Short Note

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

Riccardo Castiglia[1,*], Flavia Annesi[1], Giovanni Amori[2], Emanuela Solano[1], Gaetano Aloise[3]

[1]*Department of Biology and Biotechnology "Charles Darwin", University of Rome "La Sapienza", Via Borelli 50, 00161 Rome, Italy*
[2]*National Research Council, Institute of Ecosystem Study, Largo Tonolli 50, 28922 Verbania Pallanza, Italy*
[3]*Museo di Storia Naturale della Calabria e Orto Botanico, University of Calabria, Via Savinio s.n., 87036 Rende, Italy*

Abstract

A molecular phylogeographic study using a fragment of the mitochondrial gene for cytochrome b (cytb) was performed on the lesser white-toothed shrew, *Crocidura suaveolens*, from seven localities in central and southern Italy. Comparison with cytb European haplotypes revealed the absence of endemic lineages in the region, in contrast to what has been observed for many other Italian terrestrial vertebrates. Indeed all the Italian specimens results nested with Balkanic conspecific within an Italo-Balkan clade. Historical demography of this clade showed a scenario of expansion which preceded the LGM. This evidence of glacial persistence indicates a certain flexibility of the classic models of Pleistocene biogeography.

Southern European peninsulas (Italy, Balkans and Iberia) played an important role as refugial areas during Pleistocene cold periods, shaping the distribution of genetic variation of autochthonous European species (Hewitt, 2000). The lesser white-toothed shrew *Crocidura suaveolens* (Soricomorpha, Soricidae) is widely distributed in the Palearctic, extending from the Atlantic coast of Spain eastwards through Europe and Asia to Siberia. The phylogeography of the species has been thoroughly studied (Dubey et al., 2006, 2007) and it reveals the presence of ten parapatric clades across the species range. In Europe two divergent mtDNA lineages, can be found, the first derived from the Iberian refugium and the second of Italo-Balkan origin (Fig. 1). The latter lineage comprises haplotypes from Switzerland, Austria, Hungary, Bulgaria, Greece and localities in northern and central Italy (Dubey et al., 2007).

It is now well established that the southernmost part of the Italian Peninsula represents a "hotspot" of intraspecific diversity for many terrestrial vertebrates, with the presence of endemic genetic lineages (Amori et al., 2009; Vega et al., 2010; Bertolino et al., 2015). The most likely source of historical barriers causing this biogeographic pattern are the glacio-eustatic sea level oscillations throughout the Pleistocene, with consequent insularization of southern Italy during the multiple interglacial transgressions. Among mammals, species rank has been proposed for some of these southern Italian populations (*Sciurus vulgaris*, Grill et al., 2009; *Myodes glareolus*, Colangelo et al., 2010; *Microtus savii* "species group" Castiglia et al., 2008 and Bezerra et al., 2016).

For this reason, a genetic study of *C. suaveolens* from the southern part of the Italian refugium was carried out to provide a complete picture of its genetic diversity and to determine if it belongs to the Italo-Balkan clade, as previously suggested by Dubey et al. (2007), or if it represents an additional endemic lineage for southern Italy.
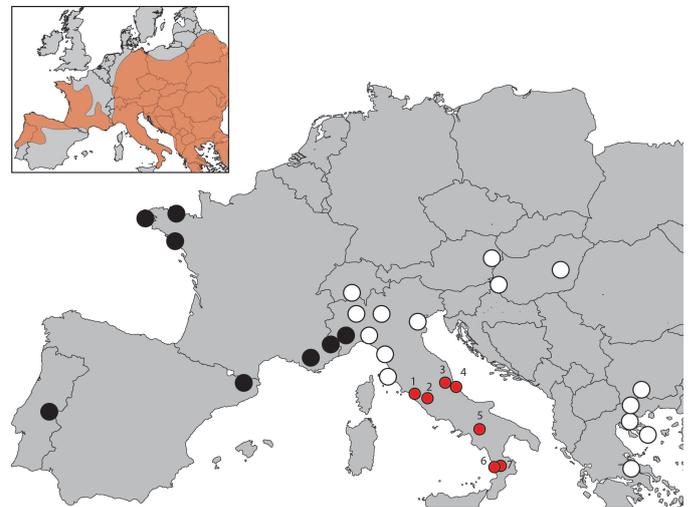
Figure 1 – Distribution of *Crocidura suaveolens* in Europe (top left, orange). The small red circles indicate the 7 localities from central and southern Italy studied in this work. The large circles indicate the localities belonging to the Iberian (black) and to the Italo-Balkan clade (white) studied in previous works (Dubey et al., 2006, 2007).

