mtDNA phylogeography and postglacial patterns of subdivision in the meadow grasshopper *Chorthippus parallelus*

DAVID H. LUNT†, KAMAL M. IBRAHIM & GODFREY M. HEWITT*

Population Biology Sector, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, U.K.

A 300 bp portion of the mitochondrial cytochrome oxidase subunit I (*COI*) gene was used to investigate the intraspecific genetic structure of the European meadow grasshopper *Chorthippus parallelus*. Levels of genetic subdivision between geographical regions in Europe were assessed using the K_{ST} statistic and revealed patterns indicative of the postglacial history of this organism. Northern European populations are shown to share a very recent common ancestor with, and hence originate from, a Balkan expansion. Genetic distances between these areas and both southern Italy and central Spain are shown to be approximately equal, even though only the Spanish form is recognized as a distinct subspecies. Distance estimates indicate that the common ancestor of the two subspecies lived more than one glacial cycle ago, and the refugial populations have probably been diverging in isolation for five or six glacial cycles ($\approx 550\,000$ years). Comparisons are made to the analysis of a similar data set using an anonymous nuclear DNA marker (Cooper *et al*., 1995). Although the structure revealed is generally very similar, differences in the location of transition zones between different genetic forms of *C. parallelus* may reflect the different historical assortment of these molecules.

Keywords: *Chorthippus parallelus*, European phylogeography, glacial refugia, mitochondrial DNA, postglacial expansion.

Introduction

The flightless meadow grasshopper *Chorthippus parallelus* (Zetterstedt) is one of the most common and widespread grasshoppers in Europe. Its range stretches from southern Sweden to southern Spain, from western France to as far east as the Urals, and possibly beyond. Yet *C. parallelus* is not a highly mobile species; Virdee & Hewitt (1990) have estimated dispersal distances at only 30 m per generation. Thus, gene flow between different regions of the species range is likely to be low and patterns of genetic differentiation will reflect historical patterns of range expansion.

During the last glaciation, when ice, steppe and tundra covered most of central Europe, *C. parallelus* could only have survived in warmer southern refugia. The number and locations of these refugia are uncertain, though the positions of refugia for deciduous trees in Spain, Italy and the Balkans

(Huntley & Birks, 1983) indicate that the conditions there were also suitable for *C. parallelus*. It is most likely that there were at least two refugia, and that one of these was in Iberia. The evidence for this comes from the location of the hybrid zone in the Pyrenees between the Spanish subspecies *C. p. erythropus* and the French subspecies *C. p. parallelus*. Because *C. p. parallelus* could not have survived the glaciation in its current position north of the Pyrenees it must have colonized this territory by postglacial expansion from a second refugium. Genetic evidence for this secondary contact has been presented by Cooper & Hewitt (1993). Furthermore, data from nuclear DNA indicate significant population differentiation between several distinct geographical regions of Europe, which may correspond to separate ice-age refugia (Cooper *et al*., 1995).

An important consideration in understanding the patterns of range expansion that led to the present-day distribution of *C. parallelus* in Europe is to discover the refugial origins of the European populations. By using genetic markers it may be possible

*Correspondence. E-mail: g.hewitt@uea.ac.uk
†Present address: Biology Department, University of California, Riverside, CA 92521, U.S.A.

633

to test if these populations are of single or multiple refugial origin and to identify the refuge or refugia from which they expanded. This study sets out to test various hypotheses of refugial locations and postglacial expansion routes of *C. parallelus* using inter-regional analysis of mtDNA sequence data. Mitochondrial sequence divergence levels are interpreted in terms of both extent and age of separation of these groups. The results are compared to those obtained from a recent survey of an anonymous nuclear DNA marker in the same species (Cooper *et al.*, 1995).

Materials and methods

Grasshoppers were collected from across the European range of *C. parallelus* and those locations analysed here are shown in Fig. 1. Identification, tissue preservation and DNA isolation were carried out as described by Cooper *et al.* (1995).

Mitochondrial DNA sequences from part of the cytochrome oxidase subunit I were obtained by direct sequencing of PCR-amplified DNA. This region corresponds to bp 4158–4458 in the published sequence (Szymura *et al.*, 1996). Template was generated by a primary amplification with primers S3825 (5′-CTTTATATTTGGAGCAT-GAGCAGG-3′) and A4785 (5′-CCTGTTAATCC-TCCAACTGTAAATA-3′), followed by reamplification of a 10^{-2} dilution of the product with primers S4134B (biotin-5′-GGAACAGCATGAACAGTT-TACCC-3′) and A4688 (5′-GCTAATCATCTAAA-AATTTTAATTCCTGTAGG-3′). Reaction condi-

