Competitive exclusion in phylogeography: *Crocidura suaveolens* (Soricidae) patterns in Iberia shaped by the arrival of *C. russula*

Running title: Competitive exclusion of *C. suaveolens*

Luis Biedma¹, Jacinto Román², Javier Calzada¹, Guillermo Friis³,* and José A. Godoy³

¹Department of Integrated Sciences, Faculty of Experimental Sciences, University of Huelva, Avenida de las Fuerzas Armadas, S/N, 21007 Huelva, Spain.

²Department of Conservation Biology, Doñana Biological Station, CSIC, C. Americo Vespucio 26, 41092 Sevilla, Spain.

³Department of Integrative Ecology, Doñana Biological Station, CSIC, C. Americo Vespucio 26, 41092 Sevilla, Spain.

* Current address: National Museum of Natural Sciences, Spanish National Research Council (CSIC), Madrid 28006, Spain

e-mails: luiseduardo.biedma@dbasp.uhu.es, jroman@ebd.csic.es, javier.calzada@dbasp.uhu.es, gfrisis@mncn.csic.es, godoy@ebd.csic.es.

Correspondence: José A. Godoy, Department of Integrative Ecology, Doñana Biological Station, CSIC, C. Americo Vespucio 26, 41092 Sevilla, Spain. E-mail: godoy@ebd.csic.es
Abstract

Despite their potential importance, biological processes like competitive exclusion have been mostly neglected in phylogeographic studies. Here we analyze the role of glacial events and competitive exclusion in the evolutionary history of the lesser white-toothed shrew, *Crocidura suaveolens*, in Iberia based on cytochrome b sequences. All the Iberian samples grouped together with the rest of western European populations within the previously described clade IV. We identified three distinct evolutionary lineages within this major clade, two of them occurring exclusively in Iberia. Iberian lineage B extends throughout the northwest with a continuous distribution and diversity values from moderate to high, whereas Iberian lineage C has a highly-patched distribution and is structured in four sublineages, all having low diversity values. No signs of demographic growth were detected for any of the lineages. The evolutionary history of *C. suaveolens* in Iberia supports the refugia-within-refugia scenario, but ecological studies in areas of sympatry, molecular and fossil datings, and contrasting patterns in the Italian Peninsula, suggest that competitive exclusion exerted by *C. russula* since its arrival to Iberia has been the main factor shaping the distribution, phylogeography and population genetics of lineage C.

Keywords: Competitive exclusion, *Crocidura suaveolens*, cytochrome b, genetic diversity, glacial refugia, phylogeography
Introduction

Distribution ranges and patterns of genetic structure and diversity of most European temperate species have been shaped by Quaternary climatic oscillations (Hewitt, 2000; Hewitt, 2004; Taberlet, Fumagalli, Wust-Saucy & Cosson, 1998). During glacial periods, species survived in refuge areas located mainly in the peninsulas of Iberia, Italy and the Balkans (Hewitt, 1999; Taberlet et al., 1998). Glacial periods of contraction and isolation were followed by interglacial periods during which recolonization of the continent occurred through northward expansion from southern refugia (Hewitt, 1999; Hewitt, 2001). Repeated glacial cycles have given rise to divergent phylogeographic clades typically associated to separate glacial refugia (Hewitt, 2004; Hewitt, 2011), and to a progressive reduction in diversity with increasing latitude (Hewitt, 1999; Hewitt, 2000).

Recent evidence shows that southern refugia were indeed a heterogeneous mosaic of suitable habitat isolated by unsuitable habitat that allowed the evolution of separate sublineages within each peninsula (refugia-within-refugia scenario; Abellán & Svenning, 2014; Gómez & Lunt, 2007; Weiss & Ferrand, 2007). This scenario resulted in the high levels of intra-specific diversity, but also of species diversity and endemism when divergence progressed towards speciation, that are characteristic of southern refuges (Abellán & Svenning, 2014; Vega, Amori, Aloise, Cellini, Loy & Searle, 2010). The description of these intraspecific lineages within European refugia is key for the setting of conservation priorities and for the elaboration of effective management plans.

With the generalized emphasis on the role of the Quaternary climatic oscillations, other factors that could also promote intraspecific divergence have been overlooked. In particular, competitive exclusion (CE) by ecologically similar species, or even conspecific populations, has
been a process almost completely neglected in phylogeography (but see Ranjard, Welch, Paturel & Guindon, 2014; Waters, 2011). CE is a basic principle of ecological theory, whereby two species cannot coexist in a stable habitat if they compete for the same resources (Ayala, 1971; Gause, 1932; Hardin, 1960). The successful competitor may drive the other towards extinction or promote its evolutionary shift towards a different ecological niche. Recent evidence suggests CE is a widespread phenomenon that can explain the observed phylogenetic overdispersion of mammal communities (Cooper, Rodríguez & Purvis, 2008). Interspecific CE can become an agent of vicariance and cause isolation, eventually resulting in genetic differentiation and lineage divergence (Gutierrez, Boria & Anderson, 2014). Even intraspecific CE can contribute to the maintenance of spatially-segregated divergent lineages, which explains the genetic homogeneity of the recolonized northern European areas (Hewitt, 1996; Ranjard et al., 2014). Furthermore, CE during range expansion can create patterns at neutral loci that mimic adaptive processes and resemble post-glacial segregation of clades from distinct refuge areas (Excoffier & Ray, 2008).

The lesser white-toothed shrew, *Crocidura suaveolens* (Pallas, 1811), represents a suitable case study to analyze the relative roles of the Quaternary climatic oscillations and CE in the genetic divergence of temperate species in Europe. *C. suaveolens* is widely distributed throughout the Palearctic, extending from the Atlantic coasts of Europe to Siberia (Hutterer, 2005; Palomo, Kryštufek, Amori & Hutterer, 2016). From Central Europe to Asia its distribution is continuous and the species is described as abundant and ubiquitous (Palomo et al., 2016). However, in western Europe, *C. suaveolens* is less common and is absent from large areas, including most of the Iberian Peninsula (Libois, Ramalhinho & Fons, 1999).

