The uncertainty of Late Pleistocene range expansions in the western Mediterranean: a case study of the colonization of south-eastern Spain by the spur-thighed tortoise, *Testudo graeca*

Eva Graciá¹, Andrés Giménez¹, José Daniel Anadón¹,², D. James Harris³, Uwe Fritz⁴ and Francisco Botella¹

**ABSTRACT**

**Aim** Recent biogeographical studies have postulated a North African, Late Pleistocene, origin for some species of the Iberian Peninsula. However, a robust assessment of such range expansions requires high-resolution molecular tools to resolve overlapping biogeographical and cultural processes. Here we aim to determine whether the spur-thighed tortoise, *Testudo graeca*, arrived in south-eastern Spain during historical or prehistoric times, and whether its dispersal to the Iberian Peninsula was human-mediated.

**Location** The western Mediterranean Basin (south-eastern Spain, northern Algeria and north-western Morocco).

**Methods** Using 428 samples from 19 sites in North Africa and 18 in south-eastern Spain, we obtained mitochondrial sequences from the cytochrome *b* gene and genotypes derived from seven microsatellite loci. These data were employed to obtain population genetics descriptors, haplotype networks, Bayesian cluster analyses and isolation-by-distance patterns. Moreover, we used a Bayesian demographic approach to delimit the dates involved in the range expansion.

**Results** We found lower levels of genetic variability and weak mitochondrial differentiation in the south-eastern Spanish tortoises compared with the North African ones. However, exclusive haplotypes occurred in the Iberian samples and microsatellite cluster analyses revealed moderate levels of admixture across both sides of the Mediterranean. A coastal area in the west of Algeria and the central-southern region in south-eastern Spain are suggested as the most probable founder and arrival places, respectively. Finally, we identified signatures of an ancient bottleneck event approximately 20,000–30,000 years ago.

**Main conclusions** The spur-thighed tortoise probably arrived in south-eastern Spain during Late Pleistocene sea-level low stands. The role that humans may have played as dispersers across the Mediterranean remains unclear. Our results are in accordance with other recent findings of trans-Mediterranean expansions during this period and highlight the importance of employing precise methodological approaches before a species can be considered as historically introduced.

**Keywords** Cytochrome *b*, demographic history, human-mediated dispersal, Mediterranean Basin, microsatellites, North Africa, population structure, species conservation, Strait of Gibraltar, transmarine dispersal.
INTRODUCTION

Mediterranean biodiversity has been influenced by humans for at least the last 10,000 years (Blondel, 2006). As a consequence, it is challenging to distinguish between a human-mediated or natural origin for some Iberian populations. This is particularly true for organisms that crossed the Strait of Gibraltar in the last 100,000 years or so, as any imprecision in dating or calibration could change the evidence favouring, or not, human-mediated transportation. Some Iberian species, including the common genet, *Genetta genetta*, the Mediterranean chameleon, *Chamaeleo chamaeleon*, and the Egyptian mongoose, *Herpestes ichneumon*, have traditionally been considered as having been introduced by humans (Morales, 1994; Pleguezuelos et al., 2008; Riquelme-Cantal et al., 2008). Their status as ‘historically introduced species’ was a consequence of their cultural linkage to historical civilizations, having their main ranges in Africa and the lack of any fossil record in Europe. However, although these species share the same genetic pattern of low mitochondrial differentiation across both sides of the Strait of Gibraltar and lower genetic variation in the Iberian populations, two alternative prehistoric scenarios could be responsible for the current pattern (Paulo et al., 2002; Gaubert et al., 2010).

1. Natural transmarine dispersals during periods of low sea level of the last glaciation (the Würm period, from 125,000 to 18,000 years ago), when the Mediterranean sea dropped by approximately 120–140 m relative to the present, reducing the sea strait between Iberia and North Africa to a width of only 3.5 km (Zazo, 1999; Flemming et al., 2003).

2. Dispersal as travelling companions with the first human migrations across the Strait of Gibraltar; as could be the case for species consumed by prehistoric cultures, including tortoises or hedgehogs (Fernández-Jalvo et al., 1999; Blasco, 2008; Vamberger et al., 2011), or those linked to human activities, such as has been proposed for the white-toothed shrew, *Crocidura russula* (Cosson et al., 2005), and the Moorish gecko, *Tarentola mauritanica* (Harris et al., 2004).

However, the assessment of when humans crossed the Strait of Gibraltar for the first time is a complex issue with many uncertainties. Although it is generally thought that human movements can be traced back to the mid- or Late Pleistocene, the first definite evidence for deep-sea fishing and navigation within the region dates to only the 12th millennium BC (Straus, 2001; Broodbank, 2006). Consequently, due to the complexity in clarifying which scenario is associated with each expansion event, there is a need to deepen our knowledge of the genetic spatial structure of populations across their Mediterranean range, by comparing their genetic signatures with both conservative and rapidly evolving markers.

Here we use this approach to unravel the case of a North African expansion: the origin of the endangered spur-thighed tortoise, *Testudo graeca* Linnaeus, 1758 (Testudinidae), in south-eastern Spain. The majority of the western Mediterranean distribution of this species lies in North Africa, with a few small and isolated western European populations confined to the Iberian Peninsula (Doñana and south-eastern Spain) and on a few islands (Mallorca, Sardinia, Sicily; Fig. 1). In general, western European populations are thought to be the result of historical introductions, which is supported by the absence of any fossil record in western Europe (Morales-Pérez & Sanchis-Serra, 2009) and low mitochondrial differentiation between the European and North African populations (Álvez et al., 2000; Harris et al., 2003;

Figure 1 Approximate distribution of the five western Mediterranean mitochondrial lineages of *Testudo graeca* as described in Fritz et al. (2009). I corresponds to *T. g. soussensis*, II to *T. g. marokkensis*, III to *T. g. graeca*, IV to *T. g. nabeulensis* and V to *T. g. cyrenaica*. Grey areas represent the global range of the species, while black dots correspond to the western European populations of North African origin. The rectangle shows the study area.
Fritz et al., 2009; Salinas et al., 2011; Vamperger et al., 2011). However, for the south-eastern Spanish tortoises an incipient evolutionary isolation has recently been suggested by the discovery of unique haplotypes and an explicit genetic pattern with spatial coherence (Fritz et al., 2009; Graciá et al., 2011; Salinas et al., 2011), suggesting a more ancient origin for this expansion.

