DOTTORATO DI RICERCA IN ETOLOGIA, ECOLOGIA, ANTROPOLOGIA E BIOSISTEMATICA
(XXVIII CICLO)

Phylogeography and population genetics of the European pond turtle *Emys orbicularis* and the Sicilian pond turtle *Emys trinacris*

Tesi di

Thania Manfredi

Coordinatore Prof. Alberto Ugolini
Tutors Prof. Guido Chelazzi e Dott. Claudio Ciofi

(2015)

UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN ETOLOGIA, ECOLOGIA, ANTROPOLOGIA E BIOSISTEMATICA

CICLO XXVIII

COORDINATORE Prof. Ugolini Alberto

Phylogeography and population genetics of the European pond turtle *Emys orbicularis* and the Sicilian pond turtle *Emys trinacris*

Settore Scientifico Disciplinare BIO/07

Dottorando
Dott.ssa Manfredi Thania

Tutore
Prof. Chelazzi Guido

Tutore
Dott. Ciofi Claudio

Coordinatore
Prof. Ugolini Alberto

Anni 2012/2015

SUMMARY

ABSTRACT

Global environmental change is widely well-know. It consists in marked alteration of the abiotic environment due to massive land use, chemical pollutants and emerging climate change, but also consists in the transformation of the composition of biotic communities. Biodiversity is undergoing deep and complex effects which are difficult to oppose for economic, cultural and aesthetics reasons.

The European pond turtle is the only Palearctic representative of Emydidae that inhabits the Americas, and despite its broad range, is a threatened species. In most countries, pond turtles had a breakdown in the 18th or 19th century, above all caused by human activities, and are often already extinct locally. Knowledge on the status of *Emys orbicularis* and its conservation actions is varied. The European pond turtle has so far been studied at the European level but only recently have fine-scale phylogenetic studies been conducted at just a few Italian sites.

I studied phylogeography and population genetics of the genus *Emys* in Italy (the sole endemic turtle) using a comprehensive sample set. The aim of this work is to reconstruct patterns of dispersal and colonization of Italian turtles following the last glacial period that brought their current distribution to our country and characterized levels of genetic divergence among the main Italian natural populations, through the analysis of the mitochondrial DNA control region and microsatellite markers. This knowledge is essential for identifying genetically diversified units, evolutionary significant units (ESU) and to encourage conservation plans, given that, despite the rise in attention and actions for safeguarding this species in recent years, there are many populations living in unsuitable areas. Moreover, acquiring this information allowed us to build a genetic database used to define the origin on Italian territory of individuals kept in captivity and then assist reintroduction programs or demographic reinforcement of natural populations. We also performed a study on a Greek population of 314 *Emys orbicularis* in the nature reserve of Lake Kerkini (Macedonia), highly impacted by human activities and inhabiting a series of ponds along a river within a few kilometers from each other to assess a fine-scale population genetic structure and movements of turtles among ponds to provide additional information on patterns of dispersal along linear habitats.

My results demonstrated the division of Italian turtles into three clusters, mostly identifiable with the subspecies of *E. o. galloitalica*, *E. o. hellenica* and the species *Emys trinacris*, which, however, showed various degrees of admixture and certain past gene flow (increasing towards the south) which nowadays has become difficult due to high fragmentation. The populations that merit immediate attention, having the lowest values of genetic diversity, were in Sardinia, Sicily and the recent rediscovered population of Albenga. Phylogeographic data prove the existence not only of a primary

refugia in the South but also of a secondary refugia, in a less rigid glacial period, in Tuscany and Latium, sources from which the dispersion started via wide wetlands that characterized the Italian coastal areas, and have disappeared today, as have the populations who inhabited them, because of the great reclamations started in the early twentieth century. Another unproven fact to date is the differentiation of Sardinian populations notwithstanding the past connection with Tuscany, and we also consider the hypothesis for introduction to be likely unreliable. The extensive genotyping of *Emys* Italian populations done through our research is important given that translocations of turtles were and are at present not negligible, and given the vulnerability of the species. Our results have been successfully used in planning the reintroduction of several animals from recovery centers back to nature and will surely continue to contribute to the awareness of reintroducing individuals. Finally, a study on Kerkini Lake populations, that display medium high values of genetic diversity, showed significant differences between males and females in their movements and relations between different ponds, and also that gene flow is favored by streams. This highlights the importance of land reinstatement that separates the ponds from cattle invasion and the importance of riverine connections especially during the breeding season which coincides with the period of greatest water withdrawal by surrounding farms.

Population status of the European and Sicilian pond turtle is highly diversified and, although still not critical in many countries, requires further investigation that includes using molecular techniques in order to understand the subtle dynamics, and to intervene in the best way possible for the conservation of the species.

INTRODUCTION

The European pond turtle *Emys orbicularis* (Linnaeus,1758) and the Sicilian pond turtle *Emys trinacris* (Fritz *et al.* 2005b), are the only representatives in Europe of the 51 species of the Emydidae family, distributed exclusively in the New World.

E. orbicularis inhabits a wide area (Fig. 1), approximately 6000 km in a west-east direction, from North Africa Atlantic coast, southern and western Europe, to Asia Minor and Caspian and Aral Seas, and 2000 km in a north-south direction from the region of Moscow to the Turkish-Syrian border, while *E. trinacris* is found exclusively in Sicily (Fritz *et al.* 2005b).

Until 1989 the European pond turtle was considered a monotypic species, currently 14 subspecies are recognized (Fritz 1998; Jesu *et al.* 2004). Northern areas host *E. o. orbicularis*, as does the Iberian Peninsula which is also populated by *E. o. hispanica* and *E. o. fritzjuergenobsti*, North Africa by *E. o. occidentalis*, Mediterranean west coast by *E. o. galloitalica*, Mediterranean east coast by *E. o. hellenica* and eastern range by *E. o. iberica*, *E. o .persica*, *E. o. colchica*, *E. o. luteofusca* and *E. o. eiselti*. Confined to the Liguria region (Italy) is *E. o. ingauna*, believed extinct in the second half of XXth century and rediscovered near Albenga in 2004 (Jesu *et al.* 2004), *E. o. capolongoi* in Sardinia (Italy) and *E. o. lanzai* in Corsica (France). A more specific range of subspecies of the European pond turtle are shown in Table 1.

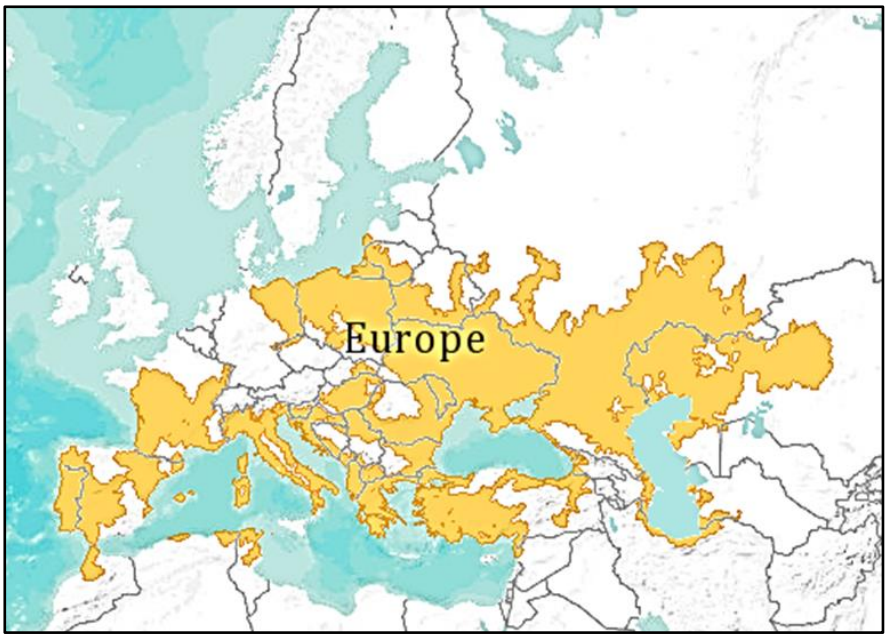

Figure 1. Distribution of *Emys orbicularis.*

The individual subspecies differ in size, shell, head proportion and color. The turtle is medium sized from 12 to 35 cm of carapace length and females are larger than males. Plastra is mainly yellow with black or dark green plastron and skin with lots of tiny yellow dots. Females have a more pronounced color than males; the head, in particular, is more speckled in yellow and the jaw is often completely yellow. Males have yellowish, reddish or white irises.

Freshwater ecosystems, that constitute the 0.01% of the world's water, with one third of all vertebrate species and high levels of endemism, are the most threatened environments (Dudgeon *et al.* 2006; Mooney & Cleland 2001; Sala *et al.* 2000; Vitousek *et al.* 1996; Vitousek *et al.* 1997). Biodiversity decay in freshwater areas, due to overexploitation, water pollution, flow modification, destruction or degradation of habitat and invasion by exotic species, is greater than in terrestrial surroundings. *Emys orbicularis* live in various water bodies: ponds, lakes, swamps, wetlands, canals and slow streams. Indeed it is on the IUCN Red List of threatened species as "Near threatened", in the Bern Convention and Habitat Directive. Terrestrial environment still plays an important role for all freshwater turtles (Ficetola & De Bernardi 2006; Ficetola *et al.* 2004). Usually, *E. orbicularis* is considered to be an aquatic species since some of their most conspicuous behaviors, such as basking or mating, occur in or very close to the water. However, *E. orbicularis* is, among Emydid turtles, the one that can move furthest from wetlands. It has been reported that European pond turtles can move up to 4 km away from their wetland for nesting and that terrestrial migrations exceeding 1 km are not uncommon (Lebboroni & Chelazzi 1998; Rovero & Chelazzi 1996; Schneeweiss & Steinhauer 1998). Upland environment is also used for aestivation, hibernation and spring dispersion that is predominated by males presumably to enhance reproductive opportunities. Moreover, the abundance of terrestrial insects in the diet of *E. orbicularis* found by Ottonello *et al.* (2005) suggests that upland environment can also be important for feeding.

For these species the main causes of danger are habitat deterioration and fragmentation, due to human impact like urbanization, farming and agricultural purposes, but also due to predation of eggs and juveniles or competitive pressure from exotic species, including the raccoon (*Nyctereutes procyonoides*), black bass (*Micropterus salmoides*), carp (*Cyprinus carpio*) and red eared slider (*Trachemys scripta*). The latter, native to southern Midwest U.S.A, included in the 100 World's Worst Invasive Alien Species of IUCN, is more aggressive than the native turtle. Recent studies suggest acclimatization and successful breeding in nature of introduced sliders that, above all, show a greater competitive ability in vying for feeding resources, occupying preferred basking sites and they are a sizable carrier parasite transmitter (Cadi & Joly 2004; Ficetola *et al.* 2009; Perez-Santigosa *et al.* 2008; Polo-Cavia *et al.* 2010, 2011; Verneau *et al.* 2011). Another source of risk seems to be restocking with genetically incompatible individuals.

In summarizing the current efforts in distinct European countries of Fritzand Chiari (2013) the different status of this species is evident. In Switzerland and the United Kingdom natural populations do not exist, in Germany and Slovakia only one reproducing population survives in Brandeburg and Tajba National Pond Reserve, where no conservation activities are reported. In other regions the situation is more stable. A number of recent management plans and European Union conservation and land management projects try to buffer the impact of habitat enchroacement. These projects promote studies on distribution and activities, like identifying unknown populations, restoring and creating new suitable habitats, captive breeding programs and also eradication of alien turtles. A good example is Spain where *E. orbicularis* is not on the IUCN Red List.

Unfortunately in Italy, in spite of several conservation projects financed by local authorities and Life projects (DUNETOSA-LIFE05, Co.Me.Bis.-LIFE06, CILENTO in rete-LIFE06, LIFE FRIULI FENS-LIFE06, DINAMO-LIFE08, ORISTANESE-LIFE08, SORBA-LIFE08, Re.S.C.We-LIFE09, NATURA 2000 in the Po Delta-LIFE09, MAESTRALE-LIFE10, LIFEEMYS-LIFE12), almost all in the northern region, the European pond turtle currently has a fragmented and uneven distribution, it is listed as "Endangered" in the IUCN Red List of threatened species.

The largest populations occur in protected areas of the Po River Delta, Tuscany, Latium, Campania and Calabria (Mazzotti & Zuffi 2006) and generally inhabit lowland areas, but can also be found at up to 1500 m of altitude like in Calabria and Sicily. On a small geographic scale it shows the more complicated diversification of the entire areal.

These data highlight the importance of an accurate study of genus *Emys* in Italy. We focused on genetic characteristics to define genetic differences between and within main Italian populations, their structure and on phylogeography (Avise 2000) to reconstruct dynamics of dispersion and colonization. Patterns of neutral genetic variation among individuals carry the signature of a species' demographic past. The degree and type of divergence in the DNA sequence of populations of different geographical regions may reflect the pattern in which geological events, climatic and ecological, also led to their current distribution on the territory. This knowledge is essential in identifying genetically diversified units, to encourage conservation plans for reintroduction and demographic reinforcement and to determine the origin of unknown individuals, thanks to the creation of a databank. We also assessed fine-scale population genetic structure and movements of turtle among ponds to provide additional information on patterns of dispersal along linear habitats by freshwater turtles.

Table 1. Areal of *Emys orbicularis* subspecies.

Emys orbicularis orbicularis	Central Europe, from eastern Germany through Poland, Baltic States and Russia eastwards to the Aral Sea, as well as central France and the Danube basin; apparently disappeared from the Rhine basin
Emys orbicularis hispanica	Donana of South-west Spain and probablly all other Atlantic drainages on the Iberian Peninsula
Emys orbicularis fritzjuergenobsti	Mediterranean drainages of Spain
Emys orbicularis occidentalis	Northern Africa: northeastern Morocco, northern Algeria and northern Tunisia
Emys orbicularis galloitalica	Mediterranean coastal region from Catalonia (Spain) through southern France and the Tyrrhenian coast of Italy
Emys orbicularis ingauna	Liguria
Emys orbicularis lanzai	Corsica
Emys orbicularis capolongoi	Sardinia
Emys orbicularis hellenica	Adriatic coastal region southwards to Peloponnesus and Bootia (Greece); possibly including populations of the southern Crimea and western European Turkey. This includes a population on the island of Limnos (Greece)
Emys orbicularis iberica	Southern Dagestan (Russia), Kura drainage basin and Kura-Araksinker depression of Georgia and Azerbaijan
Emys orbicularis persica	Northern Iran and adjoining Turkmenistan
Emys orbicularis eiselti	Gaziantep and southeastern Anatolia
Emys orbicularis colchica	Southeastern Balkans, Turkish Black Sea coast and parts of western and central Anatolia, Colchis, and possibly various areas in eastern South Ukraine
Emys orbicularis luteofusca	Southern Central Anatolian high plain

REFERENCES

Avise JC (2000) Phylogeography:The history and formation of species. Harvard University Press, Cambridge, MA.

*Cadi A, Joly P (2004) Impact of the introduction of the red-eared slider (Trachemys scripta elegans) on survival rates of the European pond turtle (Emys orbicularis). Biodiversity and Conservation **13**, 2511-2518.*

*Dudgeon D, Arthington AH, Gessner MO, et al. (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews **81**, 163-182.*

*Ficetola GF, De Bernardi F (2006) Is the European "pond" turtle Emys orbicularis strictly aquatic and carnivorous? Amphibia-Reptilia **27**, 445-447.*

*Ficetola GF, Padoa-Schioppa E, Monti A, et al. (2004) The importance of aquatic and terrestrial habitat for the European pond turtle (Emys orbicularis): implications for conservation planning and management. Canadian Journal of Zoology-Revue Canadienne De Zoologie **82**, 1704-1712.*

*Ficetola GF, Thuiller W, Padoa-Schioppa E (2009) From introduction to the establishment of alien species: bioclimatic differences between presence and reproduction localities in the slider turtle. Diversity and Distributions **15**, 108-116.*

Fritz U (1998) Introduction to zoogeography and subspecific differentiation in Emys orbicularis (Linnaeus, 1758). In: eds Fritz U, Joger U, Podloucky R, Servan J) Proceedings of the Emys Symposium Dresden 96. Mertensiella 10, pp. 1-27. Warlich, Rheinbach.; IUCN Red List of threatened species; Fritz 2001.

*Fritz U, Chiari Y (2013) Conservation actions for European pond turtles - a summary of current efforts in distinct European countries. Herpetology Notes **6**, 105-105.*

*Fritz U, Fattizzo T, Guicking D, et al. (2005b) A new cryptic species of pond turtle from southern Italy, the hottest spot in the range of the genus Emys (Reptilia, Testudines, Emydidae). Zoologica Scripta **34**, 351-371.*

Jesu R., Piombo R., Salvidio S., Lamagni L., Ortale S., Genta P. (2004) Un nuovo taxon di testuggine palustre endemico della Liguria Occidentale. Ann. Mus. Civ. St. Nat. "G. Doria" Genova.

Lebboroni, M., Chelazzi, G. (1998) Habitat use, reproduction, and conservation of Emys orbicularis in a pond system in Central Italy. In: Ponds & Pond Landscape of Europe. Proceedings of the International Conference of the Pond Life Project. Boothby, J. Ed., Maastricht.

Mazzotti S. & Zuffi M., (2006) Emys orbicularis pp. 377-381. In: F. Barbieri, G. Doria & R. Sindaco (eds), Atlante degli Anfibi e dei Rettili d'Italia. Societas Herpetologica Italica. Polistampa, Firenze, pp. 376-381.

*Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. Proceedings of the National Academy of Sciences of the United States of America **98**, 5446-5451.*

*Ottonello D, Salvidio S, Rosecchi E (2005) Feeding habits of the European pond terrapin Emys orbicularis in Camargue (Rhone delta, Southern France). Amphibia-Reptilia **26**, 562-565.*

*Perez-Santigosa N, Diaz-Paniagua C, Hidalgo-Vila J (2008) The reproductive ecology of exotic Trachemys scripta elegans in an invaded area of southern Europe. Aquatic Conservation-Marine and Freshwater Ecosystems **18**, 1302-1310.*

*Polo-Cavia N, Lopez P, Martin J (2010) Competitive interactions during basking between native and invasive freshwater turtle species. Biological Invasions **12**, 2141-2152.*

*Polo-Cavia N, Lopez P, Martin J (2011) Aggressive interactions during feeding between native and invasive freshwater turtles. Biological Invasions **13**, 1387-1396.*

*Rovero F, Chelazzi G (1996) Nesting migrations in a population of the European pond turtle Emys orbicularis (L) (Chelonia Emydidae) from central Italy. Ethology Ecology & Evolution **8**, 297-304.*

*Sala OE, Chapin FS, Armesto JJ, et al. (2000) Biodiversity - Global biodiversity scenarios for the year 2100. Science **287**, 1770-1774.*

Schneeweiss, N., Steinhauer, C. (1998): Habitat use and migrations of a remnant population of the European pond turtle, Emys o. orbicularis (Linnaeus, 1758), depending on landscape structures in Brandenburg, Germany. In: Proceedings of the EMYS Symposium Dresden 96. Mertensiella 10, p. 235-243. Fritz, U., Joger, U., Podlucky, R., Servan, J., Buskirk J.R., Eds., DGHT, Rheinbah.

*Verneau O, Palacios C, Platt T, et al. (2011) Invasive species threat: parasite phylogenetics reveals patterns and processes of host-switching between non-native and native captive freshwater turtles. Parasitology **138**, 1778-1792.*

*Vitousek PM, Dantonio CM, Loope LL, Westbrooks R (1996) Biological invasions as global environmental change. American Scientist **84**, 468-478.*

*Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of Earth's ecosystems. Science **277**, 494-499.*

1. PHYLOGEOGRAPHY AND EVOLUTIONARY HISTORY OF EUROPEAN POND TURTLES (*Emys orbicularis and E. trinacris*) IN ITALY

ABSTRACT

We investigated the genetic diversity distribution of the European pond turtle *Emys orbicularis* and the Sicilian pond turtle *Emys trinacris* in Italy, one of the three main refuges during the Quaternary Ice Age, of which little is known, to infer the evolutionary history and processes that have contributed to shaping the patterns of diversity, and the current distribution of native populations within the southern European Peninsula. A total of 784 turtles were analyzed for the mitochondrial DNA control region and 1029 for 16 microsatellite loci, covering the main Italian populations. Three principal lineages were observed: the Sicilian and other two each characterized one of the Italian coasts, confirming the Apennines, especially the northern, as the main barrier to movements. As expected, the South has the typical features of refugial areas with high variability, but also the central areas of the Tyrrhenian coast and Sardinia shows a richness of diversity, identifiable as secondary refugia. The nested clade phylogeographic analysis, which especially solved the history of the Tyrrhenian and Sicilian cluster, show a scenery of restricted gene flow with some long-distance dispersal followed by fragmentation and/or extinction of intermediate populations. On the Italian Peninsula the greatest impact on these species is certainly due to human impact and therefore to the reduction and fragmentation of their habitat. Populations with the lowest observed heterozygosity were found are in Liguria and on the islands, showing greater diversification from the Peninsula, particularly the Sardinia populations, which deserve particular conservation efforts.

INTRODUCTION

The Earth's climate became cooler in the Tertiary (65-2.6 Ma), when tectonic activity lead continents and oceans to similar present day configuration. This time was characterized by a series of major ice ages with a 100 ka periodicity and relatively short warm interglacial spells, like at present. The Croll-Milankovitch theory proposed orbital eccentricity of the earth around the sun as the start of the cycles (100 ka cycle), which caused major changes in insolation, combined with axial tilt (41 ka cycle) and precession (23 ka cycle) (Hewitt 1999). Within major 100 ka cycles rapid oscillations of changes of 7–15°C can take place in a few decades that persist for centuries, as happened 11 ka. At the Last Glacial Maximum LGM (23-18 ka) the European ice sheet extended south to 52° N and permafrost south to 47 ° N.

These severe climatic oscillations produced great changes in species distribution, latitudinally as the ice sheets advanced and retreated, longitudinally when new dispersal routes became available, like the Bering Land Bridge and Danube River Valley, and altitudinally. For temperate species in Europe, the southern peninsulas of Iberia, Italy, the Balkans and to a smaller extent the Ponto-Caspian region, were important refugia from which recolonization of

northern regions started. These refugial zones contributed differently to the colonization of Europe by different species. Balkan lineages predominate the Iberian and Italian. This is probably due to the major mountain barrier of Alps and Pyrenees. It is possible to summarize a pattern of colonization in three examples: "grasshopper", "hedgehog" and "bear" (Fig. 1). *Chorthippus parallelus* populated most of Europe from the Balkan refugium, with Iberian and Italian genomes blocked at the Pyrenees and Alps (Lunt *et al.* 1998). *Erinaceus* recolonized with three lineages from the three principal areas (Santucci *et al.* 1998; Seddon *et al.* 2002). *Ursus arctos* originated from Iberia and the eastern refugia of Ponto-Caspian (Sommer & Benecke 2005).

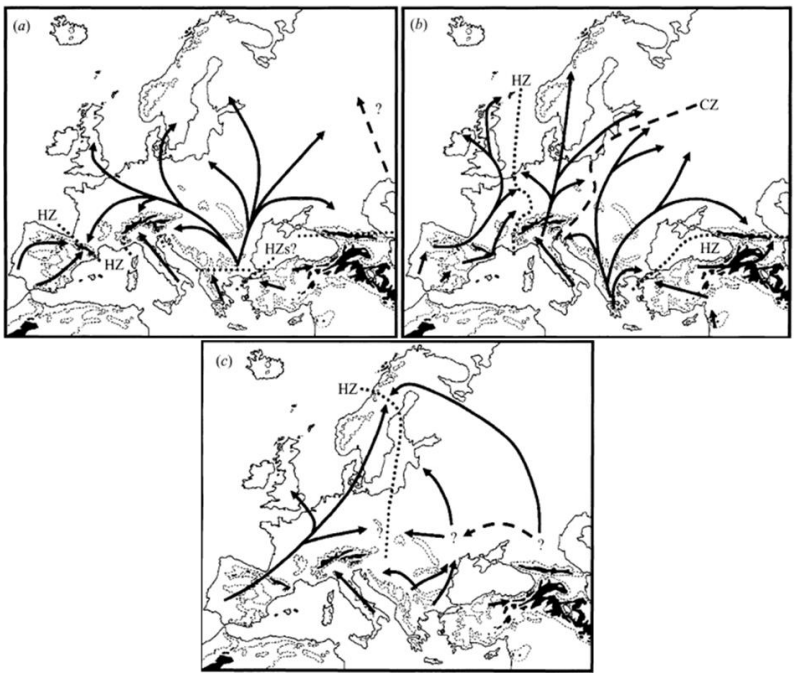

Figure 1. Possible postglacial colonization routes of Europe from refugial areas for three exemplar species, (a) meadow grasshopper—Chorthippus parallelus, (b) hedgehog— Erinaceus europaeus/concolor, (c) bear—Ursus arctos. Altitudes over 900 m are indicated with dotted lines and over 1800 m in black. Contact Zones (CZ) between taxonomic species are shown as heavy dashed lines and Hybrid Zones (HZ) between distinct genomes are shown as heavy dotted lines (Hewitt 2004).

European regions are inhabited by mixed biotas of species each derived from different refugia. Repeated ice ages in the Pleistocene caused range contractions and expansions and brought many times species in and out of refugia in the South (Hewitt 2011). Sometimes more adverse periods may have

cause extinction in any of those with consequent restocking from more distant refugia.

Genetic consequences of climatic oscillations in the Quaternary usually led to a south–north gradient in genetic diversity; Northern Europe shows less genetic variety than southern Europe in numbers of species, subspecific division and allelic diversity, "southern richness to northern purity" (Hewitt 1999). The northward postglacial expansion, often very rapid, led to a reduction in genetic diversity, accentuated by central clade that expanded range in any available space blocking the southern, where instead there are several allo-parapatric lineages. Another emerging scenario that often best explains the current redistribution of genetic structure at inter and intraspecific levels is "refugia-within-refugia" (Gomez & Lunt 2007), with several allopatric lineages that in subsequent interglacials periods can admix due to secondary contacts.

History of *Emys orbicularis*

Reptiles, which have moderate mobility and are effectively blocked by geomorphological barriers, are good indicator organisms for phylogeographic patterns which are not restricted to a particular taxonomic group. For them extrinsic factors predominate over intrinsic. Moreover, no evident differences were found in patterns of aquatic reptiles compared to terrestrial (Joger *et al.* 2007).

Biogeography, fossil, paleogeological data and the latest molecular techniques have allowed us to reconstruct the evolutionary history of a variety of species.

Both mithocondrial and nuclear genomes suggest that modern testudinoids started to diverge before the Cenozoic, about 170.0 Ma (Lourenco *et al.* 2012). The first divergences within this group, Cryptodire radiation, seemed to occur in mid-Cretaceus and between families through Paleocene and Early Eocene. Modern testudinoid families evolved from a common ancestor in Asia (Claude *et al.* 2003) and then dispersed to other continents probably via the Bering Land Bridge in Late Cretaceous and Paleocene and via the North Atlantic Land Bridge in Greenland. The first certain Emydid fossil in North America dates from Early Eocene. Divergence among ancestral *Emys* lineages occurred about 29-23 Ma, corresponding to the late Oligocene warming period, and the simultaneous presence in Europe and North America dates from about 15 Ma, consistent with the split of European orbicularis from its American sister group of 17 Ma, time supported by both mtDNA and nuDNA (Spinks & Shaffer 2009). The immigration from the New to the Old World presumably occurred in the Middle or Upper Tertiary, as out of Asia, mainly via the Beringian Land Bridge that was exposed about 24–14 Ma (Sanmartin *et al.* 2001) or the Thulean and DeGreer Land Bridges, but these routes were open much earlier (~65–39 Ma), and the climate during this period was probably too cold for northerly turtle migrations (Burbrink & Lawson 2007). The oldest fossils of

Emys in Europe, originally described as Emydoidea, *E. tarashchuki* (Chkhikvadze 1980) Middle to early Upper Miocene of Kazakhstan, *E. sukhanovi* (Chkhikvadze 1983) Upper Miocene of East Europe and *E. orbicularis antiqua* (Khosatzky 1956) Middle Pliocene to Villafranca of East and Central Europe, represent most probably chronotaxa leading to *E. orbicularis* which is known from Pleistocene (Fritz 1998).