Like many other temperate species, *C. suaveolens* retreated to southern Eurasian refugia during Pleistocene glaciations (Hewitt, 1999; Taberlet et al., 1998). As a result of these processes, ten phylogeographic clades have been identified across its wide distribution, the
Iberian populations grouping in clade IV, which is the most western and basal clade in Europe (Dubey, Cosson, Magnanou, Vohralík, Benda, Frynta, Hutterer, Vogel & Vogel, 2007; Dubey, Zaitsev, Cosson, Abdukadier & Vogel, 2006). However, these previous studies included only two Iberian samples, so that any possible internal genetic structure within Iberia might have remained unnoticed. In addition, the most isolated continental populations of European C. suaveolens are located in southwestern Iberia, being they separated more than 300 km from any other populations (Palomo et al., 2016; Román & Ruiz, 2003)(Fig. 1). A plausible hypothesis is that these marginal populations were unintentionally introduced by humans, as demonstrated for some Mediterranean island populations (Dubey et al., 2007); shrews are small and very discrete animals and southwestern Iberia has been a place with a huge commercial marine traffic for millennia (Vives, 2015).

An intriguing possibility is that the recent history and current distribution of C. suaveolens in western Europe, and especially in Iberia, have been influenced by its congeneric the greater white-toothed shrew, C. russula. C. russula managed to reach Iberia from Africa 100 kya (Arsuaga, Baquedano & Pérez-González, 2006; Laplana & Sevilla, 2006; López-García, 2008; Ruiz-Bustos, 1997), being currently a very widespread and common species throughout western Europe (Aulagnier, Hutterer, Amori, Kryštufek, Yigit, Mitsain & Palomo, 2016). Several lines of evidence suggest that C. russula competitively exclude C. suaveolens in areas of sympatry (Cosson, Pascal & Bioret, 1996; Kraft, 2000; Libois et al., 1999; Niethammer, 1979; Pascal, Lorvelec & Vigne, 2006; Poitevin, Catalan, Fons & Croset, 1987; Román & Ruiz, 2003). CE could thus explain the fragmented distribution of C. suaveolens and, if so, it may have also caused genetic isolation leading to intraspecific divergence in this area.

Therefore, in the present study we aimed to: (1) infer the evolutionary history of C. suaveolens in the Iberian Peninsula by identifying the main mitochondrial evolutionary lineages and estimating divergence times among them; (2) clarify the origin of the southwestern Iberian
populations of *C. suaveolens* and (3) discuss the phylogeographic and demographic processes that have shaped the species genetic variation and its current distribution, with explicit consideration of the possible contribution of competitive exclusion by *C. russula*.

**Materials and Methods**

**Sample collection**

Samples of 119 specimens of *C. suaveolens* were collected, 55 as bone samples (mandibles or skulls) obtained from owl pellets, and 64 as tissue samples from live specimens. Sampling was designed to cover most of the Iberian distribution range of the species, however, our attempts to sample the species in Portugal were unsuccessful (Fig. 1, see also Supporting Information, Table S1, for a full description of the collection localities).

**DNA Extraction**

DNA was extracted from tissue samples using the NucleoSpin® 96 Tissue Kit (MACHEREY-NAGEL GmbH & Co. KG) or a ‘salting-out’ protocol (Müllenbach, Lagoda & Welter, 1989), as modified in Centeno-Cuadros, Delibes and Godoy (2009). Bones samples were first frozen under liquid nitrogen and homogenized to powder with a ball-mill or with mortar and pestle (RETSC Mod.MM301). Then, DNA was extracted using a guanidinium silica protocol [method C in Rohland and Hofreiter (2007)]. The final elution volume was 50µl in all cases. New disposable material was used for extractions of each specimen and the work surface was cleaned with bleach between samples to prevent contamination. Extraction blanks were included in every round of extraction and bones samples were extracted in a laboratory used exclusively for low-copy materials.

**PCR amplification and sequencing**
A fragment of 1251 base pairs (bp) including the mitochondrial cytochrome b gene (cyt b) was PCR amplified and sequenced. For tissue samples, primers L14734/H15985 (Ohdachi, Dokuchaev, Hasegawa & Masuda, 2001) were used, however, for degraded bone extracts we designed five pairs of internal partially overlapping primers (Table S2). Sequences were trimmed, edited and assembled using SEQUENCHER 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA), and aligned with Geneious alignment in GENEIOUS version 8.1.5 (http://www.geneious.com, Kearse, Moir, Wilson, Stones-Havas, Cheung, Sturrock, Buxton, Cooper, Markowitz, Duran, Thierer, Ashton, Mentjies & Drummond, 2012). All novel sequences were deposited in GenBank (Accession numbers: MF987937 - MF988055).

**Phylogenetic and network analyses**

We reconstructed a Bayesian phylogenetic tree in BEAST 2.4.1 (Bouckaert, Heled, Kühnert, Vaughan, Wu, Xie, Suchard, Rambaut & Drummond, 2014) using 119 sequences obtained in this study (Supporting Information, Table S1), together with 122 additional sequences downloaded from GenBank (Supporting Information, Table S3) (Dubey et al., 2007; Dubey et al., 2006). Sequences of *C. nigripes* and *C. brunnea* were used as outgroup. We also obtained a maximum-likelihood (ML) tree for this same alignment and assessed support with 100 non-parametric bootstrap replicates in MEGA 7 (Kumar, Stecher & Tamura, 2016). To further explore the internal structure of clade IV, a phylogenetic network was constructed in NETWORK 4.6., using a median-joining (MJ) algorithm (Bandelt, Forster & Röhl, 1999) and the haplotypes identified by the program DNASP 5.10 (Librado & Rozas, 2009) from all sequences grouping within clade IV.

