and Russia would have been from the southern Carpathians or from the east and the Caucasus. Population of the third clade (Italian, Balkan, Greek and Bulgarian) did not expand out, probably prevented by the large expansions of the other two lineages. Interestingly, the western and eastern expansions both colonized Sweden, the eastern from Russia and the north and the western from the south via Denmark. These now form a hybrid zone in central Sweden (Taberlet *et al.*, 1995). The origin of these differences was roughly dated from the CR sequence divergence at some 0.85 Ma for the two major clades that meet in Sweden. While it was 0.32 Ma for the Iberian/Balkan split, which would have been three or four ice ages ago. A subsequent report by Kohn *et al.* (1995) using excremental PCR on 98 bears confirmed this general picture of western and eastern clades, but provided evidence of more subdivision in the southern regions with a hybrid zone in eastern Romania between Balkan and eastern lineages (see Fig. 4).

#### *Black alder*

*Alnus glutinosa* occurs all over Europe and Mediterranean as far as the Caucasus, usually in wet places with its seeds mainly dispersed by water. Pollen records indicate a number of possible southern glacial refugia including W France, S Italy, Corsica, Greece, Carpathians and SW Turkey. King and Ferris (1998) analysed some 217 individuals from the entire range by PCR-RFLP of cpDNA using 13 haplotypes. Spain, S Italy, Corsica, Greece, Bulgaria, Crimea and Turkey each show distinct cytotypes, while the rest of Europe contains two haplotypes from the northern Balkans. This clearly demonstrates that genotypes from Balkan refugium expanded after the ice age to fill all Northern Europe including France and Britain. The several southern refugia did not expand far north, probably prevented by the more rapid expansion of the Balkan genome that filled the northern areas first. As King & Ferris (1998) point out, in addition to the main refugial areas indicated by the fossil pollen, the cpDNA data clearly demonstrate two more in Iberia and Turkey. Furthermore, examination of the haplotype distribution in Turkey would suggest that the two distinct haplotypes meeting on the north coast with the Black Sea are from two distinct refugia themselves. In this species the molecular data are revealing considerable range and structuring, possibly due to the patchy distribution of its watery habitats that may restrict exchange. Both the fossil pollen and genetic data agree in indicating a very rapid colonization of Northern Europe from a Northern Balkan refugium, and the genetic data also show that the Romanian and Greek populations to the south remained near their refugia. This pattern of refugia and expansion is very similar to that of the grasshopper *C. parallelus* (Fig. 4).

#### *The oaks*

The deciduous oak trees of Europe comprise what appears to be a syngameon of several morphological species, amongst which hybridization has and still does occur. As major components of temperate woodland they have a detailed pollen record, with ice age refugia in Iberia, Italy, the Balkans and possibly the Caucasus. *Quercus robur* and *Q. patraea*, the pedunculate and sessile oaks, are most widely distributed, with *Q. robur* ranging from N Spain to the Caspian and S Italy to S

Finland. Other morphological species like *Q. pubescens* occur in Southern Europe, with some taxa relatively restricted. The two common species have been extensively investigated, initially using allozymes and more recently chloroplast DNA (Ferris *et al.*, 1993, 1995, 1998; Petit, Kremer & Wagner, 1993; Zanetto & Kremer, 1995; Dumolin-Lapeque *et al.*, 1997). The cpDNA variation has proven more discriminating amongst populations and regions of Europe than the allozyme polymorphism, as might be expected of a haploid maternally inherited genome in a heavy seed when compared with nuclear markers also carried in widely dispersable pollen (e.g. Ennos, 1994). Nonetheless, both allozymes and cpDNA reveal an east/west division in both taxa in central Europe. Ferris *et al.* (1993) used sequences of cpDNA regions of tRNA<sup>Leu1</sup> to tRNA<sup>Phe</sup> to identify four cytotypes. The east/west divide is distinguished by a unique substitution in a highly conserved part of the tRNA<sup>Leu1</sup> intron. The course of this divide has been recently plotted northwards east of Denmark (Johnk & Siegismund, 1997) to Finland (Ferris *et al.*, 1998) where a hybrid zone has been located. This western cytype is divided further by a transition in the IGS region into two haplotypes, which produce another east/west divide between France and Germany. Another distinct variant of the most western cytype is found in East Anglia. Western and central Europe are thus divided into three major cytypes running N-S in strips. These results may be interpreted as the reflections of the post glacial expansion north of three genomes, one from Iberia, one from Italy and one from the Balkans. Despite the ability to hybridize and possibility of mixing, the cpDNA signal is still clear.

This is encouraging since the generally slow rate of cpDNA sequence evolution has discouraged its use for investigating Pleistocene events. Several other recent studies have revealed intraspecific geographic structure of cpDNA which is useful for Pleistocene and post-glacial interpretations (Soltis, Soltis & Milligan, 1992; Sewell, Parks & Chase, 1996). Using restriction enzyme analysis of PCR fragments (RFLP-PCR), 14 PCR-fragment/enzyme compositions of cpDNA were examined in eight species of oaks totalling 1412 individuals from across Europe (Dumolin-Lapèque *et al.*, 1997). This massive data set revealed some 23 haplotypes, many of which appeared restricted to parts of Europe. From this the authors constructed a map of possible post-glacial colonization routes for the haplotypes (Dumolin-Lapèque *et al.*, 1997, fig. 5). Some of these routes cross over, leading them to postulate overlapping of distributions during post glacial expansion, early migration by some haplotypes with later colonization by other haplotypes across their path, and artificial introduction in recent times. These are possible, and there is particularly good evidence for human transport in English oaks (Ferris *et al.*, 1995, 1997), but other explanations are possible which are compatible with the three clear major genome expansions northward indicated previously (Ferris *et al.*, 1993, 1998).