In the present study, we aim to determine whether the spur-thighed tortoise colonized south-eastern Spain in historical or prehistoric times. To do so, we estimate the time interval since the species’ arrival using Bayesian demographic analyses. In addition, we characterize this range expansion by assessing differences in the spatial genetic patterns among the North African and south-eastern Spanish tortoises using a clustering approach to estimate the areas of origin and arrival of founder populations. Furthermore, we apply common exploratory methods of population genetics to assess the role of humans. The absence of a signature of recent bottlenecks, an isolation-by-distance pattern across the entire Iberian study area, a decrease of expected heterozygosity with increasing geographical distance from the putative centre of origin and genetic units established in a spatially coherent context would be evidence against recent human-assisted dispersal.

MATERIALS AND METHODS

Sampling and genetic analyses

This study broadly covers the range of the subspecies Testudo graeca graeca, being based on 428 samples of wild tortoises from the Moulouya Valley in Morocco (seven from a single location), the Tell and Saharan Atlas in Algeria (97 samples from 18 sites) and south-eastern Spain (325 samples from 18 sites). Additionally, we considered 37 previously published sequences from known localities (see Appendix S1 in Supporting Information).

The complete cytochrome b gene and 25 bp of the adjacent tRNA-Thr gene (1164 bp; cyt b) was amplified by polymerase chain reaction (PCR) and subsequently sequenced. Additionally, seven microsatellite loci previously used for these studied populations, namely GmuD16, GmuB08, GmuD51, Test 76, Test 21, Test 71 and Gp96 (Graciá et al., 2011; Salinas et al., 2011), were analysed. All loci were examined for genotyping errors, allelic dropout and null alleles using micro-checker 2.2.3 (van Oosterhout et al., 2004). For further extended information regarding the marker selection and laboratory procedures see Appendix S2.

Mitochondrial data analyses

Sequences were aligned with Clustal W using default parameters as implemented in MEGA4 (Tamura et al., 2007) and collapsed into haplotypes using tcs 1.2.1 (Clement et al., 2000). We used the same software to calculate a parsimony network, which connects the different haplotypes via a minimal number of mutational steps. With this approximation it is possible to have information about the relative age of each haplotype; those that have more connections and are located in the interior of the network are considered to be older than tip haplotypes (Posada & Crandall, 2001). Haplotype diversity (Hd), nucleotide diversity (π) and the average number of nucleotide differences (Pi) were obtained for the North African and south-eastern Spanish groups using DNAsP 5.10 (Librado & Rozas, 2009).

Microsatellite data analyses

Genotypic linkage disequilibrium between each pair of microsatellite loci was tested using the Markov chain Monte Carlo (MCMC) method implemented in genepop 4.0 (Rousset, 2007). To discard the linkage of the used loci is a premise for the use of the two Bayesian clustering methods we employed, structure 2.3.3 (Pritchard et al., 2000; Falush et al., 2003; Hubisz et al., 2009) and Geneland 3.3.2 (Guillot et al., 2005a,b). These clustering programs identify groups (K) that maximize the Hardy–Weinberg equilibrium (HWE) and the linkage equilibrium within them.

First, we ran structure from K = 1 to K = 10, without prior population information, using the admixture model and the option ‘correlated allele frequencies’. The burn-in was set to 10,000 and the number of iterations to 100,000. We repeated each analysis 10 times to evaluate the convergence and to estimate the optimal value of K following the method proposed by Evanno et al. (2005). Given that it is known that in south-eastern Spanish populations conservation actions including introductions and translocations have been carried out (Graciá et al., 2011; Salinas et al., 2011), we used the ‘locprior’ option, considering all the different sampling sites, and the optimal obtained value of K as priors, with the aim of detecting animals potentially introduced from North Africa (Hubisz et al., 2009). In order to corroborate these results we compared multilocus information in an assignment/exclusion test using GeneClass2 software (Piry et al., 2004), selecting the Bayesian method for computation suggested by Rannala & Mountain (1997) and the algorithm by Paetkau et al. (2004) for 1000 simulated individuals (α = 0.05). Those individuals detected as introduced in the south-eastern Spanish range were removed from further analyses. All these results were graphically displayed by Distruct (Rosenberg, 2004).

The second clustering method, Geneland, has some similarities to structure but allows the use of the geographical location of each sample as prior information. We ran the MCMC five times from K = 1 to K = 10, with the following parameters: 500,000 MCMC iterations, a thinning of 100, maximum rate of Poisson process fixed to 100, uncertainty attached to spatial coordinates fixed to 1 km, minimum K fixed to 1 and maximum fixed to 10, maximum number of nuclei in the Poisson–Voronoi tessellation fixed to 300 and a Dirichlet model for allelic frequencies (given that it has been demonstrated that this performs better than the alternative model; Guillot et al., 2005b). We inferred the optimal value
of $K$ from its modal value in these first five runs and ran the software five additional times with $K$ fixed to this obtained value. The posterior probability of population membership for each pixel in the spatial domain ($x = 50; y = 50$) was computed for these 10 runs after a burn-in of the first 10% of saved iterations. Once again the consistency of results was evaluated among the different runs.