tions were: 10 ng DNA, 2.5 mM MgCl$_2$, 200 μM dNTP, 60 nM each primer, 0.75 units Taq polymerase (Promega) and $1\times$ PCR reaction buffer (50 mM KCl, 10 mM Tris-HCl (pH 8.6), 0.1% Triton X-100) in 100 μL volume overlaid with mineral oil. Cycle conditions were: 3 min at 94°C, (40 s at 94°C, 1 min at 60°C, 2 min at 72°C) $\times 35$, 10 min at 72°C. Single-stranded template was prepared by solid-phase strand separation using streptavidin-coated magnetic beads (Dynal). Sequencing reactions were performed using the AUTOREAD dideoxy chain termination kit (Pharmacia) with the fluorescently labelled primer A4460F (5′-TAAAATATAAACTT-CAGGATGTCC-3′). Resolution of fragments was carried out on the ALF sequencer (Pharmacia), with all polymorphic sites double-checked manually. Sequences were pruned to the 300 bp obtained in all reactions and aligned using the Lasergene multiple alignment program MEGALIGN (DNAstar, West Ealing). Sequences were pooled into groups representing geographical regions defined by distance and physical barriers. Although a larger set of individuals was used to characterize the nature of the sequence evolution, suitably located geographical subsets were chosen to examine the specific phylogeographical questions. Abbreviated names of the regions and the sample sizes used are shown in Fig. 1. Abbreviations used in Fig. 2 are as follows: UK, Britain; NF, northern France; CAF, central and Alpine France; PF, Pyrenean France; NIt, northern Italy; CIt, central Italy; SIt, southern Italy; PSp, Pyrenean Spain; NSp, northern Spain; CSp, central Spain; SSp, southern Spain; EB, eastern Balkans; SWB,
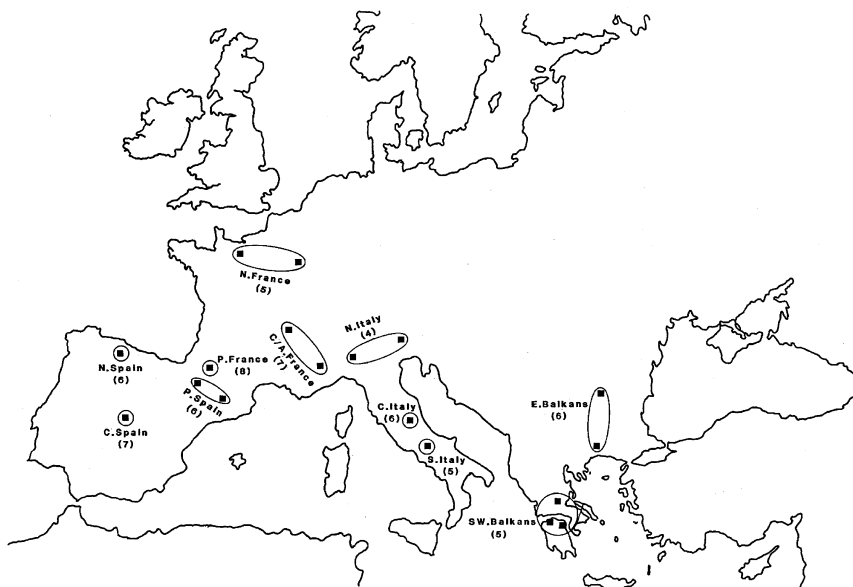

Fig. 1 A map of western Europe showing the location of the *Chorthippus parallelus* populations sampled to represent different regions.

Fig. 2 Sequence variation of the 31 mitotypes observed in *Chorthippus parallelus* and their geographical distribution. Positions varying from that of mitotype 1 are boxed. Base positions are the number of nucleotides upstream of the 3′ end of the sequencing primer.