Firstly, since cpDNA is generally slowly evolving most of the mutations producing these haplotypes will not have occurred in the last glaciation but have accumulated over millions of years. They will have been subject to many climatic oscillations and range changes. Consequently most of the haplotypes present in the refugia of the last ice age could have entered when they were previously colonized from elsewhere. Secondly, of the three certain refugia, Italy is the most sampled and contains at least nine haplotypes, Spain is less sampled and contains five haplotypes, and the Balkans are unfortunately little sampled. Even so, some haplotypes are present in both Iberia and Italy. Such shared haplotypes probably come from older range changes and could spread north from all their refugia in the last post-glacial advance.

Considering the distribution of individual haplotypes in this light (see Dumolin-Lapègue *et al.*, 1997, fig. 2), it seems that Haplotype 7 was present in Iberian, Italian and Balkan refugia and spread in the western, central and eastern European tracks, to be present over much of Europe. Haplotype 5 was apparently present in the Italian and Balkan refugia, but expanded essentially up the eastern European track, as probably did Haplotype 2, and possibly Haplotypes 4 and 6, although these are rarer and localized. Haplotype 1 was probably present in both Iberia and Italian refugia but expanded along the central European track and not the western one. Such allele loss is expected for rapid expansion (Hewitt, 1989, 1993a, b). A strong western expansion track from an Iberian refugium is indicated by Haplotypes 10, 11 and 12, which are closely related and possibly more recently derived. This distribution agrees closely with the western marker of Ferris *et al.* (1993, 1998). Interestingly, the distribution of Haplotypes 13, 14, 15 and 16 from the Balkans to Russia, north of the Caucasus, indicates a second distinct expansion route from the Balkans, suggesting that two distinct refugial populations in the NE and NW Balkans contributed significantly to the last recolonization.

The data of these two studies (Ferris *et al.*, 1998; Dumolin-Lapègue *et al.*, 1997) thus appear compatible with the presence of three or four refugia in Iberia, Italy, the Balkans and possibly also the Caucasus. These expanded northward producing tracks of distinct mixtures of cpDNA haplotypes, rather like those imagined on theoretical grounds (Hewitt, 1996, fig 4). Pollen analysis, showing that the spread of oak northward was as fast in the west as the east of Europe, supports this (Huntley & Birks, 1983). There would possibly have been some mixing and reassorting with the Younger Dryas retreat 10 500 BP, and more recently man's progressive reduction of forest cover since the Iron Age will have aided mixing of these primary partitions. Another interesting deduction from these data indicate that the hybrid zone in Finland (Ferris *et al.*, 1998) was probably formed by the meeting of a western genome advancing through Denmark and Sweden and an eastern one expanding through Estonia and Karelia. The more easterly Russian expansion appears on present information not to have travelled along this track and into this region (Fig. 4).

### *Shrews*

The *Sorex araneus* group of shrews shows exceptional centromeric chromosomal variation across Europe (Zima *et al.*, 1996). Recently sequences of mtDNA *cyt b* gene have clarified the phylogenetic relationship of its component taxa and allow some deductions about its ice age range changes (Taberlet, Fumagalli & Hausser, 1994; Fumagalli *et al.*, 1996). The different species and races within this group have a complex patchy distribution in Europe with several taxa located in mountains and some related forms distantly allopatric. Most of Northern Europe from the Cantabrians to Russia is populated by *S. coronatus* in the west and *S. araneus* in the east, with a contact zone running from the western Alps northwards to the Baltic. These differ in their *cyt b* sequence by some 3.7% (Kimura 2 parameter) and probably had Iberian and E Balkan refugia in the last ice age from which they expanded northwards. The Alps and Appenines house a distinct form of *S. araneus*, that differs from its northern neighbour by 1.8%, and probably had refugia in or near these mountains. The rest of Italy contains *S. samniticus* which differs from the northern *S. araneus* by 9% and had an Italian ice age refugium, but its origins are before all the Pleistocene range changes. Indeed there are *Sorex* species more closely related in mtDNA to *S. araneus* from Kazakhstan, the Yukon and Quebec!

Less simple explanations are required for other shrews in this group. *S. granarius* occurs in NW Iberia south of *S. coronatus*, but is quite closely related (1% divergence) to *S. araneus* in NE Europe, and probably had its last refugium in Iberia. With 1% mtDNA divergence it probably arrived there several ice ages ago. More intriguing is the presence of isolated slightly different forms of *S. araneus* in the Pyrenees, Massif Central, Switzerland and France with 0.26–0.65% divergence. These may have survived the last ice age in local refugia as remnants of a greater expansion of *S. araneus* in a previous inter-glacial, or possibly in an expansion before the Younger Dryas cold spell 11 000 BP.

Also of great interest is the presence of a hybrid zone identified on chromosomal rearrangements in the centre of Sweden (Fredga, 1996) near the location of the contact zone for the brown bear. The karyotype of this northern race shows that it is related to populations in northern, central and eastern Finland, eastern Poland and Russia, while the southern Swedish race has a karyotype that places it with Jutland and Western European shrews. This indicates that Sweden was colonized after the ice age from the south by western shrews and from the North by ones from Finland and the east. Significantly the mtDNA sequence from central Finnish shrews is of the highly distinct (13.5%) *S. caecutiens*, which has a Russian distribution, while *S. araneus* mtDNA occurs in SW Finland (Fumagalli *et al.*, 1996). It seems that *S. araneus* expanded from the Balkans through Poland as far as south Finland, while an eastern species colonized eastern Poland, eastern, central and northern Finland, and into Sweden (Fig. 4). Clearly more data will clarify these relationships and possibilities.