**Genalex 6.0** (Peakall & Smouse, 2006) was employed to estimate diversity and differentiation indices among and within clusters inferred by **structure** and **Geneland**. Allele frequencies, the mean number of alleles and effective alleles, the number of private alleles and expected heterozygosities were obtained using this software. Moreover, we used this program to obtain pairwise $F_{ST}$ values as a measure of genetic distance among clusters and between sampling sites within them. These genetic distances were employed using this software in two subsequent analyses: (1) in hierarchical analysis of molecular variance (AMOVA) to determine how genetic variation is distributed among regions, populations and individuals, when considering the groups suggested by the clustering analyses results; and (2) in two Mantel tests to detect the relationship between genetic isolation and geographical distance of North African and south-eastern Spanish populations (among sampled sites). In these last analyses, genetic distance was described by $F_{ST}(1-F_{ST})$ and geographical distance by considering the natural logarithm of the Euclidean distance among sampled sites. Finally, we correlated the expected heterozygosities at sites with the geographical distances to the putative point of arrival of the species to south-eastern Spain, which was obtained as the centroid of those sites sharing at least a 5% of their assignment with the North African clusters in the **Geneland** analyses. In this correlation, in AMOVA analyses and in Mantel tests, only those sites with more than seven individuals were considered to avoid problems due to small sample sizes.

**Demographic history analyses: dating the expansion process**

To infer the colonization process in south-eastern Spain, we explored the demographic history of the species there through different analyses of the microsatellite data. Tortoises from south-eastern Spain were treated as a single population, with the exception of one analysis, which was repeated for the location with the highest number of samples (see below).

First, we performed two different analyses using the software **Bottleneck** 1.2.02 (Piry et al., 1999). We tested the temporary excess of heterozygosity that results from a reduction of the population size under the stepwise mutation model (SMM), and the two-phase mutation model (TPM) with 95% of SMM and 12% of variance (Cornuet & Luikart, 1996). To evaluate the significance of this analysis, we used the Wilcoxon signed-rank test based on 10,000 repetitions. Using this software we also carried out a mode-shift test, which discriminates bottlenecked populations from stable populations, to detect a distortion of the expected L-shaped distribution (Luikart et al., 1998).

Secondly, **msvar** 1.3 (Beaumont, 1999; Storz & Beaumont, 2002), based on a Bayesian coalescent method, was used to infer historical changes in the effective population size of the south-eastern Spanish populations. This method is based on the observed allele distributions and allele frequencies, assuming that a stable population with a determinate size ($N_0$) started to decrease (or increase) linearly or exponentially some time ago ($T$) towards the current population size ($N_0$). Based on this assumption and using the MCMC approach the method not only allows the estimation of $N_0$ and $N_1$, but also $T$. Mutation rates ($\theta = 2N_0\mu$) were assumed under the SMM model and, like the rest of the parameters, were inferred separately across loci using priors assumed to be under a lognormal distribution. Because we were testing founder and bottleneck processes, we ran the program under the exponential growth model, using different wide uninformative priors. The generation time of $T. g. graeca$ was set at 17.72, the mean value taken from those estimated by Díaz-Paniagua et al. (2001) from different years and scenarios. To evaluate the stability of the results we ran the program five times for one of the employed priors (with mean 100,000), with the lowest interval being $2 \times 10^4$ iterations. We discarded the first 10% of total iterations to avoid bias in parameter estimation due to starting conditions. The distribution of the remaining data was plotted against prior distributions to see the consistency of the results over the different runs and used to obtain the lower (10%), the median (50%) and upper (90%) quantiles of the posterior distributions. Given that genetic structure can generate false bottleneck signals when using this method (Chikhi et al., 2010) or that multiple arrivals at different places could have given rise to south-eastern Spanish populations with different past population demographic histories, we repeated these analyses for the sampled site with the highest sample size, Galera ($n = 74$). We note that although it seems counterintuitive to use monomorphic loci in demographic analyses, it has been reported that their exclusion could lead to an over-estimation of the decrease in population (Beaumont, 1999). Therefore, in order to estimate the possible bias of the obtained results, the analyses were run a further three times but excluding the two monomorphic loci.

To delimit further the dates and agents involved in this species’ expansion, we used the data obtained from the **msvar** analyses including the two monomorphic loci. Specifically, we tested the subsequent hypotheses: (1) arrival before the last glacial period (more than 125,000 years ago); (2) arrival by transmarine dispersal during the last glacial period (18,000–125,000 years ago); (3) arrival by prehistoric dispersal during the last glacial period (18,000–30,000 years ago, assuming that humans could have crossed during conditions of low sea level); (4) arrival by introduction after the development of prehistoric navigation (3100–14,000 years ago); (5) arrival by introduction by
Arabs or Phoenicians (520–3100 years ago); and finally (6), although not very likely, arrival by recent introduction as a consequence of the international pet trade during the last 100 years.

We calculated the density of iterations of T accumulated until the beginning and end of each hypothesis and, after this, the odds in favour of the time (T) periods that delimit each hypothesis (e.g. for the fifth hypothesis the inferior limit is T ≤ 3100 and a superior limit is T > 520). The odds in favour of an event permit the comparison of two complementary alternatives, being expressed as the ratio between them \( P/(1-P) \). It is usually considered that ratios greater than 3 represent positive evidence in favour of the considered alternative, and greater than 7 are considered significant (as in Sousa et al., 2008). Consequently, values below 0.33 and 0.14, respectively, were considered as positive evidence and significantly in favour of its complementary alternative (in relation to the previous example, the complementary T periods would be T > 3100 and T < 520 for the lower and upper limits, respectively).

