Heredity (2006) 97, 398–408

www.nature.com/hdy

An unusual phylogeography in the bushcricket *Ephippiger ephippiger* from Southern France

LJ Spooner and MG Ritchie
School of Biology, Environmental and Evolutionary Biology, Dyers Brae House, University of St Andrews, St Andrews, Fife KY16 9TH, UK

Pleistocene glaciations have played a major role in species divergence. The bushcricket *Ephippiger ephippiger* shows unusual patterns of intraspecific variation in multiple traits across Southern Europe. This is centred in Southern France, and evidence implies that it results from secondary contact after differentiation in Pleistocene refugia. However, the possible time scales involved, locations of the refugia and patterns of expansion remain obscure. This study sequenced the *COII* (507 BP) and *cyt b* (428 BP) mitochondrial genes to examine the intraspecific phylogeography of Western European samples of *E. ephippiger*. A minimum evolution tree revealed little resolution between described subspecies of *E. ephippiger*. Strikingly, populations from the Pyrenees and Mediterranean coastal region contained a complex genetic structure corresponding to major river valleys, independent of the traditional taxonomy. Samples of the subspecies *E. e. vitium* formed a distinct clade, perhaps supporting their taxonomic status. However, other forms (cruciger and cunii) were not genetically distinct, which is surprising given differences in their morphology and behaviour. The extent of the genetic divergence between Pyreneen valleys is unexpectedly deep, with average Tamura-Nei distances of around 14% (net distances of 11%) separating the main clades of coding *COII* sequences. *Cyt b* showed a similar pattern, but was confounded by some non-coding probable pseudogenes. If a conventional insect molecular clock is applied, these cryptic clades must predate the Pleistocene, and hypotheses for their history are discussed. However, mtDNA divergence in *Ephippiger* is not evolving in a clock-like manner, because a likelihood ratio test rejects clock assumptions for the *COII* sequences.
Heredity (2006) **97,** 398–408. doi:10.1038/sj.hdy.6800884; published online 6 September 2006

Keywords: mtDNA; geographic variation; phylogeography; refugia; Tettigoniidae; *Ephippiger*

Introduction

Understanding the origin of intraspecific genetic variation is an essential step in the study of speciation. Recent (Pleistocene) ice ages have been a major factor shaping patterns of genetic variation within European species (Hewitt, 1996, 1999, 2000, 2004; Avise *et al*, 1998; Taberlet *et al*, 1998), and glaciations are thought to have played a substantial role in patterning population divergence and speciation (Hewitt, 1996, 2004). Numerous taxa show extensive intraspecific genetic variation, and many species consist of a mosaic of subspecies connected by areas of hybridisation (Hewitt, 1988, 1989).

Oscillations in the Earth's orbit throughout the Quaternary period (1.6 Myr ago to the present) have led to repeated cycles of ice sheet advance and retreat that have prompted major range changes for most species in temperate climates (Hewitt, 1996; Taberlet *et al*, 1998). This has created complex patterns of genetic variation in extant species, with genetic drift and selection in isolated refugia causing population divergence during cycles of allopatry (Hewitt, 1996). There is evidence that subsequent postglacial expansion progressed via successive population bottlenecks, which can result in considerable genome reorganisation and divergence among populations (Armbruster *et al*, 1998; Hewitt, 2000; Clegg *et al*, 2002; but see Tregenza *et al*, 2002). Secondary hybrid zones between populations may maintain this divergence through being partial barriers to gene flow, and mountain ranges often act as major barriers to migration. These processes can potentially lead to speciation over one or a few glaciations (Hewitt, 2000). MtDNA analyses typically find genetic distances of a few per cent between sister species and subspecies, across mountain ranges or other barriers, conforming to Pleistocene or Pliocene divergence (Avise *et al*, 1998; Lovette, 2005).

The saddlebacked bushcricket, *Ephippiger ephippiger*, shows unusual intraspecific variation in behaviour, morphology and genetic markers across Southern Europe. This variation is centred in Southern France (the Eastern Pyrenees and Mediterranean coast), and recent evidence suggests that it results from secondary contact after differentiation in Pleistocene refugia (Kidd and Ritchie, 2000; Ritchie *et al*, 2001). However, the time scale, locations of refugia and patterns of expansion remain obscure. The extensive variation within the species has led to a series of taxonomic revisions, and the current taxonomy and evolutionary history of *E. ephippiger* is unclear. Three forms of *E. ephippiger* from Southern France were originally described as separate species: *E. vitium* (Serville 1831) in Western Europe, *E. cunii* (Bolivar 1837) in Catalonia (North-eastern Spain

Correspondence: MG Ritchie, School of Biology, Environmental and Evolutionary Biology, Dyers Brae House, University of St Andrews, St Andrews, Fife KY16 9TH, UK. E-mail: mgr@st-andrews.ac.uk
Received 3 March 2006; accepted 12 July 2006; published online 6 September 2006

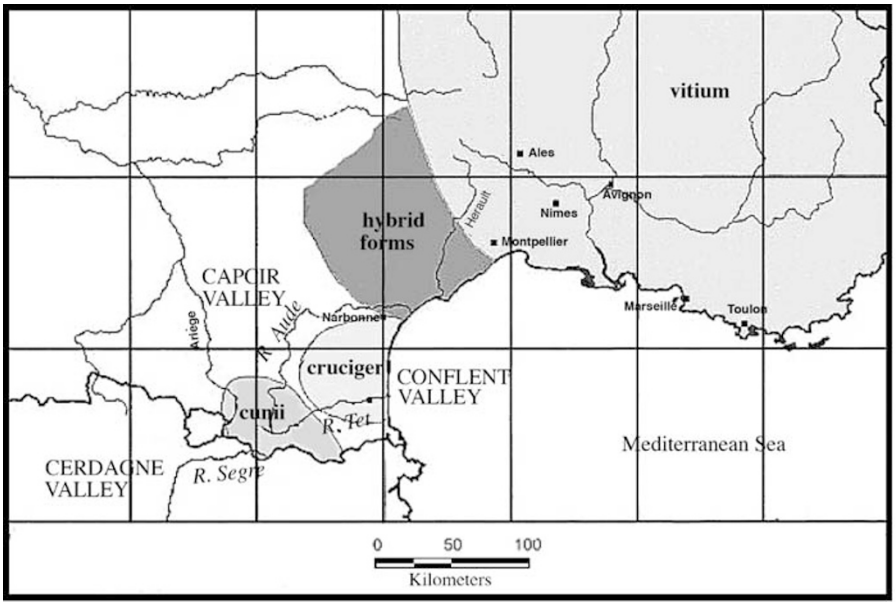

Figure 1 Approximate geographic distributions of the subspecies of *E. ephippiger* (vitium, cruciger and cunii) in Southern France. The three main rivers and river valleys in the eastern Pyrenees are also indicated.

and South-west France) and *E. cruciger* (Fiebeg 1853) in the Languedoc (Mediterranean coast) (Harz, 1969). Figure 1 shows their distributions. Subsequent studies found that none of these forms were reproductively isolated Duijm *et al*, 1983; Ritchie, 2000), and combined them (and additional subspecies) into one superspecies, *E. e. diurnus* (Dufour 1841). The cunii form was given subspecies status (Duijm and Oudman, 1983; Duijm *et al*, 1983; Hartley and Warne, 1984). Later studies questioned whether cunii was a valid subspecies: allozymes showed very low genetic divergence between the three forms, and any potential diagnostic traits in these studies were linked by non-coincident, extremely wide clines (Oudman *et al*, 1989, 1990; Duijm, 1990).