#### *Common beech*

*Fagus sylvatica* is distributed across most of Europe from Sicily to Sweden and Galicia to the Caucasus. There is a good pollen record of its post-glacial advance (Huntley & Birks, 1983) which was relatively slow compared with most widespread trees like oak. For example, in Britain, oak became apparent around 9500 BP while beech began to establish around 3500 BP (Birks, 1989). Pollen levels show that it first became common after the ice age in the Carpathians and in southern Italy, indicating that these were two major refugia. There is debate concerning the species presence in Iberia at the end of the ice age; earlier pollen records of the late glacial period from the eastern Pyrenees may have been contaminated (Reille & Lowe, 1993). Allozyme studies have not showed much differentiation among populations, but recent work using cpDNA has provided clear evidence of geographic structure (Demesure, Comps & Petit, 1996). Using a PCR-RFLP approach these authors analysed 399 individuals covering a large part of Europe, identifying some 11 chloroplast haplotypes. Populations from southern Italy were most diverse with three distinct haplotypes not found outside Italy. Those from the Caucasus were also distinct, and three distinct haplotypes were also found in northern Spain and the Pyrenees. Unfortunately there are few samples from the Balkans, but one distinct haplotype was recorded in Hungary. Just one haplotype, Haplotype 5, covers Northern Europe from northern Spain to the Black Sea. Interestingly this haplotype is also present in northern Italy. On the basis of these genetic data, Italy is clearly confirmed as a refugium and the Balkans could possibly be the refugial source of the post-glacial expansion that carried Haplotype 5 all over Europe except southern and central Italy. This deduction rests on the pollen records because of the lack of genetic data from the Balkans. On the genetic data, the diversity in the Pyrenees



and northern Spain would argue for an Iberian refugium, which also contained Haplotype 5. As mentioned, the pollen data for such a scenario is disputed. If the proposed expansion route from the Balkans to Spain, with no Iberian refugium (Demesure *et al.*, 1996) is accepted, then the unlikely carriage of three rare haplotypes from the Balkans with no trace in between requires an explanation. Happily the pollen data are in complete accord to explain the genetic distinction between northern and central Italy; an early post glacial increase in Beech pollen in North Balkans spilled through the Slovenian gap to colonize northern Italy with Haplotype 5, blocking the advance of the distinct Italian cytotypes from the south of the peninsula (Fig. 4).

#### *Other cases*

A number of other cases have been reported using molecular markers which show geographic structure across Europe but where colonization routes can be less clearly inferred. Some however provide useful supportive evidence. In a pioneering study, Wallis & Arntzen (1989) used a RFLP analysis of mtDNA to examine the relationships of the *Triturus cristatus* superspecies of newts across Europe. *T.c. cristatus* is distributed all across Europe north of the Alps; its mtDNA is nearly homogeneous. Adjoining it along the Danube is *T.c. dobrogicus*, which is homogeneous for mtDNA. To the south, *T.c. camifex* from Italy and Slovenia and *T.c. karelini* from Greece and the Balkans contain more mtDNA diversity. These taxa form hybrid zones where they meet, yet their mtDNA has diverged by 3.9–7.1%. The authors argue that *T.c. cristatus* expanded from a small glacial refuge north of the Danube to colonize N. Europe. In the west of France it meets and forms a patchy contact zone with a more distant relative *T. marmoratus*, which probably colonized from an Iberian refugium. This expansion pattern is also similar to that of *Chorthippus parallelus* (Fig. 4).

Taberlet *et al.* (1998) recently compared the phylogeographies of 10 taxa across Europe and included previously unpublished work on *Crocidura suaveolens*, the lesser white toothed shrew, using mtDNA *cyt b* sequences. This showed some 6.5% sequence divergence between a western form in Spain and France and an eastern form in Italy, the Balkans and Turkey. The Turkish form was itself 3.2% divergent from its Balkan relative. Taberlet *et al.* (1998) also report unpublished work on *Arvicola terrestris*, the water vole, which has three subspecies in Italy, Western Europe, and Eastern Europe with Scandinavia. These differ in their mtDNA *cyt b* sequence by 5% and 3.8% respectively. The shrews would seem to have colonized northward from Spain and the Balkans and also possibly Italy. The water voles likewise probably had Iberian and Balkan expansion across Europe, but the Italian subspecies remained in Italy.

As mentioned, allozyme studies have in general proved less useful than DNA studies for the investigation of post-glacial colonization. However, Konnert & Bergmann (1995) reporting on silver fir, *Abies alba*, indicated five Southern European refugia in Iberia, S France, S Italy, N Italy and S Balkans. This mountain tree now occurs in the Pyrenees (Iberian genome), southern Italy (Calabrian genome), the Appenines, the Alps and north into Saxony, Bohemia and Silesia (Italian genome), Balkans, Transylvanians and Carpathians to southern Poland (Balkan genome), Massif Central, Jura Alps and Vosges (French × Italian genome). The allozyme data indicate hybridization and introgression between the Italian and Balkan genomes

from Slovenia to southern Poland. A DNA study would be most interesting, but again northward colonization from three refugia in S France, Italy and the Balkans is the most likely explanation.

One of the classic cases of post-glacial colonization of Europe is that of the house mouse *Mus musculus*, which is commensal with man. The fossil record shows that it entered Europe from the east as *M.m. musculus*, and the south and west from Africa and the Mediterranean as *M.m. domesticus* during the Bronze Age 4000–2800 BP. The two species form a hybrid zone down the centre of Europe from Denmark to the Black Sea and their mtDNA difference indicates a 0.5 Myr divergence that occurred in India. Molecular and fossil studies have provided a particularly full understanding of the evolution of these subspecies and related taxa over several million years in which they spread all over the Old World (Boursot *et al.*, 1993) and which involved many major climatic oscillations.

#### GENERAL FEATURES

These recent studies using molecular markers augment previous palynology and biogeography and a number of features are becoming apparent.

##### *Refugia*

As expected the southern peninsulas of Iberia, Italy, the Balkans and Greece are confirmed as distinct major refugia in the last ice age for most species examined here. Where examined, Turkey also has its distinct genotypes and there is evidence in a number of cases of more easterly genomes from the Black Sea to the Caspian. These eastern regions need to be examined more thoroughly (Fig. 4). In some cases such as *Alnus* and *Fagus* the DNA data provides evidence of refugia not seen in the pollen record. The fossil record for animals is of little use in this context and so molecular markers provide the primary evidence.