**RESULTS**

**Mitochondrial genetic structure**

Our dataset for the North African and south-eastern Spanish T. g. graeca included a total of 36 different haplotypes for cyt b. Nine of these haplotypes were previously reported by Fritz et al. (2009) and given that we amplified the same fragment, we followed the nomenclature of these authors (from B1.8 to B1.34; see Appendix S1 for GenBank accession numbers). The parsimony network was fully connected under the 95% criterion, with few missing node haplotypes. With two exceptions, all haplotypes belonged to the B1 clade, corresponding to the subspecies T. g. graeca (Fig. 2a).

The remaining two haplotypes, A6 and A7, were found at El Kala, an Algerian site close to the Tunisian border, where another subspecies (T. g. nabeulensis) is present (Fritz et al., 2009). Within the B1 clade, the two haplotypes B1.2 and B1.3 were shared among Spanish and North African tortoises, while 28 haplotypes were only found in North Africa and four were exclusive to south-eastern Spain. For North Africa, an internal substructure was revealed in which the different haplotypes were mostly shared among neighbouring sites, with the exception of some common haplotypes with broad geographical distribution (such as B1.2, B1.3 or B1.29; Fig. 2b). The six haplotypes present in south-eastern Spain were distinct by one mutation step each (Fig. 2a). Additionally, a strong spatially coherent structure was revealed in which: (1) haplotype B1.4 was restricted to the central-southern part of the study area (Vera Basin and the Bédar and Cabrera mountains); (2) haplotypes B1.3 and B1.1 were common in the central and northern study area (from the Los Lobos area and the Almagro Mountains to the north of Lorca and the Almenara Mountains); (3) at Marinica, a site located in a central coastal place of the study area, the three unique haplotypes B1.2, B1.15 and B1.16 were found, the last two of which had not previously been reported for south-eastern Spain (Fig. 2b).

Haplotype diversity, nucleotide diversity and the average number of nucleotide differences were all higher in North African tortoises (\( H_d = 0.854; \pi = 0.0029; \Pi = 3.4 \)) than in south-eastern Spain (\( H_d = 0.715; \pi = 0.0013; \Pi = 1.5 \)). From these analyses North African samples belonging to the A clade were excluded, given that this lineage does not represent the T. g. graeca subspecies.

**Genetic diversity and structure inferred from microsatellites**

Two loci (Test 76 and Gp96) were monomorphic. For the other five, we detected a total of 83 different alleles, ranging from 6 (Test 71) to 28 (GmuD51) per locus (Table 1). We did not find evidence for genotyping errors or allelic dropouts and no significant linkage disequilibrium was found for any of the 10 pairwise locus combinations. However, MICROCHEKER showed potential null alleles for Test 71 at Mouloya and Sotomayor sites and for GmuB08 at Galera.

Population structure inferred from STRUCTURE analyses indicated that K = 2 was the most appropriate cluster solution using the \( \Delta K \) statistic proposed by Evanno et al. (2005) (modal value of \( \Delta K = 975.0 \) when \( K = 2; \Delta K = 49.68 \) when \( K = 3 \)). Although this method is not able to detect when \( K = 1 \) is more appropriate than \( K = 2 \), we considered this result reliable given that the different runs of \( K = 2 \) were consistent over repetitions and divided North African and south-eastern Spanish individuals in the two different clusters without any sampling prior information (99% and 92.3% of North African and south-eastern Spanish individuals, respectively, were assigned to their expected geographical group; Fig. 3a). After using the sampling locations as prior information, we detected in south-eastern Spain two potentially recently introduced tortoises, which were assigned to the North African cluster with membership probabilities > 0.90. An assignment test in GENEClass corroborated these results, and these two individuals were removed from further analyses.

North African tortoises are characterized by a higher mean level of allelic diversity, but a similar level of expected heterozygosities (number of alleles, \( N_A = 12 \); number of effective alleles, \( N_E = 6.21 \)); expected heterozygosity, \( H_E = 0.54 \) compared with south-eastern Spanish tortoises (\( N_A = 8.43; N_E = 3.30; H_E = 0.49 \); Table 1). Additionally, 26 private alleles were found in North Africa, but only one in south-eastern Spain. After sequential Bonferroni correction, the two loci GmuB08 and GmuD51 did not conform to HWE in south-eastern Spanish tortoises, showing homozygote excess (\( P < 0.003 \)). Pairwise \( F_{ST} \) values were in general significant (data not shown). The weakest differentiation (0.04) was found between a western coastal site in Algeria (Saf Saf) and one of the southernmost sites in south-eastern Spain (Teresa, located in the Cabrera Mountains), while the highest level of differentiation (0.24) was found between the National Park El Kala, in the east of Algeria, and Malacate, a site in the
**Figure 2** Mitochondrial structure. (a) Parsimony network for mitochondrial DNA haplotypes of North African and south-eastern Spanish spur-thighed tortoises (*Testudo graeca*). Each line joining haplotypes indicates one nucleotide substitution, except when hatches across lines are present; then each hatch indicates one step. Colours indicate the haplotypes found in south-eastern Spain. The hatched area within each symbol represents the proportion of south-eastern Spanish haplotypes against the proportion of North African haplotypes and symbol size corresponds approximately to haplotype frequency. Thick lines connect the described haplotypes for south-eastern Spain, shaping a perfectly connected subnet. Only those haplotypes present in south-eastern Spain or mentioned in the text are highlighted. (b) Haplotype composition of each sampled site. Site abbreviations are defined in Appendix S1. Colours correspond to those in panel (a) and represent south-eastern Spanish haplotypes, while white symbols indicate haplotypes which occur only in North Africa. Symbol size corresponds approximately to the individual sample size for each place. The black line encloses the approximate distribution of the subspecies *T. g. graeca*.