In order to estimate the arrival date of *C. russula* to Iberia and to evaluate the possible influence of competitive exclusion on the evolutionary history of *C. suaveolens* in this area, we constructed a Bayesian phylogenetic tree in BEAST 2.4.1 (Bouckaert et al., 2014) with a set of 81 cyt b published sequences of *C. russula* (Brändli, Handley, Vogel & Perrin, 2005; Cosson,
Divergence time estimates: StarBEAST2

We dated the times of divergence between lineages within clade IV in StarBEAST2, using the same sequences, partitions and substitution models previously employed in the phylogenetic tree reconstruction, and by setting the different clades found there as independent populations. As mean clock rate we used the cyt b substitution rate that we estimated for the *C. suaveolens* branch with a Bayesian relaxed clock analysis in BEAST2 (Fig. S1), as described in Igea, Aymerich, Bannikova, Gosálbez and Castresana (2015) with some minor modifications, including additional calibrations (Table S4). Similarly, we estimated the time of split of Marrocan and European populations of *C. russula* from the phylogenetic tree of *C. russula*, setting the European and Moroccan samples of *C. russula* as predefined populations and the substitution rate estimated for the *C. russula* branch as the mean clock rate (Fig. S1), which happened to be the same than that of *C. suaveolens*.

Genetic structure and diversity

We used spatial analysis of molecular variance (SAMOVA2.0; Dupanloup, Schneider & Excoffier, 2002) to infer groups of populations. Samples were grouped *a priori* in 12 geographical populations generally corresponding to separate sampling localities (Fig. 1; Tables S1 and S3). SAMOVA was run for 10000 iterations from each of 100 random initial conditions, and tested a predefined number of groups (K) ranging from 2 to 12. The inferred hierarchical structure was then used for the analysis of molecular variance (AMOVA), genetic diversity estimates and pairwise comparisons based on $F_{ST}$ (using only haplotype frequencies) or $\Phi_{ST}$ (using also nucleotide distances among haplotypes) in ARLEQUIN v3.11 (Excoffier, Laval & Schneider, 2005). Number of sequences, polymorphic sites, haplotypes and average number of nucleotide differences, as well as the haplotype and nucleotide diversities were estimated for
each population, population groups suggested by SAMOVA, and for lineages of the clade IV using ARLEQUIN v3.11 (Excoffier et al., 2005).

Demographic history

We investigated the demographic history of clade IV lineages with the mismatch distribution of pairwise nucleotide differences, the Harpending’s raggedness index (HRI) (Rogers & Harpending, 1992), and Tajima’s D (Tajima, 1989) and Fu’s FS (Fu, 1997) neutrality statistics in ARLEQUIN v3.11 (Excoffier et al., 2005). In addition, Fu and Li’s \( F^* \) and \( D^* \) (Fu & Li, 1993) and \( R_2 \) (Ramos-Onsins & Rozas, 2002) statistics were estimated and tests were performed using DNASP 5.10 (Librado & Rozas, 2009). We also generated Bayesian Skyline Plots (BSPs; Drummond, Rambaut, Shapiro & Pybus, 2005) with BEAST2.

A more detailed version of methods can be found in Supplementary Information.

Results

All 119 specimens sequenced in this study grouped within clade IV of the \( C. \ suaveolens \) group (Dubey et al., 2007; Dubey et al., 2006), including those sampled in the southern populations.

All the novel sequences form a highly supported clade together with the other twelve preexisting sequences of this clade (pp=posterior probability = 1.00 and 100 % bootstrap; Figs. S2 and S3). The 131 cyt b sequences in clade IV included 95 polymorphic sites, 74 of which were phylogenetically informative, and defined a total of 45 haplotypes (Supporting Information, Table S5).

Both phylogenetic analyses (Bayesian and ML) revealed three main phylogenetic lineages within clade IV occurring in different areas of the species range (see Figs. S2 and S3 for Bayesian and ML phylogenetic trees, Fig. 2 for a dated phylogeny of Clade IV and Fig. 1 for the geographical distribution of each lineage). One lineage was located mainly in France (Lineage
A), within which two sublineages were differentiated. Sublineage A1 was found on the French Mediterranean coast, including samples ranging from the province of Gerona, in Spain, to the northwest of Italy (Sublineage A1; pp=0.90/66 % bootstrap), and sublineage A2 grouped shrews from the different islands of the northwestern coast of France and the Channel Islands (Sublineage A2; pp=1.00/70 % bootstrap). The other two main lineages within the clade IV (lineages B and C) were strictly Iberian. Lineage B was widely distributed and located in the northwest of the Iberian Peninsula (from the province of Oviedo, in the north of Spain, to the mountain ranges of the Iberian Central System (Lineage B; pp=1.00/74 % bootstrap). The other Iberian lineage, Lineage C, was more structured, with four differentiated sublineages. One of these sublineages was located in the north-central area of Iberia (Sublineage C1; pp=0.99/72 % bootstrap). Interestingly, a second sublineage was found in the locality of Candelario (locality 9 in Fig. 1 and Tables S1 and S3), at the Eastern edge of the Iberian Central System (Sublineage C2; pp=0.99/65 % bootstrap); the co-occurrence of haplotypes of sublineage C2 and lineage B identifies this locality as a secondary contact zone. The other two remaining sublineages were located in southwestern Iberia. One of them grouped individuals sampled in several river mouths in the province of Huelva (Sublineage C3; pp=1.00/90 % bootstrap), whereas the other included individuals sampled exclusively at the Guadalquivir River mouth (Sublineage C4; pp=0.99/42 % bootstrap).