##### *Expansion routes*

These distinct southern refugia have contributed to the post-glacial colonization of Europe very differently. In many cases (8/11) the Alps posed an initial barrier to the expansion of Italian genomes so that Northern Europe was colonized from the other refugia thus blocking later Italian genome expansion. The Pyrenees were apparently less of an impediment (4/11), but since colonization from the Balkans blocked that from Italy before reaching the Pyrenees, this may have contributed along with the greater physical barrier posed by the Alps. Refugia in the Balkans have provided Northern European colonization for most species (9/11), except where some more easterly invasion blocked their advance. This could have come from the southern Carpathians, Black Sea, Caucasus or Caspian (2/11). Greek and Turkish genomes were blocked in their regions by more northerly neighbours. It is possible to describe three broad patterns of colonization from the present data, which can be termed, after their exemplars, 'grasshopper', 'hedgehog' and 'bear' (Fig. 4). (1) The grasshopper *Chorthippus parallelus* populated most of Europe from a Balkan refugium, with Iberian and Italian genomes blocked at the Pyrenees and Alps. *Fagus sylvatica* and *Alnus glutinosa* have a similar pattern. (2) The hedgehog

*Erinaceus* populated most of Europe with three genomes advancing north from Iberian, Italian and Balkan refugia, and *Quercus robur* and *Abies alba* probably did similarly given their current distribution. The Italian expansion beyond the Alps may be due to slower dispersal from the Balkans. (3) The bear *Ursos arctos* colonized from Iberia and eastern refugia in a pincer movement which prevented further expansion of Italian and western Balkan genomes. *Triturus* appears similar with its Balkan expansion reaching NW France, so it begins to resemble the grasshopper pattern. *Sorex*, *Arvicola* and possibly *Crocidura* also underwent Iberian and Balkan expansions northwards, with restricted Italian ones. It will be interesting to have such information for more species to look for explanatory commonalities among members of each general pattern.

#### *Hybrid zones*

The different genomes expanded from their refugia, and where they met they formed a hybrid zone. Their differences often meant that there was some hybrid unfitnes and this would tend to produce a narrow tension zone. Otherwise somewhat broader zones could result from low density intermixing on contact (Nichols & Hewitt, 1994). Such hybrid zones have been identified in several of the cases discussed, i.e. grasshopper, hedgehog, bear, newt, shrew and mouse, and may well exist in the other contacts between the expanded refugial genomes. A number of other species show hybrid zones in Europe that are the products of post-glacial colonization of the North from southern regions (ref. Hewitt, 1989). The locations of the better-studied ones occur from the Pyrenees across to the Black Sea (Fig. 5). There is clearly a cluster running down the centre of Europe, produced by eastern and western expansions. Another cluster falls along the Alps and several zones run through the northern Balkans to the Black Sea. The Alpine cluster reflects blocked Italian genomes and similarly the Balkan ones. Some central European zones continue up into Scandinavia, and there is another small cluster in central Sweden, produced by the contact of western genomes from the south and eastern genomes from the north, which is located where the last ice melted from the Baltic.

Such clusters and hybrid zones have been noted before (ref. Hewitt, 1993a, 1996) and as more cases are studied the picture becomes clearer (Taberlet *et al.*, 1998). Indeed Remington (1968) described “suture-zones of hybrid interaction between recently joined biotas” from his review of animals and trees in North America, and he identified six major suture zones and seven minor ones. He also suggested the Urals and Central Europe as possible candidates, but was concerned that long term extensive human cultivation may have blurred the picture. In the light of more recent evidence most of these zones are seen to be more ancient and the products of post-glacial expansion, and many are rather narrow (Hewitt, 1993a). Each zone follows its individual track, which is the product of its adaptive niche, habitat availability and climatic oscillations inducing range changes. The clustering into suture zones results from commonalities of refugia, rate of climate warming and barriers to expansion.

Many more hybrid zones—too many to mention here—occur in Europe in addition to the well-studied ones listed, and specialists in particular groups know of some of them. For example, the Pyrenees and Alps have many records of subspecific contacts between Spanish/French or Italian/Northern forms. Furthermore, both the Iberian and Italian peninsulas have reports and indications of local divergent forms, hybrid zones and subspecies within them. The same indications are found in the Balkans and Greece, and all three peninsulas need more study.