However, a mtDNA RFLP study supported the existence of *E. e. vitium* as a valid subspecies (Ritchie *et al*, 2001). A previous RAPD-based phylogeny also showed a genetic subdivision within the species, although this analysis was geographically limited (Ritchie *et al*, 1997). Partial Mantel tests and other analyses implied that most intraspecific variation originated through isolation in refugia, although environmental factors were found to influence body size (Kidd and Ritchie, 2000) and the pattern of mtDNA variation (Ritchie *et al*, 2001). *E. e. vitium* was thought likely to have migrated west across Europe from a refugium in the east, for example, the Balkans or Caucasus, whereas the other forms (cunii and cruciger) showed some geographic structuring and were thought to have possibly arisen from multiple Iberian refugia (Ritchie *et al*, 2001). However, the phylogenies produced by these studies were not well resolved, and the markers were not fully characterised.

This study uses mtDNA sequencing to construct an intraspecific phylogeny for *E. ephippiger* from more extensive samples of the cricket from Southern France. DNA sequencing is expected to be more phylogenetically informative than the RAPD or RFLP analysis used previously, and allows us to consider the potential time scale involved. A well-resolved phylogeny for the species is necessary to interpret the intraspecific behavioural variation observed in previous studies of sexual selection and communication in *E. ephippiger* (Busnel, 1963; Duijm, 1990; Ritchie, 1996, 2000; Greenfield *et al*, 2004; Berg and Greenfield, 2005; Spooner, 2005). We found an unusual distribution and genetic divergence of mtDNA haplotypes among Southern French populations, which is largely independent of their usual taxonomy and morphological and behavioural variation.

Materials and methods

Hind leg tissue samples were taken from 132 *E. ephipigger* in 28 field populations across Southern France between 2001 and 2003 (Figure 2). Samples were also collected from *E. provincialis*, *E. terrestris* and *Steropleurus catalaunicus* for outgroups. Each sample was initially identified to subspecies in the field based on its morphology and calling behaviour (a few were intermediate and classified as hybrids). Our samples incorporate all the subspecies known to occur in this region. *E. e. moralesagacinoi*, described from the Western Pyrenees, was suppressed by Oudman *et al* (1990), who could not distinguish it from *E. e. cunii*. Other subspecies (*E. e. vicheti* and *E. e. ephippiger*) are from a disjunct eastern distribution. The vitium form is found in South-eastern and central France, and is characterised by monosyllabic song, small size and green colour. Cunii is polysyllabic (3–5 syllables per chirp), is almost black in colour and is found in Catalonia (the eastern Pyrenees). Our cunii samples are mainly from three Pyreneen valleys: *Capcir*, formed by the river Aude (which continues north into Ariège then the Languedoc), *Conflent*, formed by the river Têt, and *Cerdagne*, formed by the river Segre. Cruciger is from the *Languedoc* region (along the Mediterranean coast) and is much larger and lighter in colour than cunii and vitium. It has a polysyllabic calling song (4–8 syllables per chirp) and

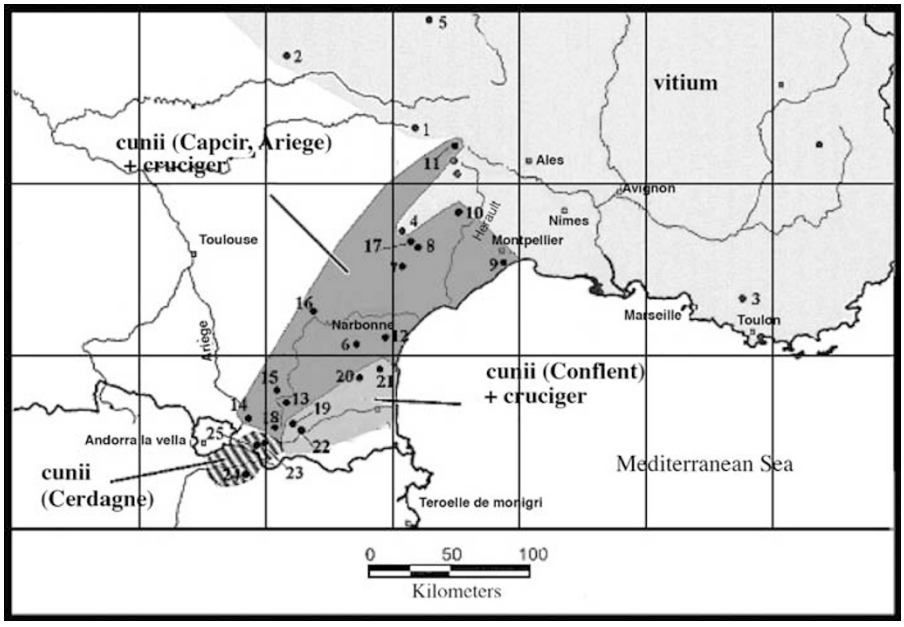

Figure 2 The geographic distribution of *E. ephippiger* in Southern France from each clade of the *COII* intraspecific phylogeny. Each shaded region indicates the distribution of a different clade sampled here (the species has a wider distribution, for example, see Duijm and Oudman, 1983). The black circles show the locations of each collecting site: (1) Severac (44.48°N, 3.19°E); (2) Viellemard (45.40°N, 2.52°E); (3) Plan d'Aups (43.33°N, 6.02°E); (4) Lodève (43.74°N, 3.36°E); (5) St Mary (45.22°N, 3.19°E); (6) Thezans (43.12°N, 2.78°E); (7) Faugeres (43.56°N, 3.17°E); (8) Clermont l'Herault (43.63°N, 3.43°E); (9) Montpellier (43.60°N, 3.71°E); (10) St Jean (43.83°N, 3.62°E); (11) LMK005 (44.11°N, 3.49°E); (12) Prats de C'Est (43.12°N, 2.95°E); (13) Escoloubre (42.73°N, 2.11°E); (14) Merens (42.66°N, 1.84°E); (15) Aunat (42.76°N, 2.11°E); (16) Carcassonne (43.18°N, 2.52°E); (17) Salagou (43.65°N, 3.43°E); (18) Llagonne (42.52°N, 2.11°E); (19) Railleu (42.53°N, 2.13°E); (20) Feuilla (42.93°N, 2.90°E); (21) Fitou (42.90°N, 2.96°E); (22) Sauto (42.51°N, 2.15°E); (23) Via (42.48°N, 2.06°E); (24) Viella (42.40°N, 1.81°E) and (25) Estavar (42.46°N, 2.01°E).

has a melanised cross on the pronotum (Hartley and Warne, 1984; Oudman *et al*, 1990). Figure 1 shows the regions and distributions.

