Phylogeography and demographic history of Shaw’s Jird (Meriones shawii complex) in North Africa

AUDE LALIS1*, RAPHAEL LEBLOIS2, EMMANUELLE STOETZEL1,3, TOURIA BENAZZOU4, KARIM SOUTTOU5, CHRISTIANE DENYS1 and VIOLAINE NICOLAS1

1Institut de Systématicque, Evolution, Biodiversité, ISYEB-UMR 7205 CNRS, MNHN, UPMC, EPHE, Muséum national d’Histoire Naturelle, Sorbonne Universités, 57 rue Cuvier, CP 51, 75005, Paris, France
2INRA-UMR1062 CBGP, F-34988 Montferrier-sur-Lez, France
3Histoire Naturelle de l’Homme Préhistorique, HNHP-UMR 7194 CNRS, Muséum national d’Histoire naturelle, Département de Préhistoire, Sorbonne Universités, Musée de l’Homme, Palais de Chaillot, 17 place du Trocadéro, 75016, Paris, France
4Département de Biologie, Faculté des Sciences, BP1014, Rabat, Morocco
5Laboratoire d’Ornithologie, Département de Zoologie, Institut d’Agronomie, Hacen badi 16200, El Harrach, Alger, Algeria

Received 4 August 2015; revised 7 October 2015; accepted for publication 7 October 2015

Palaeoenvironmental data and climatic reconstructions show that the Mediterranean ecoregion of North Africa underwent drastic ecological changes during the Pleistocene. Given its rich palaeontological record, North Africa is a pertinent region for documenting the role of climate change and human mediated-habitat changes on the demography and genetic structure of faunal species. In the present study, we collected data from this species in Morocco, Algeria, and Tunisia, and we combined molecular (mitochondrial and nuclear DNA sequences, microsatellites), fossil, palaeoenvironmental, and human context data to propose an explanation for the fluctuations of populations belonging to the Meriones shawii complex in the past. Genetic and fossil data both indicate a strong bottleneck in Moroccan populations at the Middle Holocene (last interglacial optimum) compared to the Late Pleistocene. Our mitochondrial DNA data suggest a diversification event within Morocco corresponding to the 130–125 kya interglacial optimum. Given that (1) major demographic changes in the M. shawii complex coincide with the interglacial optums, and (2) the impact of human activities on the landscape and faunal communities was moderate during the Middle Holocene (beginnings of the Neolithic culture), our results demonstrate that climate, rather than anthropogenic influences, likely explains the M. shawii complex population decline in the Holocene. © 2015 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 262–279.


INTRODUCTION

The Mediterranean ecoregion of North-western Africa is the biogeographical area extending westerly to the Atlantic Ocean, north to the Mediterranean Sea, and east and south to the Sahara desert. Palaeoenvironmental data and climatic reconstructions show that this region underwent drastic ecological changes during the Pleistocene, mainly related to the alternation of expansion/reduction in the size of the Saharan desert, reduction/expansion of the Mediterranean vegetation, and reduction/development of water ponds, lakes, and rivers (Brun, 1989, 1991; Hooghiemstra et al., 1992; Jolly et al., 1998; Trauth, Larrasoana & Mudelsee, 2009; Stoetzel et al., 2011). These environmental changes could have led to allopatric differentiation in separate refuges for many species. However, the impacts of climate change on the phylogeographical patterns of species remain poorly understood in the region.

*Corresponding author. E-mail: lalis@mnhn.fr
Rodents are good models for historical reconstructions of environment influences on biota (Wang et al., 2013). This is a result of their short generation time, rapid mitochondrial (mt)DNA substitution rate, relatively limited dispersal ability, and strong associations with particular habitats. This leads to informative contemporary patterns of genetic variation (Tolley et al., 2006; Fedorov et al., 2008; Nicolas et al., 2008; Bryja et al., 2010; Russo, Chimimba & Bloomer, 2010; Wang et al., 2013; Boratynski et al., 2014). The Jird Meriones is an appropriate model for investigating the roles of humans and climate change in shaping faunal genetic diversity and distribution. The Meriones shawii complex belongs to the Muridae family and Gerbillinae subfamily. This complex is widely distributed from Morocco to Algeria, Tunisia, Libya, and Egypt, and along the west side of the Nile (Aulagnier et al., 2008 in IUCN 2012). It is not found in mesic environment, such as forests, grasslands, wetlands, lakes, and rivers (Aulagnier et al., 2008 in IUCN 2012), and it avoids rocky basins. The M. shawii complex remains poorly known, according to its systematics, ecology, and geographical distribution. Based on external and cranial measurements, some studies recognize two valid subspecies (Petter, 1961; Aulagnier & Thevenot, 1986), with the question remaining unresolved.