Mitotype	52	127	129	138	149	162	166	168	169	181	196	202	225	244	252	256	Regional Distributions
Mitotype 1	G	T	A	G	G	C	C	A	T	C	T	T	A	A	C	G	NF (3) CSp PSp NSp (3) SSp (2) Tu (2)
Mitotype 2	G	T	A	**T**	G	**T**	C	A	T	**T**	T	T	A	A	C	G	SWB
Mitotype 3	G	T	A	**T**	G	C	C	A	T	**T**	T	T	A	A	C	G	PF (3) CAF NF NIt (2) CIt (2)
Mitotype 4	G	T	A	G	**G**	C	C	A	T	**T**	T	T	A	A	C	G	Wie (5) EB (2) SG PF SWB
Mitotype 5	**A**	T	A	G	**A**	C	C	A	T	C	T	T	A	A	C	G	CAF (2) Po
Mitotype 6	**A**	T	A	G	**A**	C	C	A	T	C	T	**C**	A	A	C	G	CAF Po (2) PSp
Mitotype 7	**A**	T	A	G	G	C	C	**G**	T	C	T	T	A	A	C	G	SIt (4)
Mitotype 8	G	T	A	G	G	C	C	A	T	C	**C**	T	A	A	C	**T**	NSp
Mitotype 9	G	T	A	G	G	C	C	A	T	C	T	T	A	A	C	**T**	CSp (6) NSp SSp
Mitotype 10	G	T	A	G	G	C	C	A	T	C	**C**	**C**	A	A	C	G	Rus PSp
Mitotype 11	G	T	A	G	G	C	**T**	A	**C**	C	T	T	A	A	C	G	SWB
Mitotype 12	G	T	A	G	G	C	C	A	T	C	T	T	A	A	**T**	G	SWB Hg (2) SFi
Mitotype 13	**A**	T	A	G	G	C	C	A	T	C	T	**C**	A	**A**	C	G	NIt (2) PSp (2)
Mitotype 14	G	T	A	G	G	C	C	A	T	C	T	**C**	A	**T**	C	G	CIt
Mitotype 15	G	T	A	G	G	C	C	A	T	C	T	**C**	A	A	C	**A**	NSp Tu
Mitotype 16	G	T	A	G	G	C	C	A	T	C	T	**C**	A	A	C	G	EBK CIt
Mitotype 17	**A**	T	A	G	G	C	C	A	T	C	**C**	T	A	**G**	C	G	PF
Mitotype 18	**A**	T	A	G	G	C	C	A	T	C	**C**	C	A	A	T	**A**	CAF (2)
Mitotype 19	G	T	A	G	**A**	C	C	A	T	C	**C**	C	A	**G**	C	G	SWB PSp
Mitotype 20	**A**	T	A	G	G	C	C	A	T	C	**C**	C	A	A	C	G	EB PF SIt SFi
Mitotype 21	G	T	A	**T**	G	C	C	A	T	C	**C**	T	A	A	C	G	CIt
Mitotype 22	G	T	A	**T**	G	C	C	A	T	**T**	T	**T**	A	A	C	G	UK
Mitotype 23	G	T	A	**T**	G	C	C	A	T	**T**	T	**T**	**G**	A	C	G	UK
Mitotype 24	G	T	**G**	G	G	C	C	A	T	C	C	T	A	**T**	C	G	EB
Mitotype 25	**A**	T	A	G	G	C	C	A	T	C	**C**	C	A	A	**T**	G	SFi
Mitotype 26	G	T	A	G	G	C	C	A	T	C	**C**	C	A	A	C	**T**	SFi
Mitotype 27	G	T	A	G	G	C	C	A	T	C	T	**C**	A	A	**T**	G	EB CAF
Mitotype 28	G	T	A	**T**	G	C	C	A	T	C	T	T	A	A	C	G	CIt
Mitotype 29	**A**	T	A	G	G	C	C	A	T	C	**C**	**G**	A	A	C	G	PF
Mitotype 30	**A**	**C**	A	G	G	C	C	A	T	C	**C**	**G**	A	A	C	G	PF
Mitotype 31	G	T	A	G	G	C	C	A	T	C	**C**	T	A	A	C	G	NF
Mitotype 32	G	T	A	G	G	C	C	A	T	**T**	**T**	C	A	A	**T**	G	Hg (2)

south-western Balkans; SG, southern Germany; Hg, Hungary; Rus, Russia; SFi, southern Finland; Tu, Turkey; Po, Poland; Wie, eastern Austria. The exact locations of all individuals sequenced may be obtained from the authors.

Genetic differentiation between geographical regions was estimated using the K_{ST} statistic (Hudson *et al*., 1992). The statistical significance of the K_{ST} estimates was tested following the permutation method suggested in their paper. One thousand random permutations were carried out to estimate

P, the probability of obtaining by chance a K_{ST} value equal to or greater than that observed for the actual data. Significant K_{ST} values are shown in Table 1 along with the intraregional diversity statistic, K_s. Phenograms were constructed, using a matrix of K_{ST} values as distance measures, by the Neighbour-Joining method (Saitou & Nei, 1987) as implemented in PHYLIP (Felsenstein, 1991).

To estimate the percentage sequence divergence between *C. p. parallelus* and *C. p. erythropus* mtDNA, and hence to gain information about the

Table 1 Estimates of differentiation (K_{ST}) between a range of European populations of *Chorthippus parallelus*. Only K_{ST} values significantly different from zero ($P < 0.05$) are shown

	Inter-regional K_{ST} values										
Region	N. Fran	P. Fran	C./A. Fran	N. Spain	P. Spain	C. Spain	N. Italy	C. Italy	S. Italy	SW. Balkan	E. Balkan
N. Fran	—	0.0522	0.1538	NS	0.2063	0.3623	NS	NS	0.4130	NS	NS
P. Fran		—	NS	0.1286	0.0874	0.2986	NS	NS	0.1854	NS	NS
C./A. Fran			—	0.1400	NS	0.3027	NS	NS	0.1619	NS	NS
N. Spain				—	0.1787	NS	NS	0.1295	0.4000	NS	NS
P. Spain					—	0.4286	NS	0.1659	0.2537	NS	NS
C. Spain						—	0.3269	0.3589	0.6480	0.1839	0.2683
N. Italy							—	NS	0.2653	NS	NS
C. Italy								—	0.3718	NS	NS
S. Italy									—	0.2406	0.2405
SW. Balk										—	NS
E. Balkan											—
	Intraregional K_s values (% substitution per bp)										
	0.4000	1.1190	1.2381	0.4666	0.7556	0.0952	1.0000	0.7556	0.4000	1.4000	1.0000