**Table 1** Statistical descriptors by locus for the south-eastern Spanish and the North African spur-thighed tortoises (*Testudo graeca*).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele size range</th>
<th>NA</th>
<th>NE</th>
<th>NP</th>
<th>HO</th>
<th>HE</th>
<th>FIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmuD16</td>
<td>186–254</td>
<td>17</td>
<td>6.76</td>
<td>0</td>
<td>0.84</td>
<td>0.85</td>
<td>0.01</td>
</tr>
<tr>
<td>GmuB08</td>
<td>206–242</td>
<td>13</td>
<td>5.36</td>
<td>0</td>
<td>0.76</td>
<td>0.81</td>
<td>0.07*</td>
</tr>
<tr>
<td>GmuD51</td>
<td>143–239</td>
<td>21</td>
<td>5.06</td>
<td>0</td>
<td>0.76</td>
<td>0.8</td>
<td>0.06*</td>
</tr>
<tr>
<td>Test 76</td>
<td>114</td>
<td></td>
<td></td>
<td></td>
<td>Monomorphic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 21</td>
<td>200–214</td>
<td>3</td>
<td>1.87</td>
<td>0</td>
<td>0.46</td>
<td>0.47</td>
<td>0.02</td>
</tr>
<tr>
<td>Test 71</td>
<td>128–132</td>
<td>3</td>
<td>2.02</td>
<td>1</td>
<td>0.53</td>
<td>0.51</td>
<td>−0.05</td>
</tr>
<tr>
<td>Gp96</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td>Monomorphic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>–</td>
<td>8.43</td>
<td>3.30</td>
<td>–</td>
<td>0.48</td>
<td>0.49</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele size range</th>
<th>NA</th>
<th>Ne</th>
<th>NP</th>
<th>HO</th>
<th>HE</th>
<th>FIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>186–270</td>
<td>200–257</td>
<td>19</td>
<td>12.61</td>
<td>6</td>
<td>0.89</td>
<td>0.92</td>
<td>0.03</td>
</tr>
<tr>
<td>139–251</td>
<td>190–214</td>
<td>9</td>
<td>2.59</td>
<td>6</td>
<td>0.58</td>
<td>0.61</td>
<td>0.05</td>
</tr>
<tr>
<td>128–138</td>
<td>128–138</td>
<td>6</td>
<td>1.74</td>
<td>3</td>
<td>0.33</td>
<td>0.42</td>
<td>0.22</td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Monomorphic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>6.21</td>
<td>0.51</td>
<td>0.54</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant deviation from Hardy–Weinberg equilibrium after Bonferroni sequential correction (*P* < 0.005).

NA, number of alleles; NE, number of effective alleles; NP, number of private alleles; HO, observed heterozygosity; HE, unbiased expected heterozygosity; FIS, fixation index.
interior of the south-eastern Spanish study area, in the Almagra Mountains. The AMOVA analysis statistically supported the two clusters and indicated that 8.5% of the global genetic variation was distributed between them, 5.6% among sampled sites within these two groups, 5.2% among individuals within each site and 80.7% corresponded to individual variation \( \left( F_{ST} = 0.14; F_{IS} = 0.06; F_{IT} = 0.19; \right) \) all \( P \)-values = 0.01). Mantel tests reported significant correlations between genetic and geographical distances both within North Africa \( r = 0.56; P = 0.03 \) and within south-eastern Spain \( r = 0.37; P = 0.01 \).

In the five GENELAND runs, \( K = 4 \) was the suggested number of different populations when the whole study area was considered. The next five posterior runs with the number of populations fixed to four showed good consistency: south-eastern Spanish individuals were gathered in a single cluster, while the three remaining clusters were located within North Africa, differentiating the individuals from the Moulouya Valley, from Ain Chorfa (a coastal sampling site in the west of Algeria) and from the remaining sampling sites in Algeria (Fig. 3b). These clusters exhibited pairwise \( F_{ST} \) values which ranged from 0.03 (between Ain Chorfa and the remaining samples in Algeria) to 0.16 (between south-eastern Spain and the Moulouya Valley). The other two \( F_{ST} \) values with respect to the south-eastern Spanish group were 0.12 and 0.10 for Ain Chorfa and the rest of Algeria, respectively. The AMOVA test, using the proposed clusters as regions, revealed that 9.2% of genetic variation was distributed among regions, 5.2% among sampled sites within regions, 5.2% among individuals within sites, and 80.4% within individuals \( \left( F_{ST} = 0.14; F_{IS} = 0.06; F_{IT} = 0.20; \right) \) all \( P \)-values = 0.01). It is noteworthy that within the south-eastern Spanish range, only the cluster which mainly comprised Ain Chorfa appeared to be represented, although in a low proportion and concentrated in the centre-south of the studied area (Vera Basin and the Bédar and Cabrera mountains; Fig. 3b). Therefore, the locations of Chinas, Centinares, Teresa, Judío and Cintas (with at least 5% of their genetic variation assigned to the North African cluster) were used for calculating a centroid between them, considered as the putative arrival of the

---

**Figure 3** Genetic structure of North African and south-eastern Spanish spur-thighed tortoises (*Testudo graeca*) inferred from microsatellite loci. (a) Bar plots generated by **STRUCTURE** (Pritchard et al., 2000). Vertical bars represent individuals, while the amount of each colour indicates the proportion of each inferred cluster. The initial model assumed admixture and correlated allele frequencies. The arrow indicates the two individuals sampled in the south-eastern Spanish group significantly identified as introduced from North African when using the 'locprior' option (see Results). (b) Proportion of each site attributed to each of the four **GENELAND** clusters (Guillot et al., 2005a,b) using the geographical location of each sample as prior information. Site abbreviations are defined in Appendix S1. The black line encloses the approximate distribution of the *T. g. graeca* subspecies.
species to south-eastern Spain. We detected a significant decrease of the expected heterozygosities at Iberian sites with increasing distance from this point \( (P = 0.08; \ r = -0.62) \).