The Meriones genus originated in North Africa at the end of the Middle Pleistocene, and has remained well represented in fossil assemblages subsequent to the beginning of the Late Pleistocene (approximately 130 kya) (Ouahbi, Aberkan & Serre, 2003; Reed & Barr, 2010; Stoetzel et al., 2010; Lopez-Garcia, Agusti & Aouraghe, 2013; Stoetzel, 2013). All fossil Meriones remains of ‘modern’ morphology are assigned to M. shawii, without any mention of M. grandis. Exceptional fossil records covering the last 120 000 years are available from the El Harhoura 2 cave, in the region of Temara on the north-Atlantic coast of Morocco, a few kilometres to the south of Rabat (Stoetzel et al., 2010, 2011, 2014). This site has recorded the entire last climatic cycle along 11 archaeological levels, from approximately 120 kya to 5.8 kya (Jacobs et al., 2012; Stoetzel et al., 2014). All Meriones remains belong to the M. shawii complex, and represent the most abundant small vertebrate taxon in the El Harhoura 2 assemblage, whereas it is now poorly represented in the septentrional Atlantic plains, including the Rabat-Témara region (Aulagnier & Thevenot, 1986; Aulagnier, 1992). The human context in which this settlement occurred is well known: several cultures succeeded during Late Pleistocene and Holocene (Nespoulet et al., 2008; El et al., 2012; Stoetzel et al., 2014). Moreover, the palaeoenvironmental context is also well known: palaeoecological analysis has shown an alternation of arid and more humid periods during the Late Pleistocene, ending with a humid period during the Middle Holocene, corresponding to the last climatic optimum (Stoetzel et al., 2011, 2014). These environmental changes are accompanied by differences in the relative proportion of small vertebrate species between the El Harhoura 2 levels. Despite these palaeoenvironmental changes, Meriones are dominant in relatively stable proportions all along the Late Pleistocene record. However, a significant decrease in the proportion of Meriones is observed in the assemblage of level 1 (Holocene: approximately 5.8 kya BP) compared to Pleistocene levels (Stoetzel et al., 2011). It is important to note that the taphonomic study has shown no or only very low biases or perturbations, indicating that the relative proportion of species along the archaeological sequence is neither related to the type of predators at the origin of the fossil accumulations, nor other taphonomic agents (Stoetzel et al., 2011).

In the present study, we combined mitochondrial and nuclear data to assess the taxonomic status of the M. shawii complex, and also whether significant demographic changes occurred during the last...
120,000 years in this taxa. The combination of molecular, fossil, palaeoenvironmental, and human context data should allow us to propose a scenario to explain the historic fluctuations of populations within this taxon in North Africa.

**MATERIAL AND METHODS**

**SAMPLING AND LOCALITIES**

Samples of *M. shawii* complex were collected between January 2010 and April 2012 in Morocco and Algeria (Fig. 1; see also Supporting information, Table S1) with permission from the ‘Haut Commissariat aux Eaux et forêts et à la Lutte contre la désertification’ (autorization no. 15 HCEFLCD/DLCDPN/DPRN/CF) in Morocco and the Ministry of Forestry in Algeria. Animals were live-captured using Sherman traps and handled in accordance with the guidelines of the American Society of Mammalogists (http://www.mammalogy.org/committees/index.asp; Animal Care and Use Committee, 2011) and also in accordance with standard procedures for BSL3 work in the field (Federal Guidelines for Field Work, CDC 1997). All manipulations of animals were made in Morocco in agreement with the global law 11-03 relative to the protection and the development of the environment. Alive animals were euthanized by the injection of a lethal dose of isofluorane, followed by cervical dislocation. The protocol was approved by Comité Cuvier (permission no. 68.009). Moreover, four specimens from Tunisia housed in the collections of the National Museum of Natural history of Paris (France) were included in the present study (MNHN). In total, we used 178 samples representing 14 localities (details of sampling localities are provided in Table 1). Total genomic DNA was extracted from ethanol-fixed intercostal muscle tissues using NucleoSpin® 96 Tissues (Macherey-Nagel) in accordance with the manufacturer’s instructions. The carcasses were fixed in formalin for later preparation as skin and skull specimens. Specimens from Morocco are temporarily housed at the MNHN and will be deposited at the Institut Scientifique de Rabat (ISR, Morocco). Specimens from Algeria are housed at the Institut of Agronomy of El Harrach (Alger, Algeria).