The three lineages and their respective sublineages were also neatly distinguished in a MJ network (Fig. 3). Sublineages A1 and A2 are clearly separated, and the rather large haplotype divergence within sublineages could indicate further subdivision in France, although the sparse sampling there impedes any strong conclusion on this issue. Lineage B included a large number of haplotypes (h = 22) with similar frequencies, whereas lineage C altogether presented less haplotypes (h = 13) with more heterogeneous frequencies (Tables 1 and S5). Indeed, haplotypes H15, H31 and H3 of the sublineages C3, C4 and C1 were clearly predominant within the clade IV, with 25, 15 and 11 occurrences, respectively.
Divergence times

The substitution rate for the cyt b gene in *C. suaveolens* was estimated in 0.0377 substitutions/site/Myr (95% confidence interval: 0.0220 – 0.0577; Fig. S1). Using this rate, the splits that separate the French lineage A and Iberian lineages B and C, were dated around 0.37 Ma (95% confidence interval: 0.16 – 0.64 Ma) and 0.32 Ma (95% confidence interval: 0.14 – 0.57 Ma), respectively, although the low support of this latter node suggests a simultaneous division for the three lineages in clade IV (Fig. 2). The split between sublineages A1 and A2 occurred around 0.24 Ma (95% confidence interval: 0.10 – 0.42 Ma), and subsequent splits giving rise to sublineages C1, C2, C3 and C4 were dated around 0.22 Ma (95% confidence interval: 0.08 – 0.39 Ma; pp = 0.99), and 0.11 Ma (95% confidence interval: 0.03 – 0.19 Ma; pp = 0.59), respectively.

The cyt b substitution rate estimated for the *C. russula* branch was identical to that of *C. suaveolens* (Fig. S1). The application of this rate on the phylogenetic tree obtained for *C. russula* yielded an estimate for the split of European and African populations of *C. russula* in 126 Ka (41 - 200 Ka; Fig. S4).

Genetic structure and diversity

SAMOVA analyses suggested a spatial subdivision of the Iberian distribution area into eight population groups (Table 2; see also Table S1). Only two groups were formed by more than one population, the north-central Iberian Peninsula (NC-IP), formed by the populations of Burgos and the combined population of País Vasco and Navarra, and the southwestern Huelva (SW-Huelva), grouping together all populations of southwestern Iberia, except Guadalquivir. Groups suggested by SAMOVA generally corresponded to lineages and sublineages, with two exceptions: i) Eastern-CS included samples of lineages B and C2, and ii) lineage B was subdivided in four genetic groups (Oviedo, Galicia, Zamora and Western-CS). The first case is representing the only occurrence of secondary contact and the second indicates some
hierarchical genetic structure within lineage B, which could be due both to historical isolation
and current restrictions to gene flow.

The AMOVA with this hierarchical structure suggested by SAMOVA showed that the majority
of the total mtDNA variation was attributed to differences among groups (77.45%), whereas a
very low and negative percentage of variation was due to differences among populations
within groups (-1.96%). Slightly negative variance can be due to random variance or to genes
from different populations being more related to each other than genes from the same
population, as it happens in Eastern Iberian Central System (Eastern-CS), where a secondary
contact area was detected (locality 9 in Fig. 1 and Table S1). Pairwise comparisons based both
on differences between sequences ($\Phi_{ST}$) and on haplotypes frequencies only ($F_{ST}$) were
consistent in showing that neighboring populations or populations grouped within the same
phylogenetic lineage displaying lower differentiation between them (Table S6; see also Table
S1 and Fig. 1). However, differentiation between populations of the Iberian Central System
(Eastern-CS and Western-CS, populations 3 and 4 in Table S6, respectively) was somewhat
larger than expected given its proximity ($\Phi_{ST}$ = 0.307; $F_{ST}$ = 0.476), which may again be due to
the co-occurrence of haplotypes of different lineages in Eastern-CS. Guadalquivir also
presented a large differentiation with respect to the other populations in southwestern Huelva
(populations 8 and 9-12 in Table S6, respectively), despite relative proximity between them,
which is consistent with both long term isolation resulting in divergent sublineages (C3 and C4,
Figs. 1 and 2).

Lineage B of northwestern Iberia showed the highest values of haplotype and nucleotide
diversity among Iberian lineages ($H_d$=0.948 and $P_i$=0.693), whereas all C sublineages showed
low values of diversity (sublineage C2 of Candelario was too poorly sampled to draw any
conclusion) (Table 1). The lowest values were obtained for sublineage C3 of southwestern
Huelva ($H_d$=0.306 and $P_i$=0.049), followed closely by sublineage C1 of the north central Iberia
(Hd=0.396 and Pi= 0.063) and sublineage C4 of the Guadalquivir River (Hd= 0.502 and Pi= 0.071).

Very similar patterns were obtained for population groups suggested by SAMOVA, with population groups of the northwest and center of the Iberian distribution range of C. suaveolens showing moderate or high diversity values (population groups 1-5 in Table 2 and localities 1-9 in Fig. 1), whereas the rest of the Iberian population groups presented low diversity values. Galicia showed the highest values of haplotype diversity (H = 0.910) and Eastern Central System the highest values of nucleotide diversity (Pi = 0.968), the latter reflecting the co-occurrence of highly divergent haplotypes in Candelario (locality 9 in Fig. 1 and Tables S1 and S3). On the contrary, C. suaveolens populations in Piedras and Tinto rivers, belonging to the southwestern Huelva group, were fixed for a single haplotype each. In fact, southwestern Huelva (SW-Huelva) was the population group with the lowest values of diversity (H = 0.306 and Pi = 0.049), followed closely by the north central Iberian Peninsula group (NC-IP), within which highlights the low values for the combined population of País Vasco and Navarra (H = 0.222 and Pi = 0.029).

**Historical demography**

No sign of demographic expansion was detected for clade IV or any of their sublineages (Table 1), with the only exception of sublineage C1 in north central Iberia, which showed significant negative values of Tajima’s D and Fu’s FS test, as well as not significant values for both, Fu and Li’s test (F* and D*). Nevertheless, the $R^2$ statistic, considered one of the most reliable statistics for small populations (Ramos-Onsins & Rozas, 2002) was not significant and the Bayesian skyline plot (BSP) did not reveal a clear sign of demographic growth for this sublineage (Fig. S5). On the other hand, sublineage C3 of southwestern Huelva showed a significant value of the $R^2$ statistic, however, all other calculated statistics were not significant.
and, similarly to sublineage C1, the BSP did not show clear evidence of population expansion (Fig. S5).