Twelve lesser white-toothed shrews from seven localities in central-southern Italy were analyzed (Fig. 1 and Tab. 1). Procedures for DNA extraction, amplification and sequencing followed Castiglia et al. (2007). A fragment (997 bp) of the mitochondrial gene for cytochrome b (cytb) was sequenced in six individuals (cro1, 3, 11–13, 19), using a combination of the universal primers L14723, L15408, H15553 and H15915. For the remaining specimens it was possible to sequence only a smaller fragment (476 bp) with primers L15408 and H15915. To place the southern Italian haplotypes in the European phylogenetic

*Corresponding author

Email address: riccardo.castiglia@uniroma1.it (Riccardo Castiglia)

Table 1 – Voucher (ID), name of region and locality (Locality), reference number to map in Fig. 1, Haplotype and GenBank accession number are reported for each specimen.

ID	Locality	Map	Haplotype	GenBank
cro1	Calabria, Figline (CS)	7	crs 1	KY677734
cro3	Calabria, Fiumefreddo (CS)	6	crs 1	KY677734
cro9	Basilicata, Muro Lucano (PZ)	5	crs 2	KY677735
cro11	Latium, Settebagni (RM)	2	crs 3	KY677735
cro12	Abruzzo, Pescara (PE)	4	crs 4	KY677737
cro13	Abruzzo, Gran Sasso (PE)	3	crs 5	KY677738
cro14	Latium, Settebagni (RM)	2	crs 6	KY677739
cro15	Latium, Settebagni (RM)	2	crs 7	KY677740
cro16	Latium, Macchiatonda (RM)	1	crs 7	KY677740
cro17	Latium, Macchiatonda (RM)	1	crs 6	KY677739
cro18	Latium, Macchiatonda (RM)	1	crs 6	KY677739
cro19	Latium, Macchiatonda (RM)	1	crs 3	KY677735

context, the 997 bp sequences were aligned with cytb haplotypes downloaded from GenBank (all the 113 haplotypes from Dubey et al., 2007; accession number references herein). A Bayesian analysis (BA) was conducted with MrBayes v. 3.2.1 (Huelsenbeck and Ronquist, 2001) under the General Time Reversible (GTR) model of evolution. The BA inference was performed by running 1000000 generations, with Markov chains sampled every 1000 generations. A burn-in of 10% was applied and the remaining trees were used to compute a 50% majority rule consensus tree and posterior probabilities. The mean pairwise distances (K2P model) between and within the clades of the phylogenetic analyses were estimated using MEGA v. 5.05 (Tamura et al., 2011).

For a detailed analysis of the relationships between our sample (476 bp) and the related haplotypes, a parsimony network of Clade VII were built (from Dubey et al., 2007), with the inclusion of our twelve individuals, with TCS v. 1.21 (Clement et al., 2000); default settings and connection limits at 95% were used to describe the relationships within the clade. On the same dataset, the population expansion by Mismatch Distribution (MMD) analysis was assessed to test if the current genetic variation best fits an equilibrium model, an expansion hypothesis or a recent bottleneck. The significance of the difference between the expected and observed MMD was tested by goodness-of-fit with the parametric bootstrap approach (Schneider and Excoffier, 1999). Rozas's R2 (Ramos-Onsins and Rozas, 2002) indices were used to test the population expansion under the hypothesis of neutral evolution.