**DNA AMPLIFICATION SEQUENCING AND GENOTYPING**

We amplified and sequenced 178 individuals (1–33 per population) (Table 1) for the cytochrome *b* gene (*cytb*) fragment of the mitochondrial DNA (mtDNA) genome using primers L14723 (Ducroz, Volobouev & Granjon, 2001) and H6 (Mongelard *et al.*, 2002). Intron 7 of the nuclear fragment, the $\beta$-fibrinogen

---

**Figure 1.** Map of the sampled geographical localities.
Table 1. *Meriones shawii* sampling locality details and genetic diversity showing the number of individuals analyzed (*N*), averaged values of observed heterozygosity (*H*<sub>O</sub>) and expected heterozygosity (*H*<sub>E</sub>), mean number of alleles (*N*<sub>A</sub>), allelic richness (A) within-population coefficient of inbreeding (FIS), nucleotide diversity (Pi), and haplotypic diversity (*H*<sub>D</sub>)

<table>
<thead>
<tr>
<th>Locality</th>
<th>Code</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Microsatellite diversity</th>
<th>mtDNA diversity: cytb</th>
<th>nDNA diversity: BFIBR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>N</em></td>
<td><em>H</em>&lt;sub&gt;O&lt;/sub&gt;</td>
<td><em>H</em>&lt;sub&gt;E&lt;/sub&gt;</td>
</tr>
<tr>
<td>Morocco</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BenGuerir</td>
<td>BG</td>
<td>32.19</td>
<td>-7.97</td>
<td>29</td>
<td>0.488</td>
<td>0.462</td>
</tr>
<tr>
<td>Guenfouda</td>
<td>GF</td>
<td>MM</td>
<td>-2.01</td>
<td>18</td>
<td>0.528</td>
<td>0.556</td>
</tr>
<tr>
<td>Ifrane</td>
<td>IF</td>
<td>33.42</td>
<td>-5.04</td>
<td>11</td>
<td>0.518</td>
<td>0.493</td>
</tr>
<tr>
<td>Sour El Az</td>
<td>$EA$</td>
<td>31.84</td>
<td>-7.01</td>
<td>33</td>
<td>0.509</td>
<td>0.464</td>
</tr>
<tr>
<td>Ouled Boughadi</td>
<td>OB</td>
<td>33.07</td>
<td>-6.72</td>
<td>33</td>
<td>0.355</td>
<td>0.411</td>
</tr>
<tr>
<td>Aglou</td>
<td>AG</td>
<td>29.83</td>
<td>-9.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 km $$$Essaoula</td>
<td>E$</td>
<td>31.32</td>
<td>-9.71</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 km $N$Manakech</td>
<td>MA</td>
<td>31.83</td>
<td>-7.97</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qued Souss</td>
<td>OS</td>
<td>3040</td>
<td>-8.92</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sidl El Moctar</td>
<td>$EM$</td>
<td>31.53</td>
<td>-9.00</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algeria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSila</td>
<td>MS</td>
<td>35.7</td>
<td>4.5</td>
<td>13</td>
<td>0.43</td>
<td>0.558</td>
</tr>
<tr>
<td>Hadjeb El Djmel</td>
<td>HEA</td>
<td>34.3</td>
<td>2.97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garat An Njila</td>
<td>GAN</td>
<td>35.07</td>
<td>9.6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sidi Bouzid</td>
<td>SB</td>
<td>35.03</td>
<td>9.5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total/mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