**Past demographic history and the arrival of the species in south-eastern Spain**

The results obtained with **bottleneck** did not suggest any evidence of a recent bottleneck event for the south-eastern Spanish tortoises. However, using **msvar** 1.3 we found a clear signal of an ancient population collapse both for the Galera site and for all the south-eastern Spanish tortoises (Fig. 4, Appendix S3). For both datasets, the effective population \( N_1 \) decreased over 100 times from a large ancestral effective size, estimated at 39,810 individuals (10th–90th quantiles = 7943–125,892 and 5011–100,000, for Galera and south-eastern Spain, respectively), to the current effective size \( N_0 \) of a few hundred (199 individuals, 10th–90th quantiles = 25–631 for Galera; and 251 individuals, 10th–90th quantiles = 63–1000, for the whole Spanish dataset; Fig. 4a). The bottleneck was dated to approximately 32,000 years ago (31,622 years, 10th–90th quantiles = 3981–100,000) when the whole Spanish dataset was considered, and to approximately 20,000 years ago when only the data from Galera were used (19,952 years, 10th–90th quantiles = 2511–79,432; Fig. 4b). The obtained posterior densities of \( N_0 \), \( N_1 \) and \( T \) were consistent over the different runs and the different priors used. The similarity in the estimated parameters both for the whole Spanish dataset or Galera alone is notable. Additionally, the obtained values were also similar when the two monomorphic loci were not considered. However, as was reported by Beaumont (1999), when these markers were excluded, the distributions were discretely shifted to the left, suggesting a slightly more recent process (Appendix S3).

To summarize these results, the hypothesis of an arrival before the last glaciation can be rejected. Likewise, the hypotheses of a very recent introduction as a consequence of the pet trade or the introduction of the species in historical times are not supported. The remaining three hypotheses (Table 2: trans-marine dispersal, prehistoric introduction during the last glacial period, or prehistoric introduction by prehistoric seafarers) are suggested as the most plausible ones. The relative temporal proximity of these three hypotheses (in part even overlapping) did not allow a decision regarding which is the most probable.

**DISCUSSION**

**Dating the arrival of the species in south-eastern Spain**

Different classes of molecular markers provide different insights into the temporal and geographical scales of the life history of a species. Whereas the study of genealogies can elucidate population processes, phylogeographical events or even speciation, allele frequencies can yield information on demographic trends (Sunnucks, 2000). The combination of these data types provides a more complete understanding of how an organism has responded to changes in the biogeographical landscape in a broad temporal spectrum, and the use of complementary analyses allows the study of different components of a species’ evolutionary history (Garrick et al., 2010). In this sense, the scale of resolution achieved by previous studies which employed mitochondrial DNA (mtDNA) only allowed rejection of the hypothesis that *T. graeca* arrived in the Iberian Peninsula before the Strait of Gibraltar opened (Alvarez et al., 2000; Fritz et al., 2009; Salinas et al., 2011). By contrast, studies based on allele frequencies of microsatellites suggested that Spanish tortoises were not introduced very recently (Gracía et al., 2011; Salinas et al., 2011). Our results, based on the Bayesian analyses of the historical demography of the species, allowed us to delimit the interval of time for its arrival in south-eastern Spain to prehistoric times, to a period pre-dating the maritime expansion of the Greeks and Phoenicians.
in the western Mediterranean, and after the beginning of the last glaciation process (between approximately 3100 and 125,000 years ago). Our analyses for Spanish tortoises revealed a bottleneck approximately 20,000–30,000 years ago with a strong reduction of the ancestral effective population size, coinciding with one of the minimum Mediterranean sea level phases (Flemming et al., 2003; Massetti, 2009). These results basically keep open two different possibilities for the species’ arrival: natural transmarine dispersal, which has been reported for other species of tortoise (e.g. Caccione et al., 1999; Gerlach et al., 2006), or its introduction as a consequence of the first cultural expansions of humans from North Africa (late Palaeolithic or early Neolithic; Broodbank, 2006).

Unfortunately, the role that humans could have played in the species’ dispersal across the Mediterranean remains unclear. Although a strong connection between tortoises and primitive Mediterranean cultures due to consumption habits is well known (e.g. Blasco, 2008; Morales-Pérez & Sanchis-Serra, 2009), it remains unclear exactly when humans crossed the Strait of Gibraltar for the first time (Flemming et al., 2003; Broodbank, 2006). One of the main traditional arguments claiming the recent arrival of *T. g. graeca* to the Iberian Peninsula is the lack of fossils, in contrast with the abundant fossil remains of another Spanish tortoise species, *Testudo hermanni*. Nevertheless, in a recent work Morales-Pérez & Sanchis-Serra (2009) reported the absence of *Testudo* fossils in the Iberian Mediterranean region for the last 30,000 years, due to the species’ rarity or even its local extinction. The rarity of *T. hermanni* as a consequence of climatic amelioration occurring during this period, or even due to its possible overexploitation, could have led to a change in the consumption habits of humans (Morales-Pérez & Sanchis-Serra, 2009). Therefore, the absence of fossil remains of *T. graeca* could be explained as a consequence both of the low initial size of the population and of a change in prey selection of modern humans, similar to the situation Stiner & Kuhn (2006) reported for several sites from Italy, in which ‘fast prey’ replaced ‘slow prey’.