Time since expansion was estimated by the τ value of the MMD (as $\tau = 2\mu t$) with two cytb divergence rates: 5.5% per Myr estimated for *Sorex* (Hope et al., 2010) and 2.3% per Myr estimated for *Neomys* (Igea et al., 2015).

Seven different haplotypes were identified (crs1–7, Tab. 1) among the twelve specimens, all new for *C. suaveolens* (Tab. 1). All the obtained haplotypes are nested in Clade VII, which shows maximum support (posterior probability=1) (tree not shown). Hereafter this Clade will be named as the Italo-Balkan clade. The mean pairwise *p* distances between the Italo-Balkan and the other clades range from a minimum of 2.8% with Clade VI to a maximum of 9.7% with Clade I in sequence divergence.

No geographic pattern can be observed in the parsimony network of the Italo-Balkan clade (Fig. 2). In fact, haplotypes from neighboring localities generally do not group together. Specifically, the Balkan and Italian haplotypes do not form separate groups. The number of substitutions separating adjacent haplotypes ranges from one to seven. The highest number of connections (three to five) is shown by three haplotypes from northern Italy, Switzerland and Bulgaria. The shape of the mismatch distribution is unimodal (Fig. 3). Rozas's R2 index is statistically significant (R2=0.0588).

Therefore, a scenario of population expansion in the past was inferred. This expansion, not documented for this clade by previous works, was possibly related to the recovery of the populations after a glacial period. The time since the start of expansion is indicated by the

modal value t=3.290 (in mutational units), calculated as 149000 ya for the slowest mutation rate and 60000 ya for the faster one.

The shared matrilineal genealogy of *C. suaveolens* specimens from Italy and the Balkans, but not from other European areas, supports the presence of a common Italo-Balkanic refugium. This pattern, seldom observed in mammals (e.g., the wood mouse *Apodemus sylvaticus*, Michaux et al., 2003), is believed to be caused by the frequent drops in the level of the Adriatic Sea during the Quaternary ice periods (Bortolami et al., 1977), which probably allowed the Italian and Balkan populations to remain in contact. From the available data, it is not possible to determine if one of the two peninsulas played a major role as a refugial area in this species. Central haplotypes of the network (i.e. the ones with the greatest number of connections) do not belong to southern Italy or the southern Balkans, as would be expected if one of these areas were the centre of expansion of the Italo-Balkan clade, and it cannot be excluded that the two peninsulas acted as a single refugium. The time of expansion of the *C. suaveolens* populations (60000–149000 ya) does not match the recovery of populations following the last glacial phases, but is more ancient with both the mutation rates used. This finding indicates that pre-LGM events such as glacial persistence were important in shaping the genetic variation of species, revealing a cer-
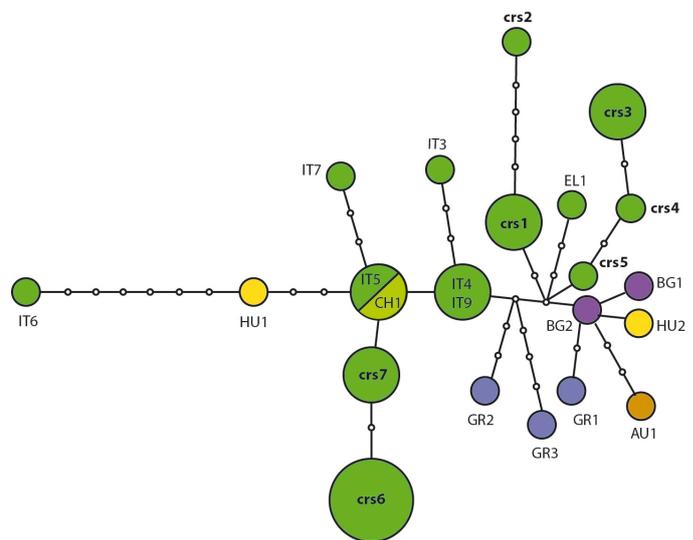
Figure 2 – Parsimony Network performed on the cytb sequences of *Crocidura suaveolens* from the Italo-Balkan Clade. Acronym "crs" indicate the central-southern Italian hapotypes (see Tab. 1 for details). The other haplotypes are named following Dubey et al. (2007) (AU, Austria; BG, Bulgaria; CH, Switzerland; EL, Elba; GR, Greece; HU, Hungary; IT, Northern Italy). Colors are as follows: Italy and Switzerland, green; Hungary, yellow; Austria, brown; Greece blue and Bulgaria violet.
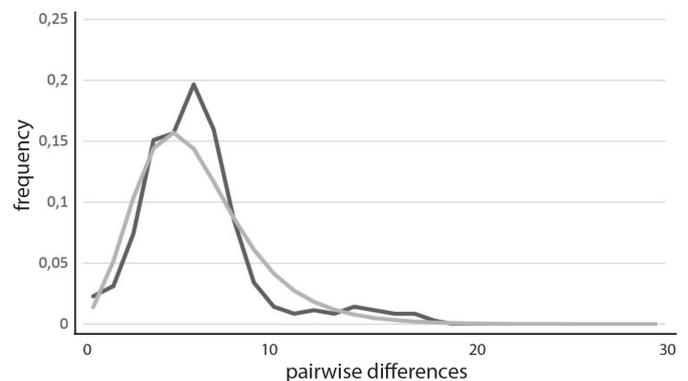
Figure 3 – Mismatch Distribution. Observed and expected distributions of pairwise differences under the model of population expansion for the Italo-Balkan Clade (see text for parameter details).