likely age of divergence, the following procedures were employed. French *C. parallelus* were taken to represent *C. p. parallelus* because it is by comparison to French populations that *C. p. erythropus* has been described. All the mitotypes found in Spain and all those in France were aligned and pairwise sequence comparisons performed by DNADIST (PHYLIP V. 3.4) with a Jukes–Cantor correction for multiple hits. Because there are indications in the data that Pyrenean Spain populations may be influenced by French populations (see Discussion) the analysis was repeated with these excluded. Furthermore, recognizing that the impact of a single base error can be significant when comparing sequences that are not highly diverged, we obtained an estimate of the error in the measure of the distance that is associated with this by jackknifing across mitotypes.

Results

Sequence variation

A 300 bp segment of the mtDNA *COI* gene was sequenced and aligned from 90 European grasshoppers (of which 65 are included in the K_{ST} analyses, described below, to address the specific phylogeographical questions). The base composition of this region varied slightly depending on the individual but was commonly found to be: 35.33% A; 15.67% C; 17.67% G; 31.67% T. The sequence contains 16 (5.3%) polymorphic sites (Fig. 2) generating 31 different mitotypes representing a minimum of 53 transitions and 11 transversions (4.8:1) from a common sequence. The 16 variable sites were not evenly distributed among the three codon positions, with nine substitutions being detected at codon position three, six at the first codon position and only a single substitution at position two. Nine of the polymorphic sites were informative in that the nucleotides were shared by at least two of the unique haplotypes which formed the taxonomic units of the phylogenetic analysis. No additions, deletions or duplications were observed. Of the 19 different substitutional changes 13 were silent and six caused amino acid replacements.

Phylogenetic analyses

The 31 unique mitotypes were analysed using distance and maximum parsimony methods in order to determine the phylogenetic relationships between them. However, both methods produced ambiguous results. PAUP (Swofford, 1993) found over 28 000 trees of equal minimal length and consequently the consensus tree only resolves the relationship

between a very small fraction of the mitotypes (Lunt, 1994). Analysis by phenetic methods, in contrast, typically produced a fully bifurcating tree; however, support is weak with only one bootstrap value greater than 50%. Of the 16 variable positions in this sequence only nine were informative. Such a data set will always cause problems for parsimony methods in resolving the relationships between this many taxa because tree bifurcations are dependent on these shared, derived characters. Although quantifying how many informative characters are sufficient depends on the nature of the dataset, it should at least equal the number of taxa being analysed (Stewart, 1993). Thus, it may be that for this data set both these standard methods of phylogenetic analysis are inappropriate.

Analysis of population genetic structure using the K_{ST} statistic has been described by Hudson *et al.* (1992). Table 1 shows the K_{ST} estimates of differentiation between the 11 regions from which the samples were collected (Fig. 1). The three phenograms in Fig. 3 are essentially graphical representations of these distances that highlight patterns of
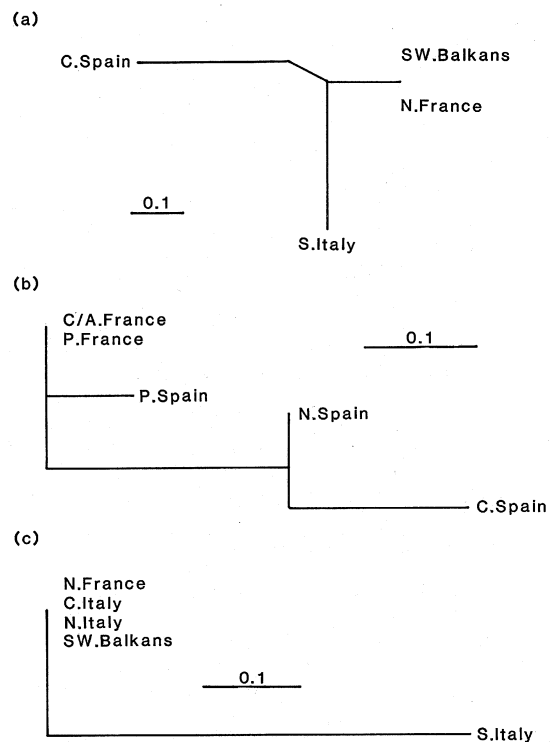
Fig. 3 Neighbour-joining phenograms based on the K_{ST} distance matrix given in Table 1, depicting the relationships between (a) the regions representing the putative refugia, (b) French and Spanish regions, and (c) French and Italian regions.

specific interest. Figure 3(a) depicts the genetic distances between populations from northern France, taken to be representative of all the northern European populations, and the putative refugial populations in the southern parts of Spain, Italy and the Balkans. Similarly, the divergences between French and all the Spanish populations and between the French and all the Italian populations are shown in Fig. 3(b) and 3(c), respectively.