Two mitochondrial sequences, cytochrome oxidase II (*COII*, 530 bp) and cytochrome *b* (*cyt b*, 450 bp), were sequenced for every individual. The primers used for *COII* were: forward, C2-J-3279 (5′-GGACAAACAATT GAGTTAATTGGAAC-3′) and reverse, TD-N-3862 (5′-TTTAGATTGACATTCTAATGTTAT-3′), developed from universal insect primers (Simon *et al*, 1994), which amplify most of *COII* and the adjoining tRNA (lysine) genes (figures are in the appendix of Simon *et al*, 1994). PCRs for *COII* were carried out in 25 μl volumes containing 25 pM each of forward and reverse primers, 1.5 mM $MgCl_2$, 1 × Bio*Taq* PCR NH_4^+ buffer, 0.04 mM of each dNTP, 0.5 μl DNA template (approximately 20 ng) and 2 μ Bio*Taq* DNA polymerase. The PCR reactions were 2 min at 95°C, followed by 40 cycles at 94°C for 30 s, 48°C for 45 s and 72°C for 1 min, and a final step at 72°C for 10 s.

The *cyt b* primers used were: forward (5′-CTACCAT GAGGACAAATATC-3′) and reverse (5′-ATTACACCTC CTAATTTATTAGGAAT-3′). These were designed for *E. ephippiger* following sequencing using primers for the *cyt b* gene from Simon *et al* (1994). PCR conditions for *cyt b* were the same as those for *COII*, except that 30 pM of each of the forward and reverse primers were used, along with 0.08 mM of each dNTP and 0.5 μ Bio*Taq*. DNA amplification conditions were as above, except the annealing stage was at 45°C for 45 s.

PCR products were purified, then sequenced on an ABI 3730 capillary DNA sequencer. Sequences were checked using Chromas 1.45 and MEGA 2.1 (Kumar *et al*, 2001), and aligned with ClustalX (Thompson *et al*, 1997) using a gap penalty of 15% and a gap extension penalty of 0.20. All sequences were deposited in Genbank under the accession numbers DQ106622-DQ106738 (*COII*) and DQ106739-DQ106849 (*cyt b*).

The optimum models for nucleotide substitution analyses were found separately for the *COII* and *cyt b* data sets using Modeltest 3.06 (Posada and Crandall, 1998). Separate intraspecific phylogenies were estimated for the *COII* and *cyt b* sequences using the minimum evolution algorithm in PAUP* 4.0 (Swofford, 1998; Nei and Kumar, 2000). A heuristic search was used for each analysis, with MaxTrees set to 1000. One thousand bootstrap replicates (resampling with replacement) were used for each data set to obtain a 50% majority-rule consensus tree with compatible groups. To test for substitution rate variation, likelihood scores were calculated for the shortest minimum evolution tree obtained with or without a clock assumption and the scores compared using a likelihood ratio test.

In order to further examine the relationships between haplotypes, a haplotype network analysis was carried out for the *COII* data. The statistical parsimony approach (Templeton *et al*, 1992) implemented in TCS (Clement *et al*, 2000) failed to resolve the relationships between clusters, producing unlinked clusters and several independent haplotypes, even with the connection limit dropped to 90%. The median-joining network approach of Bandelt *et al* (1999), which creates median vectors of consensus sequences to aid network spanning, produced

a linked network uniting all haplotypes. All outgroup sequences were removed before constructing any network. Simulations suggest that median-joining is as reliable as maximum parsimony in resolving networks (Cassens *et al*, 2005). Genetic diversity and distance measures were made with MEGA (Kumar *et al*, 2001).

Results

COII

All the *COII* samples obtained, from ingroups and outgroups, were aligned without indels and translated into proteins. Modeltest (Posada and Crandall, 1998) found the best-fit model was Tamura-Nei + I + G with unequal base frequencies, unequal rates of change according to substitution type and among-site substitution rate variation. The proportion of invariant sites (I) was 0.3245 and the distribution of rates at variable sites (gamma or G) was 0.7422. The base frequencies used were: A = 0.3093, C = 0.2264, G = 0.1379, T = 0.3264, and a substitution rate matrix was also calculated ([A-G] = 5.4341, [C-T] = 5.9749, all other values = 1.0000).

The resulting phylogeny (Figure 3) revealed substantial geographic structuring, with four main clades. The haplotype network (Figure 4) confirmed the four distinct genetic clusters, with 'loops' only present within clusters. The four groups tended to follow geography rather than traditional phylogenies. In Figure 3, Clade 1 contained samples from Provence, the Massif Central and the Cévennes (South-eastern and central France, all vitium); Clade 2 mainly contained samples from the course of the river Aude (Capcir and the adjoining Ariège valley and Languedoc). Clade 3 contained samples from Conflent and from the Languedoc; and Clade 4 contained samples from Cerdagne. These four main clades were seen in each of the analyses. However, their relationships to each other were poorly resolved. Two samples were on long branches.

Individuals identified in the field as 'cunii' were present in three of the four clades. The three cunii-containing clades correspond to the three main river valleys in the eastern Pyrenees (Capcir, Conflent and Cerdagne), with crickets from only an occasional population being found in more than one clade. These exceptions are crickets from Llagonne, which lies at the junction of all three valleys. Strikingly, the bushcrickets from each valley were as distinct from each other as vitium individuals were from other forms of *E. ephippiger*, and the 'cunii' bushcrickets from the Capcir valley appeared to be more closely related to vitium than to the cunii populations in the other two Pyrenean valleys. Sequences from individuals identified as cruciger were present in two clades, and vitium formed a distinct clade.

The mean Tamura-Nei genetic distance among the four main lineages was almost 14%. The net distances were lower, around 11%. Table 1 gives genetic diversity measures and distances between each clade (defined in Figure 3). The likelihood (-lnL) score for our shortest ME tree with a clock assumption was 4997.37, without a clock assumption of 4768.93. A likelihood ratio test rejects the clock assumption ($\chi^2 = 456.88$, df = 106, $P < 0.0001$).

Cyt b

Aligning the *cyt b* sequences produced indels and non-coding sequences, which are probably nuclear copies (see Discussion). The phylogenetic tree was less geographically coherent than that for *COII*. The best-fit model found by Modeltest was Tamura-Nei + G, with a gamma value of 0.6771. The base frequencies used were: A = 0.2603, C = 0.2468, G = 0.1637 and T = 0.3292. A substitution rate matrix was also calculated ([A-G] = 2.5964, [C-T] = 2.6607, all other values = 1.0000). Although the resultant minimum evolution tree (Figure 5) showed the same major genetic divisions as the *COII* phylogeny, many of the clades had weak bootstrap support. In particular, the tree could not be rooted to give a monophyletic clade for the ingroups. Fourteen *E. ephippiger* individuals (from Languedoc and Capcir) were grouped with two of the outgroup species. Ten of the samples in this group possessed a frameshift indel consisting of a 5 BP deletion and a 1 BP insertion. The other four samples that clustered with the outgroup species all had a single base pair deletion between 100 and 110 BP, and long branch lengths that indicated a large number of base pair changes compared to the other *cyt b* sequences.

Discussion

The *cyt b* results reported here are probably compromised owing to the presence of nuclear pseudogene copies (Numts) of the *cyt b* gene (Lopez *et al*, 1994). Numts are non-functioning copies of mitochondrial genes that have been transferred to the nucleus (Zhang and Hewitt, 1996; Mirol *et al*, 2000) and seem to be common in orthoptera (Bensasson *et al*, 2001). They usually represent an ancestral form of the gene (Zhang and Hewitt, 1996) that has subsequently degenerated, so often cluster with an outgroup species. Indels are also common in pseudogenes (Mirol *et al*, 2000). In fact, it appears that two nuclear copies of *cyt b* were amplified (one of the numts includes an indel, and the other does not). All of the *COII* sequences were potentially coding in that they translated into protein sequences using the insect mtDNA code. Given the extent of sequence divergence among these samples, for them to remain coding strongly implies that these are functional mitochondrial copies of this gene. All subsequent discussion of the *E. ephippiger* phylogeny has therefore been limited to the *COII* results.