The Tethys was an open equatorial sea separating Eurasia and Africa during the Jurassic and Cretaceous, joining the ancient Atlantic and Indo-Pacific oceans. As a result of the movement of the Eurasian and African plates, the seaways between the oceans and the Tethys were reduced or interrupted and, in the Upper to Middle Miocene Age, two principal inner marine basins were formed: the Mediterranean and the Paratethys. In the Middle Miocene the orogeny of the Alps, the Dinaric, Balkanic and Pontic mountain ranges, separated these seas. The pond turtle, after the invasion of Asia, supposedly spread through the north of the Paratethys Sea. Though the latter being mainly oligosaline or even a freshwater sea could have allowed colonization of western Europe, taking into account that marine straits, like the Strait of Gibraltar, aren't absolute barriers for *E. orbicularis* (Lenk *et al.* 1999a). During Messinian (6.1-5.1 My) two important geologic events, a major build-up of the West Antarctic ice sheet and the approach of the African and European plates, resulted in a world-wide eustatic Oceanie level lowering and in the isolation of the Mediterranean from the Atlantic; the Mediterranean became an evaporative basin "Messinian salinity crisis". At the end of Messinian, the freshwater Paratethys drained into the still arid Mediterranean and both seas were reduced to a network of lakes. This event termed "Lago Mare" phase which persisted about 0.1 My played an essential role in the early penetration by Paratethyan freshwater organisms and in dispersion into peri-Mediterranean river systems (Bianco 1990).

Numerous fossil records of Early Pleistocene from almost the entire Italian Peninsula, reveal the existence of pond turtles since approximately 1.8 Ma, while fossils from Sicily are younger and date back to early Middle Pleistocene. Lanza (1983) attests that *Emys* reached Corsica and Sardinia likely during the Cassian regression, Middle Pleistocene, when the two islands were certainly in contact with the continent via a bridge through Elba (Tuscany). Therefore, ancestors of recent taxa occurred in Italy prior to the Pleistocene climatic oscillations, that started about 700,000 years BP, prove that this region not only acted as a glacial refugia but also as a suitable habitat (Fritz *et al.* 2005b).

In the Pleistocene, the alternation of glacial and interglacial phases caused eustatic Oceanic level fluctuations, about 100-200 m sea lowering in glacial-maxima stages, and allowed connections between lands and islands and to other territories through numerous coastal corridors, such as Italy-France by the Ligurian coast or Italy-Balkans by the dry Adriatic basin. Furthermore, river captures or inversions of flow direction, due to ice dams may have occurred

between low valleys joining opposite sides of mountain ranges such as the Alps, where glaciers were building up. Italy in the Neozoic Era, during various climate cycles and tectonic processes, underwent significant changes (Amorosi *et al.* 2004; Correggiari *et al.* 1996; Ferranti *et al.* 2006; Giraudi (2004). For the duration of the coldest ages, ice covered northern Italy, pushing up the Po Valley; the Apennine was occupied by scattered glaciers making it a historical–geographical barrier to dispersal along the north–south axis of the Peninsula and a lowering of the sea level, about 100-130 m was estimated during the last Würm glacial phase (15-18 000 years ago); the Adriatic sea level reached its lowest point, about 1/7 of its current extent, and was almost completely in subaerial conditions. In warm periods Italian lands were submerged by sea and the various islands were isolated. The maximum marine ingression was about 5 ka BP. Figure 2 shows an example of the two different stages.

Emys orbicularis reached its maximum extent range in the Early Holocene, during the Boreal (8600-7100 cal BC) and Atlantic (7100-3750 cal BC) ages, thanks to favorable climatic conditions (Sommer *et al.* 2007). Increasing annual mean temperatures, 1-2.5 °C higher than today, allowed fast immigration of trees and development of rich hydrogeographic systems. Presumably pond turtles had a rapid range expansion by long distance dispersal events using rivers rising in the south and debouching in the north, like the Danube and Oder Rivers. In this period *E. orbicularis* even inhabited Britain, Denmark, Sweden and Estonia. Figure 3 shows Holocene fossil records and the current northern distribution border of native populations.

Extinction in northern parts of the range was mainly caused by climatic deterioration and, taking into account that *E. orbicularis* is quite cold-tolerant in winter—extent northernmost populations reached 57-56°N in Russia—it is plausible that lower summer temperatures predominantly affect successful breeding. This is proved by early extinction in Sweden which happened around 500 years before the end of the assumed climatic optimum of the northern hemisphere. Indeed pond turtles survived the cold winters in 8.2 ka, but not the cooler summers, nearly 1 °C lower, about 5.5 ka, as indicated by the temporary decline of the pine tree line below the 100 m level (Sommer *et al.* 2009). In Central Europe anthropogenic impact acted as the main force behind extinction.

Phylogeographic studies

The majority of vertebrate phylogenetic and phylogeographic analyses have been based almost exclusively on mitochondrial DNA, mainly because it is non-recombining and has a rapid coalescent time, and nuclear DNA markers have been available for only a handful of model organisms. Avise *et al.* (1992) was the first to suggest lower levels of variability and differentiation in turtles and lower microevolutionary rates in mithocondrial DNA (0.4% sequence divergence/million years) than in other vertebrate, probably due to a lower mutation rate or to higher generation times.

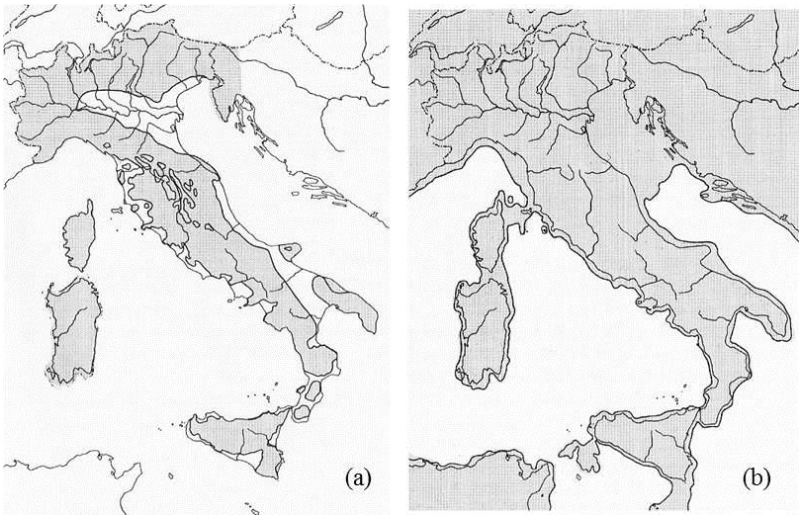

Figure 2. (a) Italy at the beginning of Quaternary 2 Ma presented only emerged Alps and Apennines ridges and (b) during maximum withdrawal of sea during LGM.

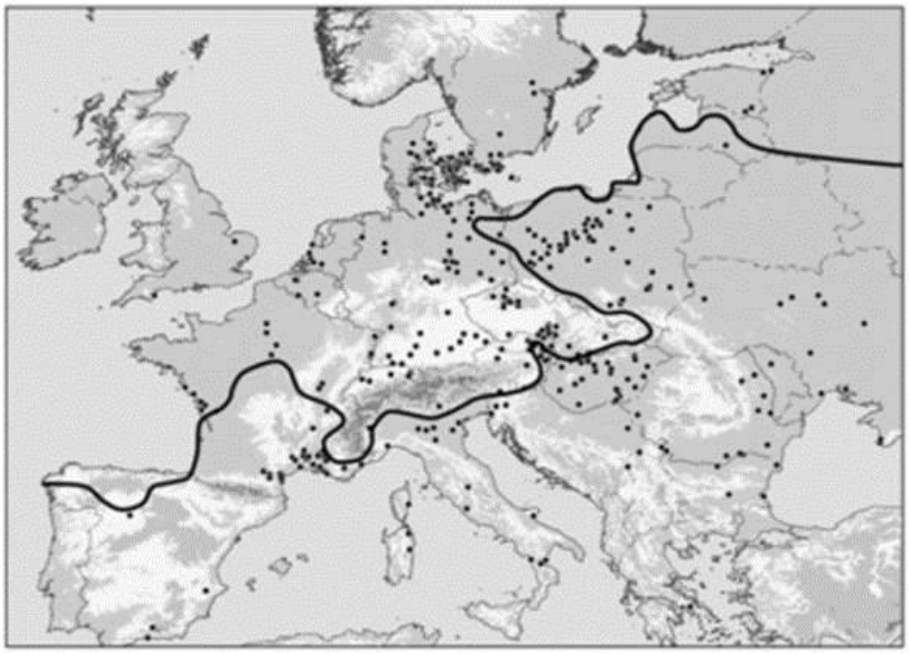

Figure 3. Holocene fossil records of *Emys orbicularis* in Europe. The black line displays current northern distribution border of native populations (Sommer *et al.* 2007).

Recent studies (Lourenco *et al.* 2013) show that latitude and habitat, and not body mass or longevity, are factors more significantly correlated with neutral rates of molecular evolution in aquatic turtles, with expected lower rates for species living in higher latitudes. A positive correlations between mtDNA substitution rates and diversification rates, and contemporary species richness was also shown in reptiles and birds (Eo & DeWoody 2010). Complete mtDNA genomes exhibited a mean rate of 3.35 x 10^{-9} substitutions per site per year (Tamura–Nei method), that varies among families. Snake and lizard families generally exhibited faster substitution rates (5.29 x 10^{-9}), whereas turtle (2.01 x 10^{-9}) and bird (2.56 x 10^{-9}) genomes evolved more slowly.

To date, phylogographical studies on the European pond turtle were carried out by analysis of the mitochondrial cytochrome *b* gene, however, for our analyses the mitochondrial DNA control region (CR) was chosen because it is among the most variable and informative regions of the mitochondrial genome in turtles and is reliable in analysis of intraspecific variations in many vertebrates (Lamb *et al.* 1994; Starkey *et al.* 2003). A comparative analysis of nucleotide divergences based on CR, ND4, Cyt b and CO I gene, in Geoemydidae turtles, established that CR evolved 2.6- to 5.7-fold compared to other genes of mtDNA (Jiang *et al.* 2011). This allowed us to further investigate the status of Italian populations and bring to light any dynamics not found thus far. Lourenco *et al.* (2012) sequenced *Emys orbicularis* mitochondrion partial genome (GenBank: JN999703.2) for a total of 16770 bp. The control region of the European pond turtle is located between tRNApro and tRNAphe and show a TCATATA repeat motif at the end. Unfortunately the start is not determined due to a sequencing gap.

Previous studies investigated population genetics of *Emys orbicularis* on the European scale (Fritz *et al.* 2009; Fritz *et al.* 2005a; Fritz *et al.* 2006b; Fritz *et al.* 2005b; Fritz *et al.* 2007; Fritz *et al.* 2004; Lenk *et al.* 1999a; Pedall *et al.* 2011; Prusak *et al.* 2013). These morphological and genetic surveys demonstrated that pond turtles represent a highly differentiated complex with a clear phylogeographic structure that corresponds to general phylogeographic patterns identified for Western Palaearctic biota, previously mentioned, with two to three endemic lineages occuring on each of the southern European peninsulas and in Asia Minor. Most of these lineages diverged during the Pliocene and their current distribution was shaped during Pleistocene and Holocene range fluctuations. The northern part of the distribution range has been recolonized in post-glacial times from two refuges: the south-eastern Balkans and the Black Sea region. Despite these results, the situation of the European pond turtle in Italian territory is still unclear.

Recent studies have shown the existence of distinct Italian microrefugia for a great variety of organisms: the Italian crayfish (Fratini *et al.* 2005), scorpions (Salomone *et al.* 2007), salamanders (Canestrelli *et al.* 2006), frogs (Canestrelli *et al.* 2007a, 2008; Canestrelli & Nascetti 2008; Santucci *et al.* 1996), several reptiles (Barbanera *et al.* 2009; Bohme *et al.* 2007; Lenk *et al.*

2001; Lenk & Wuster 1999; Podnar *et al.* 2005), the hedgehog (Seddon *et al.* 2001), the greater mouse-eared bat (Ruedi *et al.* 2008), Talpa romana (Canestrelli *et al.* 2010) and plants (Brewer *et al.* 2002; Konnert & Bergmann 1995;), and that a few glacial refugia in Italy were certainly located outside the southernmost part of the Peninsula, highlighting the potential role of central Italy, particularly the Tyrrhenian side, as a supplementary source of geographical divergence of refugial populations. It is possible to recognize two different types of refugia: the primary refugia, usually in the extreme south, that are situated in areas able to sustain species even during the glacial maximum, when cold climatic conditions, 15°C lower than today, severely limited species range distribution to restricted areas where favourable microclimates existed, and the secondary refugia found in areas able to provide temporary refugia during the shorter climatically adverse periods. It has been discovered, for example, that in Italy Latium maintained residual forests important for the survival of thermophilous and mesophilous trees during the whole last glacial period (Follieri & Magri 1997) and that the center of the Peninsula, such as Tuscany and even Sardinia, may have had suitable conditions for persistence of *Quercus* in some glacial periods (Brewer *et al.* 2002).

For species inhabiting wetlands, Pleistocene alluvial plains, like the substantial widening of the Po plain due to marine regression, have certainly played a positive role in the expansion in suitable habitat. Bianco (1995) distinguished three main ichthyogeographic districts in Italy: the Padan-Venetian district, which includes the rivers flowing into the upper and middle Adriatic Sea; the Tuscan-Latium district, ranging from the rivers Serchio and Arno, in Tuscany, up to the Tiber, flowing into in the Tyrrhenian Sea; and the Southern Italy district, including all the southern rivers flowing into both the eastern and western Italian coasts. It should also be considered that these systems have experienced huge changes; for example, there is geological evidence of a repeated connection between the rivers Arno and Tiber since the Lower Miocene but not with the Po rivers.

As rivers had a major impact on connection and facilitated expansion into new areals, the Italian Apennines promoted genealogical divergence, especially the northern chain (Canestrelli & Nascetti 2008; Canestrelli *et al.* 2007b; Mattoccia *et al.* 2011). Going down the Peninsula, the Matese Massif (northern Campania) and Pollino Massif (northern Calabria) could have represented a strong barrier to dispersal. These studies also widely discuss the role of South Italy in speciation events. Tectonic activity, resulting from the compression of the Apennine arch and subsidence movements caused by the collisions of the European and African plates, during late Pliocene-early Pleistocene times, producing in Calabria a system of grabens (rift areas between adjacent tectonic faults), the Crati-Sibari Plain, the Catanzaro Plain and the Strait of Messina. Northern and central Calabria were connected during the early Pleistocene by the emerging Catena Costiera; this connection was represented by a narrow land bridge owing to the submersion of the Crati-Sibari

plain, which gradually emerged during middle-late Pleistocene. The Catanzaro graben, which substantially separated central Calabria from southern Calabria, emerged later than the Crati-Sibari graben, during the late Pleistocene. The high genetic variability observed among Calabrian populations, particularly the south-central ones, probably did not originate in a context of prolonged stability, but instead, may be due to repeated fragmentations, divergence into separate 'mountain-islands' and subsequent secondary admixtures.
The Messina strait has been recognized as an important faunal corridor between southern Calabria and Sicily, until the Late Pleistocene, (Canestrelli *et al.* 2007a; Santucci *et al.* 1996), definitively separated by submersion about 50,000 years ago (Caloi *et al.* 1989).

The aim of the study is to investigate and reconstruct the phylogeographic history of Italian turtles following the last glacial period and characterize the population structure and levels of genetic divergence among the main Italian natural populations, through the analysis of the mitochondrial DNA control region and microsatellite markers.

MATERIALS AND METHODS

Study area and Sampling

In this study we have taken into consideration the main populations of *Emys orbicularis* and *Emys trinacris* in Italian territory. Turtles were captured at two types of sites: in natural reserves and along river bights or ponds in the countryside, often among cultivated fields or near crops fields. Following is a brief description of sampling sites proceeding from the north to the south of Italy (Fig. 4).

1. Friuli Venezia Giulia - Riserva Naturale della Foce dell'Isonzo: the Reserve lies in the far eastern part of the Po valley with still largely marshy lands at the mouth of the river. Temporary and perennial fresh water swamps, virtually disappeared from the territory due to land reclamation, have been rebuilt since the nineties.
2. Piemonte - Riserva Naturale Speciale e Zona di Salvaguardia della Palude di San Genuario: small wetland surrounded by cultivated fields. Located at 4 km from the Po River.
3. Lombardia - Cremona: ponds at 1 km in a south-east direction from Cremona in a highly urbanized zone (close to the highway), at only 1 km from the Po River.
4. Veneto - Riserva Naturale Integrale Bosco Nordio: A remnant of the wide forest and wetland zone that in the past characterized much of the Veneto coast. It is on the Adige River.
5. Emilia Romagna - Riserva Naturale Bosco della Mesola: natural coastal environments with coastal forest, swamps with typical marsh vegetation

and grasslands. In the south central part of the forest an internal network of canals has been restored and expanded, and the Elciola pond of 7 hectares was dug.

Sites 4 and 5 are located at opposite ends of the part of the coast characterized by the Po River delta which is the largest protected wetland in Italy.

6. Liguria - Albenga: Stagno Salea and Rio Carenda are an artificial pond and creek, at about 100 m above sea level, in areas of high human impact where a population of *Emys orbicularis* believed to be extinct was recently rediscovered in the region (Jesu *et al.* 2004).
7. Toscana - Parco Regionale di Migliarino, San Rossore e Massaciuccoli -Tenuta di San Rossore: coastal landscape characterized by a great variety of environments with dunes and depressed areas, meadows, wetland forests, pine forests and wetlands. Bordered on the north by the Serchio River and south by the Arno River.
8. Toscana - Parco Regionale della Maremma: typical coastal environment with a series of artificial and natural canals with slow-flowing water in which the movement of the aquifer towards the shore is hampered by the reverse saltwater current.
9. Lazio - Riserva Naturale Monte Rufeno: the environment is characterized by extensive forests in a hilly landscape and geographically limited ponds, formed as a result of landslides, located inside the forest at altitudes between 350 and 560 m above sea level. The Paglia River cuts the park in half.
10. Lazio - Riserva Naturale Statale Tenuta di Castelporziano: the estate, founded as a hunting and agricultural reserve, has wetlands and marshes that stretched south to the Pontine plain and north to the Maremma, thanks to a series of natural basins.
11. Lazio - Parco Nazionale del Circeo: behind and parallel to the coastal dune, is a wetland and lagoon consisting of four coastal lakes in succession and seasonally flooded ponds in the ancient Selva of Terracina that was extended for 11000 hectares.
12. Marche - Ascoli Piceno: ponds on San Benedetto del Tronto's hills among cultivated fields at about 150 m a.s.l..
13. Abruzzo - Riserva Naturale Regionale Lago di Serranella: the artificial lake of Serranella was created in 1981 for irrigation and industrial purposes. Located at the confluence of the rivers Sangro and Aventino, at about 60 m above sea level, is one of the last wetlands of Abruzzo.
14. Puglia - Parco Nazionale del Gargano: the lake of Lesina is located along the north shore of Gargano; it flows to the Adriatic sea through three channels, and various streams ensure a fair contribution to the freshwater lake. Lake Salso, on the other side of the Park, is a coastal wetland, one of the most important in southern Italy, consisting of large retention basins created during the reclamation of the '50s.

15. Puglia - Riserva Naturale Le Cesine: wetland on the coast, remnant of an ancient swamp that until 1800 extended from Otranto to Brindisi, mainly characterized by a pine forest.
16. Campania - Oasi WWF di Persano: in the Riserva Naturale Foce Sele, the Oasi is located in the upper part of the Sele Plain, about 50-70 m a.s.l., in the internal vertex of the fan-shaped plain that opens to the sea among the Picentini and Alburni Mountains. Rich in vegetation, thanks to reforestation started in 1980, it can be divided into three different environments: the lake, reeds and hydric wood.
17. Basilicata - Oasi WWF Policoro Herakleia: is in the Riserva Naturale Bosco Pantano on the coast, that contains one of the last coastal flooded forests of our country with several ponds that are suffering more and more prolonged periods of desiccation.
18. Calabria - Lamezia Terme: Palude dell'Imbutillo swamp on the coast near an industrial area.
19. Calabria - Crotone: River Neto near the mouth with slow flow among cultivated fields.
20. Sardegna - Berchidda: Stream Riu Mannu nearby artificial lake Coghinas at 160-180 m above sea level in an open area of countryside.
21. Sardegna - Pattada: ponds among sparse vegetation due to the numerous fires on Monte Lerno at about 800 m a.s.l..
22. Sardegna - Oristano: River Tirso at 6 km north-east from Oristano surrounded by cultivated fields.
23. Sardegna - River Flumendosa: pond adjoining the mouth of the river and totally isolated from the sea.
24. Sicilia - Lago Spartà: small pond in the province of Messina at 650 m above sea level in Mediterranean vegetation.
25. Sicilia - Riserva Naturale Orientata Monte Capodarso e Valle dell'Imera Meridionale: stream with loops of standing water among field crops at about 250 m a. s. l..
26. Sicilia - Riserva Naturale Orientata Bosco della Ficuzza: Reserve rich in torrential streams forming several natural temporary lakes at about 800-1000 m a.s.l..
27. Sicilia - Gallitello: tiny ponds in the countryside at 100 m a.s.l..
28. Sicilia - Riserva Naturale Integrale Lago Preola e Gorghi Tondi: the reserve runs parallel to the sea at about 1 km, rich with several brackish ponds surrounded by low Mediterranean vegetation.
29. Sicilia - Siculiana Marina: basin adjoining a main street on the coast.
30. Sicilia - Riserva Naturale di Vendicari: coastal wetland in contact with the Ionian sea and consequently with medium salt content.
31. Sicilia - Noto: creek between the city of Noto and the sea in a high human impact area.

Individuals were captured with traps placed in ponds, after which they were marked with engraved notches on the marginal plates of the carapace, following a scheme similar to the method described by Ernst *et al.* (1974). Morphological measurements were also taken for each individual, such as length and width of the carapace, plastron length and weight. Then blood samples were taken from the brachial or subcarapacial vein using syringes of 1-2 ml with 26G 0.45x13 mm or 25G 0.5 x 16 mm needles, depending on the turtle's size. Once collected, blood samples were preserved in a buffer solution (100 mM Tris pH 8, 100 mM EDTA pH 8, 200 mM NaCl, 1% SDS pH 7.2) which enables the maintenance of nucleic acids at room temperature, before transport to the laboratory and subsequent storage of the samples at - 80°C.

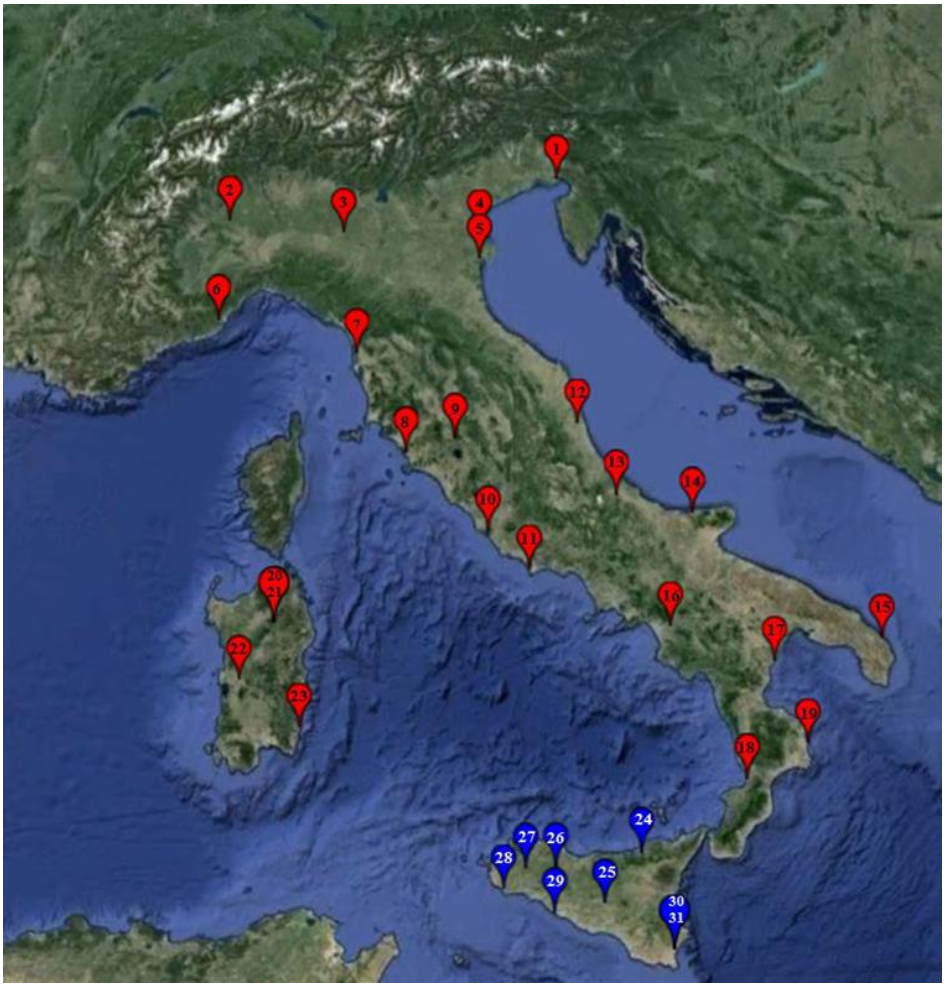

Figure 4. Map of the sampling sites.

Molecular techniques

Total genomic DNA was extracted from blood samples by an overnight incubation at 37 °C in lysis buffer (100 mM Tris pH 8, 5 mM EDTA pH 8, 100 mM NaCl, 0.5% SDS pH 7.2) including 200 µg of proteinase K, followed by a standard phenol/chloroform extraction (Sambrook *et al.* 1989). DNA was precipitated from the supernatant with two volumes of cold 100% Ethanol, centrifugated, dried and resuspended in TE buffer. We also used a membrane-based technique, a high-throughput system, Wizard SV 96 Genomic DNA Purification System (Promega). With these two systems we obtained DNA concentrations from 2 to 100 ng/µl. A qualitative and quantitative control of the extract was performed by agarose gel electrophoresis, images were analyzed and documentated with Gel Doc XR System (Bio-Rad), and Infinite 200 PRO NanoQuant (Tecan), a microplate reader developed specifically for absorbance applications with small sample size. Figure 5 shows gels from the two different extraction methods. The second system provides more standardized and cleaner results although in a lower concentration.

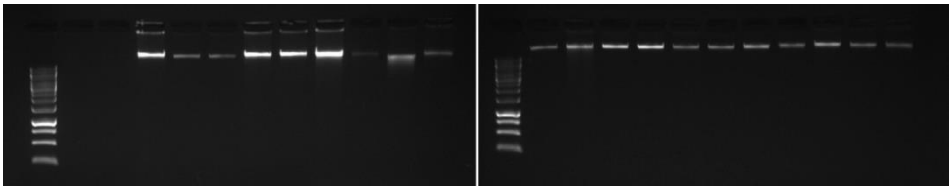

Figure 5. Left: phenol/chloroform extraction. Right: Wizard SV 96 Genomic DNA Purification System.