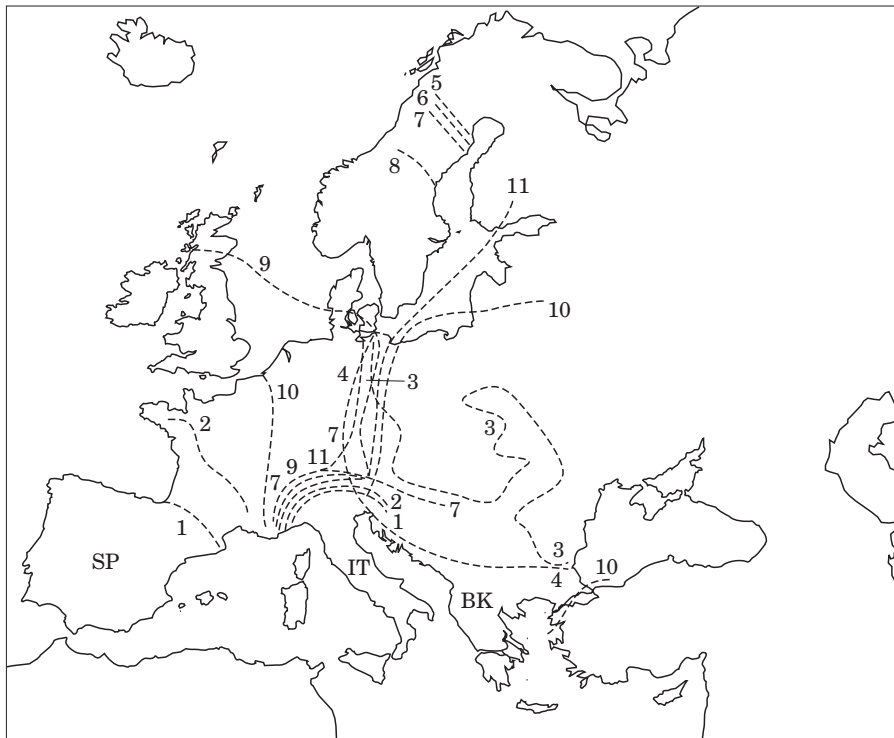


Figure 5. Positions of well studied hybrid zones in Northern Europe. Current evidence indicates a complex subdivision of taxa in the refugial areas of Iberia (SP), Italy (IT) and the Balkans & Greece (BK). 1. *Triturus marmoratus/cristatus/carnifex*, 2. *Chorthippus erythropus/parallelus/italicus*, 3. *Bombina variegata/bombina*, 4. *Mus domesticus/musculus*, 5. *Clethrionomus glareolus*, 6. *Microtus agrestis*, 7. *Sorex coronatus/araneus*, 8. *Ursus arctos*, 9. *Corvus corone/cornix*, 10. *Erinaceus europaeus/concolor*, 11. *Quercus robur-patraea*.

#### *Southern richness and northern purity*

Reduction in genetic diversity from Southern to Northern Europe can be seen at three levels: the number of species, the extent of subspecific division and the allelic variation. The ice ages and subsequent post glacial colonization have been implicated in all three. A loss of alleles as a result of rapid expansion by the leading edge has been predicted, modelled and noted in several species (ref. Hewitt, 1989, 1993a, 1996; Ibrahim *et al.*, 1996). In the European species considered such a pattern has been reported in *Triturus*, *Chorthippus*, *Fagus* and *Alnus*, which interestingly all have the fan-like expansion of the grasshopper *Chorthippus* as compared with the column-like advance of the hedgehog *Erinaceus* (Fig. 4). It may be that expansion was faster, which would predict greater loss of alleles (see Hewitt, 1996, fig. 4), but clearly more data are needed to test such ideas.

The higher number of sister species, subspecies and hybrid zones in southern regions may be explained by both population genetic and geographic factors. As mentioned earlier since post-glacial colonization occurs from the northern populations, those behind are hindered or blocked in their advance by the residents, particularly if they have genetic differences producing even a weak tension zone (Hewitt, 1993). In addition, the varied topography of Southern Europe provides warm valleys for the ice age and cool peaks for warm interstadia. While most

northern expansions are driven extinct by each ice age, populations in southern parts may survive several ice ages by ascending and descending mountains. In this process they may diverge, form hybrid zones, and ultimately speciate (Hewitt, 1996, fig. 6). In this present interglacial one might predict more or less diverged genomes on different mountains with stronger or weaker hybrid zones where they make contact. The current taxon and genetic diversity apparent in the Alps, Appenines, Balkan and Iberian mountain ranges supports this explanation.

#### *Age of divergence*

The extent of divergence of the genomes that have expanded from their ice age refugia to recolonize Europe and form hybrid zones allows a rough estimation of the maximum time they have been effectively separated (Hewitt, 1989). The rate of divergence and its variations are better known for some sequences, with certain mt DNA regions being perhaps best understood at present. Of course, two recently separated populations may, due to chance sorting, differ for two sequence haplotypes that diverged some time ago but have remained polymorphic through gene flow. Such lineage sorting is relatively quicker in mitochondrial than nuclear genomes and is of less concern further back in the phylogeny (Moore, 1995).

Viewed in the light of repeated ice ages in the Pleistocene, with major climatic oscillations recorded back through the Tertiary, Europe's species will have undergone many range contractions and expansions in and out of refugia in the South. In some adverse periods they may go extinct in some or all the three major refugia of Iberia, Italy and the Balkans. Europe would then be recolonized from more distant refugia such as Turkey, the Middle East, and possibly the Caspian region or Africa. The percentage sequence divergence amongst refugial populations of European species thus provides a maximal estimate of the age of this last major recolonization—or possibly the first colonization. Some divergences between subspecific genomes are surprisingly large and some very small (Table 1). As discussed earlier, such data indicate that the grasshopper *C. parallelus* last entered its European refugia no more than five ice ages ago and possibly less. The Iberian *C.p. erythropus* has diverged to almost specific status in that time. The brown bear divergence is of a similar age. At the other end of the scale the hedgehogs, toads and newts could have been present in or near their refugia for several million years and almost certainly for all the Pleistocene ice ages. Over this time the refugial genomes will have repeatedly been the source of northern genomes in warm interstadials. This would provide the opportunity for hybridization between genomes and the evolution of recombinant forms and species. Meanwhile the refugial populations may diverge effectively allopatrically, and possibly speciate (Hewitt, 1989).

#### CONCLUSIONS

Molecular genetic data, particularly those from DNA sequences, are proving very useful in identifying the locations of refugia from which various taxa colonized Europe after the last ice age. In particular they tell us whether different parts of Europe were filled by distinct genomes of a species from refugia in Iberia, Italy, Balkans, Greece or further east. In this they can be more discriminating than fossil pollen records, e.g. alder, beech and oak. Combined, the two approaches are quite

TABLE 1. DNA sequence divergence and estimated maximum time of separation in species groups colonizing Europe after last ice age. The southern refugia of distinct genomes are given: S=Iberia, I=Italy, B=Balkans, W=west, E=east. Those not expanding out of their peninsula are in brackets