The patterns found are unusual given the typical phylogeographic history of species from this region, and raise novel questions about the colonisation history or persistence of organisms in the Pyrenees during the Pleistocene. We have found that *E. ephippiger* from Southern France broadly divides into four geographically coherent deep mtDNA clades ('phylogroups', Avise *et al*, 1998) which do not follow current taxonomies (Duijm *et al*, 1983; Hartley and Warne, 1984; Ritchie *et al*, 2001). The vitium form, thought by Ritchie *et al* (2001) to originate from an eastern refugium, is probably distinct (and is also usually behaviourally different, although the monosyllabic song does not precisely match our 'Clade 1', Duijm, 1990). Whether vitium is a valid subspecies is difficult to infer with confidence from this mtDNA data alone. It almost perfectly corresponds to Clade 1, but there are other behaviourally and morphologically

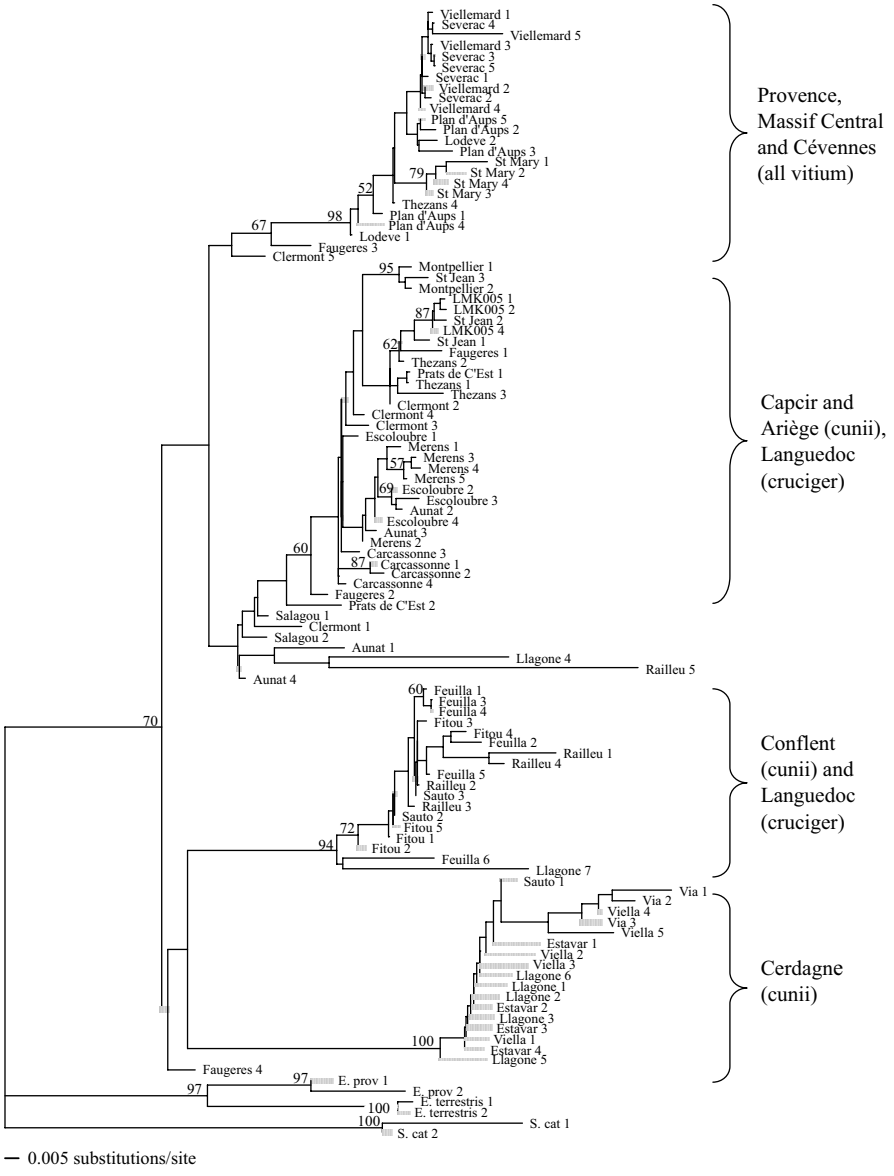

Figure 3 Intraspecific phylogeny for *E. ephippiger*. Minimum evolution tree produced from mtDNA sequencing of a 507 BP fragment of *COII*. The tree presented is the best-scoring tree after 1000 bootstrap replicates, and has a similar topography to the consensus tree. The geographic providence of each clade is indicated (and the type of cricket found there in brackets).

similar subspecies (*E. e. vichetti* and *E. e. ephippiger*) from eastern Europe which we were not able to include in this survey. We can conclude that the cruciger and cunii forms are not taxonomically valid as they are polyphyletic and consist of a number of divergent and possibly ancient haplotypes that correspond to major river valleys in the region rather than taxonomy, behaviour or morphology. These conclusions conform closely to those of Oudman *et al* (1990) who suggested that the whole complex be termed the variable species *E. e. diurnus*, with broad and relatively independent clinal variation, although a firm conclusion should await further sampling (and Kidd and Ritchie, 2000 seemed to detect some coincidence in these clines around the area of the Aude river).

The three well-supported clades that contain cunii and cruciger individuals correspond to the three major river valleys in the eastern Pyrenees (Cerdagne, Conflent and Capcir, Figures 1, 2 and 3). The mtDNA of bushcrickets from these adjoining valleys is very distinct, with net genetic distances over 10%. Similar geographically structured mtDNA variation within Pyrenean samples was suggested in a previous mtDNA RFLP analysis (Ritchie *et al*, 2001), although a phylogeny based on RAPDs did not show these divisions (Ritchie *et al*, 1997). The extent of sequence divergence is surprising, as bushcrickets from all three valleys are morphologically and behaviourally very similar and have previously been shown to have low genetic distances at allozyme loci (Oudman *et al*, 1990; Kidd and Ritchie, 2000; Spooner, 2005). The mountain ranges (~2500 m in altitude) between the three river valleys appear to be severely limiting gene flow between them. *E. ephippiger* usually has a much lower maximum altitude, so the mountains act as effective barriers to gene flow among the three valleys. However, the top of each valley meets at