Mitochondrial DNA

We amplified a 659 bp fragment of mitochondrial DNA control region for a total of 784 turtles using primers DES-1 (5'-GCA TTC ATC TAT TTT CCG TTA GCA-3') and DES-2 (5'-GGA TTT AGG GGT TTG ACG AGA AT-3'), developed for *Chrysemys picta* (Starkey *et al.* 2003). The control region was amplified by PCR (polymerase chain reaction) with the following thermal profile: initial denaturation step at 95 °C for 5 min, 35 cycles of 45 s at 95 °C, 45 s at 50 °C annealing temperature and 90 s at 72 °C extension temperature, with a final extension step of 10 min at 72 °C. Reactions were performed in 15 µL volume cointaing 0.5 U Taq DNA Polimerase Cloned (5 U/µL Life Technologies), 1x Buffer (Life Technologies), 1.5 mM $MgCl_2$ (Life Technologies), 0.2 mM dNTPs (Life Technologies) and 0.5 µM primers. After checking on agarose gel electrophoresis quality and quantity, using known concentrations of DNA phage λ and Quantity One 4.6.7 program, PCR products were directly sequenced on both DNA strands with BigDye Terminator v. 3.1 Cycle Sequencing Kit (Life Technologies). Sequencing reaction (10 µL) involves the use of 5-20 ng of template for 500-1000 bp. The cycle sequencing was carried out with an initial denaturation step at 96 °C for 3 min and 25 cycles at 96 °C for 10 s, 50°C for 5 s and 60 °C for 4 min. Sequencing products,

before the capillary electrophoresis on 3130xl Genetic Analyzer (Applied Biosystems), were purified using the following procedure: add 90 μL of 63% 2-propanolo, stir and let stand 15 minutes at room temperature, 45 min of centrifuge at 3700 rpm, remove supernatant, add 150 μL of 70% 2-propanolo, remove supernatant again, dry and resuspend in 15 μL of high density formamide. Sequences obtained are displayed and analyzed with Sequencing Analysis 5.2 (Applied Biosystems) and aligned with Geniuos 8.0 (Biomatters).

The mitochondrial genome is a single, large genetic locus that can provide a single perspective on the evolutionary history of a group (Ballard & Whitlock 2004; Zhang & Hewitt 2003). Thus, mtDNA alone is often inadequate for phylogeographic/phylogenetic analyses, especially in the face of complex evolutionary scenarios, and simultaneous analysis of mitochondrial and multiple nuclear markers is required (Spinks & Shaffer 2009). Consequentely, to understand as thoroughly as possible the history, structure and current state of *Emys orbicularis* and *Emys trinacris* populations in Italy, we added the study of 16 nuclear microsatellite loci to the analysis of the mtDNA CR.

Nuclear DNA

1029 turtles were analyzed at 16 microsatellite loci shown in Table 1 (Ciofi *et al.* 2009b; Pedall *et al.* 2009), to study the degree of diversity and the structure of the Italian populations. We performed multiplex PCR, reactions of 10 μL volume containing 1 U Taq DNA Polimerase Cloned (5 U/μL Life Technologies), 1x Buffer (Life Technologies), 1.5 mM MgCl2 (Life Technologies), 300 μM dNTP (Life Technologies) and 0.5 μM primers. After an initial denaturation step at 94 °C for 5 min, 35 cycles of 40 s at 94 °C, 40 s at annealing temperature (Tab. 1) and 1 min at 72 °C extension temperature, with a final extension step of 5 min at 72 °C, were performed. Fragment lengths were determined on a 3130xl Genetic Analyzer (Applied Biosystems) with the GeneScan 500 LIZ dye Size Standard and allele size was assessed using Genemapper version 5 (Applied Biosystems).

Statistical analysis

Mitochondrial DNA

jModelTest 2.0 (Darriba *et al.* 2012; Guindon & Gascuel 2003) was used to carry out statistical selection of best-fit models of nucleotide substitution with three different model selection strategies, Akaike and Bayesian information criteria (AIC and BIC) and a decision theory method (DT). AIC is the amount of information lost when we use a specific model to approximate the real process of molecular evolution, therefore, the model with the smallest AIC is preferred. BIC is an estimate of a function of the posterior probability of a model being true, under a certain Bayesian setup, so that a lower BIC means that a model is considered more likely to be the true model. DT is a novel approach that selects models on the basis of their phylogenetic

performance, measured as the expected error on branch lengths estimates weighted by their BIC.

Table 1. Oligonucleotides used in the study. The first eight are from Ciofi *et al.* (2009) and the remaining eight are from Pedall *et al.* (2009). Primers used jointly show the same multiplex letter. The annealing temperature of each multiplex corresponds to that of the primer with the lowest Ta.

Locus	Repeat motif	Primer sequence (5′–3′)	Ta (°C)	Multiplex
Emys 1	$(TG)_{30}$	F: AACACTCAAAGTTTTGAACAAACTAC R: GAGTTAGCCCTGAAGACAGTGTTT	56	A
Emys 2	$(TC)_{14}(AC)_{9}$	F: AAGGTGGAAGCTCCTTGATGC R: GCATAGTGCTTTGTGGCATTTCGA	54	A
Emys 4	$(CA)_{22}$	F: TGGTGCTTCTGCCTCTTGTC R: CCTTCTGTTTGTTCCAGAATAATTTG	54	B
Emys 5	$(CA)_{22}$	F: TGGTGCTCATCTGATGCCTTATAC R: ACATTAAGTGCTGCTCATTTGATG	54	/
Emys 6	$(CA)_{17}$	F: CAGGGGAGGAAAAATAAACCA R: AGGGACAGAGGTTCTTATTTTTAC	54	B
Emys 7	$(GA)_{23}$	F: CTCTCTCACTGTCTAGCCATCCCCAT R: TTCGGTCCCAACAGGAAATAA	56	C
Emys 8	$(TG)_{33}(AG)_{11}$	F: CCCATAAACTGTCCCCTTGTT R: GCTGTTAAGAACAGAACATTATAAAG	56	C
Emys 11	$(GA)_{33}$	F: AGCCAATTGTGCAATGATTTT R: CACTTTTGCTTTAGACCATATGTTCC	54	C
msEo2	$(CA)_{15}$	F: TTCAAACCAATCCGATGAGG R: GCCTTTCTATGAAATGCTACATG	54	D
msEo4	$(CA)_{9}$	F: ATCCTGCTCACAAGAAGGC R: CGCAGTGATACAGATTCTTC	54	E
msEo21	$(GA)_{11}$	F: GTAGTAACCCACTTGATGAG R: TTACCTGGCAATTACCTGGC	54	D
msEo22[a]	$(CA)_{13}$	F: GGTAATCAAGCAGCTACAGG R: CTCTGCCAAGTCGAATCAG	55	F
msEo25[a]	$(CA)_{34}$	F: GTGACGTGTGTAACCAATGTG R: TAGAGAATGTCTGCCTGTCC	55	F
msEo29	$(CT)_{14}$	F: ACTTCATCGGATGCATGAAG R: ACTTTTGGACTACTGCAGCC	54	D
msEo32	$(TG)_{4}(CA)_{12}$	F: CGAGTCTTTGGATTACACCG R: GTTGAGGTGACTGTGATTGC	55	E
msEo54	$(CA)_{7}$	F: CAGATATCCACCCCTAGCC R: GAAGTGTCTCTGTACAGACC	55	E

Using ARLEQUIN 3.5 (Excoffier & Lischer 2010), the mitochondrial sequence variation was estimated by number of polymorphic sites, haplotypic (gene) diversity (the probability that two randomly chosen haplotypes are different in the sample), nucleotide diversity (the probability that two randomly chosen nucleotide sites are different) and mean number of nucleotide differences between haplotypes. The latter software was also used to compute pairwise F_{ST} for all pairs of populations, as well as different indexes of dissimilarities (genetic distances) between pairs of population. Parameters of an

instantaneous demographic expansion were estimated from the mismatch distribution (the distribution of the observed number of differences between pairs of haplotypes), usually multimodal in samples drawn from populations at demographic equilibrium and unimodal in populations having passed through a recent demographic expansion (Rogers & Harpending 1992; Slatkin & Hudson 1991) or through a range expansion with high levels of migration between neighboring demes (Excoffier 2004; Ray *et al*. 2003)) using a generalized least-square approach, as described in (Schneider & Excoffier 1999).

The mismatch analysis was also combined with neutrality tests, Tajima's test and Fu's FS test, based on the infinite-site model without recombination. The significance of statistics is tested by generating random samples under the hypothesis of selective neutrality and population equilibrium, using a coalescent simulation algorithm adapted from Hudson (1990). The P value of statistics is then obtained as the proportion of random F statistics less or equal to the observation. Significant D test values can be due to factors other than selective effects, like population expansion, bottleneck, or heterogeneity of mutation rates and Fu (1997) has noticed that the FS statistic was very sensitive to population demographic expansion, which generally lead to large negative F values (ArisBrosou & Excoffier 1996; Tajima 1996).

Nested clade analysis (Templeton 1998; Templeton *et al*. 1995) provides a highly flexible method for testing associations between a haplotype tree and other types of data. Initially it was developed for phenotypic associations but another type of data is the geographical location where a haplotype is found. The nested clade phylogeographic analysis (NCPA) can be used to test the null hypothesis that there is no association between the haplotype tree and geography. Geographical associations with haplotypes can arise for two reasons. The first arises from a species' demographic structure and history. When a mutation first occurs to create a new haplotype, that haplotype is obviously restricted to its geographical site of origin. However, once this starts to replicate, it can spread through space and time, and the dynamics of this spread depend upon the amount and pattern of gene flow within the species and historical factors such as fragmentation that would prevent a haplotype originating in one region from spreading into another or range expansion that could place the haplotype into a new geographical area. Geographical associations arising from a species' demographic structure and history are not expected to be locus specific, as these demographic and historical factors should affect all loci. The second cause for geographical association instead is locus-specific. If natural selection is occurring on a haplotype or haplotype clade at a locus, selection can influence its spatial distribution, either by accelerating its spread throughout the species or by restricting the selected haplotype to certain areas where it is locally adaptive. Intraspecific phylogeography focuses upon a species' historical demography and events that influence how genetic variation is distributed in space and time. Therefore, the haplotype tree/geography associations that are general and not locus specific are the ones that are

informative about a species' phylogeography. One of the most unique features of nested clade phylogeographic analysis is that it uses the coalescent information contained in the genetic data to infer and test phylogeographic events and processes. Haplotype trees are the units of analysis, and such trees represent the estimable portion of the coalescent process at a given DNA region. In a haplotype tree each current DNA sequence and each node represents a distinct and sole sequence, and every branch is marked by one or more mutational events. Hence, the haplotype tree is a map that shows how all the current array of genetic variation found in the diverse haplotypes in the sample arose by the accumulation of mutations in DNA lineages over evolutionary history. In contrast, most other phylogeographic techniques make no use of the historical information contained in the genetic data but rather merely look at goodness of fit of phylogeographic scenarios to genetic measures such as heterozygosity, number of alleles, etc., that contain little to no historical information.

ANeCA v1.2 package (Panchal & Beaumont 2007), which include TCS v2.1 (Clement *et al.* 2000a) and GeoDis v2.5 programs (Posada *et al.* 2000), was used for the nested clade phylogeographic analysis. TCS software estimates genealogical relationships among sequences using the method of (Templeton *et al.* 1992a), that links first haplotypes with the smaller number of differences as defined by a 95% confidence criterion. An unrooted network of haplotypes was then obtained using the method of statistical parsimony. NCPA uses this haplotype tree to define a series of hierarchically nested clades (branches within branches) using a set of explicit nesting rules (Templeton *et al.* 1987; Templeton *et al.* 1992a). Haplotypes are the lowest units of analysis, being nested together into mutationally close subsets called one–step clades. The one–step clades in turn are nested into two–step clades, and so on, until a nesting level is reached such that the next higher nesting level would result in only a single clade spanning the entire original haplotype network. The nested design captures the most reliable temporal information contained in the haplotype network. When the haplotype network is properly rooted, the oldest clade is known in any given nesting category. Even if the haplotype tree were unrooted, coalescent theory predicts that clades on the tips of the tree are highly likely to be younger than the interior clades to which the tips are connected within a population or set of populations well connected by gene flow (Castelloe & Templeton 1994a). Within a nesting category, contrasts of interiors clade (oldest) vs. the tips clades (younger) therefore constitute a temporal contrast that does not depend upon a molecular clock or any sort of rate calibration. There is a role for making use of a molecular clock in multilocus NCPA studies (Templeton 2002), but the NCPA of any individual DNA region uses the nesting design to make temporal contrasts of spatial information in a manner independent of a clock.

Association between haplotype variation and geography was initially tested by performing an exact permutational contingency analysis for each clade

and by calculating a chi-squared statistic from the contingency tables (clades vs. geographical locations). This analysis reveals whether haplotypes are significantly associated geographically, but does not provide information on whether they are more dispersed or geographically restricted than the random expectation (Templeton et al. 1995). To further test the effects of historical population events such as range expansions, colonization, or fragmentation, an additional analysis was conducted using information on geographical distances. For each clade at each hierarchical level, departures from expectation under the null hypothesis of no geographical association were determined for two distances: the 'clade distance' Dc, which measures the mean distance of members of a given clade from its geographical centre, and the 'nested clade distance' Dn, which measures the mean distance of members of a given clade from the geographical centre of the next higher-level clade within which that clade was nested (Templeton et al. 1995). It has been suggested that riverine, riparian, or coastal species may be better examined using river (or coastal) distances (Fetzner & Crandall 2003a) rather than the standard geographic (great circle) distances among populations. Given the habits of the semi-aquatic European pond turtle, that despite the fact that it sometimes lives along the coast or rivers, we chose to use standard geographic distance because it is always circumscribed within nature reserves in ponds or bends of streams where the current is slow. In addition, the mean distance between the tip and interior clades within the nested group (I–Tc) and the tip to interior distance for the nesting clade (I–Tn) were measured. Calculations were performed with 100000 permutations in GeoDis. Observed distance patterns were interpreted using the inference key provided in Templeton (2004).

Nuclear DNA

Errors due to large allele dropout or stutter bands were evaluated using MICROCHECKER 2.2.3 (Van Oosterhout *et al.* 2004). The first may cause deviations in Hardy-Weinberg proportions due to preferential shorter alleles (short allele dominance), which in turn cause an apparent deficiency in heterozygotes, while stutter bands could make it more difficult to recognize heterozygote samples with similar alleles. Evidence for the presence of null alleles (non-amplified alleles) at each locus which can arise when mutations occur in flanking regions preventing primers from binding was also tested.

GENALEX 6.5 (Peakall & Smouse 2012) was used to estimate the basic diversity parameters for microsatellite loci, mean number of alleles per locus, observed heterozygosity for a single locus within a population (where the number of heterozygotes is determined by direct count) and unbiased expected heterozygosity.

GENEPOP 4.3 (Rousset 2008b) was used to test for departure from the Hardy–Weinberg equilibrium, testing null hypothesis "random union of gametes", applying a Markov Chain exact test for each locus per population (Guo & Thompson 1992). The Bonferroni correction, a multiple-comparison

correction used when several dependent or independent statistical tests are being performed simultaneously, that divide the critical P value (α) by the number of comparisons being made, has also been applied.

FSTAT 2.9 (Goudet 1995) was used to estimate and test gene diversities and differentiation, as F_{ST} measures heterozygote deficit among populations and F_{IS} measures heterozygote deficit within population.

Population structure was assessed using the Bayesian clustering analysis implemented in STRUCTURE 2.3.4 (Pritchard *et al.* 2000a). We ran the program without prior population information with a burn-in period of 10000 iterations, and we estimated the probability of the observed genotypes given a number of populations K ranging from 1 to 32 (the number of sampling sites plus one) by Markov Chain Monte Carlo (MCMC) methods using 10000 repetitions. The program STRUCTURE HARVESTER v 0.3 (Earl & Vonholdt 2012) was used to process the STRUCTURE result files and determine the best K, as inferred from the ΔK method of Evanno *et al.* (2005a). The K value was then used as prior information to estimate the probability that an individual belongs to a given population.

To identify recent events of demographic reduction based on allele frequencies BOTTLENECK 1.2.02 (Cornuet *et al.* 1999; Piry *et al.* 1999) was used, which tests across all loci for a significant observed heterozygosity excess relative to heterozygosity expected for the number of observed loci. Populations that have experienced a recent reduction in their effective population size exhibit a correlative reduction of the allele number and heterozygosity at polymorphic loci. Allele number is reduced faster than heterozygosity, thus heterozygosity becomes larger than the expected at mutational drift equilibrium, because expected heterozygosity is calculated from the allele number. When testing microsatellite loci the one-step stepwise mutation model (SMM) is generally appropriate and the two-phase mutation model (TPM) is even better. The Wilcoxon's test is recommended when few (<20) polymorphic loci are used, because it is the most powerful and robust.

RESULTS

Genetic diversity

The three model selection strategies implemented in JMODELTEST, AIC, BIC and DT, inferred the Hasegawa-Kishino-Yano (HKY) model as the most appropriate model of evolution with respect to estimates of nucleotide substitutions for our data set, according to which rates of substitution differ between each nucleotide especially between the rate of transitions and trasversion. In softwares, such as ARLEQUIN, where this model is not in the options we chose the Tamura and Nei (TN) model, which outputs a corrected percentage of nucleotides for which two haplotypes are different. The correction allows for different transversion and transition rates, but a distinction is also

made between transition rates between purines and between pyrimidines. MCROCHECKER software did not discover evidence of null allele.

There were 20 haplotypes of the mitochondrial DNA control region defined by 28 polymorphic sites in 784 sequences of *Emys orbicularis* and *Emys trinacris*, respectively 16 different sequences in 663 individuals of the European pond turtle and 4 sequences in 121 individuals of the Sicilian pond turtle. Tables 2a and 2b show the distribution of haplotypes in sampling populations. The more common haplotypes in the Peninsula are E01, E02 and E03, in Sicily E11. Moreover, Table 3 shows the frequency of different nucleotides in all polymorphic sites, and in Table 4 there are nucleotides of polymorphic sites for each haplotype. See Figure 6 for a simple vision of the distribution of haplotypes and their frequencies over the territory. Sequence variation was estimated by haplotypic (gene) diversity and nucleotide diversity (Tab. 5). As expected, South Italy exhibited greater variability, especially populations in Calabria, Campania and Puglia at Le Cesine. Nevertheless, relatively higher values of h, π, and p were also found in Maremma and the North of Sardinia, thanks to the presence of unique haplotypes (E08, E18, E19 and E20), and in Castelporziano. *Emys trinacris* has a very low value, except for the population of Noto, with 6/7 samples characterized by a unique haplotype (E10), and Lago Preola.

All 16 microsatellites were polymorphic, ranging from 7 to 33 alleles per locus, those characterized by Ciofi *et al.* (2009) with the greatest number. Populations showed medium values of observed heterozygosity (H_O); the lowest value was recovered in Albenga and island populations. The highest values were recorded in Calabria, Campania, Abruzzo and especially the central/north populations where heterozygosity is higher than expected (H_E) (Tab. 6). Six populations (Isonzo, San Rossore, Monte Rufeno, Castelporziano, Policoro and Lamezia Terme) showed significant departures from the Hardy–Weinberg equilibrium (HWE) ($P < 0.01$) at one to five loci. It should be considered that the above-mentioned are not populations with lower values of H_O. No heterozygote excesses were observed at any locus. Given the large sample size of population in Monte Rufeno, analyses were repeated by dividing individuals for the eight ponds where they had been sampled. Only one pond, located to the northern dise of the reserve, showed significant deviation from the equilibrium. This pond is isolated from the others and far from a series of streams in the central area that could facilitate travel in hilly surroundings which would otherwise be unlikely. The Calabrian population of Lamezia Terme indeed is considered at equilibrium being that its deviation is due to a sole locus.

F_{ST} statistic values were statistically significant for both mtDNA and nDNA, and showed a clear separation of *E. trinacris* from the Peninsula's populations. Persano was the only location where a Sicilian haplotype E13 was found. A moderate degree of genetic differentiation was found for mtDNA on the west coast, between the west coast and Sardinia, among east coast populations and among southern populations. Low values among populations

were recorded across central-western populations, in Tuscany and Lazio, and across the Po Valley, except San Genuario. Analysis of microsatellites showed lower genotypic differentiation, especially among the central and southern areas. Significant inbreeding coefficients were recorded in the populations of Latium, Ascoli Piceno, Lamezia Terme and in four sites of Sicily (Tab.6).

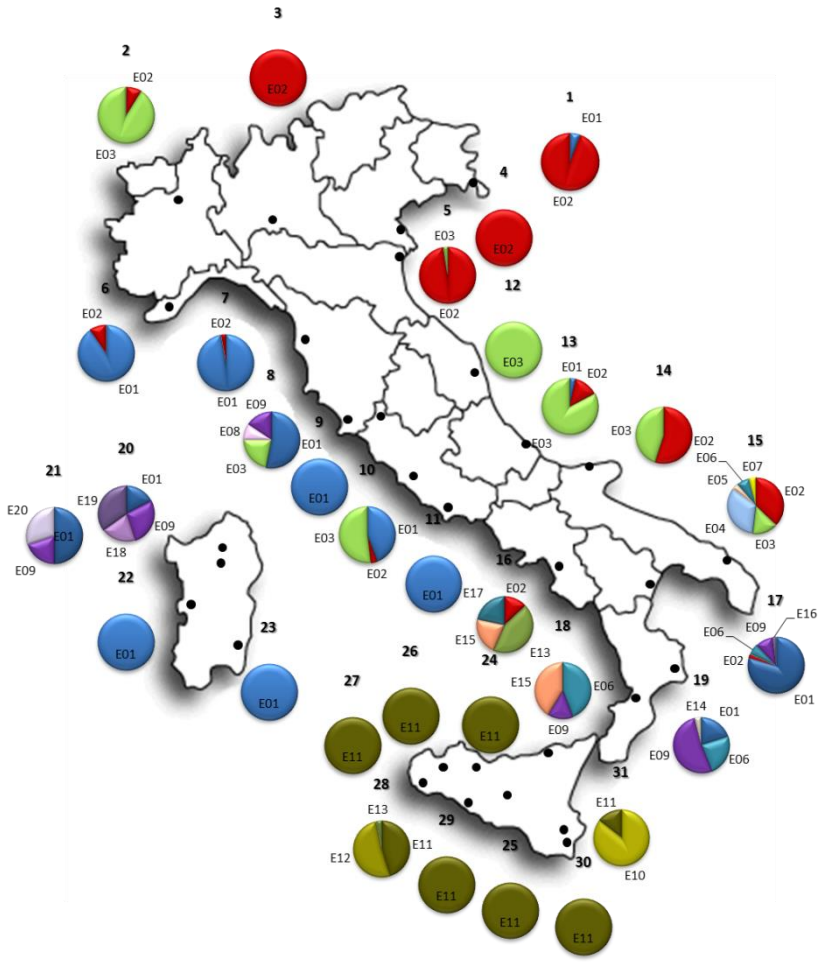

Figure 6. Pie charts of haplotypes' frequency in relation to geographic position. Pie charts of each population, numbered as the sampling site at page 18, are close to their locations, identified by a dot.

Table 2a. Haplotype frequency distribution for each site and total haplotype frequency across sites for *Emys orbicularis* and *Emys trinacris* sampled in Italy.

		Haplotype																			
Site	*N*	E01	E02	E03	E04	E05	E06	E07	E08	E09	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
Isonzo	35	0.06	0.94																		
San Genuario	34		0.09	0.91																	
Cremona	33		1.00																		
Bosco Nordio	28		1.00																		
Bosco Mesola	39		0.97	0.03																	
Albenga	21	0.90	0.10																		
San Rossore	37	0.97	0.03																		
Maremma	32	0.53		0.22					0.09	0.16											
Monte Rufeno	26	1.00																			
Castelporziano	63	0.52	0.03	0.44																	
Circeo	12	1.00																			
Ascoli Piceno	36			1.00																	
Lago Serranella	29	0.03	0.14	0.83																	
Gargano	20		0.55	0.45																	
Le Cesine	27		0.37	0.15	0.33	0.04	0.07	0.04													
Persano	23		0.13											0.43		0.22		0.22			
Policoro	49	0.80	0.02				0.06			0.10							0.02				
Lamezia Terme	34						0.44			0.15						0.41					
Crotone	25	0.20					0.24			0.52					0.04						

Table 2b. Haplotype frequency distribution for each site and total haplotype frequency across sites for *Emys orbicularis* and *Emys trinacris* sampled in Italy.

		Haplotype																			
Site	*N*	E01	E02	E03	E04	E05	E06	E07	E08	E09	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
Berchidda	40	0.18								0.28									0.20	0.35	
Pattada	10	0.50								0.20											0.30
Flumendosa	5	1.00																			
Tirso	5	1.00																			
Valle Imera	19											1.00									
Bosco Ficuzza	18											1.00									
Gallitello	8											1.00									
Lago Preola	31											0.45	0.52	0.03							
Siculiana	10											1.00									
Vendicari	19											1.00									
Noto	7										0.86	0.14									
Total	784	0.270	0.216	0.179	0.011	0.001	0.033	0.001	0.004	0.052	0.008	0.125	0.020	0.014	0.001	0.024	0.001	0.006	0.010	0.018	0.004

Table 3. Frequency of different nucleotides in all polymorphic sites; n, number of different bases for each polymorphic site.

Polymorphic sites	n	nucleotide	nucleotide	nucleotide
35	2	T : 0.98	C : 0.02	
51	2	A : 0.64	G : 0.36	
56	2	G : 0.15	A : 0.85	
61	3	C : 0.17	T : 0.39	A : 0.45
62	2	G : 0.17	A : 0.83	
63	2	T : 0.97	C : 0.03	
64	2	A : 0.97	G : 0.03	
66	2	T : 0.61	C : 0.39	
68	2	A : 0.61	T : 0.39	
112	2	G : 0.99	A : 0.01	
181	2	G : 0.97	C : 0.03	
194	2	G : 0.99	A : 0.01	
248	2	C : 0.17	T : 0.83	
304	2	G : 0.61	A : 0.39	
312	2	G : 0.17	A : 0.83	
367	2	A : 0.55	G : 0.45	
389	2	T : 0.99	C : 0.01	
415	2	C : 0.97	T : 0.03	
457	2	C : 0.94	T : 0.06	
462	2	A : 0.35	G : 0.65	
486	2	G : 0.99	A : 0.01	
490	2	A : 0.55	G : 0.45	
495	2	T : 0.17	C : 0.83	
536	2	A : 0.95	G : 0.05	
557	2	T : 0.17	C : 0.83	
573	2	C : 0.61	T : 0.39	
584	2	G : 0.61	A : 0.39	
587	2	C : 0.99	A : 0.01	

Table 4. Polymorphic sites for each haplotype. .*Emys trinacris*'s haplotypes are below dotted line.