### Disentangling an expansion process

We found an incipient differentiation process between North African and Spanish tortoises that is easily explained by the North African origin of the south-eastern Spanish populations (Álvarez et al., 2000; Harris et al., 2003; Fritz et al., 2009; Graciá et al., 2011; Salinas et al., 2011). Microsatellite cluster analyses revealed moderate levels of admixture across the Mediterranean Basin and, when the spatial component was taken into account, allowed geographical location of the potential source and entry point of the species to south-eastern Spain. The most probable founder area in North Africa corresponds to some coastal sampled sites located in the west of Algeria, while the most plausible arrival point in south-eastern Spain is located in the central-southern part of the studied area. These results were additionally supported by the fact that the lowest pairwise *F*~ST~ values across both sides of the Mediterranean Basin corresponded to sampled sites located within these two areas.

Besides the differentiation among these populations, important different genetic patterns were found within them. The six haplotypes present in south-eastern Spain appeared clearly spatially distributed, showing a north–south differentiation pattern through the studied area. This spatial structure is in strong agreement with that reported by Graciá et al. (2011) using microsatellites. However, the North African mitochondrial genetic structure seems more diffuse across the region with some common haplotypes widely spread over the whole subspecies’ range, some of them located at internal positions in the parsimony network (B12 and B13). This North African pattern could be explained by ancient contrac-

---

**Table 2** Inference of the plausible hypotheses for the arrival of the spur-thighed tortoise (*Testudo graeca*) to south-eastern Spain using Bayesian demographic data from **MSVAR 1.3** and the **odds** analyses in favour of the different periods of time (*T*; see Materials and Methods).

<table>
<thead>
<tr>
<th>Hypotheses for the species’ arrival</th>
<th>Inferred limit of the hypothesis</th>
<th><em>Odds</em> in favour of the <em>T</em> period</th>
<th>Superior limit of the hypothesis</th>
<th><em>Odds</em> in favour of the <em>T</em> period</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Transmarine dispersal before the LGP</td>
<td>South-eastern Spain</td>
<td>Galera</td>
<td><em>T</em> &gt; 125,000</td>
<td>0.05§</td>
</tr>
<tr>
<td>(ii) Transmarine dispersal during the LGP</td>
<td><em>T</em> ≤ 30,000</td>
<td>1.75</td>
<td>3.00*</td>
<td>0.82</td>
</tr>
<tr>
<td>(iii) Facilitated by humans during the LGP</td>
<td><em>T</em> ≤ 30,000</td>
<td>1.75</td>
<td>3.00*</td>
<td>0.82</td>
</tr>
<tr>
<td>(iv) Introduced by prehistoric seafarers</td>
<td><em>T</em> &gt; 12,500</td>
<td>19.29†</td>
<td>24.14†</td>
<td>0.04§</td>
</tr>
<tr>
<td>(v) Introduced in historical times</td>
<td><em>T</em> ≤ 100</td>
<td>0.00001§</td>
<td>0.0003§</td>
<td>–</td>
</tr>
</tbody>
</table>

LGP, last glacial period.

*Ratios greater than 3, which suggest positive evidence in favour of the *T* period.*

†Ratios greater than 7, considered significant in favour of the *T* period.

§Ratios below 0.14, considered significant in favour of the complementary *T* period.

---

in the western Mediterranean, and after the beginning of the last glaciation process (between approximately 3100 and 125,000 years ago). Our analyses for Spanish tortoises revealed a bottleneck approximately 20,000–30,000 years ago with a strong reduction of the ancestral effective population size, coinciding with one of the minimum Mediterranean sea level phases (Flemming et al., 2003; Massetti, 2009). These results basically keep open two different possibilities for the species’ arrival: natural transmarine dispersal, which has been reported for other species of tortoise (e.g. Caccione et al., 1999; Gerlach et al., 2006), or its introduction as a consequence of the first cultural expansions of humans from North Africa (late Palaeolithic or early Neolithic; Broodbank, 2006).

Unfortunately, the role that humans could have played in the species’ dispersal across the Mediterranean remains unclear. Although a strong connection between tortoises and primitive Mediterranean cultures due to consumption habits is well known (e.g. Blasco, 2008; Morales-Pérez & Sanchis-Serra, 2009), it remains unclear exactly when humans crossed the Strait of Gibraltar for the first time (Flemming et al., 2003; Broodbank, 2006). One of the main traditional arguments claiming the recent arrival of *T. g. graeca* to the Iberian Peninsula is the lack of fossils, in contrast with the abundant fossil remains of another Spanish tortoise species, *Testudo hermanni*. Nevertheless, in a recent work Morales-Pérez & Sanchis-Serra (2009) reported the absence of *Testudo* fossils in the Iberian Mediterranean region for the last 30,000 years, due to the species’ rarity or even its local extinction. The rarity of *T. hermanni* as a consequence of climatic amelioration occurring during this period, or even due to its possible overexploitation, could have led to a change in the consumption habits of humans (Morales-Pérez & Sanchis-Serra, 2009). Therefore, the absence of fossil remains of *T. graeca* could be explained as a consequence both of the low initial size of the population and of a change in prey selection of modern humans, similar to the situation Stiner & Kuhn (2006) reported for several sites from Italy, in which ‘fast prey’ replaced ‘slow prey’.