Discussion

Evolutionary history of *Crocidura suaveolens* in Iberia

Our phylogenetic results revealed that all *C. suaveolens* populations in Iberia belong to the clade IV of western Europe (Figs. 2 and Figs. S2, S3), discarding a human introduction in this area as the origin of southern populations, and evidenced a sharp internal phylogeographic substructure.

According to our analysis, the split that separated clade IV of the nucleus conformed by clades V-X took place 1.41 Ma (95% confidence interval: 0.71 – 2.45 Ma; Fig. S2a), a date similar to that provided by Dubey et al. (2006) (1.72 Ma [95% confidence interval: 1.40 – 2.23 Ma]), and overlapping the Donau glaciation (1.4 - 1.8 Ma) (Penck & Brückner, 1909). Concordantly, paleontological studies estimate the arrival of *C. suaveolens* to Iberia at least 1 - 1.3 Ma, in the Lower/Middle Pleistocene (Montoya, Alberdi, Barbadillo, van der Made, Morales, Murelaga, Peñalver, Robles, Ruiz-Bustos, Sánchez, Sanchiz, Soria & Szyndlar, 2001; Rofes & Cuenca-Bescós, 2011). Once established in Iberia, the combination of subsequent glacial cycles, separate refugia and competitive exclusion by *C. russula* has probably been responsible for further intraclade divergences as discussed below.

Indeed, we report a well-defined internal structure within clade IV of western Europe with three main lineages, two of them occurring exclusively in Iberia (lineages B and C, Figs 1 and 2). Lack of a more intensive sampling and the uncertain distribution of the species in France hampered the assessment of the internal structure of lineage A. However, the two Iberian
lineages are neatly distributed in different portions of the distribution range and show contrasting internal patterns. Lineage B in the northwest forms a single widespread monophyletic group, concordant with its more continuous distribution in this area, whereas lineage C is divided into several sublineages, each corresponding to separate distribution patches within Iberia. Interestingly, sublineage C2 coexists with lineage B in a single locality (Candelario, Salamanca province; locality 9 in Fig. 1), providing the only instance of secondary contact between divergent lineages within Iberia. Therefore, observed phylogeographic discontinuities within Iberia provide support to the “refugia within refugia” hypothesis (Abellán & Svenning, 2014; Gómez & Lunt, 2007; Weiss & Ferrand, 2007).

Given the similarity of their dating, it is likely that a same event produced the two oldest split within clade IV (Fig. 2), the split that gave rise to the French lineage A, estimated at 370 Ka (95% confidence interval: 160 – 640 Ka), and the split that separated the Iberian lineages B and C, dated at 320 Ka (95% confidence interval: 140 – 570 Ka), most likely the Mindel glacial period (390 – 580 Ka). During glacial maxima, habitats favorable for temperate species like C. suaveolens would be in the lower zones separated by inhospitable habitats in high altitude zones (Gómez & Lunt, 2007). Therefore, Lineage A would take refuge in the south of France, a known refuge area described for other temperate species of small mammals (Feuda, Bannikova, Zemlemerova, Di Febbraro, Loy, Hutterer, Aloise, Zykov, Annesi & Colangelo, 2015; Vega, Fløjgaard, Lira-Noriega, Nakazawa, Svenning & Searle, 2010; Yannic, Basset & Hausser, 2008), whereas lineage B and C might have diverged within Iberia in subrefugia separated by the Central Iberian Plateau. This elevated and large plateau, located in Central Iberia, could have acted as a barrier for this temperate small mammal because it was dominated by cold and arid steppe landscapes during the Pleistocene glaciations (González-Sampérez, Leroy, Carrión, Fernández, García-Antón, Gil-Garcia, Uzquiano, Valero-Garcés & Figueiral, 2010). In fact, cold and dry climate limits the current distribution of the species in northern Eurasia (Palomo et al., 2016). Interestingly, a similar phylogeographical divide has been recently