Organism	DNA Sequence	Diverg/Max Age	Refugia	Authors
<i>Bombina bombina</i> (fire bellied toad)	mt RFLP	9.4% 5M	(I) B B	Szymura <i>et al.</i> , 1985
<i>Erinaceus europaeus</i> (hedgehog)	mt cyt b	6–12% 3–6 M	S I B	Santucci <i>et al.</i> , 1998
<i>Triturus cristatus</i> (crested newt)	mt RFLP	4–8% 2–4 M	S I B	Wallis & Arntzen, 1989
<i>Arvicola terrestris</i> (water vole)	mt cyt b	4–7.6% 2–4 M	S I B	Taberlet <i>et al.</i> , 1998
<i>Crocidura suaveolens</i> (white toothed shrew)	mt cyt b	3–6.4% 1.5–3.2 M	S (I) B	Taberlet <i>et al.</i> , 1998
<i>Mus musculus</i> (house mouse)	mt RFLP	3.4% 1.7 M	W & E	Ferris <i>et al.</i> , 1983
<i>Microtus agrestis</i> (field vole)	mt RFLP	2% 1 M	W & E	Jaarola & Tegelstrom, 1995
<i>Sorex araneus</i> (red toothed shrew)	mt cyt b	1–3.8% 0.5–2 M	S (I) B	Taberlet <i>et al.</i> , 1994
<i>Ursus arctos</i> (brown bear)	mt contr reg	2.7–7% 0.35–0.85 M	S (I) B	Taberlet & Bouvet, 1994
<i>Chorthippus parallelus</i> (meadow grasshopper)	mt 6.7 kb	0.7–0.9% 0.3–0.5 M	(S) (I) B	Szymura <i>et al.</i> , 1996

revealing. For animals, DNA data are more valuable, since the fossil record is sparse or non-existent. Significantly, species inhabiting Northern Europe are different for the southern refugia from which their component genomes expanded, thereby producing three broad patterns of subspecific parapatry. These substructures and routes of expansion are termed ‘grasshopper’, ‘hedgehog’ and ‘bear’ after paradigm species patterns.

The extent of genetic diversity and further structure within these broad range subdivisions is beginning to emerge as more DNA approaches are employed. Already a number of species show lower genetic diversity in northern populations that expanded rapidly from the Balkans, e.g. alder, beech, grasshoppers and newts. In contrast, southern populations in refugial regions show considerable diversity both in terms of alleles and distinct genomes. This extends biogeographical observations of greater subspecific and specific diversity in Mediterranean than northern temperate parts. Such southern richness may be seen as a product of repeated major climatic oscillations, which genomes were able to survive in the southern mountainous peninsulas where suitable habitats could be found continuously by small range changes. Meanwhile, interglacial northern expansions were extinguished by each cold period. This ensuing repeated allopatry could allow their divergence to subspecific and even specific status. The maximum time that has been spent in a refugium is reflected in the sequence divergence of a genome.

To allow such deduction of past events it is necessary to use a suitable DNA sequence that incorporates genealogy and sufficient variability. For Pleistocene and Holocene phylogeography it has been suggested that sequence evolution was too slow to provide regional variation. However, mtDNA in animals and even cpDNA with its much slower divergence, are proving useful, seemingly because refugial

populations have diverged over a sufficient period of time before the last ice age and have retained more anciently divergent haplotypes that were sorted differentially among them by drift. Where a severely bottlenecked population has recently colonized an area there may be little or no signal from these sequences, and other genetic markers will be required for such studies. Population diversity and genome subdivision are better understood by determining the history of their construction, since this allows present events to be studied in a more meaningful context.

## ACKNOWLEDGEMENTS

I am most grateful to the people who have discussed these ideas with me, to the EC, SERC and NERC who funded my work, to Kamal Ibrahim, Fiammetta Santucci, David Rees and James Hewitt for help with the manuscript, and to Louis Bernatchez for his helpful review.

## REFERENCES

- Angers B, Bernatchez L. 1998.** Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related Brook Charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. *Molecular Biology and Evolution* **15**: 143–159.
- Atkinson TC, Briffa KR, Coope GR. 1987.** Seasonal temperatures in Britain during the past 22,000 years, reconstructed using beetle remains. *Nature* **325**: 587–592.
- Beaulieu JL de, Reille M. 1992.** The last climatic cycle at La Grande Pile (Vosges, France): A new pollen profile. *Quaternary Science Reviews* **11**: 431–438.
- Bennett KD. 1986.** The rate of spread and population increase of forest trees during the post-glacial. *Philosophical Transactions of the Royal Society, London* **B 314**: 523–531.
- Bennett KD. 1997.** *Evolution and Ecology: The Pace of Life*. Cambridge: Cambridge University Press.
- Birks HJB. 1989.** Holocene isochrone maps and patterns of tree-spreading in the British Isles. *Journal of Biogeography* **18**: 103–115.
- Böhle UR, Hilger HH, Martin WF. 1996.** Island colonization and evolution of the insular woody habit in *Echium* L. (Boraginaceae). *Proceedings of the National Academy of Sciences USA* **93**: 11740–11745.
- Boursot P, Auffray J-C, Britton-Davidian J, Bonhomme F. 1993.** The Evolution of House Mice. *Annual Reviews of Ecology and Systematics* **24**: 119–152.
- Butler PM. 1988.** Phylogeny of the insectivores. In: Benton MJ, ed. *The Phylogeny and Classification of the Tetrapods, Volume 2, Mammals*. Oxford: Clarendon Press, 117–141.
- Butlin RK, Hewitt GM. 1985.** A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera: Acrididae) I. Morphological and electrophoretic characters. *Biological Journal of the Linnean Society* **26**: 269–285.
- Coope GR. 1977.** Fossil coleopteran assemblages as sensitive indicators of climatic change during the Devensian (last) cold stage. *Philosophical Transactions of the Royal Society, London* **B280**: 313–340.
- Coope GR. 1990.** The invasion of Northern Europe during the Pleistocene by Mediterranean species of Coleoptera. In: di Castri F, Hansen AJ, De Bussche M, eds. *Biological Invasions in Europe and the Mediterranean Basin*. Dordrecht: Kluwer, 203–215.
- Coope GR. 1994.** The response of insect faunas to glacial-interglacial climatic fluctuations. *Philosophical Transactions of the Royal Society, London*, **B344**: 19–26.
- Cooper SJB, Hewitt GM. 1993.** Nuclear DNA sequence divergence between parapatric subspecies of the grasshopper *Chorthippus parallelus*. *Insect Molecular Biology* **2**: 1–10.
- Cooper SJB, Ibrahim KM, Hewitt GM. 1995.** Post-glacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. *Molecular Ecology* **4**: 49–60.
- Corbet GB. 1988.** The family Erinaceidae: a synthesis of its taxonomy, phylogeny ecology and zoogeography. *Mammal Review* **18**: 117–172.