tain flexibility of the classic models of Pleistocene biogeography (e.g. Stewart et al., 2010).

Many of the studied species of small mammals display genetically divergent endemic lineages in the southernmost part of the Italian peninsula (Amori et al., 2009; Vega et al., 2010). Among shrews, *Neomys anomalus*, *N. fodiens* and *Sorex minutus* exhibit genetically divergent lineages in the region (Castiglia et al., 2007; Vega et al., 2010). The present study indicates that *C. suaveolens* is not one of these species. In this context it should be considered that the two Calabrian samples are settled south of the Crati-Sibari plain, an important suture zone that repeatedly acted as a barrier to dispersal of terrestrial fauna along the north-south axis (Canestrelli et al., 2010). For this reason, our sampling could be able to catch, with reasonably chance, the eventual Calabrian endemic lineages. This is somewhat surprising in view of the observed genetic differentiation in shrews promoted by ecological barriers (e.g. Jacquet et al., 2014), and the fact that *C. suaveolens* itself shows a puzzling pattern of divergent lineages in the Middle East (Dubey et al., 2007). The reasons for differences in phylogeographic patterns of co-distributed species could depend on many, still not well understood factors; however, they should be either linked to the ecological requirements of each species or be stochastic. Multi-species comparisons of LGM refugial locations inferred by Ecological Niche Modeling with those resulting from phylogeographic predictions may throw light on this question (Waltari et al., 2007). ✍

References

Amori G., Gippoliti S., Castiglia R., 2009. European non volant mammal diversity: conservation priorities inferred from mitochondrial DNA. Folia Zool. 58: 270–278.

Bertolino S., Colangelo P., Mori E., Capizzi D., 2015. Good for management, not for conservation: an overview of research, conservation and management of Italian small mammals. Hystrix 26(1): 25–35. doi:10.4404/hystrix-26.1-10263

Bezerra A.M.R., Annesi F., Aloise G., Amori G., Giustini L., Castiglia R., 2016. Integrative taxonomy of the Italian pine voles, *Microtus savii* group (Cricetidae, Arvicolinae). Zool. Scripta 45: 225–236.

Bortolami G.C., Fontes J.Ch., Markgraf V., Saliege J.F., 1977. Land, sea and climate in the northern adriatic region during late pleistocene and holocene. Palaeogeogr. Palaeoclimatol. Palaeoecol. 21: 139–156.