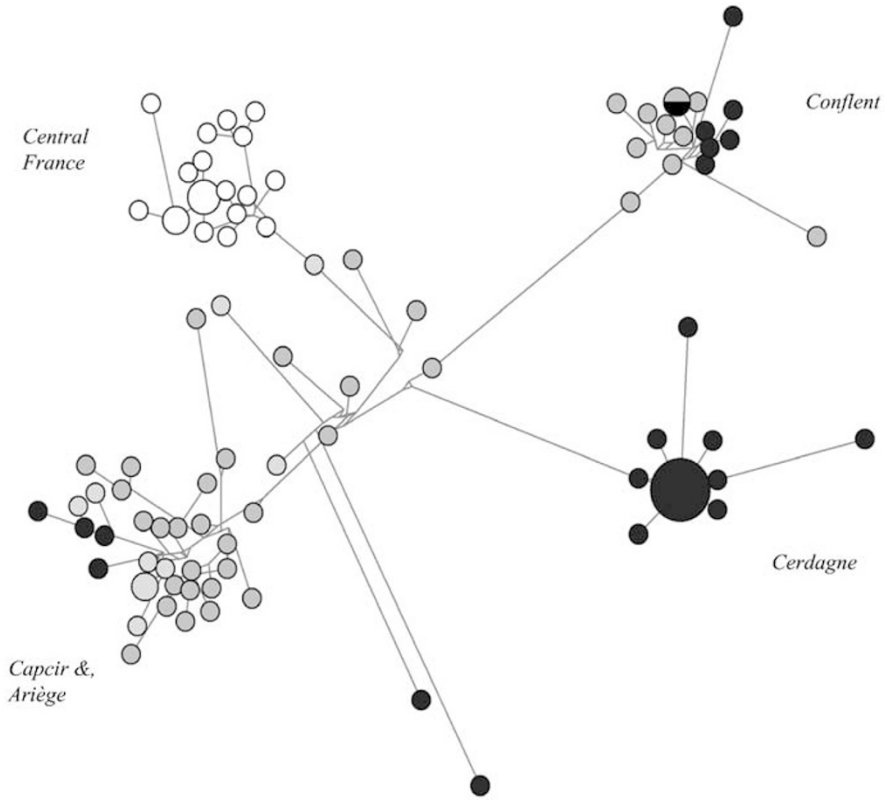

Figure 4 Median-joining network for the mtDNA haplotypes of *E. ephippiger*. Inferred median vectors are omitted for clarity, and the subspecies identification made in the field indicated by shading (blank = vitium, grey = cruciger, black = cunii, shaded = uncertain, usually an intermediate between cruciger and cunii). Branch lengths are approximately equal to inferred mutational steps, although some manipulation of short branches was carried out to reduce node overlap. Node size is proportional to the number of samples. The main geographic locations of clusters are indicated (samples from Llagonne, at the junction of the three Pyreneen valleys, are found in each cluster).

Table 1 Nucleotide diversity measures (Jukes-Cantor's π, standard deviation) for each clade are shown in the diagonals

	Clade 1	*Clade 2*	*Clade 3*	*Clade 4*
Clade 1	0.026, 0.005	0.072, 0.015	0.123, 0.022	0.121, 0.022
Clade 2	0.107, 0.019	0.047, 0.005	0.114, 0.020	0.119, 0.020
Clade 3	0.146, 0.024	0.147, 0.023	0.021, 0.002	0.131, 0.024
Clade 4	0.142, 0.023	0.150, 0.024	0.150, 0.025	0.016, 0.009

Above the diagonal is the net between-group average Tamura-Nei genetic distance (and standard error). Below the diagonal is a simple average Tamura-Nei distance (and standard error).

Llagonne (1637 m altitude), so limited gene flow between the valleys is likely to be occurring now, as Llagonne samples occurred in each of the Pyrenean clades.

The deep genetic divisions suggest that the three clades have been allopatric for some time. These phylogroups are likely to have originally migrated north from an Iberian refugium, but this must have occurred much earlier than most phylogeographic reconstructions of variation in this area suggest. The Pyrenees are a major 'suture zone' with many species or subspecies meeting in this region, but most, such as the well-characterised *Chorthippus parallelus* hybrid zone, seem to occur in relatively narrow secondary hybrid zones along the main mountain ridge (eg Butlin *et al*, 1992; Serrano *et al*, 1996). They have usually been interpreted as a single broad contact, occurring independently in different valleys, but all taking place following the last climatic amelioration. Most valleys show similar clinal patterns, although widths can vary (Butlin *et al*, 1992; Hewitt, 1993; Butlin, 1998). However, some lower valleys provide exceptions to typically simple clinal patterns (eg Buño *et al*, 1994) and perhaps differences between valleys are worth greater attention. The Pyrenees also contains narrow hybrid zones in birds (eg Bensch *et al*, 2002), and *Plantago media* shows an interesting pattern of large differentiation in cpDNA haplotypes running east–west along the Pyrenees, which is related to mating system variation and may have occurred over several glacial cycles (van Dijk and Bakx-Schotman, 1997). In probably the closest similarity to our results, the classic 'area effects' of *Cepaea nemoralis* are seen in both allozyme markers and banding patterns between Pyreneen valleys, and were partly assigned to independent invasions of valleys by already differentiated forms of this low-vagility organism (Ochman *et al*, 1983). Other studies have shown the existence of cryptic taxa with considerable genetic divergence, despite little or no morphological differentiation. For example, phylogeographic analysis of the field vole *Microtus agrestis* found a divergent mitochondrial lineage in Southern Europe that had not been previously recognised in studies of morphology or karyotype (Jaarola and Searle, 2004).

The lack of resolution of the relationship between the clades (the pattern of the deeper branches uniting them) makes it difficult to infer their evolutionary