- Dansgaard W, Johnsen SJ, Clausen HB, Dahl-Jensen D, Gundestrup NS, Hammer GU, Hvidberg CS, Steffensen JP, Sveinbjornsdottir AE, Jouzel J, Bond G. 1993. Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* **364**: 218–220.
- Demesure B, Comps B, Petit RJ. 1996. Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* **50**: 2515–2520.
- Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ. 1997. Phylogeographic structure of white oaks throughout the European continent. *Genetics* **146**: 1475–1487.
- Ennos RA. 1994. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**: 250–259.
- Ferris C, Davy AJ, Hewitt GM. 1997. A strategy for identifying introduced provenances and translocations. *Forestry* **70**: 211–222.
- Ferris C, King RA, Vainola R, Hewitt GM. 1998. Chloroplast DNA recognises three refugial sources of European oaks and shows independent eastern and western immigrations to Finland. *Heredity* **80**: 584–593.
- Ferris C, Oliver RP, Davy AJ, Hewitt GM. 1993. Native oak chloroplasts reveal an ancient divide across Europe. *Molecular Ecology* **2**: 337–344.
- Ferris C, Oliver RP, Davy AJ, Hewitt GM. 1995. Using chloroplast DNA to trace post-glacial migration routes of oaks into Britain. *Molecular Ecology* **4**: 731–738.
- Ferris SD, Sage RD, Huang C-M, Nielson JT, Ritte U, Wilson AC. 1983. Flow of mitochondrial DNA across a species boundary. *Proceedings of the National Academy of Sciences, USA* **80**: 2290–2294.
- Filippucci MG, Simson S. 1996. Allozyme variation and divergence in Erinaceidae (Mammalia, Insectivora). *Israel Journal of Zoology* **42**: 335–345.
- Flanagan NS, Mason PL, Gosalvez J, Hewitt GM. 1999. Chromosomal differentiation through an Alpine hybrid zone in the grasshopper *Chorthippus parallelus*. *Journal of Evolutionary Biology* **12**: 577–585.
- Fredga K. 1996. The chromosome races of *Sorex araneus* in Scandinavia. *Hereditas* **125**: 123–135.
- Fumagalli L, Hausser J, Taberlet P, Gielly L, Stewart DT. 1996. Phylogenetic structures of the Holarctic *Sorex araneus* group and its relationships with *S. samniticus*, as inferred from mtDNA sequences. *Hereditas* **125**: 191–199.
- Gaut BS, Morton BR, McCaig BC, Clegg MT. 1996. Substitution rate comparisons between grasses and palms: Synonymous rate differences at the nuclear gene *Adh* parallel differences at the plastid gene *rbcl*. *Proceedings of the National Academy of Science USA* **93**: 10274–10279.
- Guiot J. 1997. Back to the last interglacial. *Nature* **388**: 25–27.
- Harrison C. 1982. *An atlas of the birds of the western palearctic*. London: Collins.
- Hays JD, Imbrie J, Shackleton NJ. 1976. Variations in the Earth's orbit: pacemaker of the ice ages. *Science* **194**: 1121–1132.
- Hengeveld R. 1989. *Dynamics of Biological Invasions*. London: Chapman and Hall.
- Hewitt GM. 1975. A sex chromosome hybrid zone in the grasshopper *Podisma pedestris*. *Heredity* **35**: 375–387.
- Hewitt GM. 1988. Hybrid zones – natural laboratories for evolutionary studies. *Trends in Ecology and Evolution* **3**: 158–167.
- Hewitt GM. 1989. The subdivision of species by hybrid zones. In: Otte D, Endler J, eds. *Speciation and its Consequences*. Sunderland, Massachusetts: Sinauer Associates: 85–110.
- Hewitt GM. 1990. Divergence and speciation as viewed from an insect hybrid zone. *Canadian Journal of Zoology* **68**: 1701–1715.
- Hewitt GM. 1993a. Post-glacial distribution and species substructure: lessons from pollen, insects and hybrid zones. In: Lees DR, Edwards D, eds. *Evolutionary Patterns and Processes*. Linnean Society Symposium Series **14**: 97–123. London: Academic Press.
- Hewitt GM, 1993b. After the ice: *Parallelus* meets *Erythropus* in the Pyrenees. In: Harrison RG, ed. *Hybrid zones and the evolutionary process*. New York: Oxford University Press: 140–146.
- Hewitt GM, 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**: 247–276.
- Hewitt GM, Barton NH. 1981. The structure and maintenance of hybrid zones as exemplified by *Podisma pedestris*. In: Blackman RL, Hewitt GM, Ashburner M, eds. *Insect Cytogenetics. Symposia of the Royal Entomological Society*, vol. 10. Oxford: Blackwell, 149–170.
- Hewitt GM, Butlin RK. 1997. Causes and consequences of population structure. In: Krebs J, Davies, N, eds. *Behavioural Ecology: An Evolutionary Approach*. Oxford: Blackwell, 350–372.
- Higgins LG, Hargreaves B. 1983. *The Butterflies of Britain and Europe* London: Collins.