Figure 3(a) shows clearly that northern France and Balkan populations group together very closely and that, overall, three highly differentiated regions exist. Central Spain and southern Italy are equally differentiated from the French–Balkan clade (although the Italian and the French–Balkan populations belong to the same subspecies, *C. p. parallelus*). Again it is apparent that central Spain is very different from either Pyrenean or central and Alpine France, just as it was from northern France. Figure 3(b) shows the relationship between northern Spain and central Spain and the relationship between Pyrenean France and Pyrenean Spain. The latter, despite being from different subspecies, group together whereas the former, from the same subspecies *C. p. erythropus*, are significantly differentiated. The grouping of northern and central Italian populations with the French and Balkan regions (Fig. 3c) is an unexpected result and is discussed later in terms of the rates of postglacial range expansion from the different refugial locations.

Age of separation of C. p. parallelus *and* C. p. erythropus

Table 2 shows the distance values between all the mitotypes found in France and Spain ($N = 96$). The mean percentage sequence divergence between the two regions is $1.072 \pm 0.0232\%$, but is slightly higher when the mitotypes endemic to Pyrenean Spain are excluded ($1.102 \pm 0.0211\%$, $N = 48$). Taking a rate of

mtDNA sequence evolution of 2% per million years (see Discussion) a time of divergence for these two sequences of 500 000–550 000 years is indicated.

Discussion

The discussion will consider three major aspects of the results presented in the previous section. First, the location of European glacial refugia and the possible routes of postglacial expansion into northern Europe. Second, the extent of genetic divergence between the refugial populations and the barriers to gene flow encountered today. Third, the age of separation between subspecies will be examined and interpreted in terms of the effects of different climate cycles. Prior to discussing these, we note that it should have been possible to construct a parsimonious network of the relatively closely related mitotypes that were identified in this study. Where this appears to be confounded, mutational hotspots and the limited sample size relative to the size of the range sampled may be important factors. We argue that the lack of resolution in the parsimony tree results from the fact that only nine of the 16 variable sites that define the 31 unique haplotypes are informative. Because the topology of a tree and the bifurcation of branches is essentially dependent on there being a sufficient number of synapomorphies, it is not surprising that there is so little resolution in the parsimony analysis. A larger sample would have identified 'link' mitotypes that could have increased the proportion of informative sites and hence tree resolution. These concerns are largely ameliorated, however, by the use of frequency-based statistics such as K_{ST}.

Location of refugia and expansion routes

The lack of any significant level of differentiation between northern France and the Balkans indicates

Table 2 Matrix of Jukes–Cantor corrected distances between mitotypes of *Chorthippus parallelus* found in France (columns) and those found in Spain (rows)

	M1	M3	M4	M5	M6	M17	M18	M20	M27	M29	M30	M31
M1	0.0000	0.0067	0.0033	0.0067	0.0101	0.0101	0.0169	0.0101	0.0067	0.0101	0.0135	0.0033
M6	0.0101	0.0169	0.0135	0.0033	0.0000	0.0135	0.0135	0.0067	0.0101	0.0101	0.0135	0.0135
M8	0.0067	0.0135	0.0101	0.0135	0.0169	0.0101	0.0135	0.0101	0.0135	0.0101	0.0135	0.0033
M9	0.0033	0.0101	0.0067	0.0101	0.0135	0.0135	0.0169	0.0135	0.0101	0.0135	0.0169	0.0067
M10	0.0067	0.0135	0.0101	0.0135	0.0101	0.0101	0.0101	0.0033	0.0067	0.0067	0.0101	0.0033
M13	0.0067	0.0135	0.0101	0.0067	0.0033	0.0101	0.0101	0.0033	0.0067	0.0067	0.0101	0.0101
M15	0.0067	0.0135	0.0101	0.0135	0.0101	0.0169	0.0101	0.0101	0.0067	0.0135	0.0169	0.0101
M19	0.0135	0.0230	0.0169	0.0135	0.0101	0.0101	0.0169	0.0101	0.0135	0.0135	0.0169	0.0101

a shared glacial origin for these populations. Because we know that during the last glacial maximum *C. parallelus* would only have been able to survive at low latitudes, we can conclude that the ancestral populations were in the Balkan region and must have subsequently expanded their ranges northwards.

For *C. parallelus*, a flightless grasshopper normally not observed above 2000 m, high mountain ranges may be an effective barrier to dispersal. Thus, both the Alps and the Pyrenees could have prevented northward expansion from southern refugial populations. In contrast, the Carpathian mountains do not form a barrier that runs east–west, and thus may have permitted a low-altitude northward expansion of Balkan *C. parallelus*. These mountain ranges cannot be viewed as complete barriers to gene flow as *C. parallelus* may pass more easily through lower-level mountain cols or circumvent the ranges through lower-altitude coastal regions. It is the rate at which expansion was possible across the mountain ranges that would have been the crucial factor in the historical biogeography of this grasshopper.