mtDNA, mitochondrial DNA.
(BFIBR) gene, was also amplified and sequenced for 49 specimens representing all mtDNA groups (Table 1) using primers BFIBR1 and BFIBR2 (Seddon et al., 2001). The polymerase chain reaction (PCR) consisted of 35 cycles of 30 s at 94 °C, 40 s at 50 °C, and 90 s at 72 °C. Double-stranded PCR products were purified and sequenced at the Genoscope (Ivry/Seine, France). Chromatograms were checked and sequences were manually corrected in CODEALIGNER, version 3.5.6. Sequences of 1062 and 732 bp were obtained in the final analyses for the cytb and BFIBR genes, respectively. Newly obtained sequences were submitted to GenBank (accession numbers: KM581449 to KM581497 for the BFIBR data; KM58150 to KM581674 for the cytb data). We genotyped 137 individuals from six populations for 10 microsatellite loci (Table 1) using primers and genotyping protocols described in Lalis & Lambourdière (2014).

mtDNA and nDNA: Phylogenetic reconstruction and divergence time estimates

We constructed two median-joining haplotype networks for the cytb and BFIBR datasets using NETWORK, version 4.500 (Bandelt, Forster & Rohl, 1999). Prior to this analysis, the existence of heterozygous positions for the nuclear gene fragment was investigated in accordance with the procedure described in Nicolas et al. (2012). The results obtained from four individuals (MA97, MA127, MA848, MA853) were unclear (i.e. several possible haplotype pairs for each individual with similar probabilities). They were thus removed from all subsequent analyses. The Bayesian Markov chain Monte Carlo (MCMC) approach implemented in BEAST, version 1.8 (Drummond & Rambaut, 2007) was employed to simultaneously estimate phylogenetic relationships and time to the most recent common ancestor (TM RCA). Divergence times and their credibility intervals were estimated using a relaxed clock model with branch rates drawn from an uncorrelated log-normal prior distribution. We used the model of sequence evolution retained by MRMODELTEST, version 3.04 (Nylander, 2004) and a coalescent model with varying population sizes (skyline model). Two independent runs of 100 million iterations with burn-ins of 25% were performed. The results were inspected visually using TRACER, version 1.8 to ensure proper mixing of the MCMC. A consensus chronogram with median age estimates and 95% higher posterior density intervals was generated and visualized with TREEANNOTATOR (BEAST package, version 1.8) and FIGTREE, version 1.3.1 (Rambaut, 2009). Four fossil calibrations were used to calibrate the chronogram. All calibrations were applied as log-normal prior distributions, and the means and SDs of these distributions were chosen to construct 95% confidence intervals (CIs) spanning the 90–95% Marshal indices (Marshall, 1994) reported by the Paleobiology Database (PDB) (Jaeger, Tong & Denys, 1986; PDB 2011). These represent the estimated 95% CIs for the actual origination of a taxon based on first occurrences and stratigraphic sampling. The two first calibrations applied in the present study have been used in previous studies (Schenk, Rowe & Steppan, 2013) and the other two are derived from the PDB: (1) The Gerbillinae-Deomyinae split based on the first occurrence of Gerbillinae in the Lower Miocene fauna of Saudi Arabia (offset = 15.868, range = 16.000–23.700); (2) the Lophuromys-Acomys-Deomys split, hence the origin of Acomys, based on the earliest known Acomys fossil from Kenya (offset = 5.258, range = 5.3–29.050); (3) the Meriones-Psammomys-Rhombomys split, hence the origin of Meriones, based on the earliest known Meriones fossil from Kazakhstan in the Pliocene (offset = 2.6, range = 2.6–5.3); and (4) the Gerbillus-Sekeetamys split based on the first occurrence of the genus Gerbillus in Armenia in the Pliocene (offset = 2.6; range = 2.6–5.3). The species used as outgroup were: Mus musculus (Genbank AB819920), Acomys airesiens (AJ012021), Acomys cahirinus (AJ233953), Acomys chudeaui (FJ415538), Acomys ciliiceps (AJ233957), Acomys dimidiatus (AJ233959), Acomys ignitus (Z96064), Acomys johannis (HM635823), Acomys minous (GU046553), Acomys nesiotes (AJ233952), Acomys percivali (EF187818), Acomys russatus (FJ415485), Acomys spinossissimus (AM409396), Acomys subspinosus (JN247673), Acomys wilsoni (EF187799), Deomys ferrugineus (FJ415478), Lophuromys flavopunctatus (EU349754), L. sikapusi (AJ012023), Desmodillus auricularis (AJ851272), Gerbilliscus robustus (AM409374), Gerbilliscus guinea (AJ430562), Gerbillurus paeba (AJ430557), Gerbillurus tytonis (AJ430559), Sekeetamys calurus (AJ851276), Gerbillurus cumestriss (AJ851271), Gerbilliscus gerbillus (AJ851269), Gerbilliscus henleyi (JQ753050), Gerbilliscus nanus (JQ753051), Gerbilliscus poecilos (JQ753064), Gerbilliscus simoni (GU356577), Gerbilliscus tarabuli (GU356573), Desmodillus brueneri (AJ851273), Taterillus arenarius (AJ851261), Psammomys obesus (AJ851275), Rhombomys opimus (AJ430556), Meriones vinogradovi (VV1989001), Meriones chengi (AB381900), Meriones crassus (AJ851267), Meriones libycus (JQ927411), Meriones meridianus (AJ851268), Meriones rex (AJ851265), Meriones unguiculatus (AF119264), Meriones tristrami (KU189331), and Meriones persicus (KT949958). These phylogenetic and divergence
time analyses were only performed on the cytB dataset as a result of the lack of outgroup sequences for the BFIBR gene.