	Polymorphic sites																												
H	35	51	56	61	62	63	64	66	68	112	181	194	248	304	312	367	389	415	457	462	486	490	495	536	557	573	584	587
E01	T	G	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	C	C	G	G	A	C	A	C	T	A	C
E02	T	A	A	A	A	T	A	T	A	G	G	G	T	G	A	G	T	C	C	G	G	G	C	A	C	C	G	C
E03	T	A	A	A	A	T	A	T	A	G	G	G	T	G	A	G	T	C	C	A	G	G	C	A	C	C	G	C
E04	T	A	A	A	A	T	A	T	A	A	G	G	T	G	A	G	C	C	C	G	G	G	C	G	C	C	G	C
E05	T	A	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	C	C	A	G	A	C	A	C	T	A	C
E06	T	A	A	A	A	C	G	T	A	G	C	G	T	G	A	G	T	C	C	G	G	G	C	G	C	C	G	C
E07	T	A	A	A	A	T	A	T	A	G	G	G	T	G	A	G	T	C	C	A	G	G	C	A	C	C	A	C
E08	T	G	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	T	T	G	G	A	C	A	C	T	A	C
E09	T	G	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	C	T	G	G	A	C	A	C	T	A	C
E14	T	G	A	T	A	T	A	C	T	G	G	G	T	A	A	G	T	C	T	G	G	A	C	A	C	T	A	C
E15	T	A	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	T	C	G	G	A	C	A	C	T	A	C
E16	T	A	A	A	A	C	G	T	A	G	C	G	T	G	A	G	T	T	C	G	G	G	C	G	C	C	G	C
E17	T	A	A	A	A	T	A	T	A	G	G	G	T	G	A	G	T	C	T	A	G	G	C	A	C	C	G	C
E18	T	G	A	T	A	T	A	C	T	G	G	A	T	A	A	A	T	C	C	G	G	A	C	A	C	T	A	C
E19	C	G	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	C	C	G	G	A	C	A	C	T	A	C
E20	T	G	A	T	A	T	A	C	T	G	G	G	T	A	A	A	T	C	C	G	G	A	T	A	C	T	A	C
E10	T	A	G	C	G	T	A	T	A	G	G	G	C	G	G	A	T	C	C	A	G	A	T	A	T	C	G	A
E11	T	A	G	C	G	T	A	T	A	G	G	G	C	G	G	A	T	C	C	A	G	A	T	A	T	C	G	C
E12	T	A	A	C	G	T	A	T	A	G	G	G	C	G	G	A	T	C	C	A	G	A	T	A	T	C	G	C
E13	T	A	G	C	G	T	A	T	A	G	G	G	C	G	G	A	T	C	C	A	A	A	T	A	T	C	G	C

Table 5. Mitochondrial DNA genetic diversity measures of *Emys*; *N*, sample size; n, number of haplotypes; *h*, haplotypic diversity; π, nucleotide diversity; *p*, number of polymorphic sites.

	Mitochondrial DNA				
Site	*N*	n	*h*± *SE*	π± *SE*	*p*
Isonzo	35	2	0.111 +/- 0.070	0.002 +/- 0.001	9
San Genuario	34	2	0.166 +/- 0.080	0.000 +/- 0.000	1
Cremona	33	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Bosco Nordio	28	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Bosco Mesola	39	2	0.051 +/- 0.048	0.000 +/- 0.000	1
Albenga	21	2	0.181 +/- 0.104	0.003 +/- 0.002	9
San Rossore	37	2	0.054 +/- 0.051	0.001 +/- 0.001	9
Maremma	32	4	0.657 +/- 0.067	0.006 +/- 0.004	12
Monte Rufeno	26	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Castelporziano	63	3	0.536 +/- 0.025	0.008 +/- 0.004	10
Circeo	12	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Ascoli Piceno	36	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Lago Serranella	29	3	0.305 +/- 0.101	0.001 +/- 0.001	10
Gargano	20	2	0.521 +/- 0.042	0.001 +/- 0.001	1
Le Cesine	27	6	0.749 +/- 0.051	0.004 +/- 0.003	15
Persano	23	4	0.731 +/- 0.057	0.013 +/- 0.007	18
Policoro	49	5	0.359 +/- 0.084	0.004 +/- 0.002	15
Lamezia Terme	34	3	0.633 +/- 0.040	0.011 +/- 0.006	15
Crotone	25	4	0.657 +/- 0.069	0.009 +/- 0.005	14
Berchidda	40	4	0.750 +/- 0.027	0.002 +/- 0.001	3
Pattada	10	3	0.689 +/- 0.104	0.001 +/- 0.001	2
Flumendosa	5	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Tirso	5	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Lago Spartà	9	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Valle Imera	19	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Bosco Ficuzza	18	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Gallitello	8	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Lago Preola	31	3	0.546 +/- 0.038	0.001 +/- 0.001	2
Siculiana	10	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Vendicari	19	1	0.000 +/- 0.000	0.000 +/- 0.000	0
Noto	7	2	0.286 +/- 0.196	0.000 +/- 0.001	1

Table 6. Nuclear DNA genetic diversity measures of *Emys*. *N*, sample size; Na, mean number of alleles per locus; Pa, private alleles; H_O, mean observed heterozygosity; H_E, mean expected heterozygosity; *HWE*, significant deviation from Hardy–Weinberg equilibrium after Bonferroni correction; F_{IS}, inbreeding coefficient, bold and underline values are significant at the 5% and 1% level, respectively.

	Microsatellites						
Site	*N*	Na	Pa	*Ho*	*He*	*HWE*	*Fis*
Isonzo	42	5.6	1	0.62 ± 0.04	0.62 ± 0.04	<**0.01**	-0.004
San Genuario	36	4.8		0.59 ± 0.06	0.56 ± 0.05		-0.065
Cremona	38	5.5		0.60 ± 0.07	0.59 ± 0.06		-0.027
Bosco Nordio	26	7.3	1	0.67 ± 0.07	0.64 ± 0.05		-0.054
Bosco Mesola	43	9.0	3	0.64 ± 0.04	0.65 ± 0.04		0.005
Albenga	25	3.0		0.37 ± 0.08	0.38 ± 0.08		0.032
San Rossore	52	7.1	2	0.64 ± 0.05	0.64 ± 0.05	<**0.01**	-0.007
Maremma	38	9.2	1	0.60 ± 0.07	0.68 ± 0.05		-0.029
Monte Rufeno	206	5.9	1	0.54 ± 0.05	0.57 ± 0.05	<**0.01**	**<u>0.054</u>**
Castelporziano	45	9.6		0.62 ± 0.05	0.75 ± 0.03	<**0.01**	**0.084**
Circeo	18	4.9		0.49 ± 0.08	0.54 ± 0.06		**0.097**
Ascoli Piceno	48	5.3	1	0.56 ± 0.06	0.60 ± 0.07		**0.075**
Lago Serranella	33	6.4	3	0.70 ± 0.05	0.68 ± 0.04		-0.033
Gargano	20	5.8		0.59 ± 0.05	0.63 ± 0.05		0.060
Le Cesine	28	9.3	4	0.64 ± 0.08	0.68 ± 0.06		0.058
Persano	23	9.3		0.78 ± 0.05	0.82 ± 0.02		0.045
Policoro	43	7.4	1	0.60 ± 0.06	0.62 ± 0.06	<**0.01**	0.032
Lamezia Terme	37	8.1	5	0.57 ± 0.12	0.64 ± 0.08	<**0.01**	**<u>0.112</u>**
Crotone	25	7.4	3	0.66 ± 0.05	0.66 ± 0.05		0.013
Berchidda	40	4.5		0.36 ± 0.08	0.36 ± 0.08		-0.153
Pattada	10	3.1		0.46 ± 0.13	0.43 ± 0.10		0.167
Flumendosa	5	2.5		0.53 ± 0.07	0.47 ± 0.06		0.007
Tirso	5	2.6		0.38 ± 0.12	0.44 ± 0.09		-0.083
Lago Spartà	15	3.9	1	0.52 ± 0.07	0.58 ± 0.07		0.111
Valle Imera	22	5.0	2	0.46 ± 0.09	0.52 ± 0.10		0.112
Bosco Ficuzza	24	3.8		0.55 ± 0.10	0.50 ± 0.09		**-0.104**
Gallitello	10	3.6		0.41 ± 0.10	0.49 ± 0.09		**0.170**
Lago Preola	30	7.1	4	0.62 ± 0.07	0.65 ± 0.08		0.046
Siculiana	14	5.0		0.54 ± 0.10	0.61 ± 0.10		**0.121**
Vendicari	20	6.1	2	0.59 ± 0.07	0.64 ± 0.08		**0.091**
Noto	8	3.9		0.53 ± 0.10	0.53 ± 0.09		-0.003

Population structure

ΔK method (Evanno et al. 2005), for structure analyses, suggested three as the optimal number of clusters (K=3; Fig. 7): the Sicilian pond turtle, and the others, identifiable by geographic location, with the two subspecies *E. o. galloitalica* on the west coast and *E. o. hellenica* on the east coast. However, because STRUCTURE is known to identify only the uppermost hierarchical level of genetic partitioning (Evanno et al. 2005), we also ran STRUCTURE for *E. orbicularis* data only, confirming that two is a likely number of populations (K=2 Fig. 7).

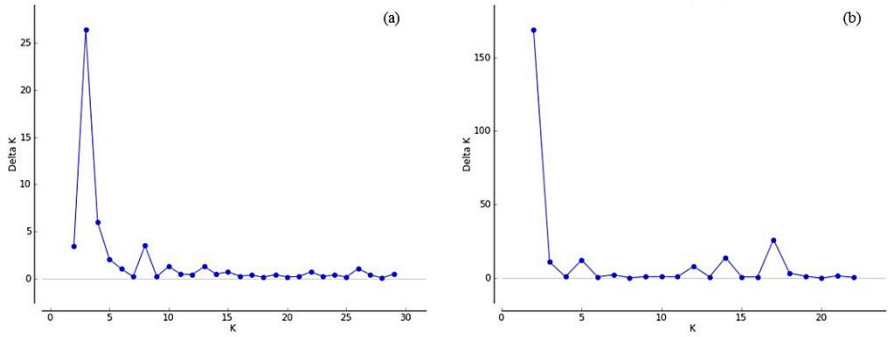

Figure 7. Evanno method plots for detecting the number of K groups that best fit the data. (a) K=3 for the entire data set; (b) K=2 for *Emys orbicularis* samples.

The results are evident in Figure 8; the colors of the three clusters are green (*E. trinacris*), red (*E. o. galloitalica*) and blue (*E. o. hellenica*). Southern Italy and the Tyrrhenian coast host turtle with largely admixed genotypes. Furthermore, the influence of *Emys trinacris* genes is observed not only in Calabria but also far away from any potential contact zone, in some individuals in Maremma and Castelporziano, and in a small percentage in individuals from Persano. Microsatellites of Calabrian individuals, despite not having Sicilian haplotypes, clearly show the mark of genetic impact by Sicilian pond turtles. On the other hand, there is a weaker introgression of *E. orbicularis* in south/eastern Sicilian *E. trinacris* populations.

Historical demography

Mismatch analysis was carried out for different hierarchical levels (Fig. 9, 10a - 10c): all data pooled, individuals that from previous analyses were assigned mainly to the *E. o. galloitalica* subspecies, to the *E. o. hellinca* subspecies, and to *Emys trinacris*, and for each sampling site. Multimodal mismatch distributions of genus *Emys* indicate stationary population sizes, but

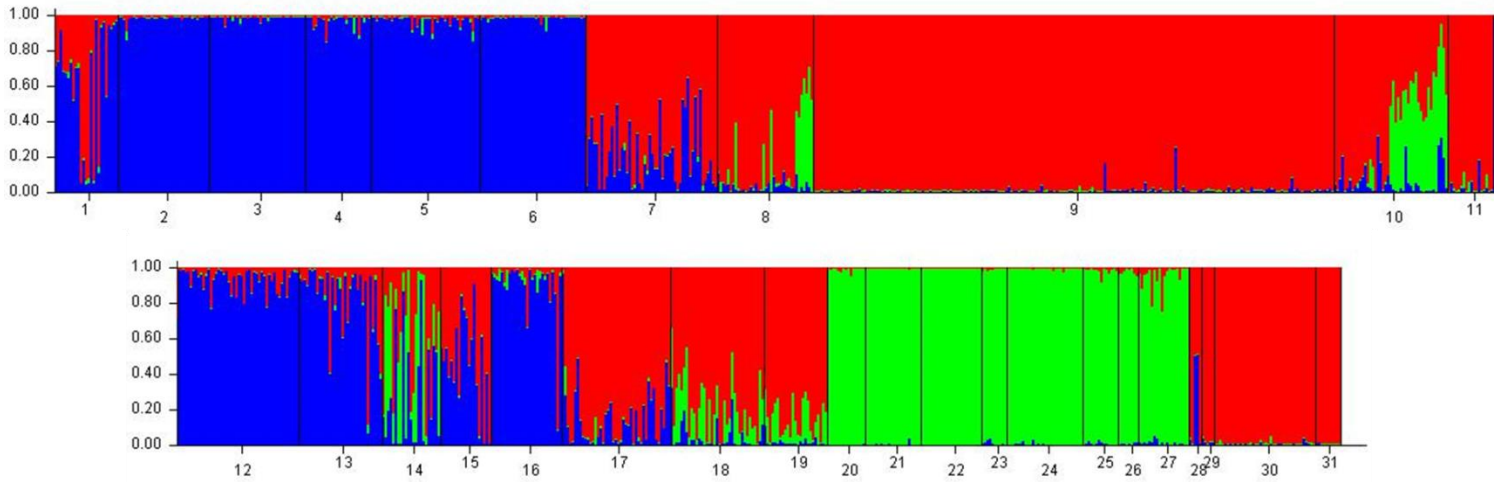

Figure 8. Results of STRUCTURE analysis. Black lines separate individuals from main different populations and each color represents a single cluster. The height of each color represents the individual's estimated membership fractions in K clusters. 1, Albenga; 2, San Genuario; 3, Cremona; 4,Bosco Nordio; 5 Bosco della Mesola; 6, Isonzo; 7, San Rossore; 8, Maremma; 9, Monte Rufeno; 10, Castelporziano; 11, Circeo; 12 Ascoli Piceno; 13, Lago di Serranella; 14, Persano; 15, Gargano; 16, Le Cesine; 17, Policoro; 18, Lamezia Terme; 19, Crotone; 20, Lago Spartà; 21 Valle dell'Imera; 22 Bosco Ficuzza; 23, Gallitello; 24, Lago Preola; 25, Siculiana; 26, Noto; 27, Vendicari; 28, Flumendosa; 29, Tirso; 30, Berchidda; 31, Pattada.

E. o. hellenica and the Sicilian pond turtle curve fit with the unimodal pattern expected under a model of population expansion. *E. trinacris* has little change between θ_0 and θ_1 , respectively the size of the population before and after the expansion event, and expansion isn't supported by SSD and Harpending test. The same applies to Tajima's *D* and Fu's F_S values that despite their negativity are not statistically significant. Instead *E. o. hellenica* shows a considerable difference in population size (θ_0=0 and θ_1=99999) and expansion hypothesis is confirmed by the other tests. For these two populations the estimated time for possible expansion that, considering a generation time of 10 years and the estimated mtDNA mutation rates of 0.4-0.5%/Myr (Avise – Lamb), was about 45,000-57,000 years, for the Sicilian pond turtle, that is congruent with the separation of Sicily and Calabria, and about 10,000-14,000 years, for populations on the Italian east coast at the end of the last glacial period.

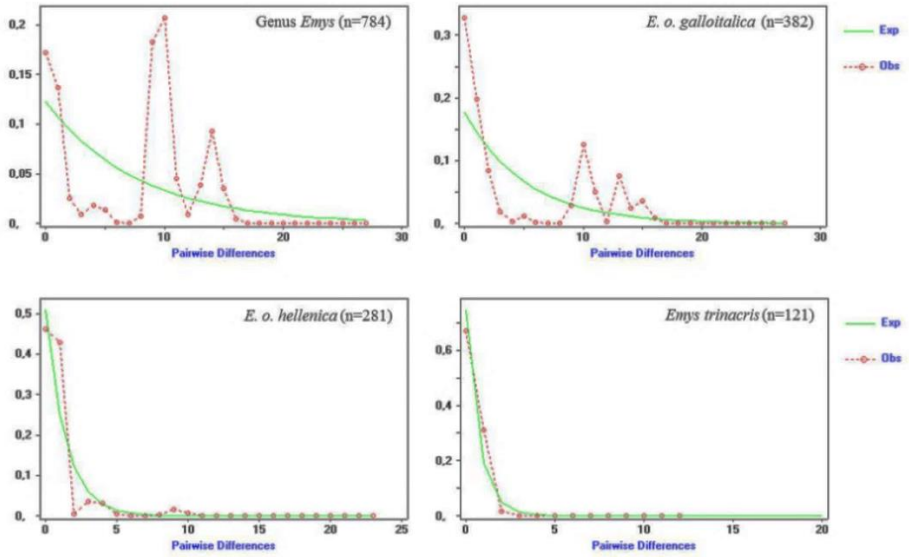

Figure 9. Pairwise distribution for all samples, *E. o. galloitalica*, *E. o. hellenica* and *Emys trinacris*. Solid green lines represent expected frequencies of pairwise differences under demographic expansion; broken red lines and circles, observed values.

At the individual sampling sites, the majority of samples show multimodal or bimodal curve, except five sites. San Genuario and Bosco della Mesola populations in spite of unimodal distribution did not show change between θ_0 and θ_1. The Gargano, Lago Preola and Noto populations did show such differences in population sizes, however, no statistical significance was recorded by the SSD and Harpending test. Also Tajima's *D* (Tajima 1989) and Fu's F_S values (Fu 1997) did not suggest population expansion, in spite of negative values for Noto.

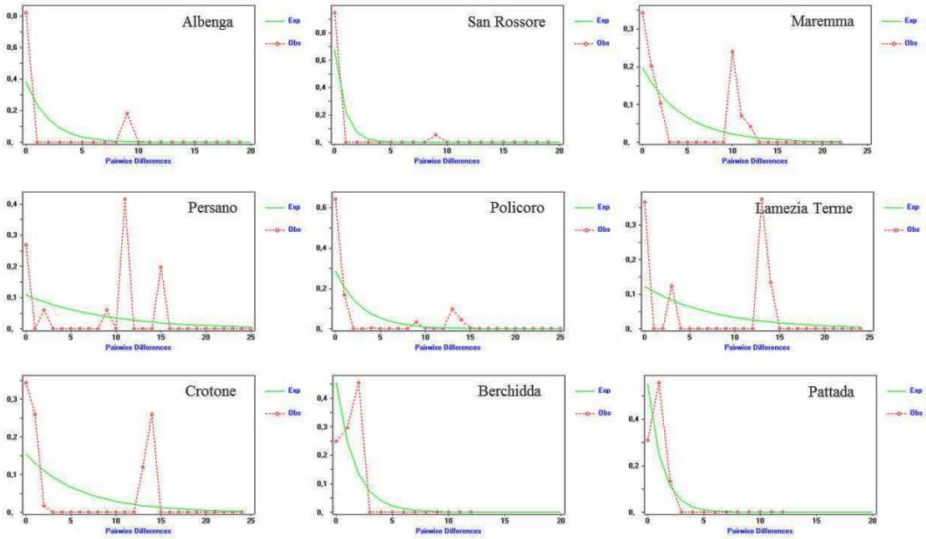

Figure 10a. Mismatch distributions of sample sites of populations grouped under *E. o. galloitalica*. Locations not shown have no polymorphism thus pairwise number of differences can't be computed.

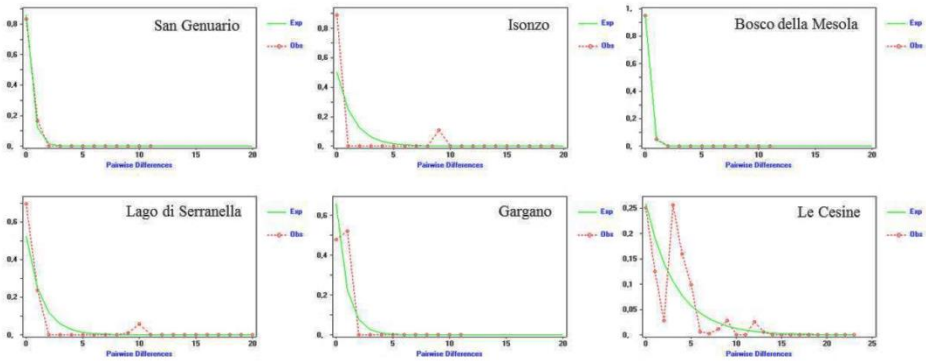

Figure 10b. Mismatch distributions of sample sites of populations grouped under *E. o. hellenica*. Locations not shown have no polymorphism thus pairwise number of differences can't be computed.

Bottleneck analysis, considering concordant results under SMM and TPM mutational models, did not find populations that experienced recent bottleneck events. Rather heterozygosity deficiency caused by allele excess due to expansion events, immigrants or populations substructure is demonstrated for Bosco Nordio, Bosco della Mesola, Maremma, Castelporziano, Le Cesine, Lamezia Terme and Noto.

Nested clade phylogeographic analysis

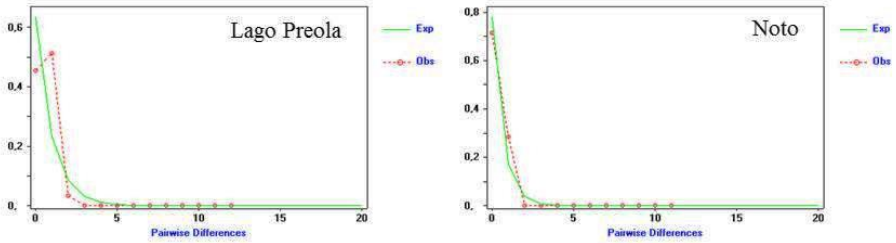

Figure 10c. Mismatch distributions of sample sites of populations grouped under *E. trinacris*. Locations not shown have no polymorphism thus pairwise number of differences can't be computed.

Figure 11 shows the estimated haplotype network. In the network, haplotypes are connected in the most parsimonious way and the overall number of putative mutations leading from one haplotype to another is minimized. Since the position of a haplotype in a network implies information about its age, the older haplotypes are thought to have a greater likelihood of being located internally; E03 has the biggest outgroup weight. Loops, an ambiguity due to occurrence of reverse or parallel mutations, formed between E15 and E08 by intermediate haplotypes that are not present in the sample, and which are in turn connected to E01, have been resolved using two criteria suggested by Crandalland Templeton (1993): (i) rare haplotypes are more likely to be found at the tip, and more common haplotypes at the interior nodes of a cladogram; and (ii) a singleton is more likely to be connected to haplotypes from the same population than to haplotypes from different populations.

The network was then partitioned into one-, two- and three-step clades (Fig. 12). The nested contingency analysis of geographical associations are shown in Table 7. Among the thirteen clades examined, ten demonstrated significant geographical association at the 0.05 level. Outcomes of the nested clade analysis of geographical distances (Fig. 13) show significant geographical associations for almost all clades. Geographical association was not observed in clades in nested contingency analysis (1-1, 1-6 and 3-4) that have haplotypes found in a single individual, and in 1-7 clade in geographical distances which present haplotype E13, that despite belonging to the Sicilian pond turtle is found almost exclusively in Campania. Biological interpretations of these statistical outcomes using the inference key of Templeton (2004) are shown in Table 8.

Despite the fact that our sampling covers much of the Italian distribution of *Emys*, the analysis was repeated by using all locations known in literature on the geographical list. The result did not change, except for clade 1-3 which showed inadequate geographical sampling. Results showed that a situation of restricted gene flow and some long-distance dispersal was the most common one.

DISCUSSION

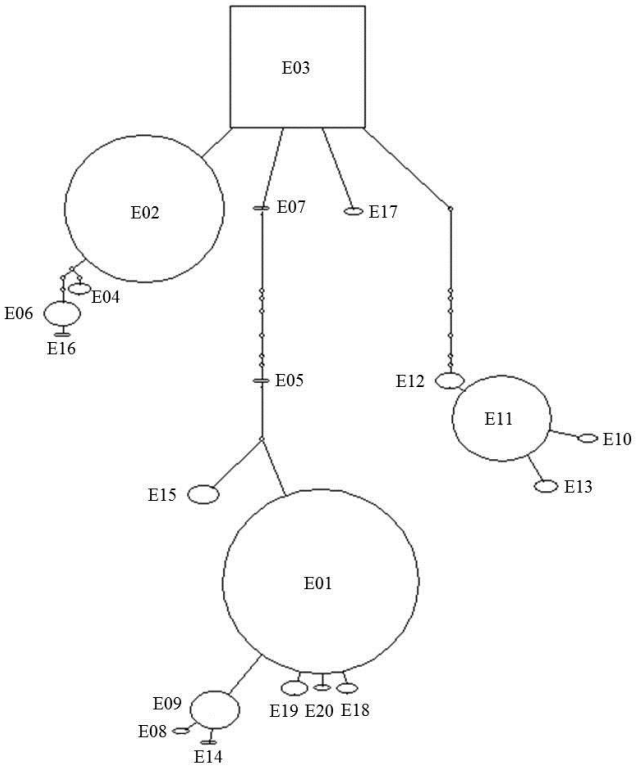

Figure 11. Mitochondrial control region haplotype network for genus *Emys*. Each haplotype is represented by a circle (or a square for the one with the biggest outgroup weight) whose area is proportional to the number of individuals bearing this particular haplotype. Small, empty circles correspond to intermediate haplotypes that are not present in the sample but are necessary to link all observed haplotypes to the network.

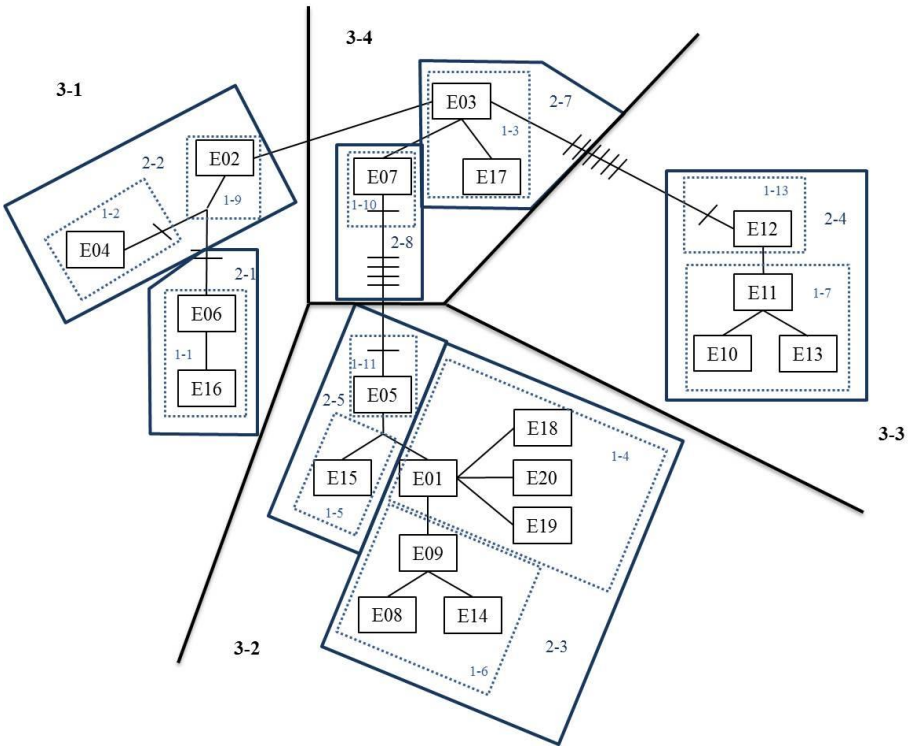

Figure 12. Nesting design of haplotype network in Fig. 11. To simplify the graphic representation, haplotypes are shown as rectangles no longer corresponding to the size of the sample, and intermediate haplotypes as small bars. Haplotypes enclosed by light blue dashed lines indicate the one-step clades, blue solid lines enclose the two-step clades and finally, a thick black solid line indicates the partitioning of the cladogram into four three-step clades.

Table 7. Nested contingency analysis of geographical associations. Values significant at the 0.05 level are in bold.

Nesting clade	Chi-square statistic	Probability
1-1	5.97	0.223
1-3	145	**0.000**
1-4	260.64	**0.000**
1-6	17.01	0.086
1-7	201.71	**0.000**
2-2	79.33	**0.000**
2-3	109.52	**0.000**
2-4	58.79	**0.000**
2-5	20	**0.049**
3-1	182.18	**0.000**
3-2	242.42	**0.000**
3-4	28.39	0.075
Total Cladogram	1800.08	**0.000**

This study provided in depth information on the genetic structure of the genus Emys in Italy, based on a comprehensive sampling of the main wild populations. We used molecular analysis of mitochondrial DNA control region and 16 nuclear microsatellite loci to recover patterns of gentic differentiation and phylogeographical history.