### Disentangling an expansion process

We found an incipient differentiation process between North African and Spanish tortoises that is easily explained by the North African origin of the south-eastern Spanish populations (Álvarez et al., 2000; Harris et al., 2003; Fritz et al., 2009; Graciá et al., 2011; Salinas et al., 2011). Microsatellite cluster analyses revealed moderate levels of admixture across the Mediterranean Basin and, when the spatial component was taken into account, allowed geographical location of the potential source and entry point of the species to south-eastern Spain. The most probable founder area in North Africa corresponds to some coastal sampled sites located in the west of Algeria, while the most plausible arrival point in south-eastern Spain is located in the central-southern part of the studied area. These results were additionally supported by the fact that the lowest pairwise *F*~ST~ values across both sides of the Mediterranean Basin corresponded to sampled sites located within these two areas.

Besides the differentiation among these populations, important different genetic patterns were found within them. The six haplotypes present in south-eastern Spain appeared clearly spatially distributed, showing a north–south differentiation pattern through the studied area. This spatial structure is in strong agreement with that reported by Graciá et al. (2011) using microsatellites. However, the North African mitochondrial genetic structure seems more diffuse across the region with some common haplotypes widely spread over the whole subspecies’ range, some of them located at internal positions in the parsimony network (B12 and B13). This North African pattern could be explained by ancient contrac-
tions and expansions of the species’ range and population following climatic changes (Fritz et al., 2009), and contrasts with the more structured mitochondrial pattern reported for the Spanish tortoises, which is probably a direct reflection of the colonization–expansion process. In any case, the finding of isolation-by-distance patterns within both ranges (identified from the microsatellite data) suggests that their present spatial structures are the result of a main differentiation–dispersal event across the space, which in the case of the south-eastern Spanish range suggests the natural expansion of the species within the newly colonized territory (as proposed by Graciá et al., 2011). This hypothesis would be additionally corroborated by the detection of a spatial negative gradient in heterozygosities from the hypothetical arrival area in south-eastern Spain. However, the possibility of different origins for Iberian populations at different sites cannot be discarded without analysing in a landscape context the effect of geographical barriers on the genetic configuration of the populations. Such a methodology could clarify whether the exclusive haplotypes found in Marinica are the result of a particular colonization history or, alternatively, are the consequence of isolation and differentiation of the site.

Understanding recent Mediterranean range expansions

In this work we have used T. g. graeca as an example of a group of organisms from the Iberian Peninsula that show minimal mtDNA differentiation from North African populations, whose estimated dates of divergence are generally less than 100,000 years and which have complex and unresolved expansion processes. Recent genetic studies corroborated the historical introduction of some such species to the Iberian Peninsula, such as the Barbary macaque, Macaca sylvanus (Modolo et al., 2005), or the common genet, Genetta genetta (Gaubert et al., 2010); or contrarily suggested an ancient origin, as in the case of the white-toothed shrew, Crocidura russula (Cosson et al., 2005), the Egyptian mongoose, Herpestes ichneumon (Gaubert et al., 2010), or the snakes Malpolon monspessulanus and Hemorrhoides hippocrepis (Carranza et al., 2006). For other species, including the Mediterranean chameleon, Chamaeleo chamaeleon (Paulo et al., 2002), and the treefrog Hyla meridionalis (Recuero et al., 2007), mtDNA genetic studies have failed to distinguish between historical and prehistoric origins of the Iberian populations. Our results highlight the importance of developing an accurate methodological approach to disentangle the particular expansion processes of each species. In fact, our assessment of T. g. graeca indicates that, even with extensive sampling and the use of microsatellites, it may be difficult or even impossible to ascertain with confidence whether such species naturally crossed during low sea level stands or were introduced by humans. In this sense, the range expansions of these species are at the limits of resolution of the tools available and constitute a current knowledge frontier.

It is clear, however, that the Late Pleistocene was an important period of faunal interchange between North Africa and Europe and this framework should be taken into account from a species conservation point of view. Much caution should be taken when labelling a species as historically introduced because, as is shown in this work, disentangling the origin of some species can be a challenging task. The consideration of the species as an historical introduction may modify the social perception of the species (Pérez et al., 2011), and could predispose to their conservation value being underestimated. Furthermore, given that conservation plans usually focus only on so-called native species, legal protection for the spur-thighed tortoise could be reduced (Colautti & McIaasac, 2004). Assuming that sometimes it is impossible to distinguish using current methods between the native or human-mediated origin of some populations due to their complex biogeographical histories, the precautionary principle should therefore apply.

ACKNOWLEDGEMENTS

Financial support was granted by the Spanish Ministry of Education and Science (project CGL2009-08251). Eva Graciá was funded by a grant from the Regional Government of the Community of Valencia (BFPI/2007/294), and this article is part of her PhD. We thank the Global Nature and Castillo de Chuecos foundations for allowing us to work in their reserves. Further thanks go to Santiago García-Martínez, for his help with laboratory work, to M. Asunción Martínez-Mayoral and Javier Morales, for their help with the Bayesian data, and to Marcos Ferrández, Said Larbes and Rachid Rouag, who helped us to obtain our North African samples. We appreciate the interest of the Regional Governments of Murcia and Andalucia in the conservation of this tortoise.

REFERENCES


**SUPPORTING INFORMATION**

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

Appendix S1 Geographical and genetic information by site.
Appendix S2 Laboratory procedures.
Appendix S3 Detailed demographic inference for the arrival of *Testudo graeca* in south-eastern Spain.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Eva Graciá is a PhD student at the Department of Applied Biology, Ecology Area, of the Miguel Hernández University, Spain, and is interested in population genetics, conservation genetics and the study of genetic and phenotypic patterns of recent range expanded species. Her PhD thesis focuses on these issues, using *Testudo graeca* in south-eastern Spain as a model.

Author contributions: E.G., A.G., J.D.A. and F.B. conceived the ideas and collected samples; E.G., J.H. and U.F. processed the samples; E.G. analysed the data; and E.G., F.B and A.G. led the writing. All authors commented on and revised earlier manuscript versions.