- Hillis DM, Moritz C, Mable BK. 1996. *Molecular Systematics* Sunderland: Sinauer Associates.
- Huntley B, Birks HJB. 1983. *An Atlas of Past and Present Pollen Maps for Europe* Cambridge: Cambridge University Press.
- Huntley B. 1990. European vegetation history: palaeovegetation maps from pollen data – 1300 yr BP to present. *Journal of Quaternary Science* 5: 103–122.
- Ibrahim K, Nichols RA, Hewitt GM. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77: 282–291.
- Jaarola M, Tegelström H. 1995. Colonization history of north European field voles (*Microtus agrestis*) revealed by mitochondrial DNA. *Molecular Ecology* 4: 299–310.
- Johnk N, Siegmund HR. 1997. Population structure and post-glacial migration routes of *Quercus robur* and *Quercus patraea* in Denmark, based on chloroplast DNA analysis. *Scandinavian Journal of Forestry Research* 12: 130–137.
- King RA, Ferris C. 1998. Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology* 7: 1157–1161.
- Kohn M, Knauer F, Stoffella A, Schröder W, Pääbo S. 1995. Conservation genetics of the European brown bear – a study using incremental PCR of nuclear and mitochondrial sequences. *Molecular Ecology* 4: 95–103.
- Konnert M, Bergmann F. 1995. The geographical distribution of genetic variation of silver fir (*Abies alba*, Pinaceae) in relation to its migration history. *Plant Systematics and Evolution* 196: 19–30.
- Lunt DH, Ibrahim KM, Hewitt GM. 1998. MitDNA phylogeography and post-glacial patterns of subdivision in the meadow grasshopper. *Chorthippus parallelus*. *Heredity* 80: 63–641.
- Mack RN. 1981. Invasion of *Bromus tectorum* L into western North America: an ecological chronicle. *Agro-Ecosystems* 7: 145–165.
- Moore WS. 1995. Inferring phylogenies from mtDNA variation – mitochondrial-gene trees versus nucleargene trees. *Evolution* 49: 718–726.
- Nichols RA, Hewitt GM. 1994. The genetic consequences of long distance dispersal during colonisation. *Heredity* 72: 312–317.
- Nichols RA, Beaumont MA. 1996. Is it ancient or modern history that we can read in our genes? In: Hochberg ME, Clobert J, Barbault R, eds. *Aspects of the Genesis of Maintenance of Biological Diversity* Oxford: Oxford University Press, 69–87.
- Orti G, Hare MP, Avise JC. 1997. Detection and isolation of nuclear haplotypes by PCR-SSCP. *Molecular Ecology* 6: 575–580.
- Petit RJ, Kremer A, Wagner DB. 1993. Geographical structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics* 87: 122–128.
- Ponel P. 1997. The response of Coleoptera to late-Quaternary climate changes: Evidence from north-east France. In: Huntley B, et al., eds. *Past and future rapid environmental changes* NATO ASI Series Vol. 147. Berlin, Springer-Verlag, 143–151.
- Reille M, Lowe JJ. 1993. A re-evaluation of the vegetation history of the eastern Pyrenees (France) from the end of the Last Glacial to the Present. *Quaternary Science Reviews* 12: 47–77.
- Remington CL. 1968. Suture-zones of hybrid interaction between recently joined biotas. *Evolutionary Biology* 2: 321–428.
- Reynolds WJ. 1980. A re-examination of the characters separating *Chorthippus montanus* and *C. parallelus* (Orthoptera: Acrididae). *Journal of Natural History* 14: 283–303.
- Rohling RJ, Fenton M, Jorissen FJ, Bertrand P, Ganssen G, Caulet JP. 1998. Magnitudes of sea-level lowstands of the past 500,000 years. *Nature* 394: 162–165.
- Santucci F, Emerson BC, Hewitt GM. 1998. Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology* 7: 1163–1172.
- Sewell MM, Parks CR, Chase MW. 1996. Intraspecific chloroplast DNA variation and biogeography of North American *Liriodendron* L. (Magnoliaceae). *Evolution* 50: 1147–1154.
- Slade RW, Moritz C, Heideman A. 1994. Multiple nuclear-gene phylogenies: Application to pinnipeds and comparison with a mitochondrial DNA gene phylogeny. *Molecular Biology and Evolution* 11: 341–356.
- Soltis DE, Soltis PS, Milligan BG. 1992. Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. In: Soltis PS, Soltis DE, Doyle JJ, eds. *Molecular Plant Systematics*. New York: Chapman and Hall, 117–150.
- Szymura JM, Lunt DH, Hewitt GM. 1996. The sequence and structure of the meadow grasshopper (*Chorthippus parallelus*) mitochondrial 12s TRNA, ND2, COI, COII, ATPase 8 and tRNA genes. *Insect Molecular Biology* 5: 127–139.

- Szymura JM, Spolsky C, Uzzel T. 1985.** Concordant changes in mitochondrial and nuclear genes in a hybrid zone between two frog species (genus *Bombina*). *Experientia* **41**: 1469–1470.
- Taberlet P, Bouvet J. 1994.** Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear (*Ursus arctos*) in Europe. *Proceedings of the Royal Society of London B* **255**: 195–200.
- Taberlet P, Fumagalli L, Hausser J. 1994.** Chromosomal versus mitochondrial DNA evolution: tracking the evolutionary history of the southwestern European populations of the *Sorex araneus* group (Mammalia, Insectivora). *Evolution* **48**: 623–636.
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cossons J-F. 1998.** Comparative phylogeography and post-glacial colonization routes in Europe. *Molecular Ecology* **7**: 453–464.
- Taberlet P, Swenson JE, Sandegren F, Bjärvall A. 1995.** Localization of a contact zone between two highly divergent mitochondrial DNA lineages of the brown bear (*Ursus arctos*) in Scandinavia. *Conservation Biology* **9**: 1255–1261.
- Wallis GP, Arntzen JW. 1989.** Mitochondrial DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43**: 88–104.
- Webb T, Bartlein PJ. 1992.** Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Reviews of Ecology and Systematics* **23**: 141–173.
- Whitlock C, Bartlein PJ. 1997.** Vegetation and climate change in northwest America during the past 125 kyr. *Nature* **388**: 57–61.
- Zanetto A, Kremer A. 1995.** Geographical structure of gene diversity in *Quercus patraea* (Matt.) Liebl.I. Monolocus patterns of variation. *Heredity* **75**: 506–517.
- Zima J, Fedyk S, Fredga K, Hausser J, Mishta A, Searle JB, Volobouer VT, Wojcick JM. 1996.** The list of the chromosome races of the common shrew (*Sorex araneus*). *Hereditas* **125**: 97–107.