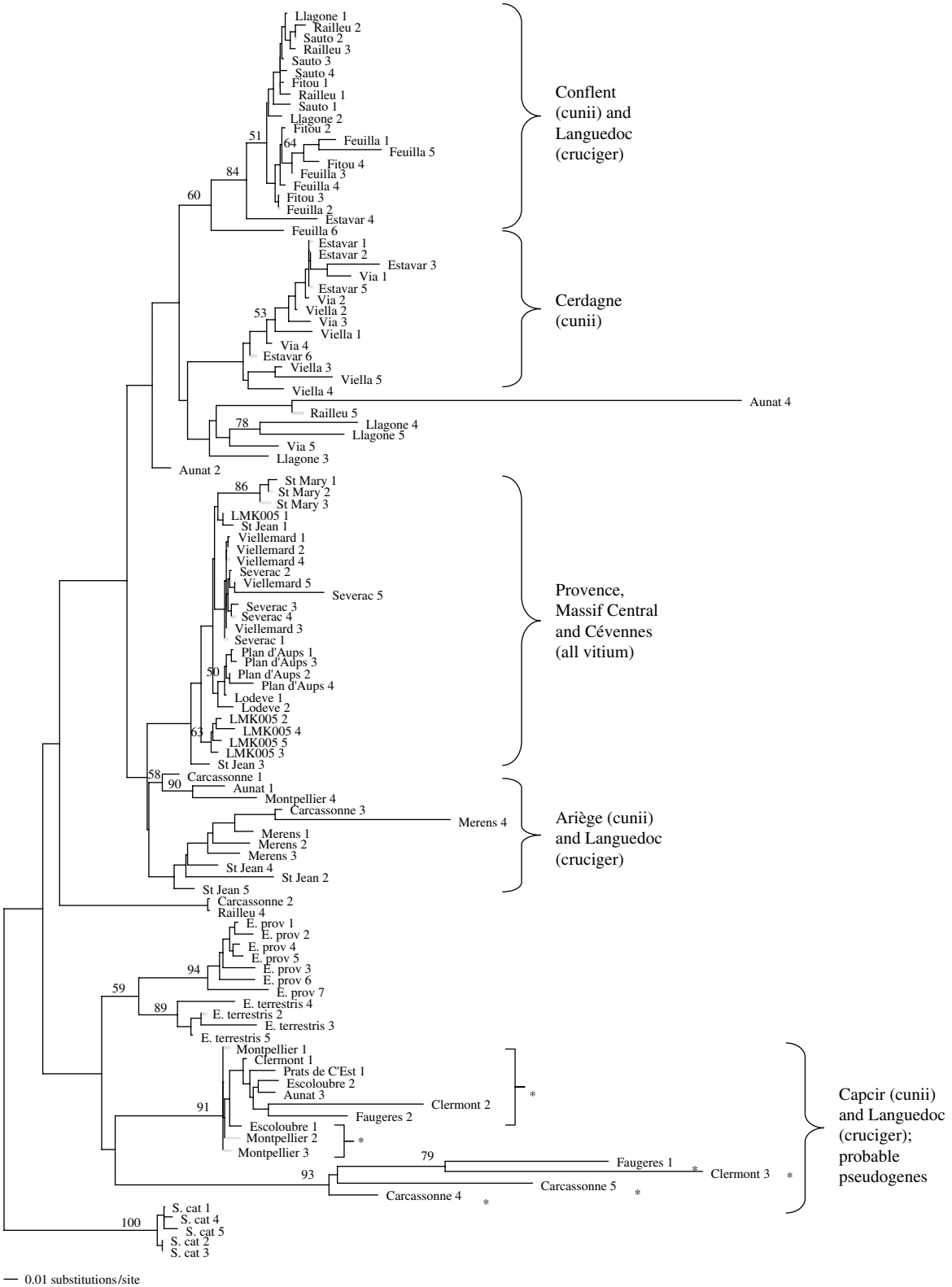

Figure 5 Intraspecific *cyt b* phylogeny for *E. ephippiger*. Minimum evolution tree produced from mtDNA sequencing of a 428 BP fragment of *cyt b*. The tree presented is the best-scoring tree after 1000 replicates, and has a similar topography to the consensus tree. Potential pseudogenes are indicated (*).

history with confidence, but several existing or novel hypotheses for their origin can be briefly discussed (Figure 6a–d):

(a) A Southern form migrated north from Iberia by two separate routes, with subsequent divergence and adaptation to mountain or lowland habitats. Bush-

crickets from Capcir and Conflent may have migrated along the valleys during colder periods and hybridised with lowland forms, but bushcrickets from Cerdagne (a more distinct clade) would be unlikely to come into contact with these. The divergence between the three Pyrenean valleys would therefore be due to long-term occupation of distinct valleys, with limited migration during recent interglacial periods. The appearance of cruciger and cunii 'forms' is consistent with adaptation to differing environments – the larger size and lighter colour of cruciger may reflect the longer season and warmer conditions in the Languedoc versus the high Pyrenees. Environmental conditions do affect body size in *E. ephippiger* (Kidd and Ritchie, 2000; Ritchie *et al*, 2001).

(b) The cruciger form is derived from cunii bushcrickets that migrated from the Capcir into the Languedoc (Ritchie *et al*, 2001). This would explain the presence of cruciger individuals within the Capcir clade, but does not account for the deep divisions between the three phylogroups.

(c) A previously distinct cruciger form is currently introgressing extensively with cunii and vitium. It may have originated from a distinct Pleistocene refugium located in Iberia or on the Mediterranean coast (Rohling *et al*, 1998; Ritchie *et al*, 2001).

(d) The phylogroups originate from separate refugia in Iberia during recent glaciations (Ritchie *et al*, 2001). They then invaded the South of France via different routes, and come into contact at higher altitudes in the Pyrenees. This is compatible with the genetic distinctiveness of bushcrickets from Cerdagne, but less so with the differences between the other valleys.

Genetic divergence between lineages

The level of divergence among the clades found here was extremely high at 10–14% (the level of divergence was similar for the coding *cytb* sequences, so this is not confined to one mitochondrial gene). In the comparative study of Taberlet *et al*, 1998, east versus west European subspecies of a broad range of organisms, including insects, mammals and plants, typically show mtDNA divergence of around 5%, compatible with Pleistocene differentiation. Given an approximate insect mtDNA divergence rate of 2% per million years (Brown *et al*, 1979; Brower, 1994; Fleischer *et al*, 1998; Lunt *et al*, 1998), the divergence between valleys seen here translates to a divergence time of 5–7 million years ago, that is, early Pliocene divergence. This is more typical of interspecific, or even higher, divergence in insects (Funk, 1999; Trewick and Morgan-Richards, 2005) and is rather

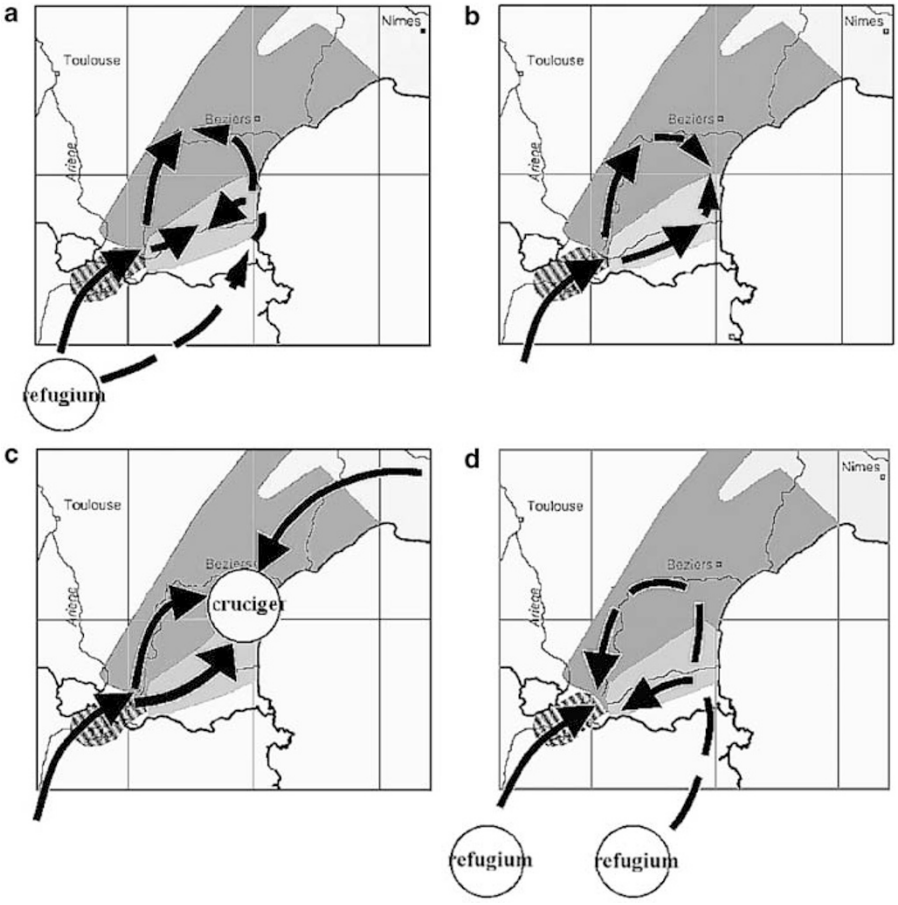

Figure 6 Hypotheses of post-Pleistocene range expansion of the Southernmost clades. (**a**) *E. ephippiger* migrated north by two separate routes; subsequent adaptation to mountain (cunii) or lowland (cruciger) habitats. (**b**) Cruciger is derived from cunii. (**c**) A distinct cruciger form is becoming extinct through introgression; its Pleistocene refugium is unknown. (**d**) Clades originate in two Iberian refugia; separate forms invade Cerdagne and Languedoc, and migrate along the major river valleys.

surprising, as bushcrickets from each clade are probably not reproductively isolated (they have similar courtship songs and preferences and some crosses between them have been made successfully, Ritchie, 2000). There seem two possible non-exclusive explanations – the phylogroups are extremely old, or *COII* evolves particularly quickly in this species.