The exact location of Iberian subrefugia cannot be directly inferred from our data, but geographical ranges of these lineages coincide with areas of refuge proposed for other species in Iberia. On the one hand, the distribution range of lineage B coincides with a well-known area of speciation in Iberia, where numerous species of endemic vertebrates occur, including the small mammals *Sorex granarius* and *Microtus lusitanicus*. Furthermore, northwestern Iberia has been previously proposed as subrefugium for small mammals, such as *Arvicola sapidus* and *Galemys pyrenaicus* (Centeno-Cuadros et al., 2009; Igea, Aymerich, Fernández-González, González-Esteban, Gómez, Alonso, Gosálbez & Castresana, 2013), whereas a region in central Portugal has been proposed as one of the main subrefugia for *Microtus agrestis* (Jaarola & Searle, 2004). So, it is very likely that these areas had also adequate habitats for *C. suaveolens* during the glaciations. Unfortunately, our lack of sampling the potential subrefugium in Portugal may have caused us to miss the origin of B clade, as suggested by the absence of a basal central haplotype for the star-like lineage B (Fig. 3). On the other hand, lineage C fragmented range suggests either several separate subrefugia in the eastern half of Iberia, like suggested for other small mammals (Centeno-Cuadros et al., 2009; Igea et al., 2013; Jaarola & Searle, 2004) or a single refugia that expanded and was subsequently isolated in at least three separate patches. In this latter case, the basal position of haplotype H31 sampled in southwestern Iberia (Fig. 3), suggests the possibility that this region - an important glacial refuge area described for other European mammals (Randi, 2007) - acted as the main eastern refugium for lineage C. In the former scenario, a subsequent climatic event, most likely the Riss
glacial period (140 - 200 Ka), may have caused a new isolation event resulting in the formation of Iberian sublineages C1, C2, C3 and C4, as well as French sublineages A1 and A2 (Fig. 2). However, the fragmentation of these sublineages may have had a non-climatic cause. Thus, according to our own dating (Fig. S4), approximately at the end of the Riss glacial period and beginning of the Riss-Wurm interglacial period (80 - 140 Ka), a new and a very competitive species of Crocidura, C. russula, colonized Iberia from Africa. Previous studies (Brändli et al., 2005; Cosson et al., 2005), estimated somewhat more recent dates for the arrival of C. russula during the Wurm glacial period (11 – 80 Ka). Nevertheless, we estimated the arrival date for C. russula at 126 Ka (41 - 200 Ka) (Fig. S4), a confidence interval with substantial overlap with the internal diversifications of lineages A and C. Furthermore, C. russula seems to be present in the Iberian fossil record since at least 100 Ka (Arsuaga et al., 2006; Laplana & Sevilla, 2006; López-García, 2008; Ruiz-Bustos, 1997). Therefore, an intriguing possibility is that the isolation driving the divergence of sublineages A and C was caused by the competitive exclusion exerted by C. russula since its arrival. C. russula would outcompete C. suaveolens in the hotter and drier habitats of Iberia, more similar to habitats in North Africa where C. russula evolved (González & Román, 1988; Poitevin, Catalan, Fons & Croset, 1986). The same seems to be happening currently in some areas of sympatry (Cosson et al., 1996; Kraft, 2000; Libois et al., 1999; Niethammer, 1979; Pascal et al., 2006; Poitevin et al., 1987; Román & Ruiz, 2003). In the Mediterranean climate region of Iberia (Fig. S6), C. suaveolens has been able to withstand the pressure of C. russula only in relatively wetter habitats (Rey, 2007), as the marshy areas in the southwest (sublineages C3 and C4, Fig. 1) (Román & Ruiz, 2003) and high-altitude wooded areas in the mountain ranges of the Iberian Central System (sublineage C2 and localities 7, 8 and 9 of lineage B, Fig. 1). On the other hand, the more humid and cold climate in the Atlantic region of Iberia (Fig. S6), might have favored C. suaveolens in its competition with C. russula, allowing the persistence of the sublineage C1 in the north-central area as well as a more extensive and continuous distribution of lineage B in the northwest (Fig. 1).
The recent contraction of the Iberian range of *C. suaveolens* is supported by numerous citations of this shrew in Pleistocene fossil deposits located in areas where the species is currently absent (Arribas, 1994; Barroso Ruiz & Desclaux, 2006; Guillem-Calatayud, 1995a; Guillem-Calatayud, 1995b; Guillem-Calatayud, 2000; Guillem-Calatayud, 2001; López-García, 2008; Montoya et al., 2001; Ruiz-Bustos, Vargas, Camprodón & Sans-Coma, 1984). Moreover, in other European areas of Mediterranean climate where *C. russula* is absent, *C. suaveolens* appears to have a more continuous distribution and to occupy a wider range of habitats (Palomo et al., 2016). Particularly relevant is the contrasting case of the Italian Peninsula, given its climate similarity with Iberia and its role as glacial refugia. Here *C. suaveolens* is distributed continuously throughout the peninsula, including the most Mediterranean and dry habitats (Mortelliti & Boitani, 2009). Furthermore, no phylogeographic substructure was found in a recent study (Castiglia, Annesi, Amori, Solano & Aloise, 2017), with all Italian populations belonging to a single evolutionary lineage of Italo-Balkanic origin (clade VII in Dubey et al., 2007; Dubey et al., 2006), which expanded around 60-149 Kya, coinciding with the arrival of *C. russula* to Iberia. Therefore, in Iberia a similar population expansion from subrefugia may have been impeded or, if not, the resultant range may have subsequently fragmented into isolated patches by CE with *C. russula*, specifically in areas of Mediterranean climate (Fig. S6), giving rise to the divergence of sublineages within lineages A and C (Fig. 2). Furthermore, a rather long history of small size and isolation caused by CE can explain the low diversity and high differentiation observed for isolated populations of lineage C when compared to populations of the widespread and more continuous lineage B (Tables 1 and 2).

**Taxonomic and conservation implications**

Despite the great divergence that clade IV presents with respect to other clades of the *C. suaveolens* group, it was classified following a conservative criterion like subspecies *C. suaveolens icuslisma*, mainly because the degree of reproductive isolation with its neighboring
clade, clade VII of the south and center of Europe, is unknown (Dubey et al., 2006). Our results reveal the existence of several lineages and sublineages within clade IV of C. suaveolens, but further analysis would be necessary to know if there is reproductive isolation between them that justifies any taxonomic split within this clade. Analysis in secondary contact areas, as the one we detected in the locality of Candelario between lineage B and sublineage C2, as well as the use of nuclear markers, could through help to clarify the taxonomy in this region.

Irrespective of the degree of reproductive isolation, lineages and sublineages of the clade IV identified here should be conservatively considered as separate Evolutionary Significant Units for conservation purposes, because of its independent evolutionary history and the potential for local adaptation.

According to the IUCN Red List of Threatened Species, C. suaveolens is a species classified as Least Concern (LC) (Palomo et al., 2016), but the species is classified as Data Deficient (DD) in the Spanish Red List (Palomo, Gisbert & Blanco, 2007), and it is Not Evaluated (NE) in Portugal (Cabral, Almeida, Almeida, Dellinger, Ferrand de Almeida, Oliveira, Palmeirim, Queiroz, Rogado & Santos-Reis, 2005). Consequently, this study provides a first approximation to the genetic status of the species in Iberia and encourages further studies of the species in western Europe.

Acknowledgements

This work was funded through project "Evolutionary history of lesser white-toothed shrew, Crocidura suaveolens, in the Iberian Peninsula and genetic status of populations in the Gulf of Cádiz" funded by the internal EBD call "Microproyectos" financed by the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Program for Centres of Excellence in R+D+I (SEV-2012-0262). Luis Biedma benefited from a FPU fellowship from the Ministerio de Educación, Cultura y Deporte. Logistical support was provided by Laboratorio de Ecología
Molecular, Estación Biológica de Doñana, CSIC (LEM-EBD). We thank I. Merino-Sainz, A. Anadón, A. Somoano, P. L. Ramos Bueno, G. Burgos, I. Elorduizapaterietxe, J. Ugarteburu and Aranzadi Society of Sciences for providing samples for this study, and Laura Soriano and Elena Marmesat for technical assistance. We also thank regional administrations for sampling permits, as well as the environmental agents who assisted us in the field work. Finally, we thank two anonymous reviewers for their helpful comments.