GEOMETRIC DIVERSITY AND POPULATION STRUCTURE

For sequence data, the nucleotide diversity, haplotype diversity, and mean number of nucleotide differences were calculated using DNASP, version 5.0 (Librado & Rozas, 2009). For microsatellite loci, the number of alleles \( N_A \) and the expected and observed heterozygosities \( (H_E \) and \( H_O \)) were calculated using the R package ADEGENET, version 1.2-7 (Jombart, 2008). Allelic richness \( (R) \) was calculated using FSTAT. All loci were tested for Hardy–Weinberg equilibrium (HWE) and the presence of linkage disequilibrium using GENEPOP, version 4.1.3 (Rousset, 2008). MICRO-CHECKER, version 2.2.1 (Van Oosterhout et al., 2004) was used to test for the presence of null alleles, large allele dropouts, and scoring errors. For the microsatellite data, we employed clustering analyses to describe the genetic structure of our sample and to estimate the most likely number of genetically homogeneous clusters \( (K) \) using STRUCTURE, version 2.3.3 (Pritchard, Stephens & Donnelly, 2000; Falush, 2003). STRUCTURE was run using the admixture model, with localities as priors and assuming correlated allelic frequencies. Final outputs were obtained for 20 independent runs, testing \( K = 1 \) to \( K = 6 \), each with a total of 250 000 iterations and a burn-in of 150 000. The number of contributing populations was tested using the ad-hoc Evanno statistic Delta \( K \) (Evanno, Regnaut & Goudet, 2005). Membership probabilities (i.e. Q-values) of the 20 runs for \( K = 2 \) and \( 3 \) were averaged using CLUMPP, version 1.2 (Jakobsson & Rosenberg, 2007) and displayed using DISTRUXT, version 1.1 (Rosenberg, 2004). Population differentiation was inferred from the mtDNA and microsatellite data sets using \( F_{ST} \) estimates. Values for mtDNA data were computed with ARLEQUIN, version 3.5.1 (Excoffier & Lischer, 2010) using populations with over 12 individuals. Population differentiation was also inferred from the microsatellite data set by comparing \( F_{ST} \) estimates (Weir & Cockerham, 1984) between all population pairs and among all populations, at each locus and over all loci, as computed by FSTAT (Goudet, 1995; Goudet et al., 1996) and GENEPOP, version 4.1.3. We tested whether \( F_{ST} \) estimates were significantly > 0 by permuting multilocus genotypes among samples with FSTAT. Genotypic differentiation between populations was investigated by exact tests using Markov chain algorithms implemented in GENEPOP, version 4.1.3. (Rousset, 2008). Pairwise \( F_{ST} \) values were used to build Neighbour-joining trees using POPULATIONS, version 1.2.30b (http://bioinformatics.org/~tryphon/populations/).