As the climate ameliorated at the end of the last glaciation, the temperature and vegetation would have changed sufficiently to open new territory to *C. parallelus* further north in Europe. The exposure of suitable habitat would have depended both on latitude and altitude, and the high mountains of Europe – the Alps, Pyrenees and Carpathians – would have retained their ice cover much longer than land at lower altitude even if it were much further north. Thus an expansion from refugia in southern Italy or Spain may well have been stopped at the mountains waiting for suitable conditions, rather than purely by an inability to cross such high terrain. An expansion from the Balkans may, however, have proceeded north much more rapidly — a pattern of colonization which would account for the genetic similarity observed between populations in northern Europe and the Balkans. Once the new habitat was colonized, later expansions from other refugia are unlikely to have been influential with the low rate of immigration from other refugia having a relatively small effect on the composition of the established populations.

Population subdivision

The extent of differentiation within and between the three refugial regions in southern Europe, as well as the postglacial populations of northern Europe, are important in understanding the historical processes of subdivision. Table 1 shows no significant differen-

tiation between northern and central Spain. Analysis of a nuclear DNA marker, however, has revealed substantial subdivision between populations from these regions. Cooper & Hewitt (1993) suggested that this may have resulted from independent expansions from a common southern refugium with lineage sorting accentuating the differences. The mtDNA data presented here do not contradict such a hypothesis of independent expansions, and support Cooper & Hewitt's assertion that the causes of the detected differentiation are associated with lineage sorting rather than multiple, independent Spanish refugia.

The populations from the Spanish and French sides of the Pyrenees as well as those from central and Alpine France cluster together when compared to the northern and central Spanish populations (Fig. 3b). Such similarities in mitotype composition between different subspecies, at locations displaced from the hybrid zone, may initially suggest introgression. The two subspecies, however, have been shown to be very distinct in many characters associated with reproduction. These include courtship song (Butlin & Hewitt, 1985), mate preference (Butlin & Ritchie, 1991) and cuticular hydrocarbons (Neems & Butlin, 1994). Hybrids have also been shown to experience F_1 hybrid dysgenesis (Hewitt *et al.*, 1987). Thus it is unlikely that neutral introgression could have led to the presence of French/Balkan mitotypes in Pyrenean Spain. Could a selective advantage exist for *C. p. parallelus* mitotypes which might lead to their spread into Spain?

There are sufficient studies of nuclear encoded traits to indicate that neither subspecies has a selective advantage over the other in the Pyrenees. The clines for many traits change in the high mountains and the hybrid zone itself follows the mountain ridge reasonably closely. This midway position is exactly as expected assuming that the two subspecies met as the summer-ice left the mountains during the last glacial amelioration, and that neither has significantly encroached on the other since. This does not, however, rule out selective advantages of *C. p. parallelus* mtDNA in *C. p. erythropus* cytoplasm, as this would have no effect on the nuclear encoded traits usually studied, but still allow mtDNA introgression. It is difficult to envisage this sort of situation occurring in practice as phenotypic body traits would be expected to coevolve for maximum efficiency. It seems unlikely that a set of mitotypes which do not cause a general increase in fitness in their own subspecies should, when transplanted into a different subspecies, be subject to a selective advantage. Indeed the opposite relationship has been

documented for several species (reviewed by Moritz *et al.*, 1987).

Sampling bias is another possible explanation, but this is unlikely because the sample sizes from the Pyrenean populations are amongst the largest in this study and are likely at the very least to reveal broad associations. It should be borne in mind that only one population (Escarilla) has been extensively sampled from Pyrenean Spain. In this well-studied region south of the Col du Portalet clines for many characters have been observed to differ in width across the hybrid zone (Butlin *et al.*, 1991; Ferris *et al.*, 1993). If the mtDNA clines were several km wider than for other characters then French mitotypes would extend into Escarilla. Cline width has been predicted to be broader for characters for which there is little heterotic selection (Barton & Hewitt, 1989). Thus, because these mitochondrial polymorphisms are likely to be neutral (Lunt, 1994), this cline might be expected to be relatively wide. It has also been reported that the cline centres of some subspecies-specific characters, such as the X chromosome nuclear organizer region (NOR) and cuticular hydrocarbon composition, are displaced away from the centres of many other characters (Hewitt, 1993; Neems & Butlin, 1994). It is possible that the mtDNA difference is displaced much further beyond Escarilla. Such noncoincidence of mtDNA has been reported in a few other hybrid zones (Hewitt, 1993) and perhaps the most likely hypothesis to explain this involves genome reassortment during colonization.

Analysis of samples from locations further into Spain and along the Pyrenees would allow one to determine whether the sample at Escarilla is a local introgression, or whether *C. p. erythropus* in the Pyrenees has more generally a *C. p. parallelus* mtDNA type, in contrast to the distinct *C. p. erythropus* mtDNA we know exists in central and northern Spain.