References


Castiglia R, Annesi F, Amori G, Solano E, Aloise G. 2017. The phylogeography of Crocidura suaveolens from southern Italy reveals the absence of an endemic lineage and
supports a Trans-Adriatic connection with the Balkanic refugium. *Hystrix, the Italian* Journal of Mammalogy **28**.


**Figure Legends**

**Figure 1.** Distribution of the clade IV of *Crocidura suaveolens* group in western Europe (striped), showing collection localities (1-18) and localities of downloaded sequences (19-25) (in colour). Localities in close proximity to each other are represented as a single location; see Tables S1 and S3 for more details. The colours represent the different lineages/sublineages identified in the phylogenetic analyses, which are also separated by thick black lines/dashed lines on the map. Altitude is shown with a grayscale, the lower areas with lighter tones and the higher areas with darker tones. Distribution of *C. russula* is also shown with different striped fills. Note how the distribution range of *C. suaveolens* in western Europe is fragmented only where both species are sympatric.

**Figure 2.** Dated tree of the clade IV of *Crocidura suaveolens* group with the main lineages (A, B and C) and sublineages (A1 and A2, C1 – C4) identified in this work. Bayesian posterior probabilities/bootstrap supports (above branch) were obtained for the Bayesian and ML trees, respectively (Figs. S2 and S3). Bootstrap supports of the nodes that were not present in the ML tree are represented by a dash (-). Times of divergence and credibility intervals (in parentheses) estimated in *BEAST* are also indicated for each node.

**Figure 3.** Median-joining network of cytochrome b clade IV haplotypes of *Crocidura suaveolens*, colored according to lineage (A1-red; A2-yellow; B-orange; C1-light green; C2-dark green; C3- light blue and C4-dark blue). The diameter of the circles represents the number of sampled individuals with that haplotype. Black dots indicate unsampled intermediary haplotypes.
### Table 1. Genetic diversity and demographic statistics for lineages of the clade IV of *Crocidura suaveolens* group.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>N</th>
<th>h</th>
<th>S</th>
<th>H</th>
<th>Pi</th>
<th>K</th>
<th>SSD</th>
<th>HRI</th>
<th>Tau</th>
<th>Tajima D</th>
<th>Fu's Fs</th>
<th>Fu and Li's D</th>
<th>Fu and Li's F</th>
<th>$R_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1*</td>
<td>8</td>
<td>7</td>
<td>26</td>
<td>0.964</td>
<td>1.026</td>
<td>9.75</td>
<td>0.034</td>
<td>0.070</td>
<td>7.607</td>
<td>-0.146</td>
<td>-0.505</td>
<td>-0.420</td>
<td>-0.413</td>
<td>0.1428</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.465)</td>
<td>(P = 0.588)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.448)</td>
<td>(P = 0.304)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.10)</td>
<td>(P = 0.211)</td>
</tr>
<tr>
<td>A2*</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>1.000</td>
<td>1.026</td>
<td>10</td>
<td>0.000</td>
<td>0.000</td>
<td>0</td>
<td>0.000</td>
<td>1.139</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.689)</td>
<td>(P = 0.478)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.06)</td>
<td>(P = 0.034)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.10)</td>
<td>(P = 0.084)</td>
</tr>
<tr>
<td>B</td>
<td>50</td>
<td>22</td>
<td>41</td>
<td>0.948</td>
<td>0.693</td>
<td>5.55</td>
<td>0.010</td>
<td>0.026</td>
<td>6.515</td>
<td>-1.342</td>
<td>-6.046</td>
<td>-0.075</td>
<td>-0.592</td>
<td>0.0672</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.211)</td>
<td>(P = 0.227)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.06)</td>
<td>(P = 0.034)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.10)</td>
<td>(P = 0.084)</td>
</tr>
<tr>
<td>C1</td>
<td>14</td>
<td>4</td>
<td>3</td>
<td>0.396</td>
<td>0.063</td>
<td>0.429</td>
<td>0.005</td>
<td>0.168</td>
<td>0.505</td>
<td>-1.671</td>
<td>-2.288</td>
<td>-2.091</td>
<td>-2.255</td>
<td>0.1368</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.684)</td>
<td>(P = 0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.026)</td>
<td>(P = 0.005)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.05)</td>
<td>(P = 0.138)</td>
</tr>
<tr>
<td>C2*</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>0.667</td>
<td>0.402</td>
<td>3.333</td>
<td>0.392</td>
<td>1.000</td>
<td>5.881</td>
<td>0.000</td>
<td>2.357</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.068)</td>
<td>(P = 0.545)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.782)</td>
<td>(P = 0.824)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>30</td>
<td>4</td>
<td>4</td>
<td>0.306</td>
<td>0.049</td>
<td>0.453</td>
<td>0.005</td>
<td>0.286</td>
<td>3</td>
<td>-1.416</td>
<td>-1.436</td>
<td>-0.608</td>
<td>-1.050</td>
<td>0.0631</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.494)</td>
<td>(P = 0.566)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.07)</td>
<td>(P = 0.087)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.10)</td>
<td>(P = 0.000)</td>
</tr>
<tr>
<td>C4</td>
<td>23</td>
<td>3</td>
<td>2</td>
<td>0.502</td>
<td>0.071</td>
<td>0.561</td>
<td>0.013</td>
<td>0.154</td>
<td>0.697</td>
<td>0.080</td>
<td>0.109</td>
<td>-0.646</td>
<td>-0.514</td>
<td>0.1564</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.272)</td>
<td>(P = 0.313)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.622)</td>
<td>(P = 0.452)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.10)</td>
<td>(P = 0.563)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>131</td>
<td>45</td>
<td>95</td>
<td>0.936</td>
<td>1.214</td>
<td>9.589</td>
<td>0.008</td>
<td>0.011</td>
<td>13.617</td>
<td>-1.441</td>
<td>-11.101</td>
<td>-1.889</td>
<td>-2.123</td>
<td>0.0447</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.547)</td>
<td>(P = 0.497)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(P = 0.054)</td>
<td>(P = 0.028)</td>
<td>(P &gt; 0.10)</td>
<td>(P &gt; 0.10)</td>
<td>(P = 0.037)</td>
</tr>
</tbody>
</table>