Large ice sheets began to form 2.4 Mya in Europe (Webb and Bartlein, 1992), and there have been climatic oscillations dating back to the Tertiary that were sufficient to produce major range changes for many species (Hewitt, 1996). It is therefore likely that the current patterns in *E. ephippiger* reflect Pleistocene glaciations acting on existing, much older, patterns of diversity within the species, which have accumulated over several interglacials. *E. ephippiger* shows characteristic signs of repeated and recent range changes, for example, the complex, non-concordant clines in different traits (Oudman *et al*, 1989, 1990; Duijm, 1990; Ritchie *et al*, 2001). The only similar results we can find are those of Trewick *et al* (2000), who found mtDNA differences of a similar magnitude between populations of the alpine scree weta, another flightless orthopteran, confirming that isolation which predates Pleistocene glaciations can result in extensive intraspecific divergence without speciation. However, the reasons for the lack of greater divergence in other traits remain curious. Hewitt (1996) discusses how cycles of isolation may accumulate differences over a great period, so we should perhaps not be surprised to find that genetic divergence often implicates a time scale much greater than the most recent ice age. However, a considerable debate ensued (Lovette, 2005) when a comparative study of North American songbirds found that the 35 sister species had a protracted history of speciation over the last 5 million years, when they were previously thought to have diverged during the late Pleistocene (Klicka and Zink, 1997).

There are two intriguing alternatives to the 'accumulation' model described above. Firstly, it is generally assumed that mountain barriers such as the Pyrenees were completely inhospitable during Pleistocene cold spells, but our results could suggest that this bushcricket somehow persisted in the lower regions of the valleys during more recent glacial periods. Alternatively, the sequence may evolve particularly quickly in *Ephippiger.* The likelihood ratio test rejects the constant substitution rate model for the tree, so there is clear evidence of rate variation. Other studies have recently questioned the use of molecular clocks; between species, divergence may reflect effective substitution rates and contrast with mutation rates seen within species (Ho and Larson, 2006), and hence the clock would be far from consistent. However, this seems unlikely to explain why *Ephippiger* shows higher within-species variation than other insect or orthopteran species, as other studies of mtDNA divergence do not imply a typically faster overall rate of divergence (eg Lunt *et al*, 1998). Alternatively, some studies have recently suggested that mtDNA may not be neutral but be under climatic selection (Mishmar *et al*, 2003; Bazin *et al*, 2006). Selection on mtDNA, acting independently in the different valleys, would exaggerate their divergence. Ritchie *et al* (2001) found that the mtRFLP variation in *Ephippiger* was partly explained by climate, a result they assumed would be indirect via an influence of climate on genetic introgression following secondary contact, but perhaps a direct role of climate on mtDNA is worth further consideration.

Conclusions

Geographic variation within Southern European *E. ephippiger* is complex, with unusual non-concordant patterns of genetic, morphological and behavioural variation in this species. MtDNA clades are suprisingly divergent. The vitium form is perhaps distinct, but the cunii and cruciger forms are polyphyletic, with their mtDNA showing three clades that are more coincident with geographic Pyrenean valleys than morphology or behaviour. Cunii bushcrickets in all three clades have similar calling songs (Duijm, 1990) and mate preferences (Ritchie, 1996, 2000; Spooner, 2005). Either this organism has persisted in different Pyrenean valleys for longer than previously expected from other phylogeographic studies of this area, or mtDNA is evolving rapidly, possibly under selection.

Acknowledgements

We are grateful to Jenny Gleason, Jeff Graves, Leon Hockham, Tanya Hamill, Ruth Hamill and David Kidd among others for their help in the lab and field. Sean Earnshaw helped with the figures. This work was funded by NERC (studentship to LJS and grant to MGR). We are very grateful to various reviewers for helpful comments.

References

Armbruster P, Bradshaw WE, Holzapfel CM (1998). Effects of postglacial range expansion on allozyme and quantitative genetic variation of the pitcher-plant mosquito, *Wyeomyia smithii*. *Evolution* **52**: 1697–1704.

Avise JC, Walker D, Johns GC (1998). Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc Roy Soc Lond B* **265**: 1707–1712.

Bandelt H-J, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**: 37–48.

Bazin E, Glémin S, Galtier N (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**: 570–572.

Bensasson D, Zhang D, Hartl DL, Hewitt GM (2001). Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol* **16**: 314–321.

Bensch S, Helbig AJ, Salomon M, Siebold I (2002). Amplified fragment length polymorphism analysis identifies hybrids between two subspecies of warblers. *Mol Ecol* **11**: 473–481.

Berg A, Greenfield MD (2005). Sexual selection in insect choruses: influences of call power and relative timing. *J Ins Behav* **18**: 59–75.

Brower AVZ (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci USA* **91**: 6491–6495.

Brown WM, George Jr M, Wilson AC (1979). Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA* **76**: 1967–1971.

Buño I, Torroja E, López-Fernández C, Butlin RK, Hewitt GM, Gosálvez J (1994). A hybrid zone between two subspecies of the grasshopper *Chorthippus parallelus* along the Pyrenees: the west end. *Heredity* **73**: 625–634.

Busnel M-C (1963). Charactérisation acoustique de populations d'*Ephippiger* écologiquement voisines. *Ann Epiphyties* **14**: 25–34.

Butlin RK, Ferris C, Hewitt GM, Ritchie MG (1992). Broad-scale mapping of a hybrid zone between subspecies of *Chorthippus parallelus* (Orthoptera: Acrididae). *Ecol Entomol* **17**: 359–362.

Butlin RK (1998). What do hybrid zones in general, and the *Chorthippus parallelus* zone in particular, tell us about speciation? In: Howard DJ, Berlocher SH (eds) *Endless Forms: Species and Speciation*. OUP: New York. pp 367–378.

Cassens I, Mardulyn P, Milinkovitch MC (2005). Evaluation intraspecific 'network' construction methods using simulated sequence data: do existing algorithms outperform the global maximum parsimony approach? *Syst Biol* **54**: 363–372.

Clegg SM, Degnan SM, Kikkawa J, Moritz C, Estoup A, Owens IPF (2002). Genetic consequences of sequential founder events by an island-colonizing bird. *Proc Natl Acad Sci USA* **99**: 8127–8132.

Clement M, Posada D, Crandall K (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**: 1657–1660.

Duijm M (1990). On some song characteristics in *Ephippiger* (Orthoptera: Tettigonioidae) and their geographic variation. *Neth J Zool* **40**: 428–453.