The existence of a Spanish ice-age refugium has been a central tenet of the interpretation of much work relating *C. p. parallelus* and *C. p. erythropus* (e.g. Hewitt, 1993). The origin of *C. parallelus* in Italy, however, has been less studied. Many parallels between Spain and Italy can be drawn which would point to a refugium in both. Although not much has been recorded about *C. parallelus* differentiation across the Alps, the work presented here indicates that the differences between southern Italian and French populations are very substantial, and at least equal to those found in central Spain. Similar results indicating a southern Italian refugium have been obtained using a nuclear DNA marker (Cooper *et*

al., 1995) and some preliminary results indicate hybrid testes dysfunction in matings between southern Italian and French *C. parallelus* (N. Flanagan and G. M. Hewitt, unpubl. obs.).

Table 1 and Fig. 3(c) show that whereas southern Italy is very distinct from the other regions, neither central nor northern Italy are significantly differentiated from the French or Balkan populations. This result is unexpected; every indication from the geography of the peninsulas and from how the *C. p. erythropus* expansion progressed in Spain would suggest that *C. parallelus* from a southern Italian refugium should expand its range as far as the Alps. Yet the conclusions from this data set are that the refugial populations are still located in the south and that northern and central Italian *C. parallelus* share a common ancestor with (and hence probably resulted in an expansion from) the Balkans. Again, comparison with the nuclear DNA study indicates differences in the associations of these populations. The nuclear DNA study places northern Italy somewhat closer to southern Italy than to the Balkans (Cooper *et al.*, 1995).

Similar considerations are appropriate here as when considering the relationship between Pyrenean Spain and Pyrenean France. However, two further perspectives may be important in this case. First, there is no hybrid zone yet described across the Alps, and thus there is no wealth of evidence with which to confirm the differentiation between the different regions. Although the processes restricting immigration to already occupied territory are the same, there is no definite picture of the situation with respect to the nuclear genome as a whole. Studies have only recently begun to investigate the possibility of song differentiation and 'hybrid' sterility across the Alps. Secondly, it should be noted that northern Italy is much closer to the Balkan refugium than is Pyrenean Spain. If, for some reason, the expansion from the Balkans had been much more rapid than that from southern Italy, then Balkan *C. parallelus* may have reached the Alps before the new territories had been colonized. This colonization would have been aided by the relatively lower altitude of the Dinaric Alps (<1000 m) between Ljubljana and Trieste on the Italian–Slovenian border. If such a rapid colonization occurred then Balkan *C. parallelus* would have met Calabrian *C. parallelus* at some point along the peninsula. Such a scenario would require a relatively slow expansion from southern Italy, the reasons for which are not clear, though either very small Italian refugial population size or an earlier warming of the Balkans could be factors.

Age of divergence of C. p. parallelus *and* C. p. erythropus

The average sequence divergence between French and Spanish mitotypes was found to be in the range 1–1.1%. This figure would correlate to an age of divergence of around 500 000 years given a rate of sequence evolution of 2% per million years. The figure of 2% per million years for mtDNA divergence comes from Brown *et al.* (1979) and reflects the initial slope of the curve of substitutions per base pair vs. divergence time for mammalian species. This rate may contain large errors arising from (i) differences between a mammalian (largely primate) rate and that operating on Orthoptera and (ii) differences in single-gene vs. average mitochondrial rate, as sampled by RFLPs, from which the estimate was obtained. Despite these cautionary points, however, molecular 'clocks' should not be dismissed, as once the errors have been acknowledged the information from rate estimates may still be informative.

Our knowledge about the glacial cycles which have shaped the historical biogeography of many species indicates that *C. parallelus* may have expanded from southern refugia many times before being pushed back by the advancing ice. Glacial maxima had a periodicity of ≈100 000 years, of which maybe around 20 000 years were temperate enough to allow *C. parallelus* its current range. There is evidence (Webb & Bartlein, 1992) that the glaciations 500 000–850 000 years before present (ybp) were especially severe, with much cooler interglacial periods. The refugia in the southern tips of Iberia, Italy and the Balkans, which persisted through the last glaciation, may not all have supported a suitable habitat for *C. parallelus* during this period. Thus, in the subsequent interglacials, any southern population which did survive (possibly in Spain, which is the furthest south, or Turkey, or perhaps the Caucusus) would have been able to expand to fill the whole of the current European range of *C. parallelus*, rather than being limited by meeting other populations as more recent expansions seem to have been. This would have brought about a relatively homogeneous population, which would have been the starting point for the patterns of differentiation described above.

If the above speculations concerning the glacial history of Europe are correct, the date of differentiation estimated above at 550 000 ybp for *C. p. parallelus* and *C. p. erythropus* would reflect the first opportunity for the two subspecies to diverge following a potentially homogenizing glacial period. Hillis & Moritz (1990) indicated that the margin of error for such calibrations can be as much as 30%. Such an error would give an age of divergence of 363 000–731 000 ybp for *C. p. parallelus* and *C. p. erythropus*. It is interesting to note that these values, even including large errors, indicate a period in the late Pleistocene as the likely point of divergence of these two subspecies, a time point that matches well to the available climatic data and to what we strongly suspect will be the important factors in the divergence of *C. p. parallelus* and *C. p. erythropus.*

To summarize, the analysis of subdivision between European regions indicates that glacial refugial populations existed in each of the three southern peninsulas (Spain, Italy and the Balkans). Analysis of the composition of northern European populations further suggests that they share a common origin with the present-day Balkan populations. This provides strong evidence that it was the Balkan refugial population, and not those in Spain or Italy, that expanded to fill northern Europe. Although only two subspecies are currently recognized (*C. p. erythropus* in Spain and *C. p. parallelus* in France) the genetic distances presented here are at least as great to southern Italian populations, indicating the potential for a third western European subspecies. This possibility is currently being investigated using a range of markers (N. Flanagan & G. M. Hewitt, pers. comm.).