N, sample size; h, number of haplotypes; S, number of polymorphic sites; H, haplotype diversity; Pi, nucleotide diversity (expressed as percentages, i.e. 0.001 = 0.1%); K, average number of pairwise nucleotide differences; SSD, sum of square deviation; HRI, Harpending’s raggedness index; Tau is the mode of the curve of the mismatch distribution when a signal of demographic expansion is detected. It is proportional to the time since expansion; n.a., not available (for lineages with low number of samples). Values in bold show significant tests ($P < 0.05$) for SSD, HRI, Tajima’s D, Fu and Li’s tests and $R_2$ statistic; ($P < 0.02$) for Fu’s Fs test. *Lineages including haplotypes downloaded from GenBank obtained from an unknown number of samples (assumed here to be one).
Table 2. Genetic diversity indices in Iberian populations of *Crocidura suaveolens*.

<table>
<thead>
<tr>
<th>Population Group ID</th>
<th>Population</th>
<th>Localities</th>
<th>N</th>
<th>h</th>
<th>S</th>
<th>H</th>
<th>Pi</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oviedo</td>
<td>1</td>
<td>13</td>
<td>6</td>
<td>15</td>
<td>0.821</td>
<td>0.483</td>
<td>5.077</td>
</tr>
<tr>
<td>2</td>
<td>Galicia</td>
<td>2, 3, 4</td>
<td>13</td>
<td>9</td>
<td>17</td>
<td>0.910</td>
<td>0.531</td>
<td>4.410</td>
</tr>
<tr>
<td>3</td>
<td>Zamora</td>
<td>5, 6</td>
<td>10</td>
<td>4</td>
<td>7</td>
<td>0.800</td>
<td>0.399</td>
<td>3.022</td>
</tr>
<tr>
<td>4</td>
<td>Western-CS</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>0.600</td>
<td>0.434</td>
<td>3.600</td>
</tr>
<tr>
<td>5</td>
<td>Eastern-CS</td>
<td>8, 9</td>
<td>11</td>
<td>3</td>
<td>21</td>
<td>0.473</td>
<td>0.968</td>
<td>8.036</td>
</tr>
<tr>
<td>6</td>
<td>NC-IP</td>
<td>10, 11, 12, 13</td>
<td>14</td>
<td>4</td>
<td>3</td>
<td>0.396</td>
<td>0.063</td>
<td>0.429</td>
</tr>
<tr>
<td></td>
<td>Burgos</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0.700</td>
<td>0.117</td>
<td>0.800</td>
</tr>
<tr>
<td></td>
<td>País Vasco and Navarra</td>
<td>11, 12, 13</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>0.222</td>
<td>0.029</td>
<td>0.222</td>
</tr>
<tr>
<td>7</td>
<td>SW-Huelva</td>
<td>14, 15, 16, 17</td>
<td>30</td>
<td>4</td>
<td>4</td>
<td>0.306</td>
<td>0.049</td>
<td>0.453</td>
</tr>
<tr>
<td></td>
<td>Guadiana</td>
<td>14</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>0.533</td>
<td>0.115</td>
<td>1.067</td>
</tr>
<tr>
<td></td>
<td>Piedras</td>
<td>15</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Odiel</td>
<td>16</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>0.667</td>
<td>0.071</td>
<td>0.762</td>
</tr>
<tr>
<td></td>
<td>Tinto</td>
<td>17</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>8</td>
<td>Guadalquivir</td>
<td>18</td>
<td>23</td>
<td>3</td>
<td>2</td>
<td>0.502</td>
<td>0.071</td>
<td>0.561</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>120</td>
<td>35</td>
<td>63</td>
<td>0.923</td>
<td>1.031</td>
<td>8.146</td>
</tr>
</tbody>
</table>

N, sample size; h, number of haplotypes; S, number of polymorphic sites; H, haplotype diversity; Pi, nucleotide diversity (expressed as percentages, i.e. 0.001 = 0.1%) and K, average number of pairwise nucleotide differences. Groups of populations suggested by SAMOVA are indicated in bold. (Western-CS = Western Central System; Eastern-CS = Eastern Central System; NC-IP = North Central Iberian Peninsula and SW-Huelva = Southwestern Huelva).
Supporting Information

Additional Supporting Information may be found in the online version of this article:

Supporting Information: Tables S1–S6, Figures S1.1–S6

Table S1 *Crocidura suaveolens* samples collected for this study.
Table S2 Primers used for the amplification of the cytochrome b gene.
Table S3 *Crocidura suaveolens* cytochrome b sequences downloaded from Genbank.
Table S4 Calibrations constrains, in Myr, used as priors in the BEAST2 analysis of cytochrome b of soricids.
Table S5 Populations of occurrence and frequencies of haplotypes found within clade IV of *Crocidura suaveolens* group.
Table S6 Matrix of population pairwise comparisons for Iberian population of *Crocidura suaveolens*.
Figure S1 Bayesian relaxed clock tree reconstructed with cytochrome b sequences of soricids.
Figure S2 a) Bayesian phylogenetic tree of the *Crocidura suaveolens* group. b) Amplified view of clade IV subtree.
Figure S3 a) Maximum-likelihood phylogenetic tree of the *Crocidura suaveolens* group. b) Amplified view of clade IV subtree.
Figure S4 Dated tree of *Crocidura russula*.
Figure S5 Bayesian skyline plots showing demographic histories of the different lineages / sublineages identified within clade IV.
Figure S6 Distributions of *Crocidura suaveolens* and *Crocidura russula* in different European biogeographic regions in western Europe.

Supporting Information: Supplementary Methods