Canestrelli D., Aloise G., Cecchetti S., Nascetti G., 2010. Birth of a hotspot of intraspecific genetic diversity: notes from the underground. Mol. Ecol. 19: 5432–5451.

Castiglia R., Annesi F., Aloise G., Amori G., 2007. Mitochondrial DNA reveals different phylogeographic structures in the water shrews *Neomys anomalus* and *N. fodiens* (Insectivora: Soricidae) in Europe. J. Zool. Syst. Evol. Res. 45: 255–262.

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

Colangelo P., Aloise G., Franchini P., Annesi F., Amori G., 2010. Mitochondrial DNA reveals hidden diversity and an ancestral lineage of the bank vole in the Italian peninsula. J. Zool. 287: 41–52.

Dubey S., Cosson J.-F., Magnanou E., Vohralík V., Benda P., Frynta D., Hutterer R., Vogel V., Vogel P., 2007. Mediterranean populations of the lesser white-toothed shrew (*Crocidura suaveolens* group): an unexpected puzzle of Pleistocene survivors and prehistoric introductions. Mol. Ecol. 16: 3438–3452.

Dubey S., Zaitsev M., Cosson J.-F., Abdukadier A., Vogel P., 2006. Pliocene and Pleistocene diversification and multiple refugia in a Eurasian shrew (*Crocidura suaveolens* group). Mol. Phylogenet. Evol. 38: 635–647.

Grill A., Amori G., Aloise G., Lisi I., Tosi G., Wauters L.A., Randi E., 2009. Molecular phylogeography of European *Sciurus vulgaris*: refuge within refugia? Mol. Ecol. 18: 2687–2699.

Hewitt G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405: 907–913.

Hope A.G., Waltari E., Dokuchaev N.E., Abramov S., Dupal T., Tsvetkova A., Henttonen H., MacDonald S.O., Cook J.A., 2010. High-latitude diversification within Eurasian least shrews and Alaska tiny shrews (Soricidae). J. Mammal. 91: 1041–1057.

Huelsenbeck, J.P., Ronquist F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17: 754–755.

Igea J., Aymerich P., Bannikova A.A., Gosálbez J., Castresana J., 2015. Multilocus species trees and species delimitation in a temporal context: application to the water shrews of the genus *Neomys*. BMC Evol. Biol. 15: 209.

Jacquet F., Nicolas V., Colyn M., Kadjo B., Hutterer R., Decher J., Akpatou B., Cruaud C., Denys C., 2014. Forest refugia and riverine barriers promote diversification in the West African pygmy shrew (*Crocidura obscurior* complex, Soricomorpha). Zool. Scripta 43: 131–148.

Michaux J.R., Magnanou E., Paradis E., Nieberding C., Libois R.M., 2003. Mitochondrial phylogeography of the woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. Mol. Ecol. 12: 685–697.

Ramos-Onsins, S.E., Rozas J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19: 2092–2100.

Schneider S., Excoffier L., 1999. Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. Genetics. 152: 1079–1089.

Stewart J.R., Lister A.M., Barnes I., Dalén L., 2010. Refugia revisited: individualistic responses of species in space and time. Proc R Soc Lond [Biol]. 277: 661–671.

Strahler A.N., Strahler A.H., 1989. Geografia física. 3ª ed. Barcelona, Omega. [in Italian]

Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739.

Vega R., Amori G., Aloise G., Cellini S., Loy A., Searle J.B., 2010. Genetic and morphological variation in a Mediterranean glacial refugium: evidence from Italian pygmy shrews, *Sorex minutus* (Mammalia: Soricomorpha) Biol. J. Linn. Soc. 100: 774–787.

Waltari E., Hijmans R.J., Peterson A.T., Nyári Á.S., Perkins S.L., Guralnick R.P., 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoSONE 2: e563. doi:10.1371/journal.pone.0000563