References

BARTON, N. H. AND HEWITT, G. M. 1989. Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

BROWN, W. M., GEORGE, M. AND WILSON, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 1967–1971.

BUTLIN, R. K. AND HEWITT, G. M. 1985. A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthip-*

pus parallelus erythropus (Orthoptera: Acrididae): behavioural characters. *Biol. J. Linn. Soc.*, **26**, 287–299.

BUTLIN, R. K. AND RITCHIE, M. G. 1991. Variation in female mate preference across a grasshopper hybrid zone. *J. Evol. Biol.*, **4**, 227–240.

BUTLIN, R. K., RITCHIE, M. G. AND HEWITT, G. M. 1991. Comparisons among morphological characters between localities in the *Chorthippus parallelus* hybrid zone (Orthoptera: Acrididae). *Phil. Trans. R. Soc. B*, **334**, 297–308.

COOPER, S. J. B. AND HEWITT, G. M. 1993. Nuclear DNA sequence divergence between parapatric subspecies of the grasshopper *Chorthippus parallelus*. *Insect Mol. Biol.*, **2**, 1–10.

COOPER, S. J. B., IBRAHIM, K. M. AND HEWITT, G. M. 1995. Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. *Mol. Ecol.*, **4**, 49–60.

FELSENSTEIN, J. 1991. PHYLIP (*Phylogeny Inference Package*). Version 3.4. University of Washington, Seattle, WA.

FERRIS, C., RUBIO, J. M., SERRANO, L., GOSALVEZ, J. AND HEWITT, G. M. 1993. One way introgression of a subspecific sex chromosome marker in a hybrid zone. *Heredity*, **71**, 119–129.

HEWITT, G. M. 1993. After the ice: *Parallelus* meets *Erythropus* in the Pyrenees. In: Harrison, R. G. (ed.) *Hybrid Zones and the Evolutionary Process*, pp. 140–164. Oxford University Press, Oxford.

HEWITT, G. M., BUTLIN, R. K. AND EAST, T. M. 1987. Testicular dysfunction in hybrids between parapatric subspecies of the grasshopper *Chorthippus parallelus*. *Biol. J. Linn. Soc.*, **31**, 25–34.

HILLIS, D. M. AND MORITZ, C. 1990. An overview of applications of molecular systematics. In: Hillis, D. M. and Moritz, C. (eds) *Molecular Systematics*, pp. 502–515. Sinauer Associates, Sunderland, MA.

HUDSON, R. R., BOOS, D. D. AND KAPLAN, N. L. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.*, **9**, 138–151.

HUNTLEY, B. AND BIRKS, H. J. B. 1983. *An Atlas of Past and Present Pollen Maps for Europe: 0–13 000 Years Ago*. Cambridge University Press, Cambridge.

LUNT, D. H. 1994 *mtDNA Differentiation Across Europe in the Meadow Grasshopper* Chorthippus parallelus *(Orthoptera: Acrididae)*. Ph.D. Thesis, University of East Anglia, U.K.

MORITZ, C., DOWLING, T. E. AND BROWN, W. M. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.*, **18**, 269–292.

NEEMS, R. M. AND BUTLIN, R. K. 1994. Variation in cuticular hydrocarbons across a hybrid zone in the grasshopper *Chorthippus parallelus*. *Phil. Trans. R. Soc. B*, **257**, 135–140.

SAITOU, N. AND NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.

STEWART, C.-B. 1993. The powers and pitfalls of parsimony. *Nature*, **361**, 603–607.

SWOFFORD, D. L. 1993. PAUP: *Phylogenetic Analysis using Parsimony*. Version 3.1.1. Smithsonian Institute, Washington, DC.

SZYMURA, J. M., LUNT, D. H. AND HEWITT, G. M. 1996. The sequence and structure of the meadow grasshopper (*Chorthippus parallelus*) mitochondrial srRNA, ND1, COI, COII, ATPase8 and 9 tRNA genes. *Insect Mol. Biol.*, **5**, 127–139.

VIRDEE, S. R. AND HEWITT, G. M. 1990. Ecological components of a hybrid zone in the grasshopper *Chorthippus parallelus* (Zetterstedt) (Orthoptera: Acrididae). *Bol. San. Veg.*, **20**, 299–309.

WEBB, T. AND BARTLEIN, P. J. 1992. Global changes during the last 3 million years: climatic controls and biotic responses. *Ann. Rev. Ecol. Syst.*, **23**, 141–173.