Duijm M, Oudman L (1983). Interspecific mating in *Ephippiger* (Orthoptera: Tettigonioidae). *Tidjschrift voor Entomologie* **126**: 97–108.

Duijm M, Oudman L, Veldstra BG (1983). Copulation in *Ephippiger* (Orthoptera: Tettigonioidae). *Tidjschrift voor Entomologie* **126**: 91–96.

Fleischer RC, McIntosh CE, Tarr CL (1998). Evolution on a volcanic conveyor belt: using phylogeographical reconstructions and K-Ar-based ages of the Hawaiian islands to estimate molecular evolutionary rates. *Mol Ecol* **7**: 533–545.

Funk DJ (1999). Molecular systematics of cytochrome oxidase I and 16S from *Neochlamisus* leaf beetles and the importance of sampling. *Mol Biol Evol* **16**: 67–82.

Greenfield MD, Siegfreid E, Snedden WA (2004). Variation and repeatability of female choice in a chorusing katydid, *Ephippiger ephippiger*: an experimental exploration of the precedence effect. *Ethology* **110**: 287–299.

Hartley JC, Warne AC (1984). Taxonomy of the *Ephippiger ephippiger* complex (*ephippiger, cruciger* and *cunii*) with special reference to the mechanics of copulation. *EOS* **60**: 43–54.

Harz K (1969). *Die Orthopteran Europas*. I. Dr W Junk, NV: The Hague.

Hewitt GM (1988). Hybrid zones – natural laboratories for evolutionary studies. *Trends Ecol Evol* **3**: 158–167.

Hewitt GM (1989). The subdivision of species by hybrid zones. In: Otte D, Endler JA (eds) *Speciation and its Consequences*. Sinauer Associates Inc.: Sunderland, Massachusetts. pp 85–110.

Hewitt GM (1993). After the ice: *parallelus* meets *erythropus* in the Pyrenees. In: Harrison RG (ed) *Hybrid Zones and the Evolutionary Process*. Oxford University Press: Oxford. pp 140–164.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc* **58**: 247–276.

Hewitt GM (1999). Post-glacial re-colonization of European biota. *Biol J Linn Soc* **68**: 87–112.

Hewitt GM (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Phil Trans Roy Soc Lond B* **359**: 183–195.

Ho SYW, Larson G (2006). Molecular clocks: when times are a-changin'. *Trends Genet* **22**: 79–83.

Jaarola M, Searle JB (2004). A highly divergent mitochondrial DNA lineage of *Microtus agrestis* in Southern Europe. *Heredity* **92**: 228–234.

Kidd DM, Ritchie MG (2000). Inferring the patterns and causes of geographic variation in *Ephippiger ephippiger* (Orthoptera, Tettigoniidae) using geographical information systems (GIS). *Biol J Linn Soc* **71**: 269–295.

Klicka J, Zink RM (1997). The importance of recent ice ages in speciation: a failed paradigm. *Science* **277**: 1666–1669.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001). *MEGA2: Molecular Evolutionary Genetics Analysis*. Arizona State University: Tempe, Arizona, USA.

Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ (1994). Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol* **39**: 174–190.

Lovette IJ (2005). Glacial cycles and the tempo of avian speciation. *Trends Ecol Evol* **20**: 57–59.

Lunt DH, Ibrahim KM, Hewitt GM (1998). mtDNA phylogeography and postglacial patterns of subdivision in the meadow grasshopper *Chorthippus parallelus*. *Heredity* **80**: 633–641.

Mirol PM, Mascheretti S, Searle JB (2000). Multiple nuclear pseudogenes of mitochondrial cytochrome *b* in *Ctenomys* (Caviomorpha, Rodentia) with either great similarity to or high divergence from the true mitochondrial sequence. *Heredity* **84**: 538–547.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S *et al* (2003). Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA* **100**: 171–176.

Nei M, Kumar S (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press: New York.

Ochman H, Jones JS, Selander RK (1983). Molecular area effects in *Cepaea*. *Proc Natl Acad Sci USA* **80**: 4189–4193.

Oudman L, Duijm M, Landman W (1990). Morphological and allozyme variation in the *Ephippiger ephippiger* complex (Orthoptera: Tettigonioidae). *Neth J Zool* **40**: 454–483.

Oudman L, Landman W, Duijm W (1989). Genetic distance in the genus *Ephippiger* (Orthoptera: Tettigonioidae) – a reconnaissance. *Tidjschrift voor Entomologie* **132**: 177–181.

Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

Ritchie MG (1996). The shape of female mating preferences. *Proc Natl Acad Sci USA* **93**: 14628–14631.

Ritchie MG (2000). The inheritance of female preference functions in a mate recognition system. *Proc Roy Soc Lond B* **267**: 327–332.

Ritchie MG, Kidd DM, Gleason JM (2001). Mitochondrial DNA variation and GIS analysis confirm a secondary origin of geographical variation in the bushcricket *Ephippiger ephippiger* (Orthoptera: Tettigonioidae), and resurrect two subspecies. *Mol Ecol* **10**: 603–611.

Ritchie MG, Racey SN, Gleason JM, Wolff K (1997). Variability of the bushcricket *Ephippiger ephippiger*: RAPDs and song races. *Heredity* **79**: 286–294.

Rohling EJ, 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.

Serrano L, de laVega CG, Bella JL, Lopez-Fernandez C, Hewitt GM, Gosalvez J (1996). A hybrid zone between two subspecies of *Chorthippus parallelus*. X-chromosome variation through a contact zone. *J Evol Biol* **9**: 173–184.

Simon C, Frati F, Beckenback A, Crespi B, Liu H, Flook P (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Ent Soc Am* **87**: 651–701.

Spooner LJ (2005). *Geographic variation in the mate preferences, calling song and mtDNA of the bushcricket* Ephippiger ephippiger. PhD thesis, University of St Andrews.

Swofford D (1998). *PAUP* Phylogenetic Analysis Using Parsimony* (version 4). Sinauer: Sunderland, MA.

Taberlet P, Fumagalli L, Wust-Saucy A, Cosson J (1998). Comparative phylogeography and postglacial colonisation routes in Europe. *Mol Ecol* **7**: 453–464.

Templeton AR, Crandall KA, Sing CF (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **24**: 4876–4882.

Tregenza T, Pritchard VL, Butlin RK (2002). The origins of postmating reproductive isolation: testing hypotheses in the grasshopper *Chorthippus parallelus*. *Popul Ecol* **44**: 137–144.

Trewick SA, Morgan-Richards M (2005). After the deluge: mitochondrial DNA indicates Miocene radiation and Pliocene adaptation of tree and giant weta (Orthoptera: Anostostomatidae). *J Biogeogr* **32**: 295–309.

Trewick SA, Wallis GP, Morgan-Richards M (2000). Phylogeographical pattern correlates with Pliocene mountain building in the alpine scree weta (Orthoptera, Anostostomatidae). *Mol Ecol* **9**: 657–666.

van Dijk P, Bakx-Schotman T (1997). Chloroplast DNA phylogeography and cytotype geography in autopolyloid *Plantago media*. *Mol Ecol* **6**: 345–352.

Webb T, Bartlein PJ (1992). Global changes during the last 3 million years – climatic controls and biotic responses. *Annu Rev Ecol Syst* **23**: 141–173.

Zhang D, Hewitt GM (1996). Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. *Mol Ecol* **5**: 295–300.