Italy provided a suitable habitat for *Emys orbicularis* both before and after the Pleistocene glaciations. Favorable characteristics, despite high mountain ranges (Alps and Apennines) being the main barriers to the dispersion of the species, included the vast floodplains, such as the Padana plain, coastal wetlands and mild climate. This was so until the XIX century, when anthropogenic impact started to considerably reduce its habitat and fragment populations. The Italian Peninsula shows the greatest diversity of all the species' range, hosting four *Emys orbicularis* subspecies, *E. o. galloitalica* on the west coast, *E. o. hellenica* on the east coast, *E. o. capolongoi* in Sardinia, the recently rediscovered *E. o. ingauna* in Liguria, and *Emys trinacris*, a species inhabiting only the island of Sicily.

Medium values of genetic variability suggested a stable situation for Italian populations, tending toward levels slightly more critical for populations in Sardinia, Sicily and Liguria. In Albenga, current conservation plans are contributing substantially to the preservation of the scarce number of *Emys orbicularis.* On the islands, especially Sardinia, the status of the species is not closely monitored. Unfortunately, in Italy there are few sites where the species is protected within the boundaries of reserves. In fact, *E. orbicularis* can also be

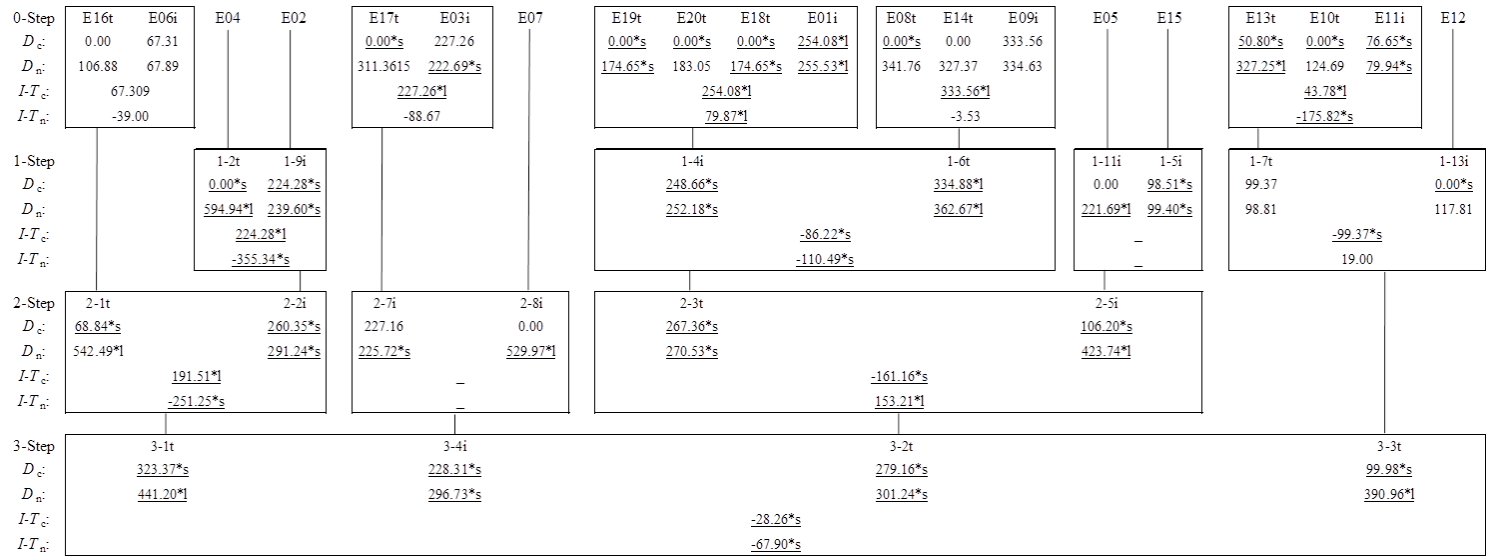

Figure 13. Nested clade analysis of geographical distances. Haplotypes (given in the first line) nested in one-step clades, as well as higher-level nested clades (portrayed as one moves down the figure), are grouped in boxes following the nesting design of Fig. 12. 'Interior' or 'tip' status is represented by 'i' and 't', respectively. Each box also contains the clade distance (*Dc*) and nested clade distance (*Dn*) calculated for each clade within the nested group, as well as the average difference in distances between interior and tip clades for *Dc* (*I–Tc*) and *Dn* (*I–Tn*). Statistically significant distance values (at the 0.05 level) are underlined and followed by a letter indicating whether the distance measure was significantly large ('l') or small ('s')

Table 8. Biological interpretations of nested clade phylogeographical analysis statistical outcomes using the inference key of Templeton (2004).

Nesting clade	Chain of iference	Demographic inference
1-1	Null hypothesis cannot be rejected	
1-3	1-19 NO	Allopatric fragmentation
1-4	1-2-3-4 NO	Restricted gene flow with isolation by distance
1-6	1-2-3-4 NO	Restricted gene flow with isolation by distance
1-7	1-2-3-5-6*-7-8 YES	Restricted gene flow/dispersal but with some long-distance dispersal over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations. Too few Clades: Insufficient genetic resolution to discriminate between range expansion/colonization and restricted dispersal/gene flow
2-2	1-2-3-5-6*-7-8 YES	Restricted gene flow/dispersal but with some long-distance dispersal over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations. Too few Clades: Insufficient genetic resolution to discriminate between range expansion/colonization and restricted dispersal/gene flow
2-3	1-2-11-12 NO	Contiguous range expansion
2-4	1-2-11-12 NO	Contiguous range expansion
2-5	1-19-20-2 IO	Inconclusive outcome
3-1	1-19-20-2-3-5-6*-7 YES	Restricted gene flow/dispersal but with some long-distance dispersal. Too few Clades: Insufficient genetic resolution to discriminate between range expansion/colonization and restricted dispersal/gene flow
3-2	1-2-11-12-13 YES	Long-distance colonisation possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
3-4	1-2 IO	Inconclusive outcome
Total	1-2-11-12-13-14 NO	Long-distance colonisation and/or past fragmenation (not necessarily mutually exclusive)

found in locations of high human impact, in small ponds close to cities and in the countryside among intensively cultivated fields. The increasing presence of syntopic alien species such as *Trachemys scripta* also represetn a threat to wild *Emys* populations. Genetic structure analysis supports division of Italian pond turtles into three clusters, compatible with the classification of the two main subspecies of *Emys orbicularis* and the Sicilian pond turtle.

Genetic diversity among populations, as expected, was lower between populations of the same ichthyogeographic areas, particularly in the Padan-Venetian area and Tuscan-Latium area. The latter, considering the information resulting from microsatellites, shows a major connection with southern areas. Furthermore, possible past-connectoin between *Emys trinacris* and *Emys orbicularis*, with preferred direction from Sicily to the Peninsula, was detected. All individuals in the two Calabrian populations exhibit a small percentage of Sicilian genotypes. This was also the case for Persano, in Campania, where 43% of turtles had the *Emys trinacris* mitochondrial haplotype E13. This population is not only the only one in the Italian mainland that bears a Sicilian haplotype, but E13 was found almost exclusively there, with a ratio of 10:1 in favor of Persano. A small number of turtles that shows an influence of Sicilian genotypes was also found in Maremma and Castelporziano populations. The latter was discussed by Fritz *et al.* (2005b) and Vamberger *et al.* (2015), who claimed that local turtles had nuclear genotypes of *E. o. galloitalica* but harbour mitochondrial haplotypes of *E. o. hellenica*, and therefore it is likely that non-native turtles from the Adriatic coast were introduced in the Castelporziano estate, a former royal hunting reserve. Our genetic data for Castelporziano were comparable with these studies but a small part of individuals bearing haplotypes considered characteristic of the east coast most likely belong to the Sicilian cluster. Moreover, the nested clade phylogeography analysis and the concept of multiple refugia by Gomezand Lunt (2007), validate a different hypothesis which is discussed below. Nonetheless Italy is a densely populated country, and introductions and relocations of turtles have probably occurred for centuries, like the well-known example of the Greek marginated tortoise population (*Testudo marginata*) established in Sardinia, which was possibly introduced in ancient or prehistoric times.

Analysis didn't point out evidence of a recent bottleneck but rather of rapid demographic expansion for populations on the east coast, over the last glacial periods, and for *Emys trinacris* about 50,000 years ago that coincides with the definite separation of Sicily from the Italian Peninsula.

Distribution of *Emys* control region haplotypes on Italian territory follow patterns found in species from temperate regions, with the highest diversity in the south, from three to six haplotypes in our study, and a northern lower diversity, with only one or two haplotypes among the most frequent ones. An exception seems to be Maremma and North Sardinia populations that have three to four haplotypes per site, of which four are unique. According to the nested clade phylogeographic analysis, associations were mostly statistically

significant in every group and E03, the haplotype suspected to be "allochthonous" on the west coast, was indicated as the oldest one. Clades showed restricted gene flow with isolation by distance and some events of long distance dispersals. A clear picture was illustrated mainly for Sicilian pond turtles (clade 2-4), that have contiguous range expansion with restricted gene flow or dispersal, with some long-distance dispersal over intermediate areas not occupied by the species or past gene flow followed by extinction of intermediate populations. Haplotypes strictly connected with E01 (clade 2-3 and 3-2), mostly from the Thyrrenian coast, showed contiguous range expansion and restricted gene flow with isolation by distance and long-distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion.

The endangered Italian pond turtle, although it has not yet reached critical levels of loss of variability, increasingly suffers decline in areas of occupancy, extent of occurrence, quality of habitat and the effects of introduced taxa, pathogens, pollutants, competitors and parasites. In ancient times, Italy, an area rich in wetlands, played a key role as a refuge during main ice ages.

The primary refugia was indubitably the extreme south, but hypothesis of secondary refugia is confirmed by our results. Repeated severe climatic oscillations produced great changes in species distribution, with many range contractions and expansions, thus subjecting various lineages to different dynamics, such as extinction, admixture or allo-parapatric speciation. Tuscany and Latium, close to the northern Apennines and principal barrier to the northern dispersal of organisms, almost certainly acted as a refugia for *Emys orbicularis* in less rigid glacial periods. Sardinia, due to isolation, hosts very differentiated turtles most likely resulting from a dispersal route from an Italian central refugia. We believe an origin isn't plausible via introduction of Sardinian populations (Vamberger *et al.* 2015). To confirm this a study on the population of Corsica would also be appropriate. The discovery of this secondary refugia and the identification of E03 haplotype as the oldest could mean the existence of an alternative route of dispersion which might have included both coastal and central corridors through the southern Apennines. This is a different theory that does not necessarily completely rule out the other, which is the introduction of Adriatic coast individuals to Tyrrhenian coast reserves. If so, we might also expect the reverse situation, but there were no recorded massive introductions of westeren individuals to the east coast. What is difficult to interpret is the main presence of *Emys trinacris* haplotype E13 in Campania compared to only one individual out of 121 turtles found in Sicily, and the absence of Sicilian haplotypes in Calabria. However, microsatellites show a non-negligible Sicilian influence in Calabria and on the west coast. Moreover, a lower concentrations of *Emys orbicularis* is found in the Sicilian populations of the south-east, evidence of a past gene flow that nowadays would be practically impossible because of substantial population fragmentation.

Samplings of other Sicilian populations and others from the Calabria and Campania areas could be considered.

The genus *Emys* on the Italian Peninsula shows dynamics found in many other organisms and areas. The Apennines obviously function as an impediment to movements, especially the northern chain, given their continuity. However, our study shows clear connection between southern, central and northern turtles, while distinct differentiation was found in other reptile studies across Italy (Barbanera *et al.* 2009; Canestrelli *et al.* 2007a; Canestrelli & Nascetti 2008; Mattoccia *et al.* 2011). This study describes for the first time the existence of a secondary refugia for *Emys orbicularis* in Italy, the diversification of Sardinian pond turtles, and that supposedly "allochthonous" individuals on the west coast could be actuality turtles harbouring an ancient lineage that took refuge, in part, in central regions. Italy is surely a key zone for survival of the endangered *Emys orbicularis* and *Emys trinacris*, the latter confined solely to Sicily. We envisage the possibility of increasing sampling sites to gather further information which may help development of conservation plans for a species that, so far, has been thoroughly studied only for taxonomic purposes.

REFERENCES

Amorosi, A., Colalongo, M.L., Fiornini, F., Fusco, F., Pasini, G., Vaiani, S.C. & Sarti, G. (2004) Palaeogeographic and palaeoclimatic evolution of the Po plain from 150-ky core records. Global and Planetary Change, 40, 55–78.

ArisBrosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Molecular Biology and Evolution 13, 494-504.

*Avise JC, Bowen BW, Lamb T, Meylan AB, Bermingham E (1992) Mitochondrial-DNA Evolution at a Turtles Pace - Evidence for Low Genetic-Variability and Reduced Microevolutionary Rate in the Testudines. Molecular Biology and Evolution **9**, 457-473.*

*Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Molecular Ecology **13**, 729-744.*

*Barbanera F, Zuffi MAL, Guerrini M, et al. (2009) Molecular phylogeography of the asp viper Vipera aspis (Linnaeus, 1758) in Italy: Evidence for introgressive hybridization and mitochondrial DNA capture. Molecular Phylogenetics and Evolution **52**, 103-114.*

Bianco PG (1990). Potential role of the paleohistory of the Mediterranean and Paratethys basins on the early dispersal of Euro-Mediterranean freshwater fishes. Icht Expl Fresh 1: 167–184.

*Bohme MU, Fritz U, Kotenko T, et al. (2007) Phylogeography and cryptic variation within the Lacerta viridis complex (Lacertidae, Reptilia). Zoologica Scripta **36**, 119-131.*

Brewer S, Cheddadi R, de Beaulieu J, Reille M & Data Contributors (2002) The spread of deciduous Quercus throughout Europe since the last glacial period. Forest Ecology and Management 156: 27-48.

Burbrink F.T., Lawson R. (2007) How and when did Old World ratsnakes disperse into the New World?. Mol. Phylogenet. Evo. 43:173–189.

Caloi, L., Malatesta, A. & Palombo, R.M. (1989) Biogeografia della Calabria meridionale durante il Quaternario. Atti Accademia Peloritana dei Pericolanti Classe I, Scienze Matematiche Fisiche e Naturali, 67, 261–278.

*Canestrelli D, Aloise G, Cecchetti S, Nascetti G (2010) Birth of a hotspot of intraspecific genetic diversity: notes from the underground. Molecular Ecology **19**, 5432-5451.*

*Canestrelli D, Cimmaruta R, Nascetti G (2007a) Phylogeography and historical demography of the Italian treefrog, Hyla intermedia, reveals multiple refugia, population expansions and secondary contacts within peninsular Italy. Molecular Ecology **16**, 4808-4821.*

*Canestrelli D, Cimmaruta R, Nascetti G (2008) Population genetic structure and diversity of the Apennine endemic stream frog, Rana italica - insights on the Pleistocene evolutionary history of the Italian peninsular biota. Molecular Ecology **17**, 3856-3872.*

*Canestrelli D, Nascetti G (2008) Phylogeography of the pool frog Rana (Pelophylax) lessonae in the Italian peninsula and Sicily: multiple refugia, glacial expansions and nuclear-mitochondrial discordance. Journal of Biogeography **35**, 1923-1936.*

*Canestrelli D, Verardi A, Nascetti G (2007b) Genetic differentiation and history of populations of the Italian treefrog Hyla intermedia: lack of concordance between mitochondrial and nuclear markers. Genetica **130**, 241-255.*

*Canestrelli D, Zangari F, Nascetti G (2006) Genetic evidence for two distinct species within the Italian endemic Salamandrina terdigitata (Bonnaterre, 1789) (Amphibia : Urodela : Salamandridae). Herpetological Journal **16**, 221-227.*

*Castelloe J, Templeton AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. Molecular Phylogenetics and Evolution **3**, 102-113.*

*Ciofi C, Tzika AC, Natali C, et al. (2009) Characterization of microsatellite loci in the European pond turtle Emys orbicularis. Molecular Ecology Resources **9**, 189-191.*

*Claude J, Paradis E, Tong H, Auffray JC (2003) A geometric morphometric assessment of the effects of environment and cladogenesis on the evolution of the turtle shell. Biological Journal of the Linnean Society **79**, 485-501.*

*Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology **9**, 1657-1659.*

*Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics **153**, 1989-2000.*

Corregiari, A., Roveri, M. & Trincardi, F. (1996) Late Pleistocene and Holocene evolution of the North Adriatic Sea. Il Quaternario, 9, 697–704.

Crandall KA, Templeton AR (1993) Empirical Tests of Some Predictions from Coalescent Theory with Applications to Intraspecific Phylogeny Reconstruction. Genetics 134, 959-969.

*Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nature Methods **9**, 772-772.*

*Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources **4**, 359-361.*

*Eo SH, DeWoody JA (2010) Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. Proceedings of the Royal Society B-Biological Sciences **277**, 3587-3592.*

Ernst CH, Barbour RW, Hershey MF (1974) A new coding system for hard-shelled turtles. Transactions of the Kentucky Academy of Science 35, 27-28.

*Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology **14**, 2611-2620.*

*Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Molecular Ecology **13**, 853-864.*

*Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources **10**, 564-567.*

Ferranti, L., Antonioli, F., Mauz, B., Amorosi, A., Dai Pra, G., Mastronuzzi, G., Monaco, C., Orru` , P., Pappalardo, M., Radtke, U., Renda, P., Romano, P., Sanso` , P. & Verrubbi, V. (2006). Markers of the last interglacial sea-level high stand along the coast of Italy: tectonic implications. Quaternary International, 145–146, 30–54.

*Fetzner JW, Crandall KA (2003) Linear habitats and the nested clade analysis: An empirical evaluation of geographic versus river distances using an ozark crayfish (Decapoda : Cambaridae). Evolution **57**, 2101-2118.*

Follieri M, Magri D. (1997) Paesaggi vegetali del Quaternario in Italia centrale. Biogeographia 19: 57–68.

*Fratini S, Zaccara S, Barbaresi S, et al. (2005) Phylogeography of the threatened crayfish (genus Austropotamobius) in Italy: implications for its taxonomy and conservation. Heredity **94**, 108-118.*

Fritz U (1998) Introduction to zoogeography and subspecific differentiation in Emys orbicularis (Linnaeus, 1758). In: eds Fritz U, Joger U, Podloucky R, Servan J) Proceedings of the Emys Symposium Dresden 96. Mertensiella 10, pp. 1-27. Warlich, Rheinbach.; IUCN Red List of threatened species; Fritz 2001.

*Fritz U, Ayaz D, Hundsdorfer AK, et al. (2009) Mitochondrial diversity of European pond turtles (Emys orbicularis) in Anatolia and the Ponto-Caspian Region: Multiple old refuges, hotspot of extant diversification and critically endangered endemics. Organisms Diversity & Evolution **9**, 100-114.*

*Fritz U, Cadi A, Cheylan M, et al. (2005a) Distribution of mtDNA haplotypes (cyt b) of Emys orbicularis in France and implications for postglacial recolonization. Amphibia-Reptilia **26**, 231-238.*

*Fritz U, d'Angelo S, Pennisi MG, Lo Valvo M (2006) Variation of Sicilian pond turtles, Emys trinacris - What makes a species cryptic? Amphibia-Reptilia **27**, 513-529.*

*Fritz U, Fattizzo T, Guicking D, et al. (2005b) A new cryptic species of pond turtle from southern Italy, the hottest spot in the range of the genus Emys (Reptilia, Testudines, Emydidae). Zoologica Scripta **34**, 351-371.*

*Fritz U, Guicking D, Kami H, et al. (2007) Mitochondrial phylogeography of European pond turtles (Emys orbicularis, Emys trinacris) - an update. Amphibia-Reptilia **28**, 418-426.*

*Fritz U, Guicking D, Lenk P, Joger U, Wink M (2004) When turtle distribution tells European history: mtDNA haplotypes of Emys orbicularis reflect in Germany former division by the Iron Curtain. Biologia **59**, 19-25.*

Giraudi C (2004) The Apennine glaciations in Italy. Quaternary glaciations – extent and chronology, Part I: Europe (ed. by J. Ehlers and P.L. Gibbard), pp. 215–224. Elsevier, Amsterdam.

Gomez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. Phylogeography of Southern European Refugia, 155-188.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of Heredity 86, 485-486.

*Guo SW, Thompson EA (1992) Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. Biometrics **48**, 361-372.*

*Hewitt GM (1999) Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society **68**, 87-112.*

*Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **359**, 183-195.*

*Hewitt GM (2011) Quaternary phylogeography: the roots of hybrid zones. Genetica **139**, 617-638.*

Hudson RR (1990) Gene genealogies and the coalescent proces, pp. 1-44 in Oxford Surveys in Evolutionary Biology, edited by Futuyama, and J. D. Antonovics. Oxford University Press, New York.

Jesu R., Piombo R., Salvidio S., Lamagni L., Ortale S., Genta P. (2004) Un nuovo taxon di testuggine palustre endemico della Liguria Occidentale. Ann. Mus. Civ. St. Nat. "G. Doria" Genova.

*Jiang Y, Nie LW, Huang ZF, et al. (2011) Comparison of complete mitochondrial DNA control regions among five Asian freshwater turtle species and their phylogenetic relationships. Genetics and Molecular Research **10**, 1545-1557.*

*Joger U, Fritz U, Guicking D, et al. (2007) Phylogeography of western Palaearctic reptiles - Spatial and temporal speciation patterns. Zoologischer Anzeiger **246**, 293-313.*

Konnert M, Bergmann F (1995) The geographical distribution of genetic variation of silver fir (Abies alba, Pinaceae) in relation to its migration history. Plant Systematics and Evolution, 196, 19–30.

Lanza B (1983) Ipotesi sulle origini del popolamento herpetologico della Sardegna. Lav. Soc. Ital. Biogeogr. (1980)8: 723-744.

*Lamb T, Lydeard C, Walker RB, Gibbons JW (1994) Molecular Systematics of Map Turtles (Graptemys) - a Comparison of Mitochondrial Restriction Site Versus Sequence Data. Systematic Biology **43**, 543-559.*

*Lenk P, Fritz U, Joger U, Wink M (1999) Mitochondrial phylogeography of the European pond turtle, Emys orbicularis (Linnaeus 1758). Molecular Ecology **8**, 1911-1922.*

*Lenk P, Joger U, Wink M (2001) Phylogenetic relationships among European ratsnakes of the genus Elaphe Fitzinger based on mitochondrial DNA sequence comparisons. Amphibia-Reptilia **22**, 329-339.*

*Lenk P, Wuster W (1999) A multivariate approach to the systematics of Italian rat snares of the Elaphe longissima complex (Reptilia, Colubridae): Revalidation of Camerano's Callopeltis longissimus var. Lineata. Herpetological Journal **9**, 153-162.*

*Li X-s, Nie L-w, Wang L, Xiong L, Zhou K (2010) The mitochondrial genome complete sequence and organization of the Pig-nosed Turtle Carettochelys insculpta (Testudines, Carettochelyidae) and its phylogeny position in Testudines. Amphibia-Reptilia **31**, 541-551.*

*Lourenco JM, Claude J, Galtier N, Chiari Y (2012) Dating cryptodiran nodes: Origin and diversification of the turtle superfamily Testudinoidea. Molecular Phylogenetics and Evolution **62**, 496-507.*

*Lourenco JM, Glemin S, Chiari Y, Galtier N (2013) The determinants of the molecular substitution process in turtles. Journal of Evolutionary Biology **26**, 38-50.*

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

Mattoccia M, Marta S, Romano A, Sbordoni V (2011) Phylogeography of an Italian endemic salamander (genus Salamandrina): glacial refugia, postglacial expansions, and secondary contact. Biological Journal of the Linnean Society ***104****, 903-922.*

Panchal M, Beaumont MA (2007) The automation and evaluation of nested clade phylogeographic analysis. Evolution ***61****, 1466-1480.*

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics ***28****, 2537-2539.*

Pedall I, Fritz U, Stuckas H, Valdeon A, Wink M (2011) Gene flow across secondary contact zones of the Emys orbicularis complex in the Western Mediterranean and evidence for extinction and re-introduction of pond turtles on Corsica and Sardinia (Testudines: Emydidae). Journal of Zoological Systematics and Evolutionary Research ***49****, 44-57.*

Pedall I, Schafer H, Fritz U, Wink M (2009) Isolation of microsatellite markers in the Emys orbicularis complex and development of multiplex PCR amplification. Conservation Genetics ***10****, 725-727.*

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity ***90****, 502-503.*

Podnar M, Mayer W, Tvrtkovic N (2005) Phylogeography of the Italian wall lizard, Podarcis sicula, as revealed by mitochondrial DNA sequences. Molecular Ecology ***14****, 575-588.*

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Molecular Ecology ***9****, 487-488.*

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics ***155****, 945-959.*

Prusak B, Mitrus S, Najbar B, et al. (2013) Population differentiation of the European pond turtle (Emys orbicularis) in Poland inferred by the analysis of mitochondrial and microsatellite DNA: implications for conservation. Amphibia-Reptilia ***34****, 451-461.*

Rogers AR, Harpending H (1992) Population-Growth Makes Waves in the Distribution of Pairwise Genetic-Differences. Molecular Biology and Evolution ***9****, 552-569.*

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources ***8****, 103-106.*

Ruedi M, Walter S, Fischer MC, et al. (2008) Italy as a major ice age refuge area for the bat Myotis myotis (Chiroptera : Vespertilionidae) in Europe. Molecular Ecology ***17****, 1801-1814.*

Salomone N, Vignoli V, Frati F, Bernini F (2007) Species boundaries and phylogeography of the "Euscorpius carpathicus complex" (Scorplones :

*Euscorpiidae) in Italy. Molecular Phylogenetics and Evolution **43**, 502-514.*

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press.

*Sanmartin I, Enghoff H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biological Journal of the Linnean Society **73**, 345-390.*

*Santucci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. Molecular Ecology **7**, 1163-1172.*

*Santucci F, Nascetti G, Bullini L (1996) Hybrid zones between two genetically differentiated forms of the pond frog Rana lessonae in southern Italy. Journal of Evolutionary Biology **9**, 429-450.*

*Sbisa E, Tanzariello F, Reyes A, Pesole G, Saccone C (1997) Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene **205**, 125-140.*

*Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates very among sites: Application to human mitochondrial DNA. Genetics **152**, 1079-1089.*

*Seddon JM, Santucci F, Reeve N, Hewitt GM (2002) Caucasus Mountains divide postulated postglacial colonization routes in the white-breasted hedgehog, Erinaceus concolor. Journal of Evolutionary Biology **15**, 463-467.*

*Seddon JM, Santucci F, Reeve NJ, Hewitt GM (2001) DNA footprints of European hedgehogs, Erinaceus europaeus and E-concolor. Pleistocene refugia, postglacial expansion and colonization routes. Molecular Ecology **10**, 2187-2198.*

*Slatkin M, Hudson RR (1991) Pairwise Comparisons of Mitochondrial-DNA Sequences in Stable and Exponentially Growing Populations. Genetics **129**, 555-562.*

*Sommer RS, Benecke N (2005) The recolonization of Europe by brown bears Ursus arctos Linnaeus, 1758 after the Last Glacial Maximum. Mammal Review **35**, 156-164.*

*Sommer RS, Lindqvist C, Persson A, et al. (2009) Unexpected early extinction of the European pond turtle (Emys orbicularis) in Sweden and climatic impact on its Holocene range. Molecular Ecology **18**, 1252-1262.*

*Sommer RS, Persson A, Wieseke N, Fritz U (2007) Holocene recolonization and extinction of the pond turtle, Emys orbicularis (L., 1758), in Europe. Quaternary Science Reviews **26**, 3099-3107.*

Spinks PQ, Shaffer HB (2009) Conflicting Mitochondrial and Nuclear Phylogenies for the Widely Disjunct Emys (Testudines: Emydidae)

Species Complex, and What They Tell Us about Biogeography and Hybridization. Systematic Biology ***58****, 1-20.*

Starkey DE, Shaffer HB, Burke RL, et al. (2003) Molecular systematics, phylogeography, and the effects of Pleistocene glaciation in the painted turtle (Chrysemys picta) complex. Evolution ***57****, 119-128.*

Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics 143, 1457-1465.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology ***7****, 381-397.*

Templeton AR (2002) Out of Africa again and again. Nature 416, 45-51.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. Molecular Ecology ***13****, 789-809.*

Templeton AR, Boerwinkle E, Sing CF (1987) A Cladistic-Analysis of Phenotypic Associations with Haplotypes Inferred from Restriction Endonuclease Mapping .1. Basic Theory and an Analysis of Alcohol-Dehydrogenase Activity in Drosophila. Genetics ***117****, 343-351.*

Templeton AR, Crandall KA, Sing CF (1992) A Cladistic-Analysis of Phenotypic Associations with Haplotypes Inferred from Restriction Endonuclease Mapping and DNA-Sequence Data .3. Cladogram Estimation. Genetics ***132****, 619-633.*

Templeton AR, Routman E, Phillips CA (1995) Separating Population-Structure from Population History - a Cladistic-Analysis of the Geographical-Distribution of Mitochondrial-DNA Haplotypes in the Tiger Salamander, Ambystoma-Tigrinum. Genetics ***140****, 767-782.*

Vamberger M, Stuckas H, Sacco F, et al. (2015) Differences in gene flow in a twofold secondary contact zone of pond turtles in southern Italy (Testudines: Emydidae: Emys orbicularis galloitalica, E. o. hellenica, E. trinacris). Zoologica Scripta 44, 233-249.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes ***4****, 535-538.*

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology ***12****, 563-584.*

Zhang Y, Nie L, Huang Y, Pu Y, Zhang L (2009) The Mitochondrial DNA Control Region Comparison Studies of Four Hinged Turtles and its Phylogentic Significance of the Genus Cuora Sensu Lato (Testudinata: Geoemydidae). Genes & Genomics ***31****, 349-359.*

2. IDENTIFICATION AND MANAGEMENT OF EUROPEAN POND TURTLES OF UNKNOWN ORIGIN BASED ON DNA ANALYSIS OF A COMPREHENSIVE ITALIAN NATURAL POPULATION DATABASE

ABSTRACT

The endangered European pond turtle *Emys orbicularis* and Sicilian pond turtle *Emys trinacris* display a fragmented distribution in Italy and unfortunately not all natural populations are protected within the borders of nature reserves. Despite Italian turtles not being very well known to the general public, they appear to greatly attract breeders, so it is not uncommon for animals to be removed from the wild and to find individuals released outside of their habitat. In this study, we used eight microsatellite loci to identify the source population of 79 pond turtles of unknown origin from two Recovery Centers in Emilia Romagna (Italy). We performed a partially Bayesian exclusion test of Cornuet *et al.* (1999) and the fully Bayesian assignment test of Pritchard *et al.* (2000). By comparing these to 1029 individuals studied in Chapter 2 the population of origin was identified for about 75% of the turtles.

INTRODUCTION

Dispersal of organisms across biogeographic barriers in the past was a low-probability event, while today it is recurrent. Most of the dispersion of species has been promoted, by international commerce (Mooney & Cleland 2001). Wildlife and their products are among the items most commonly traded illegally (Robert 2000). Reptiles are becoming increasingly popular as pets and freshwater turtles are among the most popular in the world. From the United States, between 1989 through 1997, more than 57.8 million reptiles, representing over 570 taxa, were exported, including over 53.7 million turtles and tortoises (Telecky 2001). The species that dominate the pet trade is *Trachemys scripta* with about 6 million individuals exported from the United States every year. This turtle is included in the 100 World's Worst Invasive Alien Species of IUCN and is the main competitor of the endangered European pond turtle.

Emys orbicularis has a widespread distribution from West Europe to Russia, throughout Mediterranean countries, including North Africa, and Asia Minor (Fritz 1998), but is clearly in decline in most of its range; in fact it is listed on the IUCN Red List of threatened species as "Near threatened". The main causes of danger are decrease of freshwater ecosystems which are the environments most at risk, fragmentation of habitat due to human impact like urbanization, farming and agricultural purposes, and also competitive pressure of exotic species (Herborg *et al.* 2007; Padilla & Williams 2004). In the past, *Emys orbicularis* was kept in captivity for ornamental purposes. Today, despite the fact that possession and sale are prohibited and the species is not as well known as the American slider, it is still attractive, especially to freshwater turtles farmers, so much so that the census of Telecky (2001) reported up to a few hundred marketed per year.

Removal of animals from wild populations contributes to increased population vulnerability, especially in the case of rare species, and when these animals become large as adults, there is a tendency to eitehr release them into natural habitats or collect individuals in recovery centers and zoos. Nowadays there are recurrent management plans that determine the probable region of origin of a particular specimen. Introducing individuals from very different regions, crossing species or subspecies, can cause genetic homogenization, hybridization and outbreeding depression that might results in reduced reproductive fitness. Localizing organisms inhabiting a sole region or with a unique phenotypic character can be simpler but is not always enough. For these reasons reliable methods which rely on molecular techniques are increasingly necessary to identify the source population of animals of unknown origin.

In this study, 79 *Emys orbicularis* individuals found abandoned outside of their habitat or left by the public at collection centers of freshwaters turtle in Emilia Romagna (Italy) were analyzed using eight microsatellite loci (Ciofi *et al.* 2009b) to determine the population of origin. This was possible due to the a large genetic databank created by analyzing over a thousand samples of the European pond turtle *E. orbicularis* from the Italian Peninsula and the Sicilian pond turtle *E. trinacris*, for a total of 31 populations, covering all Italian regions except Valle d'Aosta and Trentino Alto Adige, (where the species has almost completely disappeared), and Molise. Given the few individuals analyzed, the Flumendosa and Tirso populations from Sardinia (see chapter 2) were excluded.

MATERIALS AND METHODS

Molecular techniques

Total genomic DNA was extracted from blood samples by overnight incubation at 37 °C in lysis buffer (100 mM Tris pH 8, 5 mM EDTA pH 8, 100 mM NaCl, 0.5% SDS pH 7.2) including 200 µg of proteinase K, followed by a standard phenol/chloroform extraction (Sambrook *et al.* 1989). DNA was precipitated from the supernatant with two volumes of cold 100% Ethanol, centrifugated, dried and resuspended in TE buffer. Then a multiplex PCR was performed, reaction volume of 10 µL contain 1 U Taq DNA Polimerase Cloned (5 U/µL Life Technologies), 1x Buffer (Life Technologies), 1.5 mM $MgCl_2$ (Life Technologies), 300 µM dNTP (Life Technologies) and 0.5 µM primers. After an initial denaturation step at 94 °C for 5 min, 35 cycles of 40 s at 94 °C, 40 s at annealing temperature (Ciofi *et al.* 2009b) and 1 min at 72 °C extension temperature, with a final extension step of 5 min at 72 °C, were performed. Fragment lengths were determined on a 3130xl Genetic Analyzer (Applied Biosystems) with the GeneScan 500 LIZ dye Size Standard and allele size was assessed using Genemapper version 5 (Applied Biosystems).

Statistical analysis

We chose to use a Bayesian approach, that provides slightly higher assignment accuracy than the first assignation tests performed based on frequency statistics (Paetkau *et al.* 1995) which calculate the probability of drawing a single multilocus genotype from each potential source by using the observed allele frequencies at each locus in each source population. This was done by means of GENECLASS2 software (Piry *et al.* 2004b), that uses the Rannalaand Mountain (1997a) method, a Bayesian approach to estimate population allele frequencies, and the Cornuet *et al.* (1999) method that extends assignment methods including a statistical significance test based on Monte Carlo simulations. The fully Bayesian method of Pritchard *et al.* (2000b) was also used to confirm results although it requires that the true population of origin be sampled.

In the assignation test in GENECLASS2, the probability of belonging to each population is computed separately by simulating a frequency distribution of 10,000 genotypes for each population. If the genotype of the individual to assign is observed less than once in 1000 randomly simulated genotypes, then population X would be excluded as the origin ($p<0.001$). For the Bayesian method in STRUCTURE (Pritchard *et al.* 2000b), the probability that a genotype originates from a population is computed by considering all the sampled populations jointly. The fully Bayesian assignment test computes the posterior probability that the sampled individual is from one population or the other. However, this method always identifies a population with a higher probability of being the source population, assuming a priori that all possible origin populations were sampled. Notwithstanding the extensive sampling on Italian territory of main natural populations of *Emys* described in chapter 2 and therefore having most likely analyzed the true populations of origin of unknown individuals, we consider it appropriate to proceed with the comparison of all methods to get the most reliable results possible.

RESULTS

Genetic profiles of wild populations were described in chapter 2. In that survey, structure analysis showed a clear division of Italian turtles into three clusters (K=3). The first includes populations inhabiting the Adriatic coast and Padana Plain, the second comprises populations from the west coast and the last one consdiers the Sicilian pond turtles. Analysis of differentiation among population displays a definite separation of *E. trinacris* from the Peninsula's populations. The population of Albenga also show a significant degree of genetic differentiation. A moderate degree of genetic differentiation was found across the west coast, Sardinia, the east coast and in the South, and to a lower extent among populations from the mid-west, in Tuscany and Lazio, and the Po Valley, except San Genuario.

Forty-one percent of the 79 *Emys orbicularis* individuals analyzed with GENECLASS were successfully assigned to the population in which their

multilocus microsatellite genotype was most likely to occur (Table 1a-1e). Twenty-one specimens were shown to originate in Riserva Naturale Bosco della Mesola in Emilia Romagna, a natural coastal environment, six in the Riserva Naturale Integrale Bosco Nordio in Veneto, a remnant of the wide forest and wetland zone that characterized much of the Veneto coast in the past, close to the Adige River. One individual was from the Riserva Naturale Speciale e Zona di Salvaguardia della Palude di San Genuario in Piemonte, a small wetland surrounded by cultivated fields 4 km from the Po River, one from the Parco Regionale della Maremma in Tuscany, a typical coastal environment with a series of artificial and natural canals, and one from the Riserva Naturale Statale Tenuta di Castelporziano in Latium, an estate founded as a hunting and agricultural reserve, a relic of ancient wetlands and marshes that stretched south to the Pontine plain and north to the Maremma. An additional individual originated from the Riserva Naturale Le Cesine in Apulia, a remnant of an ancient swamp that until 1800 extended from Otranto to Brindisi, and one from Parco Nazionale del Gargano in Apulia, the most important coastal wetland in southern Italy. Furthermore, the three individuals marked by code "Z", that were reported likely to be born from the same nest, were also analyzed. They were assigned a 98% probability to the population of Circeo in Latium, a promontory that was part of the ancient Selva of Terracina.

STRUCTURE software, run with K=3, assigned all every individuals to the Adriatic cluster except for Z1, Z2, Z3 and EREC1 assigned to the Tyrrhenian cluster, supporting the first test, and EHY006 to the Sicilian cluster. The last two samples were not assigned using the previous techinque. We repeated the analysis with the fully Bayesian method to assign samples in specific populations (Tab. 2a-2d, 3 and 4). As expected all individuals were assigned. In general the results from the two tests are comparable. The thing that stands out is that the majority of individuals that weren't assigned in the first test with code "ERE" were found to belong with high probability values to Bosco della Mesola populations. As regards to the samples of alleged siblings, two out of three were assigned to the population of Circeo. Possible origin sites not previously mentioned were found to be the Riserva Naturale della Foce dell'Isonzo in the far eastern part of the Po valley, the artificial lake of Serranella in Abruzzo, a population from an unprotected area very close to Cremona city where animal removal could be very easy, the population of River Neto near to Crotone and the Monte Lerno (Pattada) in Sardinia.

DISCUSSION

In ecology and forensics, molecular markers provide powerful approaches for relocating individuals of unknown origin and geolocating wildlife and their products exposed to poaching. This also contemplates the creation of large genetic databanks especially for threatened species.

Table 1a. GENECLASS probabilities of 79 European pond turtles of unknown origin belonging to various Italian populations studied in chapter 2. The Table indicates only populations that show significant values. Probability higher than the defined threshold of 0.01 is in bold, and the highest probability of the specimen is underlined. The last column shows the number of loci used for analysis. EPI, San Genuario; ECR, Cremona; EBN, Bosco Nordio; EBM, Bosco della Mesola; EFR, Isonzo; EMA, Maremma; ECP, Castelporziano; ELPC, Circeo; EAP, Ascoli Piceno; EAB, Serranella lake; EPE, Persano; EGG, Gargano; ELE, Le Cesine; EPO, Policoro.

	Population														
Individual	EPI	ECR	EBN	EBM	EFR	EMA	ECP	ELPC	EAP	EAB	EPE	EGG	ELE	EPO	Nb. of loci
EHY001	**0.77**	0.00	**0.37**	**0.75**	**0.21**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	8
EHY002	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY003	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY004	0.00	0.00	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	7
EHY005	0.00	0.00	0.03	**0.47**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY006	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY007	**0.08**	**0.73**	**0.63**	**0.81**	**0.39**	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.00	8
EHY008	0.00	0.00	0.00	**0.06**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY009	**0.39**	0.04	**0.36**	**0.82**	**0.06**	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.02	0.00	8
EHY010	**0.06**	**0.11**	**0.53**	**0.84**	**0.23**	0.00	0.00	0.00	**0.27**	0.03	0.00	0.00	**0.24**	0.00	7
EHY011	0.01	0.01	0.04	**0.37**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY012	**0.07**	**0.50**	**0.86**	**0.99**	**0.77**	0.00	0.00	0.00	0.04	**0.07**	0.00	0.00	**0.31**	0.00	8
EHY013	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	8

Table 1b. Refer to Table 1a.

	Population														
Individual	EPI	ECR	EBN	EBM	EFR	EMA	ECP	ELPC	EAP	EAB	EPE	EGG	ELE	EPO	Nb. of loci
EHY014	0.00	**0.13**	0.01	**0.14**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7
EHY015	0.00	0.00	**0.50**	**0.26**	0.03	0.00	0.00	0.00	0.00	**0.06**	0.00	0.03	**0.09**	0.00	7
EHY016	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY017	**0.14**	0.01	**0.24**	**0.59**	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY018	0.00	0.00	**0.07**	**0.41**	0.00	0.00	0.00	0.00	0.00	**0.20**	0.00	**0.05**	**0.07**	0.00	8
EHY019	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	**0.45**	**0.05**	**0.53**	**0.21**	**0.25**	8
EHY020	0.00	0.00	**0.12**	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	6
EHY021	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY022	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY023	0.02	0.04	**0.74**	**0.26**	**0.11**	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	8
EHY024	0.01	0.00	0.02	**0.12**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY025	0.00	0.00	**0.06**	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.20**	0.00	8
EHY026	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY027	**0.20**	0.02	**0.34**	**0.63**	0.01	0.00	0.00	0.00	**0.31**	0.00	0.00	0.00	0.00	0.00	5
EHY028	0.00	**0.43**	**0.79**	**0.31**	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.13**	0.00	7
EHY029	0.00	**0.00**	**0.34**	**0.35**	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**	0.00	7
EHY030	0.42	**0.61**	**0.67**	**0.82**	0.41	0.00	0.00	0.00	0.71	0.00	0.04	0.14	**0.00**	0.00	6

Table 1c. Refer to Table 1a.

	Population														
Individual	EPI	ECR	EBN	EBM	EFR	EMA	ECP	ELPC	EAP	EAB	EPE	EGG	ELE	EPO	Nb. of loci
EHY031	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY032	**0.43**	0.00	**0.12**	**0.83**	**0.75**	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	8
EHY033	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EHY034	0.03	0.02	**0.77**	**0.38**	**0.20**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	8
ERE064	0.00	0.00	**0.12**	**0.32**	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.05	0.00	8
ERE065	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE066	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE001	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE002	0.00	0.00	0.01	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE003	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE004	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE005	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE006	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE007	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE008	0.00	0.00	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE009	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE010	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8

Table 1d. Refer to Table 1a.

	Population														
Individual	EPI	ECR	EBN	EBM	EFR	EMA	ECP	ELPC	EAP	EAB	EPE	EGG	ELE	EPO	Nb. of loci
ERE011	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE012	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE013	0.00	0.01	**0.24**	**0.51**	0.00	0.03	0.01	0.00	**0.06**	0.04	0.00	0.00	**0.50**	0.00	6
ERE014	0.00	0.00	0.00	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE015	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE021	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE022	0.00	0.00	0.00	**0.06**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE023	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE040	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE050	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE051	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE052	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE053	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE054	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE055	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE056	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE057	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8

Table 1e. Refer to Table 1a.

	Population														
Individual	EPI	ECR	EBN	EBM	EFR	EMA	ECP	ELPC	EAP	EAB	EPE	EGG	ELE	EPO	Nb. of loci
ERE058	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE059	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE060	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE061	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE062	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE063	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE078	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
ERE400	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
EREC1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
E101	0.00	0.02	**0.22**	**0.23**	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	8
E102	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
E103	0.00	0.02	**0.10**	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8
Z1	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	8
Z2	0.00	0.00	0.00	0.00	0.00	**0.22**	**0.51**	**0.23**	0.00	0.00	0.00	0.00	0.00	**0.15**	8
Z3	0.00	0.00	0.00	0.00	0.00	**0.26**	**0.13**	0.01	0.00	0.00	0.01	0.00	0.00	0.00	8

Table 2a. STRUCTURE assignment test of *E. orbicularis* individuals of unknown origin in Adriatic cluster populations. In bold, the most likely population of origin for each individual is highlighted. The proportion of missing data indicates cases where one or more loci failed to amplify. EPI, San Genuario; ECR, Cremona; EBN, Bosco Nordio; EBM, Bosco della Mesola; EFR, Isonzo; EAP, Ascoli Piceno; EAB, Serranella Lake; EGG, Gargano; ELE, Le Cesine.

Sample	% Missing data	EPI	ECR	EBN	EBM	EFR	EAP	EAB	EGG	ELE
EHY001	0	**0.88**	0.01	0.02	0.02	0.04	0.01	0.01	0.01	0.01
EHY002	0	0.03	0.02	**0.73**	0.05	0.11	0.03	0.02	0.01	0.01
EHY003	0	0.02	0.01	**0.89**	0.01	0.02	0.01	0.01	0.01	0.02
EHY004	-12	0.03	0.03	**0.68**	0.04	0.03	0.09	0.02	0.06	0.03
EHY005	0	0.09	0.27	0.12	**0.36**	0.09	0.03	0.01	0.01	0.01
EHY007	0	0.04	**0.79**	0.04	0.03	0.07	0.01	0.01	0.01	0.01
EHY008	0	0.06	**0.30**	0.21	0.12	0.27	0.01	0.01	0.01	0.01
EHY009	0	**0.62**	0.06	0.05	0.16	0.06	0.03	0.01	0.01	0.01
EHY010	-12	0.13	0.15	0.09	**0.27**	0.09	0.23	0.02	0.03	0.01
EHY011	0	0.12	**0.43**	0.07	0.19	0.11	0.05	0.02	0.01	0.01
EHY012	0	0.06	0.13	0.11	**0.34**	0.30	0.03	0.01	0.01	0.02
EHY013	0	0.08	0.01	**0.75**	0.02	0.03	0.01	0.06	0.03	0.03
EHY014	-12	0.02	**0.84**	0.02	0.07	0.01	0.01	0.01	0.01	0.01
EHY015	-12	0.06	0.02	**0.32**	0.09	0.18	0.02	0.06	0.23	0.02
EHY016	0	0.01	0.06	0.18	0.04	0.20	0.12	0.06	**0.21**	0.11
EHY017	0	**0.61**	0.17	0.06	0.07	0.04	0.01	0.01	0.01	0.01
EHY018	0	0.07	0.04	0.07	0.13	0.04	0.01	0.08	**0.53**	0.02
EHY019	0	0.06	0.02	0.02	0.03	0.02	0.04	0.18	**0.62**	0.02

Table 2b. Refer to Table 2a.

Sample	% Missing data	EPI	ECR	EBN	EBM	EFR	EAP	EAB	EGG	ELE
EHY020	-25	0.01	0.03	0.18	0.02	0.35	**0.36**	0.02	0.01	0.02
EHY021	0	0.02	0.01	**0.88**	0.01	0.02	0.01	0.01	0.02	0.02
EHY022	0	**0.45**	0.07	0.40	0.02	0.03	0.01	0.01	0.01	0.01
EHY023	0	0.14	0.17	**0.52**	0.05	0.06	0.01	0.03	0.01	0.01
EHY024	0	**0.63**	0.09	0.08	0.09	0.02	0.04	0.02	0.02	0.02
EHY025	0	0.06	0.03	**0.53**	0.12	0.02	0.01	0.01	0.05	0.18
EHY026	0	0.05	0.01	0.11	**0.40**	0.02	0.03	0.04	0.29	0.06
EHY027	-37	**0.32**	0.10	0.07	0.25	0.06	0.17	0.01	0.01	0.02
EHY028	-12	0.02	**0.68**	0.17	0.06	0.03	0.01	0.01	0.01	0.01
EHY029	-12	0.02	0.02	0.05	0.04	**0.82**	0.02	0.01	0.01	0.01
EHY030	-25	0.15	**0.40**	0.07	0.09	0.06	0.20	0.01	0.02	0.01
EHY031	0	0.03	0.02	**0.89**	0.03	0.01	0.01	0.01	0.01	0.01
EHY032	0	0.24	0.01	0.02	0.04	**0.66**	0.01	0.01	0.01	0.01
EHY033	0	0.01	0.01	**0.65**	0.13	0.03	0.01	0.01	0.01	0.13
EHY034	0	0.12	0.10	**0.53**	0.08	0.11	0.02	0.03	0.01	0.01
ERE064	0	0.05	0.03	**0.33**	0.18	0.14	0.02	0.03	0.20	0.02
ERE065	0	**0.67**	0.08	0.14	0.06	0.01	0.02	0.01	0.01	0.01
ERE066	0	0.12	0.06	**0.68**	0.06	0.05	0.01	0.01	0.01	0.01
ERE001	0	0.03	0.01	0.02	**0.89**	0.02	0.01	0.01	0.01	0.01
ERE002	0	0.02	0.03	0.03	**0.87**	0.02	0.02	0.01	0.01	0.01
ERE003	0	0.01	0.02	0.01	**0.90**	0.03	0.01	0.01	0.01	0.01
ERE004	0	0.02	0.02	0.01	**0.92**	0.01	0.01	0.01	0.01	0.01

Table 2c. Refer to Table 2a.

Sample	% Missing data	EPI	ECR	EBN	EBM	EFR	EAP	EAB	EGG	ELE
ERE005	0	0.03	0.02	0.01	**0.90**	0.01	0.01	0.01	0.01	0.01
ERE006	0	0.01	0.02	0.01	**0.91**	0.02	0.01	0.01	0.01	0.01
ERE007	0	0.03	0.01	0.01	**0.89**	0.02	0.01	0.01	0.01	0.01
ERE008	0	0.03	0.02	0.01	**0.91**	0.01	0.01	0.01	0.01	0.01
ERE009	0	0.02	0.02	0.01	**0.91**	0.02	0.01	0.01	0.01	0.01
ERE010	0	0.01	0.02	0.01	0.93	0.01	0.01	0.01	0.01	0.01
ERE011	0	0.02	0.01	0.47	0.42	0.01	0.01	0.01	0.02	0.04
ERE012	0	0.01	0.01	0.01	**0.91**	0.01	0.01	0.01	0.01	0.01
ERE013	-25	0.04	0.22	0.12	**0.25**	0.04	0.09	0.07	0.13	0.04
ERE014	0	0.03	0.01	0.01	**0.89**	0.02	0.02	0.01	0.02	0.01
ERE015	0	0.07	0.04	0.30	**0.41**	0.03	0.01	0.03	0.11	0.02
ERE021	0	0.03	0.15	0.01	**0.78**	0.01	0.01	0.01	0.01	0.01
ERE022	0	0.02	0.01	0.01	**0.88**	0.03	0.01	0.01	0.01	0.01
ERE023	0	0.03	0.05	0.05	**0.82**	0.01	0.01	0.02	0.01	0.01
ERE040	0	0.03	0.01	0.01	**0.91**	0.01	0.01	0.01	0.01	0.01
ERE050	0	0.01	0.04	0.18	**0.60**	0.04	0.03	0.06	0.02	0.02
ERE051	0	0.01	0.30	0.01	**0.57**	0.01	0.01	0.08	0.02	0.01
ERE052	0	0.02	0.01	0.02	**0.89**	0.02	0.01	0.01	0.01	0.01
ERE053	0	0.01	0.01	0.07	**0.78**	0.08	0.03	0.01	0.01	0.01
ERE054	0	0.01	0.02	0.02	**0.86**	0.04	0.02	0.01	0.01	0.01
ERE055	0	0.01	0.02	0.07	**0.59**	0.02	0.01	0.10	0.14	0.04
ERE056	0	0.03	0.01	0.02	0.41	0.01	0.04	0.01	**0.43**	0.03

Table 2d. Refer to Table 2a.

Sample	% Missing data	EPI	ECR	EBN	EBM	EFR	EAP	EAB	EGG	ELE
ERE057	0	0.02	**0.42**	0.05	0.07	0.11	0.03	0.02	0.28	0.01
ERE058	0	0.01	0.03	0.07	**0.80**	0.01	0.03	0.01	0.01	0.03
ERE059	0	0.03	0.02	0.02	**0.88**	0.01	0.02	0.01	0.01	0.01
ERE060	0	0.01	0.01	0.05	0.07	0.01	0.01	0.01	**0.68**	0.16
ERE061	0	0.02	0.01	0.03	**0.87**	0.04	0.01	0.01	0.02	0.01
ERE062	0	0.02	0.01	0.05	**0.75**	0.02	0.01	0.04	0.04	0.06
ERE063	0	0.01	0.01	0.01	**0.91**	0.01	0.03	0.01	0.01	0.01
ERE078	0	0.02	0.03	0.02	**0.67**	0.04	0.10	0.04	0.08	0.01
ERE400	0	0.01	0.17	0.03	**0.72**	0.02	0.01	0.02	0.01	0.01
101	0	0.03	0.19	0.16	0.10	**0.30**	0.02	0.01	0.18	0.02
102	0	0.13	0.02	0.23	0.18	0.02	0.02	0.03	**0.36**	0.01
103	0	0.05	**0.55**	0.22	0.03	0.07	0.02	0.01	0.01	0.06

Table 3. STRUCTURE assignment test of *E. orbicularis* individuals of unknown origin in Tyrrhenian cluster populations. In bold, the most likely population of origin for each individual is highlighted. The proportion of missing data indicates cases where one or more loci failed to amplify. EAL, Albenga; ESR, San Rossore; EMA, Maremma; EMR, Monte Rufeno; ECP, Castelporziano; ELPC, Circeo; EPE, Persano; EPO, Policoro;ELT, Lamezia Terme; EKR, Crotone; EBE, Berchidda; EPA; Pattada.

Sample	% Missing data	EAL	ESR	EMA	EMR	ECP	ELPC	EPE	EPO	ELT	EKR	EBE	EPA
EREC1	0	0.02	0.02	0.05	0.02	0.02	0.03	0.03	0.02	0.11	**0.63**	0.03	0.03
Z1	0	0.05	0.12	0.05	0.06	0.05	**0.53**	0.02	0.02	0.02	0.02	0.03	0.03
Z2	0	0.03	0.05	0.09	0.16	0.08	**0.30**	0.03	0.06	0.05	0.04	0.07	0.05
Z3	0	0.08	0.07	0.13	0.12	0.08	0.10	0.03	0.04	0.04	0.04	0.07	**0.20**

Table 4. STRUCTURE assignment test of *E. orbicularis* individuals of unknown origin in Sicilian cluster populations. In bold, the most likely population of origin for each individual is highlighted. The proportion of missing data indicates cases where one or more loci failed to amplify. ELS, Lago Spartà; EIM, Valle dell'Imera; ECO, Bosco della Ficuzza, EGA, Gallitello; EMZ, Lago Preola; ESI, Siculiana; ENO, Noto; EVE, Vendicari.

Sample	% Missing data	ELS	EIM	ECO	EGA	EMZ	ESI	ENO	EVE
EHY006	0	0.10	0.16	0.07	0.14	0.14	0.10	0.11	**0.18**

The currently available molecular markers and statistical assignment, exclusion and fully Bayesian methods, can allow the identification of the population of origin of individuals (or tissues) with high certainty (Manel *et al.* 2002). Assignment methods estimate the likelihood that an individual originates from a certain population. This technique should work reliably at least in cases with few candidate populations that are genetically differentiated, and that have no immigration or cryptic substructure. Individuals from recovery centers had, in some cases, lower assignation probability, clearly because sampling all possible populations is not always feasible.

This study employed two of the most reliable statistical tests, that also have certain limits. The exclusion method is less powerful than the fully Bayesian method that, nevertheless, requires the sampling of the true population of origin. The two tests gave very similar results. Of thirty turtles assigned with the exclusion test, likely deriving from the Adriatic cluster, 43.4% were assigned to the same population and 40% to populations still considered reliable using the fully Bayesian test. Most of the ERE samples unfortunately had no success with the exclusion test; of thirty-five individuals only four were assigned with the same results of the Pritchard *et al.* (2000b) method. It was also seen that the majority of turtles were found to belong to neighboring populations of the collection areas, northern and central Italy, and that only 10% originated in the South. As regards the three juveniles declared probable siblings, the most likely source population is the Circeo reserve. For animals with conflicting data, about 6%, as well as those assigned with low probability only via the STRUCTURE software, and the only one turtle belonging to the Sicilian cluster it will be necessary to conduct further analyses based on both mitochondrial DNA control region and microsatellites by increasing the number of sites sampled in order to better determine the source population. To conclude, the remaining turtles can be released in the wild, minimizing their genetic differences with the local animals.

REFERENCES

Ciofi C, Tzika AC, Natali C, et al. (2009) Characterization of microsatellite loci in the European pond turtle Emys orbicularis. Molecular Ecology Resources ***9****, 189-191.*

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics ***153****, 1989-2000.*

Fritz U (1998) Introduction to zoogeography and subspecific differentiation in Emys orbicularis (Linnaeus, 1758). In: eds Fritz U, Joger U, Podloucky R, Servan J) Proceedings of the Emys Symposium Dresden 96. Mertensiella 10, pp. 1-27. Warlich, Rheinbach.; IUCN Red List of threatened species; Fritz 2001.

Herborg L-M, Weetman D, Van Oosterhout C, Hanfling B (2007) Genetic population structure and contemporary dispersal patterns of a recent European invader, the Chinese mitten crab, Eriocheir sinensis. Molecular Ecology ***16****, 231-242.*

Manel S, Berthier P, Luikart G (2002) Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conservation Biology ***16****, 650-659.*

Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. Proceedings of the National Academy of Sciences of the United States of America ***98****, 5446-5451.*

Padilla DK, Williams SL (2004) Beyond ballast water: aquarium and ornamental trades as sources of invasive species in aquatic ecosystems. Frontiers in Ecology and the Environment ***2****, 131-138.*

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite Analysis of Population-Structure in Canadian Polar Bears. Molecular Ecology ***4****, 347-354.*

Piry S, Alapetite A, Cornuet JM, et al. (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. Journal of Heredity ***95****, 536-539.*

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics ***155****, 945-959.*

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences of the United States of America ***94****, 9197-9201.*

Robert J (2000) Dossier trafic animal: la Mafia a l'assault de la nature. Terre Sauvage 155:35–50.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press.

TeleckyT (2001) United States Import and Export of Live Turtles and Tortoises. Turtle and Tortoise Newsletter 4:8–13.

3. PATTERNS OF GENE FLOW IN LINEAR HABITATS: SCANT EVIDENCE OF GENETIC STRUCTURE IN A RIPARIAN POPULATION OF THE EUROPEAN POND TURTLE *Emys orbicularis* FROM NORTHERN GREECE

ABSTRACT

Patterns of dispersal and colonization in freshwater turtles can be affected by geographical barriers shaping phylogeographic histories of natural populations. Although variation can be structured according to distance or hierarchical system of rivers, extensive movements may also occur within drainages, advocating the importance of waterways for turtles to move across habitats. The European pond turtle, *Emys orbicularis*, is a non-migratory terrapin found in wetlands, ponds and rivers. In northern Greece, a sizeable population of *E. orbicularis* is found in seasonal and permanent ponds along the Strymon river and occasionally in the river bed. This populations are significantly affected by habitat encroachment. In this study, we assessed fine-scale population genetic structure and movements of turtles among ponds to provide additional information on patterns of dispersal along linear habitats by freshwater turtles. A relatively high genetic diversity was recovered across the five ponds anlaysed in this study. The highest and lowest number of mitochondrial DNA haplotypes and polymorphic sites were recorded in downstream and upstream ponds, respectively. A source-sink patterns of gene flow was supported by microsatellite Bayesian analysis whereby the relatively large proportion of migrants received from downstream ponds by the other ponds suggested recent upstream movements of *E. orbicularis*. Pond turtles showed low levels of genetic divergence, no significant differences between ponds for males, and a significant multilocus genotype divergence between ponds when females only were considered in the analysis, suggesting female phylopatry. For *E.orbicularis* in the Kerkini area water permanance is of paramount importance. The relatively high connections among ponds advocate the need for their conservation through proper management and minimized used for livestock.

INTRODUCTION

Landscape and environmental features can significantly affect patterns of genetic variability in natural populations, and a diversity of genetic frameworks have described the influence of landscape composition, configuration and matrix type on gene flow and spatial genetic diversity (Manel *et al.* 2003; Storfer *et al.* 2007). Terrestrial and freshwater habitats in particular present a variety of environmental features and anthropogenic discontinuities which affect gene flow and population structure at different spatial and temporal scales, with important ecological, evolutionary and conservation implications

(Segelbacher *et al.* 2010; Storfer *et al.* 2010). Genetic divergence correlates to geographical distance under an isolation by distance process (Hardy & Vekemans 1999). However, more complex patterns of genetic diversity that are generally a result of topographic and ecological barriers to gene flow may also occur in continuously distributed populations (Earnest *et al.* 2014; Robinson *et al.* 2012).

In many riverine or riparian species, stream networks can result in various degrees of patterns of isolation by distance or hierachical population structure due to seasonal barriers, natural falls, presence of tributaries, impoundments, weirs, discontinuities in the river profile as well as past climates and geological history (Fluker *et al.* 2014; Hirschmann *et al.* 2015; Meldgaard *et al.* 2003; Osborne *et al.* 2014). Stream populations may be spatially structured just over a few kilometers, because of limited dispersal, position along different river drainages, or even highly skewed reproductive success by a small number of breeders (Hughes *et al.* 2009; Kanno *et al.* 2011; Neville *et al.* 2006; Ostergren & Nilsson 2012).

In obligatory aquatic species, river distances calculated by following river courses, rather than geographic distances may better reflect the actual distances to be covered if individuals were to move among populations (e.g. Fetzner & Crandall 2003b). On the other hand, for aquatic species capable of terrestrial dispersal the effect of drainage and network arrangement on population genetic structure can be negligible, and movements may instead depend on a combination of within-network stream and euclidean distances (Hughes *et al.* 2009; Hurwood & Hughes 2001).

Patterns of dispersal and colonization of freshwater turtles in particular, can be affected by main geographical barriers shaping phylogeographic histories of natural populations (e.g. Fritz *et al.* 2006a; Lenk *et al.* 1999b) and by landscape hydrological patterns. Genetic variation often appears to be structured according to the hierarchical system of rivers, whereby little or no gene flow is recorded among turtles from different streams, while extensive movements occur within drainages, advocating the importance of waterways for turtles to move across habitats (Chelazzi *et al.* 2007; Souza *et al.* 2002) (but see Gonzalez-Porter *et al.* 2011).

The European pond turtle, *Emys orbicularis*, is a non-migratory terrapin found in wetlands, ponds and rivers from northwestern Africa through most of southern, central, and eastern Europe to Asia Minor and the Caspian and Aral Seas (Fritz 2003). Natural populations rely heavily on water for feeding and mating and may show strict site fidelity to specific drainages with enhanced homing ability (Lebboroni & Chelazzi 1991, 2000). Female *E. orbicularis* equally depends on loose, pity soil close to aquatic habitats for oviposition, and perform migrations to nesting sites in temperate forest and shrubs even across unsuitable landscape (Rovero & Chelazzi 1996).

In northern Greece, a sizeable population of *E. orbicularis* is found along the Strymon river and in the nature reserve of Lake Kerkini, an irrigation

reservoire created on former wetlands approximately 60 km north of Thessaloniki. The lake is a Natura2000 site and a wetland of international importance according to the RAMSAR convention (Matthews 1993). Turtles are mainly found in seasonal and permanent ponds along the Strymon and occasionally in the river bed, and are significantly affected by habitat encroachment (Crivelli *et al.* 1995; Pyrovetsi & Papastergiadou 1992). Assessment of fine-scale population genetic structure and movements of turtles among ponds would provide additional information on patterns of dispersal along linear habitats by freshwater turtles. The Strymon itself is known to serve as waterway to movements of syntopic turtle species (Chelazzi *et al.* 2007). In this study, we compared mitochondrial DNA control region sequences and allelic variation at species-specific nuclear DNA microsatellite loci to assess levels of genetic diversity in a substantial number of *E. orbicularis* sampled across the Strymon river, and define patterns of gene flow among ponds. Genetic analysis of population structure will also help devising management plans for natural populations of *E. orbicularis* in such an important buffer zone to the Kerkini national nature reserve. By investigating degrees of connectedness among ponds may in turn stress the importace of single reservoires and foster mitigation of anthopogenic impact on microhabitats which are vital to the viability of near threatened species.

MATERIALS AND METHODS

Study area

The study was conducted along the Strymon River, upstream of Lake Kerkini (Figure 1). The Strymon enters Greek territory from Bulgaria, and it is heavily impacted by human activities from the Sidirokastro bridge, near the Bulgarian border, to the Kerkini nature reserve. Embankments delimit the river bed, riparian forest is limited, river banks are used for livestock grazing and water is often pumped from the river to irrigate farmland. Ponds along the north bank of the Strymon were formed by excavation works during dike construction in the early 1980s and were negatively affected by river dredging in 1986 and 1999 (Crivelli *et al.* 1995, AJ Crivelli unpublished data). They are fed by rainfall and groundwater, which in turns depends on river regime. Ponds are used as a source of drinking water for livestock, they are rich in macroinverterbates, amphibians and grass snakes and represent important habitat for *E. orbicularis*, particularly in summer when water is diverted from the Strymon for farming (Pyrovetsi & Papastergiadou 1992). A total of 314 turtles were sampled in five ponds ranging from 0.03 to 3.2 ha in size and less then 2 m deep (Table 1). Ten pond turtles were also captured in the Strymon river (Figure 1).

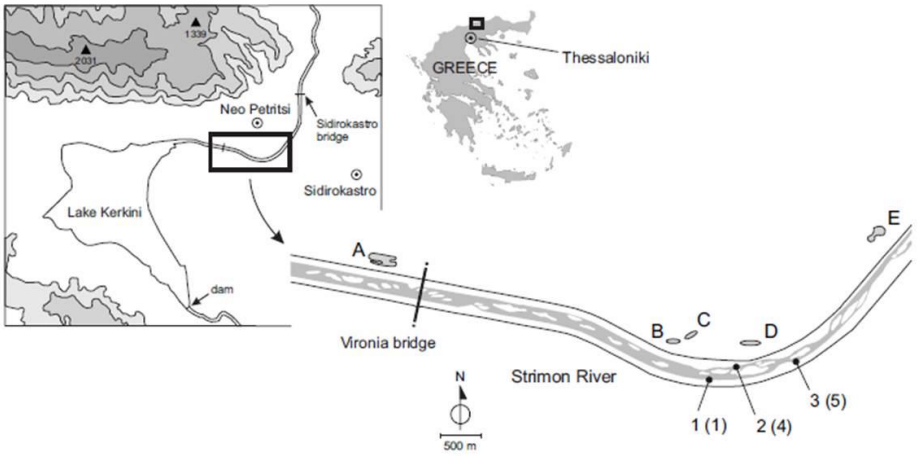

Figure 1. Map of the study area showing the Strymon River close to the northern section of Lake Kerkini. Topography is reported using different shades of grey. Triangles are elevation peaks and bullets are towns. The inset map to the right shows the Strymon (grey area) and locations of ponds A to E. Solid lines along the river represent white roads on embankments. White patches in the river course are shoals and sandbanks. Additional trapping sites along the Strymon are numbered from 1 to 3 and report the number of turtles sampled per site.

Trapping and sample collection

Pond turtles were captured using mesh fyke nets consisting of two cone-shaped netting bags linked together and held open by four plastic hoops of 50 cm diameter. The net is 150 cm long and acts as funnel to trap swimming turtles. A vertical leader of 50x200 cm (HxL) extends from the larger mouth of the net and connects to the mouth of a second, specular net to form an array. The array was anchored in place and kept tight by closing the smaller chambers at both ends, which were then fastened each to a fence post using small diameter ropes. Turtles swimming across the pond encountered the net leader and were funneled into one of the fyke nets in the array with no need of baiting the nets. We used a total of 10 custom-made fyke net arrays. Pond turtles were marked using a numerical coding system similar to the method described by Ernst et al. (1974) whereby a number is assigned to each marginal scutes of the shell. A notch is made in certain scutes using a small iron saw so that turtles can be sequentially numbered by adding the numerical values of the notched marginals. Sex was determined based on tail and plastron morphology (Rollinat 1934; Zuffi & Gariboldi 1995). Bood samples were collected from the subcarapacial sinus and stored in a lysis buffer containing 0.1 M Tris buffer, 0.1 M EDTA, 0.2 M NaCl and 1% SDS, pH 8.0. High molecular weight DNA was

isolated by phenol-chloroform extraction (Sambrook & Russell 2001) using Phase Lock Gel (Eppendorf) to improve separation between organic and acqueous phases.

Genetic analysis

The mitochondrial DNA (mtDNA) control region was amplified by polymerase chain reaction (PCR) using primers DES1 and DES2 designed for *Chrysemis picta* by Starkey et al. (2003). PCR amplification was performed in 10 µl total volume with 50 ηg of DNA, 1 x PCR buffer, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.5 µM of each primer and 0.5 units of *Taq* DNA polymerase (Invitrogen). Thermal profiles consisted of an initial denaturation step of 5 min at 95 °C, followed by 35 cycles of 30 sec at 95 °C, 30 sec at 52 °C and 1 min at 72 °C, with a final extension step of 10 min at 72 °C. Amplicons were cycle-sequenced using BigDye Terminator v3.1 chemistry (Applied Biosystems) according to the manufacturer's protocol, and sequencing products were resolved by capillary electrophoresis in an Applied Biosystems 3130xl genetic analyzer. Sequences consisted of a total of 659 nucleotides for a total of 202 individuals and were deposited in Genebank under accession numbers KT 868933–868942.

Eight species-specific polymorphic microsatellite loci were amplified using PCR conditions and thermal profiles described in Ciofi et al. (2009a). PCR products of 314 turtles were resolved using an Applied Biosystems 3130xl genetic analyser and allele sizes assessed using GENEMAPPER 5.0 (Applied Biosystems). Errors due to large allele dropout or stutter bands and evidence for the presence of null alleles at each locus were evaluated using MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004).

Analysis of mitochondrial DNA diversity

Mitochondrial control region sequence diversity was inferred by the number of haplotypes (*k*) and polymorphic sites, average number of pairwise differences among haplotypes, haplotype diversity (*h*) and nucleotide diversity (π) using ARLEQUIN 3.5 (Excoffier *et al.* 2005). We also calculated the sequence diversity estimator $\Theta = 2N_f\mu$ (female effective population size times mutation rate per site) derived from Ewens (1972) formula. The extent of mtDNA control region differentiation was assessed from the average number of pairwise differences in haplotype frequencies among sampling sites using the Φ-statistics of Excoffier *et al.* (1992) implemented in ARLEQUIN 3.5. Significance of Φ values was obtained after 10,000 haplotype permutations.

Analysis of microsatellite variation

We assessed allelic diversity, observed and unbiased expected heterozygosity using GENALEX 6.5 (Peakall & Smouse 2012). Departure from Hardy-Weinberg (HW) equilibrium and genotypic linkage disequilibrium were tested in GENEPOP 4.2 (Rousset 2008a). Number of turtles varied significantly across ponds. Because large samples are expected to have more alleles than small sample sets, we estimated allelic richness using the rarefaction technique implemented in HP-RARE 1.1 (Kalinowski 2005) to compensate for differences in sample size. Allelic diversity is also particularly biased by differences in sample sizes when highly polymorphic loci are considered (Leberg 2002). We used a standardized sample of size equal to the smallest number of genes sampled per locus and pond (locus Emys2, pond D, g = 46) to estimate the expected number of alleles and private allelic richness averaged over loci (Kalinowski 2004; Petit *et al.* 1998). Deviation from random mating and differences in allele frequencies between ponds were tested by assessing statistical significance of the F_{IS} and F_{ST} estimators f and θ, respectively, after 10,000 permutations of genotypes performed in GENETIX 4.05 (Belkhir *et al.* 2004). Genotypic differentiation was also estimated using the log likelihood ratio (G) test implemented in GENEPOP 4.2. We also performed a Mantel test for matrix correspondence to test for isolation by distance. Geographic distances were calculated using GPS coordinates taken at approximately the centre of each pond. Genetic divergence was computed as F_{ST} / (1 - F_{ST}). The Mantel statistic r_{xy} was compared to a frequency distribution of 9999 r_{xy} obtained by permutations using GENALEX 6.5.

Mitochondrial DNA network

Population level genealogies were inferred by a network of haplotypes constructed using statistical parsimony (Templeton *et al.* 1992b). The method, implemented in TCS 1.21 (Clement *et al.* 2000b) links haplotypes with the smaller number of differences as defined by a 95% confidence criterion, estimate haplotype outgroup probabilities, which correlate with haplotype age, and identify the most probable ancestral haplotype according to coalescence theory (Castelloe & Templeton 1994b).

Population structure and gene flow

Patterns of population structure were assessed by a Bayesian clustering analysis implemented in STRUCTURE 2.3.4 (Pritchard *et al.* 2000b). We used sampling locations (ponds) as prior in order to provide more accurate inference of structure in case of relatively low levels of population divergence (Hubisz *et al.* 2009). After a burn-in period of 10^5 iterations, we estimated the probability of the observed genotypes given a number of populations K ranging from 1 to 7 (the number of sampling sites plus two) by Markov Chain Monte Carlo (MCMC) methods using 10^6 repetitions. We calculated the mean likelihood

over 10 runs for each *K* and estimated the most likely number of clusters as described in Evanno et al. (2005b). The *K* value with the highest posterior probability was then used as prior information to estimate the proportion of membership of each pond genotypes in each of the *K* clusters. Results were graphically visualized using DISTRUCT 1.1 (Rosenberg 2004).

Patterns of population structure were also assessed by comparing, for each pond, within-pond pairwise relatedness with relatedness values estimated by pairwise comparisons to turtles from the other four ponds. We used Monte Carlo simulations implemented in COANCESTRY 1.0.1.2 to find the most appropriate relatedness estimator for our data (Wang 2011). Using known allele frequencies, we generated seven groups of multilocus genotypes consisting of 100 dyads each with simulated relatedness values varying from 0 (unrelated individuals) to 0.5 (parent-offspring or fullsibs). These data were comparatively analysed by two maximum likelihood estimators and five moment estimators (Wang 2011). A matrix of correlation coefficients was then calculated among means and variances of the seven relatedness estimators and the true simulated values. The Triadic likelihood estimator (Wang 2007) had the strongest correlation with the true value (r = 0.81) and was selected for subsequent analyses. Differences in average relatedness among turtles within and among ponds were tested using a bootstrapping procedure whereby observed difference in relatedness between groups was compared to a distribution of 10^5 differences in average relatedness between simulated groups. A similar procedure was used to test differences between groups made with respect to gender and assess whether males were more related to females from the same pond than to females from other ponds.

Evidence of recent migration rates between ponds and 95% confidence intervals were obtained using the Bayesian multilocus genotyping procedure implemented with MCMC methods in BAYESASS 3 (WILSON & RANNALA 2003). The MCMC was run for a total of 3.0 x 10^6 iterations, with the first 10^6 iterations discarded as burn-in to allow the chain to reach stationarity. Samples were collected every 2000 iterations to infer posterior probability distributions of parameters of interest.

A likelihood-based Bayesian genetic assignment technique implemented in GENECLASS 2.0 (Piry *et al.* 2004a) was used to assign or exclude ponds as possible origins of individual turtles sampled in the Strymon River (Figure 1). We used the assignment criterion of Rannala and Mountain (1997b) and assessed significance of likelihood values using the Monte Carlo resampling method of Paetkau et al. (2004) with 10^5 simulations.

RESULTS

Genetic diversity

A total of 10 haplotypes were defined by seven polymorphic sites. Three haplotypes were found only once in three individual turtles, respectively, sampled in separate ponds. The most common sequence was instead shared by 75 individuals from all five ponds (Table 1). The highest and lowest number of haplotypes and polymorphic sites were recorded downstream in pond A and upstream in pond E, respectively. Average number of pairwise nucleotide differences across the Strymon river was 1.28 ± 0.05 and ranged from 0.92±0.09 recorded in pond C to 1.47±0.15 from pond D. Haplotype and nucleotide diversities were similar across the study area, with the lowest *h* value recovered among turtles in pond C (Table 2). Theta values were also very similar across ponds with the lowest number recorded in pond C.

Number of microsatellite alleles per locus varied from 9 to 34. No significant linkage disequilibrium was detected for any of the 28 pairwise locus combinations, and there was no evidence of null alleles or scoring errors due to stuttering or large allele dropout. High allelic diversities were recovered in ponds A and C, characterized by the largest sample sizes. On the other hand, allelic richness, observed and expected heterozygosities were remarkably similar across ponds (Table 2). High values of observed and expected heterozigosity were recorded for all ponds. However, mean observed heterozygosities were all lower than expected values and deviation from HW proportions and significantly high inbreeding coefficients were recorded for all sampling locations (Table 2). Ponds A to D showed 4 to 5 loci deviating from HW equilibrium ($P < 0.05$), while only one locus (Emys 5) in pond E departed from HW expectations.

Mitochondrial DNA genealogical relationships

All five ponds shared haplotypes 1 to 4. Similarly, haplotypes 5 to 7 revealed a pattern of connection among ponds A to D. On the other hand, downstream pond A and ponds B and C close to the river bend all differed by a singletone. Upstream ponds D and E had no unique haplotypes and were characterized by a relatively more uniform haplotype distribution (Table 1).
Haplotypes were all closely related (0.15% to 0.61% of sequence divergence) and connected by 1 to 3 mutational steps from the sequence with the largest root probability (Figure 2). The shallow sequence divergence among the 10 haplotypes suggests short-term evolutionary isolation among ponds, and the scattered spatial pattern of the main four haplotypes (1-4) would imply recurrent interchange between ponds. Statistical parsimony identified mtDNA haplotype 1 as ancestral sequence. This haplotype was the most abundant, it was found in all five ponds but had the highest frequency in pond A, C and D.

Genetic divergence and population structure

Low levels of genetic divergence were recorded by pairwise analysis of F_{ST} values based on microsatellite loci (- 0.001 < θ < 0.009). However, statistical significance was recovered between pond A and the other upstream ponds (B to D), and between pond C and ponds D and E (P < 0.01). When males and females were considered separately, F_{ST} values were all low and not significant in males. On the other hand, a rather consistent pattern of differentiation resulted from pairwise θ values estimated for females only and Φ statistics based on mtDNA (Table 3). The exact G test for genotypic differentiation confirmed a pattern of no significant differences between ponds for males, and a significant multilocus genotype divergence between pond A and ponds B to D and between pond D and ponds B, C and E when females only were considered in the analysis.

Population structure analysis revealed that the most probable number of clusters for interpreting the obeserved genotypes was K = 3 (P (K | X) = 0.999) . The highest modal value of ΔK=19.2 for K=3 corroborated the result so that three main partitions were used as prior population information for calculating the posterior probability of individual assignment. Turtles from pond A and E had a substantial sharing of the same cluster, with an average proportion of individual genome originating from the first cluster q = 0.822 and q = 0.922, respectively (Figure 3). A significantly higher level of admixture was recorded across ponds B to D. High average proportional memberships of turtles from pond B and C (q = 0.281 and q = 0.426, respectively) was found in a second, distinct cluster. Similarly, an average proportion of individual genome q = 0.267 of turtles from pond D clustered in a third partition.

Turtles from ponds A, B and D were more related to each other than to turtles from other ponds (r = 0.085±0.002SE - 0.103±0.003SE vs. r = 0.072±0.005SE - 0.079±0.002SE; P < 0.01. The same pattern was recovered among females, while males from pond C only had a significantly higher within- than among-pond average relatedness value (r = 0.101±0.007SE vs. r = 0.076±0.005SE; P < 0.01). None of the five ponds had males showing average relatedness with females of the same ponds higher than average relatedness with females from other ponds (P > 0.05).

Pond turtles sampled at sites 1 to 3 along the Strymon River had no common origin and did not belong to ponds geographically close to the sampling sites. The sole individual caught at site 1 had a probability of 0.768 to be assigned to pond D. Turtles sampled at site 2 had either a low probability of assignment or they were assigned with probabilities between 0.800 and 0.998 to pond A, C and D. Two turtles collected at site 3 showed low probabilities of assignment, two others were allocated with similar probability (>0.775) to pond A and D, while a fifth individual had probabilities ranging from 0.706 to 0.966 to be assigned to any of the five ponds.

Gene flow

Table 1. Haplotype frequency distribution for each pond and total haplotype frequency across ponds for *Emys orbicularis* sampled along the Strymon River. N, sample size.

			Haplotype									
Pond	Pond area (m^2)	*N*	1	2	3	4	5	6	7	8	9	10
A	32,000	64	0.42	0.19	0.2	0.11	0.03	0.02	0.02	0.02		
B	300	23	0.17	0.48	0.09	0.13		0.04	0.04		0.04	
C	400	55	0.42	0.42	0.05	0.07		0.02				0.02
D	1,300	35	0.4	0.17	0.23	0.09	0.11					
E	16,000	25	0.28	0.36	0.24	0.12						
Total		202	0.371	0.302	0.158	0.099	0.03	0.015	0.01	0.005	0.005	0.005

Table 2. Genetic diversity measures in *Emys orbicularis* from five ponds along the Strymon River. *N*, sample size; P, number of polymorphic sites; *k*, number of haplotypes; h, haplotype diversity; π, nucleotide diversity; Θ, number of expected segregating polymorphic sites, *A*, mean number of alleles per locus; A_R, allelic richness; H_O, mean observed heterozygosity; H_E, mean expected heterozygosity; *f*, estimator of F_{IS} index (Weir & Cockerham 1984). Means ± standard error values.

	Mitochondrial DNA						Microsatellites					
Pond	*N*	P	*k*	h	π	Θ	*N*	*A*	A_R	H_O	H_E	*f*
A	64	6	8	0.74±0.00	0.002	2.24±0.05	85	19.1±2.4	12.9	0.76±0.06	0.86±0.01	0.11
B	23	5	7	0.74±0.02	0.002	2.23±0.21	30	14.8±1.6	13.6	0.79±0.06	0.84±0.03	0.058
C	55	4	6	0.65±0.00	0.001	1.42±0.03	100	19.7±2.4	13.2	0.75±0.06	0.86±0.02	0.128
D	35	4	5	0.76±0.00	0.002	2.46±0.11	71	17.1±1.6	13.1	0.71±0.06	0.84±0.03	0.151
E	25	3	4	0.75±0.00	0.002	2.32±0.10	28	13.6±1.1	13.1	0.78±0.06	0.86±0.02	0.096

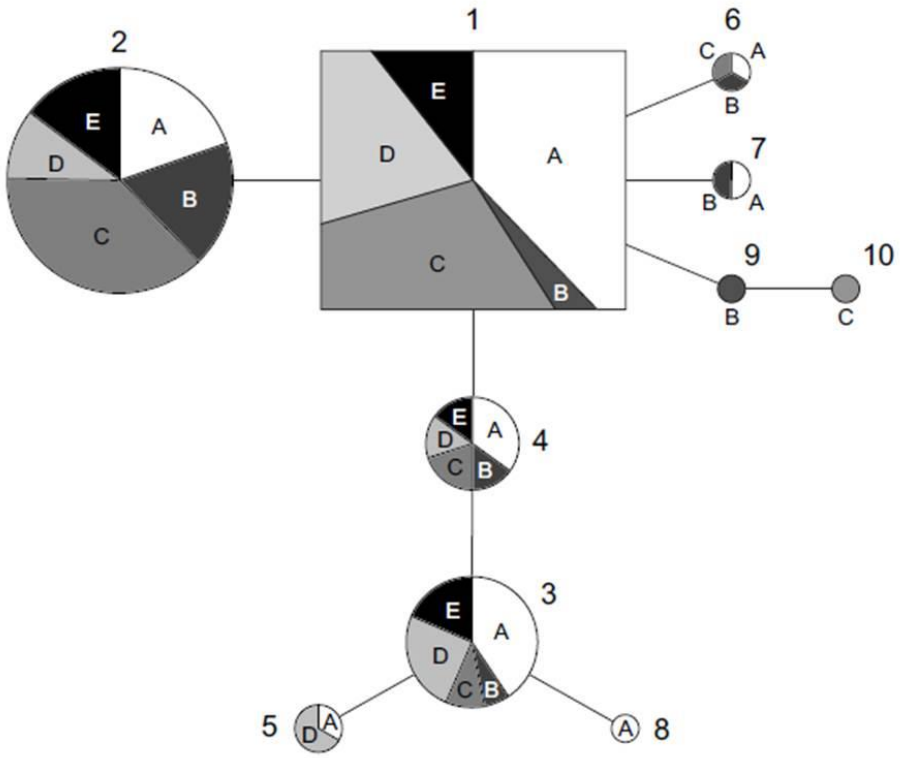

Figure 2. Unrooted haplotype network estimated by statistical parsimony analysis. Haplotype numbers (1-10) are shown for each of the five ponds in different shades of grey. The box indicates the haplotype identified by coalescence criteria as the ancestral sequence. Areas of pies and slices within pies are proportional to haplotype frequencies.

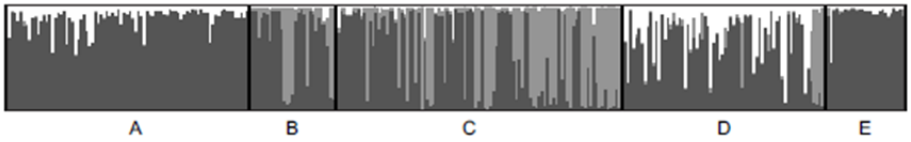

Figure 3. Plot of proportions of turtles' genome belonging to each of the three clusters considered most likely to account for the observed genotypes. Each individual is represented by a vertical line partitioned into K=3 segments (dark grey, light grey, white) with lengths corresponding to the proportion of its genome originating from each of the three clusters inferred by a model-based Bayesian method. Black lines separate turtles sampled in different ponds (A to E).

Table 3. Pairwise comparison matrix of F-statistics between ponds based on mtDNA and female E. orbicularis microsatellite analysis (Φ / θ, respectively, below diagonal) and male microsatellite analysis (θ, above diagonal). Bold and underlined values are significant at the 1% and 5% level, respectively.

	A	B	C	D	E
A		- 0.001	0.010	0.000	- 0.013
B	**0.075** / **0.030**		0.006	- 0.006	- 0.017
C	<u>0.037</u> / 0.003	0.022 / **0.015**		0.014	- 0.006
D	- 0.015 / **0.015**	<u>0.080</u> / **0.023**	<u>0.049</u> / **0.015**		- 0.013
E	0.007 / 0.004	- 0.005 / <u>0.017</u>	0.014 / 0.003	0.008 / <u>0.011</u>	

Table 4. Means and 95% confidence intervals of the posterior probabilities for migration rates of E. orbicularis between ponds. Values in bold show migration rates higher than 0.1.

Pond of origin	Recipient pond				
	A	B	C	D	E
A	**0.993** (0.976, 0.999)	**0.229** (0.170, 0.281)	**0.202** (0.165, 0.237)	**0.291** (0.073, 0.298)	**0.302** (0.265, 0.327)
B	0.001 (0.000, 0.009)	**0.677** (0.667, 0.702)	0.002 (0.000, 0.010)	0.003 (0.000, 0.013)	0.006 (0.000, 0.030)
C	0.001 (0.000, 0.008)	0.078 (0.033, 0.134)	**0.791** (0.756, 0.827)	0.027 (0.008, 0.055)	0.007 (0.000, 0.033)
D	0.002 (0.000, 0.012)	0.007 (0.000, 0.031)	0.002 (0.000, 0.011)	**0.675** (0.667, 0.694)	0.006 (0.000, 0.031)
E	0.001 (0.000, 0.008)	0.007 (0.000, 0.031)	0.002 (0.000, 0.009)	0.002 (0.000, 0.012)	**0.676** (0.667, 0.703

Mantel statistics was negative, varying from -0.36 to -0.42 when females and males only were tested, respectively, showing no significant decline in genotypic similarity with increasing geographic distance ($P > 0.05$). The mean posterior probabilities and 95% confidence intervals for migration rates between sampling sites showed a significantly high percentage of resident individuals in pond A (Table 4). Most turtles from ponds B to E were also residents, however, the relatively large proportion of migrants received from pond A (m = 0.202-0.302) suggested a recent upstream movement pattern.

DISCUSSION

The European pond turtle shows a relatively wide distribution which is, however, deceiving for natural populations show evidence of severe decline in many parts of the range. Wetlands conversion to cultivated land, replacement of earthen sides with concrete and changing agricultural practices are among the main threats to the species. As a consequence, *E. orbicularis* is now protected under the Bern Convention on European wildlife, listed in the lower risk, near threatened category of the IUCN red list of threatened species and considered a biological indicator of wetland habitat quality in many nature reserves of southern Europe.

Molecular genetic studies have so far contributed to the conservation of *E. orbicularis* by reconstructing the phylogeography history of natural populations in Europe and southwestern Asia (Kotenko 2000; Fritz et al. 2005; Fritz et al. 2006; Lenk et al. 1999) and using geographical distribution of mtDNA haplotypes to address management issues related to the pet trade (Fritz et al. 2004). The prevailing nomenclature, based mainly on morphological traits, is also being re-evaluated in regions, such as southern Italy, where strong gradients of genetic variations have been detected by mtDNA analysis and multilocus DNA fingerprinting thorugh PCR amplification of inter-microsatellite sequences with universal primers (Fritz et al. 2005; Fritz et al. 2006). However, demographic patterns as assessed by fine-scale genetic analysis of population structure that may help devising management plans for natural populations of *E. orbicularis* in protected areas are scarce. The area of Kerkini Lake and its affluent, the Strymon River, in northern Greece include wetlands of international importance. Marsh wetlands and sparse temporary and permanent ponds along the river are crucial to the survival of *E. orbicularis* and represent an area where fine scale gentic structure analysis may provide important information for species management.

A relatively high genetic diversity was recovered across the five ponds anlaysed in this study. A number of shared and unique haplotypes were recorded in downstream and middle ponds. Of particular interest, the highest and lowest number of haplotypes and polymorphic sites were recorded downstream in pond A and upstream in pond E, respectively. Moreover, upstream ponds had no unique haplotypes and were characterized by a

relatively more uniform haplotype distribution. A pattern of higher haplotipic diversity is common in populations that function as a source of gentic variation. Similarly, more uniform and less diverse haplotype distribution are expected in sites where turtles tend to disperse.

For many riverine, riparian, or coastal species, geographic distances may not adequately represent the actual distances separating populations. In such cases, user defined distances, such as river distances calculated by following river courses or coastlines may better reflect the actual distances that must be traversed if individuals were to migrate between populations (e.g. Fetzner & Crandall 2003). For aquatic species capable of terrestrial dispersal, however, the effect of landscape hydrological pattern on population genetic structure is typically negligible (Hughes et al. 2009), although individuals located in different parts of a river drainage may display greater genetic differentiation than populations from within restricted areas. (Hurwood & Hughes 2001). The connectivity among *E. orbicularis* individuals along the Strymon river seems not to be affectes by the position in the stream network. Turtles from pond A and E had a substantial sharing of the same genotypic cluster and a significantly higher level of admixture was recorded across ponds B to D. Moreover, pond turtles sampled at sites 1 to 3 along the Strymon River had no common origin and did not belong to ponds geographically close to the sampling sites.

Patterns of movement recorded in *E. orbicularis* are close to the headwater model of Hughes et al (2009), whereby the ability for some terrestrial movement may allow dispersal and gene flow among nearby headwaters, or ponds in our study, even if this entails crossing unsuitable environments. The shallow sequence divergence among the 10 haplotypes suggests short-term evolutionary isolation among ponds, and the scattered spatial pattern of the main four haplotypes (1-4) would imply recurrent interchange between ponds. Syntopic adult *Mauremys rivulata* were caught both in the ponds and in the river, although the occurrence in the two different habitats changed seasonally (Chelazzi et al. 2007). Overland movements, besides the transfer from the river to the ponds and southern marshes, were rare, but not absent, while the river served for males and females mainly as a corridor for seasonal transit between the mating and aestivation–overwintering grounds. The river was occasionally also used for longitudinal dispersal, particularly in males. Chelazzi et al (2007) also found that male *M. rivulata* were less phylopatric and moved longer distances than females, which remained confined within a very localized pond–marsh system. *Emys orbicularis* showed low levels of genetic divergence, no significant differences between ponds for males, and a significant multilocus genotype divergence between ponds when females only were considered in the analysis. Moreover, female turtles from ponds A, B and D were more related to each other than to turtles from other ponds, while males from pond C only had a significantly higher within- than among-pond average relatedness value. Lack of make phylopatry was also confirmed by the fact that none of the five ponds

had males showing average relatedness with females of the same ponds higher than average relatedness with females from other ponds.

Finally, statistically significant differences recovered between pond A and the other upstream ponds (B to D), and between pond C and ponds D and E. The higher frequency of the ancestral haplotype in pond A, C and D supported a Bayesian gene flow analysis whereby the relatively large proportion of migrants received from pond A by the other ponds suggested recent upstream movements of *E. orbicularis*.

Conservation implications

There is a growing need for natural resource managers to evaluate the impact of proposed management actions on the extent of habitats and the degree of fragmentation of these habitats. Habitat loss and fragmentation create discontinuities in the distribution of critical resources or environmental conditions. From the perspective of an organism, such discontinuities in the distribution of suitable habitat lead to a reduction of connectivity among population fragments. As habitat is lost and populations fragmented, functional connectivity through individual exchange and gene flow becomes critically important. Specifically, subdivision and isolation of populations leads to reduced dispersal success and patch colonization rates, which may result in a decline in the persistence of populations and an enhanced probability of regional extinction across a landscape (Segelbacher et al. 2010). Gene flow among populations is necessary to maintain the long-term viability of populations. Gene flow maintains local genetic variation by counteracting genetic drift and spreads potentially adaptive genes. From the perspective of conservation biology, it is thus essential to infer the functional connectivity of populations across landscapes. Understanding the landscape features that drive gene flow, the spatial scales at which they act, and the temporal dynamics of their effects on population substructure is essential to effectively use genetic data as a tool for evaluating population status and fragmentation. In addition, using this understanding to predict, localize and implement empirically based conservation corridors should greatly improve the successfulness of efforts to promote landscape connectivity of species at risk due to fragmentation.

Water permanance in ponds and proper management of river regime is of paramount importance for *E.orbicularis* in the Kerkini area. The relatively high connections among ponds advocate the need for their conservation through proper management and minimized used for livestock. Several studies have identified diverse dam-related effects on the genetic diversity and population structure of riverine species (e.g. Reid et al. 2008). A possible further threat to the species is isolation due to main geographical barriers such as weirs and the dam constructed to retain water downstream of the Strymon river, in the Kerkini wetland. Pond turtles may in fact be affected by barriers to migration via waterways. Turtles from fragmented habitats have significantly reduced home

range sizes and daily movements compared to turtles from unfragmented habitats (Bennett et al. 2010). Deviation from Hardy-Weinberg expectation was observed, on average, for 50% of the loci examined, suggesting recurrent demographic changes in *E. orbicularis* upstream of the dam. Further studies may want to compare levels of genetic diversity and connectivity among turtles sampled downstream of the Kerkini lake dam and upstream of the site considered in this study, where water is not diverted for agricultural purposes.

REFERENCES

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (2004) GENETIX 4.05, WindowsTM Software for Population Genetics. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bennett AM, Keevil M, Litzgus JD (2010) Spatial ecology and population genetics of northern map turtles (Graptemys geographica) in fragmented and continuous habitats in Canada. Chelonian Conservation and Biology 9, 185-195.

Castelloe J, Templeton AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. Molecular Phylogenetics and Evolution 3, 102-113.

Chelazzi G, Naziridis T, Benvenuti S, Ugolini A, Crivelli AJ (2007) Use of river-wetland habitats in a declining population of the terrapin (Mauremys rivulata) along the Strymon River, northern Greece. Journal of Zoology 271, 154-161.

Ciofi C, Tzika AC, Natali C, et al. (2009) Characterization of microsatellite loci in the European pond turtle Emys orbicularis. Molecular Ecology Resources 9, 189-191.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1659.

Crivelli AJ, Grillas P, Jerrentrup H, Naziridis T (1995) Effects on fisheries and waterbirds of raising water levels at Kerkini reservoir, a Ramsar site in northern Greece. Environmental Management 19, 431-443.

Earnest K, Scott J, Schaefer J, Duvernell D (2014) The landscape genetics of syntopic topminnows (Fundulus notatus and F. olivaceus) in a riverine contact zone. Ecology of Freshwater Fish 23, 572-580.

Ernst CH, Barbour RW, Hershey MF (1974) A new coding system for hard-shelled turtles. Transactions of the Kentucky Academy of Science 35, 27-28.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611-2620

Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theoretical Population Biology 3, 87-112.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479-494.

Fetzner JW, Jr., Crandall KA (2003) Linear habitats and the nested clade analysis: An empirical evaluation of geographic versus river distances using an Ozark crayfish (Decapoda: Cambaridae). Evolution 57, 2101-2118.

Fluker BL, Kuhajda BR, Harris PM (2014) The effects of riverine impoundment on genetic structure and gene flow in two stream fishes in the Mobile River basin. Freshwater Biology 59, 526-543.

Fritz U (2003) Die Europäische Sumpfschildkröte. Laurenti-Verlag, Bielefeld.

Fritz U, Barata M, Busack SD, Fritzsch G, Castilho R (2006) Impact of mountain chains, sea straits and peripheral populations on genetic and taxonomic structure of a freshwater turtle, Mauremys leprosa (Reptilia, Testudines, Geoemydidae). Zoologica Scripta 35, 97-108.

Fritz U, Cadi A, Cheylan M, et al. (2005) Distribution of mtDNA haplotypes (cyt b) of Emys orbicularis in France and implications for postglacial recolonization. Amphibia-Reptilia 26, 231-238.

Fritz U, D'Angelo S, Pennisi MG, Lo Valvo M (2006) Variation of Sicilian pond turtles, Emys trinacris – What makes a species cryptic? Amphibia-Reptilia 27, 513-529.

Fritz U, Guicking D, Lenk P, Joger U, Wink M (2004) When turtle distribution tells European history: mtDNA haplotypes of Emys orbicularis reflect in Germany former division by the Iron Curtain. Biologia 59 (Supplement 14), 19-25.

Gonzalez-Porter GP, Hailer F, Flores-Villela O, Garcia-Anleu R, Maldonado JE (2011) Patterns of genetic diversity in the critically endangered Central American river turtle: human influence since the Mayan age? Conservation Genetics 12, 1229-1242.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. Heredity 83, 145-154.

Hirschmann A, Malabarba LR, Thomaz AT, Fagundes NJR (2015) Riverine habitat specificity constrains dispersion in a Neotropical fish (Characidae) along Southern Brazilian drainages. Zoologica Scripta 44, 374-382.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9, 1322-1332.

Hughes JM, Schmidt DJ, Finn DS (2009) Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. Bioscience 59, 573-583.

Hurwood DA, Hughes JM (2001) Nested clade analysis of the freshwater shrimp, Caridina zebra (Decapoda: Atyidae), from north-eastern Australia. Molecular Ecology 10, 113-125.

Kalinowski ST (2004) Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. Conservation Genetics 5, 539-543.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes 5, 187-189.

Kanno Y, Vokoun JC, Letcher BH (2011) Fine-scale population structure and riverscape genetics of brook trout (Salvelinus fontinalis) distributed continuously along headwater channel networks. Molecular Ecology 20, 3711-3729.

Kotenko TI (2000) The European pond turtle (Emys orbicularis) in the Steppe zone of the Ukraine. In: Stapfia 69, Vol. 149. Die Europäische Sumpschildkröte (eds. Hödl W, Rössler M), pp. 87-106. OÖ Landesmuseums, Linz.

Lebboroni M, Chelazzi G (1991) Activity patterns of Emys orbicularis L. (Chelonia, Emydidae) in central Italy. Ethology Ecology & Evolution 3, 257-268.

Lebboroni M, Chelazzi G (2000) Waterward orientation and homing after experimental displacement in the European pond turtle, Emys orbicularis. Ethology Ecology & Evolution 12, 83-88.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. Molecular Ecology 11, 2445-2449.

Lenk P, Fritz U, Joger U, Wink M (1999) Mitochondrial phylogeography of the European pond turtle, Emys orbicularis (Linnaeus 1758). Molecular Ecology 8, 1911-1922.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology & Evolution 18, 189-197.

Matthews GVT (1993) The Ramsar Convention on Wetlands: its History and Development. Ramsar Convention Bureau, Gland, Switzerland.

Meldgaard T, Nielsen EE, Loeschcke V (2003) Fragmentation by weirs in a riverine system: A study of genetic variation in time and space among populations of European grayling (Thymallus thymallus) in a Danish river system. Conservation Genetics 4, 735-747.

Neville HM, Isaak DJ, Dunham JB, Thurow RF, Rieman BE (2006) Fine-scale natal homing and localized movement as shaped by sex and spawning habitat in Chinook salmon: insights from spatial autocorrelation analysis of individual genotypes. Molecular Ecology 15, 4589-4602.

Osborne MJ, Perkin JS, Gido KB, Turner TF (2014) Comparative riverscape genetics reveals reservoirs of genetic diversity for conservation and restoration of Great Plains fishes. Molecular Ecology 23, 5663-5679.

Ostergren J, Nilsson J (2012) Importance of life-history and landscape characteristics for genetic structure and genetic diversity of brown trout (Salmo trutta L.). Ecology of Freshwater Fish 21, 119-133.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13, 55-65.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28, 2537-2539.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12, 844-855.

Piry S, Alapetite A, Cornuet J-M, et al. (2004) GENECLASS2: A Software for genetic assignment and first-generation migrant detection. Journal of Heredity 95, 536-539.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

Pyrovetsi M, Papastergiadou E (1992) Biological conservation implications of water-level fluctuations in a wetland of international importance - Lake Kerkini, Macedonia, Greece. Environmental Conservation 19, 235-244.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences of the USA 94, 9197-9221.

Reid SM, Wilson CC, Mandrak NE, Carl LM (2008) Population structure and genetic diversity of black redhorse (Moxostoma duquesnei) in a highly fragmented watershed. Conservation Genetics 9, 531-546.

Robinson SJ, Samuel MD, Lopez DL, Shelton P (2012) The walk is never random: subtle landscape effects shape gene flow in a continuous white-tailed deer population in the Midwestern United States. Molecular Ecology 21, 4190-4205.

Rollinat R (1934) La cistude d'Europe. In: La Vie des Reptiles de la France Centrale, pp. 58-111. Librairie Delagrave, Paris.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4, 137-138.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8, 103-106.

Rovero F, Chelazzi G (1996) Nesting migrations in a population of the European pond turtle Emys orbicularis (L.) (Chelonia Emydidae) from central Italy. Ethology Ecology & Evolution 8, 297-304.

Sambrook J, Russell DW (2001) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Segelbacher G, Cushman SA, Epperson BK, et al. (2010) Applications of landscape genetics in conservation biology: concepts and challenges. Conservation Genetics 11, 375-385.

Souza FL, Cunha AF, Oliveira MA, Pereira GAG, dos Reis SF (2002) Estimating dispersal and gene flow in the neotropical freshwater turtle Hydromedusa maximiliani (Chelidae) by combining ecological and genetic methods. Genetics and Molecular Biology 25, 151-155.

Starkey DE, Shaffer HB, Burke RL, et al. (2003) Molecular systematics, phylogeography, and the effects of Pleistocene glaciation in the painted turtle (Chrysemys picta) complex. Evolution 57, 119-128.

Storfer A, Murphy MA, Evans JS, et al. (2007) Putting the 'landscape' in landscape genetics. Heredity 98, 128-142.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? Molecular Ecology 19, 3496-3514.

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 and Molecular Biology 132, 619-633.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4, 535-538.

Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise relatedness. Genetical Research 89, 135-153.

Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. Molecular Ecology Resources 11, 141-145.

Weir BS, Cockerham CC (1984) Estimating F statistics for the analysis of population structure. Evolution 38, 1358-1370.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163, 1177-1191.

Zuffi MAL, Gariboldi A (1995) Sexual dimorphism in Italian populations of the European pond terrapin, Emys orbicularis. In: Scientia Herpetologica (eds. Llorente GA, Montori A, Santos X, Carretero MA), pp. 124-129. Asociación Herpetológica Española, Barcelona.

CONCLUSIONS

The European pond turtle and the Sicilian pond turtle, the only two species of the genus *Emys* inhabiting Italy, are threatened species that mainly suffer exploitation and degradation of their habitat.

In recent years, interest in *Emys orbicularis* has certainly grown, causing Sicilian turtles to be identified in 2005 as a different species *Emys trinacris* (Fritz *et al.* 2005b). Phylogeographic and population genetics studies have focused on the European level, with several studies on turtles specifically found in France and Spain. To date, in Italy the monitoring of pond turtles is varied and uneven, and mainly focused in the North. Genetic investigations primarily examined a few specimens from multiple sites without going deeply into population dynamics and therefore an overall view is lacking (Fritz *et al.* 2006b; Fritz *et al.* 2005b).

In this survey we have taken into account the main Italian natural populations of *Emys orbicularis* and *Emys trinacris* with the highest number of individuals ever studied in Italy so far. Samples were analyzed by molecular techniques through the study of the mitochondrial DNA control region and sixteen microsatellite loci. The aim of this research was to describe the degree of genetic variability and divergence, and to reconstruct recent and historical dynamics of dispersal and colonization of the distribution area of *Emys* in Italy since the last glacial period. Indeed the Italian Peninsula has played a key role for the survival of the temperate species during ice ages which affected the Quaternary Period (Hewitt 2004).

Thanks to the use of classical and Bayesian statistics and analysis using complex models, such as the nested clade analysis, we described for the first time the existence of a likely secondary refugia for *Emys orbicularis* in Italy, out of the southern area, in the Tyrrhenian central Peninsula regions. This is supported by studies on organisms in different countries based on the recent theory of multiple refugia (Gomez & Lunt 2007), that has been receiving more and more attention and support. Moreover, "allochthonous" individuals considered to have been introduced by humans on the west coast (Fritz *et al.* 2005b) could actually be turtles harboring an ancient lineage that took refuge, in part, in these central zones, thereby explaining their existence on both coasts. Despite western Italy, including Sicily, having shown a scenario of contiguous range expansion, a general restricted gene flow was found among populations, in addition to some long-distance dispersal, most likely followed by fragmentation and extinction of intermediate populations.

As one priority, we highlight the requirement for management plans for island populations because of their diversity, particularly the Sardinian popualtion; island populations have the lowest values of heterozygosity and are currently not carefully monitored. Even if genetic diversity is not at such low critical levels, many of the Italian populations displayed the effects of high

fragmentation. Furthermore, some of these inhabit diminished wetlands that are increasingly surrounded by industrial and agricultural entities.

The results of this first study, as well as bringing new knowledge to the evolutionary history of the species, has provided a genetic database for the development of management programs and procedures to preserve wildlife. Surely turtles are animals that have suffered translocations by humans since prehistoric times. In fact, another point which negatively affects turtle conservation is handling by man as well as incorrect reintroductions which are as dangerous as the subtractions of wild individuals. We therefore considered it important to carry out extensive genotyping and characterization of Italian populations which provided the proper tools to proceed with the assignment of individuals of unknown origin to the source population. Comparing the genetic profile of individuals confiscated and kept in captivity with over a thousand turtles from thirty-one Italian populations in our genetic databank, using two Bayesian assignment tests, we determined with high probability the more similar natural population to each individual of unknown origin and therefore assist reintroduction plans in nature.

In this study were also evaluated the dynamics of movements on a small geographical scale of *Emys orbicularis* in northern Greece. The study area characterized by a series of ponds along the Strimon River is subject to intense human activity particulary in the use of significant amounts of river water for crop irrigation and ponds for cattle. The analysis has determined the genetic divergence degree and gene flow among turtles from the different ponds, highlighting the importance of the individual ponds for maintaining medium-high levels of intraspecific genetic diversity, and of ecological corridors, like the rivers, to facilitate movements in environments highly impacted by humans. *Emys* turtles, being a semi-aquatic species, are closely linked to the riverine compartment as much as the terrestrial one, therefore, it is clear that turtle-inhabited locations need rehabilitation. Land use for cultivation and cattle breeding includes water reserves and streams, which in the spring and summer, during the turtles' breeding season, are often reduced, and suffer from organic and fertilizers pollution. The importance of ecological corridors, demonstrated on a small scale, is essential over longer distances. A source-sink patterns of gene flow was supported by microsatellite Bayesian analysis whereby the relatively large proportion of migrants received from downstream ponds by the other ponds suggested recent upstream movements of *E. orbicularis*. Pond turtles showed low levels of genetic divergence, no significant differences between ponds for males, and a significant multilocus genotype divergence between ponds when females only were considered in the analysis, suggesting female phylopatry. For *E.orbicularis* in the Kerkini area water permanance is of paramount importance. The relatively high connections among ponds advocate the need for their conservation through proper management and minimized used for livestock.

This project on phylogeography and population genetics of the threatened *Emys orbicularis* and *Emys trinacris*, is the first of its kind on such an extensive scale in Italy, but also among nations at the European level, and is a step forward to better understanding their history and current state. This study recovered the structure and complex dynamics of main pond turtle populations in Italy. Further studies will be implemented to analyze additional Italian populations in order to obtain information on sites not considered in this research and help develop management programs for comprehensive conservation of the species in Italy.

REFERENCES

*Fritz U, d'Angelo S, Pennisi MG, Lo Valvo M (2006) Variation of Sicilian pond turtles, Emys trinacris - What makes a species cryptic? Amphibia-Reptilia **27**, 513-529.*

*Fritz U, Fattizzo T, Guicking D, et al. (2005b) A new cryptic species of pond turtle from southern Italy, the hottest spot in the range of the genus Emys (Reptilia, Testudines, Emydidae). Zoologica Scripta **34**, 351-371.*

Gomez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. Phylogeography of Southern European Refugia, 155-188.

*Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **359**, 183-195.*