We also used spatial analysis of shared alleles (SaShA) using SaShA, version 2.0 (http://sasha.stanford.edu) to test for genetic subdivision across the geographical space for cytB and BFIBR datasets. This analysis uses spatial and haplotypic information to detect nonrandom allele distribution against an expectation of panmixia (Kelly et al., 2010). The test statistic describes the observed mean distance between alleles (OM). When OM is less than the expected mean (EM), alleles are considered to be aggregated. When OM is larger than EM, alleles are considered randomly distributed. A jackknife procedure identifies which alleles are strongly influencing the observed distribution. Isolation-by-distance patterns were tested with a Mantel’s test (Mantel, 1967) with 30 000 permutations. For sequence data, the test was performed using ARLEQUIN, version 3.5.1 (Excoffier & Lischer, 2010) on the relationship between the mean number of pairwise nucleotide differences and geographical distances between sampling localities. For microsatellite data, the test was performed using GENEPOP, version 4.1.3, by regressing \( F_{ST} (1 – F_{ST}) \) between populations over the logarithm of geographical distances (Rousset, 1997).

DEMOGRAPHIC HISTORY

Demographic history was explored using the MIGRAINE (http://kimura.univ-montp2.fr/~rousset/Migraine.htm) and the model with historic variations in population size (Leblois et al., 2014). This model consists of a single, isolated panmictic population that undergoes continuous exponential change in population size starting at time \( T \) generations in the past and continuing until the moment of sampling (i.e. present). MIGRAINE uses the class of importance sampling algorithms developed by de Iorio & Griffiths (2004a, b), de Iorio et al. (2005) and extended in Leblois et al. (2014). MIGRAINE was used to test for past change in population size and to estimate current and ancestral scaled population size (\( \theta = 2N \mu \) and \( \theta_{anc} = 2N_{anc} \mu \), where \( N \) and \( N_{anc} \) are the current and ancestral haploid population size, respectively, and \( \mu \) is the mutation rate per generation of the whole locus) and \( D \), the time when the demographic change starts, scaled by population size (i.e. \( D = T/2N \)). MIGRAINE was first applied on the mitochondrial and nuclear sequence data sets separately, and only for the Moroccan samples. Prior to this analysis, alignments of nuclear genes were pruned to exclude stretches with missing data at the beginning and the end of some sequences. Also, because MIGRAINE is based in the infinitely many
sites model (ISM) for analysis of sequence data, different datasets were produced for both the mtDNA cytb region and for the nuclear BFIBR region to fit this model. There are two reasons why a sequence data set may not fit the ISM: sites can show more than two nucleotidic states or pairwise comparisons of sites may not comply to the four gamete test (Hudson & Kaplan, 1985). For one data set, we chose to systematically remove incompatible sites for all individuals. For the second, we removed haplotypes with incompatible sites. For the MIGRAINE analysis, mtDNA and nuclear DNA (nDNA) sequence data were pooled for all individuals from Moroccan populations (136 individuals). This initial data set contained 138 and 26 segregating sites, with 42 and 12 unique haplotypes in 136 and 76 individuals, for cytb and BFIBR respectively. Fitting the data sets to the ISM resulted in four modified data sets (Table 3) distinguished by the remaining number of sites, haplotypes, and individuals. All runs with MIGRAINE consisted of 1 000 000 trees, 2400 points, and two iterations. The MIGRAINE software was also used on the microsatellite data set to infer past changes in population size on the pooled Moroccan dataset. A benefit of using MIGRAINE over MSVAR (Beaumont, 1999; Storz & Beaumont, 2002), which does similar analyses, is that it allows departure from the strict stepwise mutation model by use of a generalized stepwise mutation model. All MIGRAINE analyses for microsatellites were only run on the Moroccan samples and used 20 000–200 000 trees, 2400 points, and three iterations (see the MIGRAINE manual for details concerning these settings). To convert scaled population sizes (θ and θanc) into effective population sizes (N and Nanc), and scaled times (D) into times in years (T), we used: (1) a generation time of 1 year; (2) mutation rates of $5 \times 10^{-4}$ mutations per locus per generation for all microsatellite loci (Dib et al., 1996; Ellegren, 2000; Sun et al., 2012) and $10^{-7}$ and $10^{-8}$ mutations per site per generation for the mitochondrial and nuclear locus, respectively. Those values are similar to those found in our BEAST analyses, and also similar to other values found in the literature for rodents (Gündüz et al., 2005; Nabholz, Glemin & Galtier, 2008); (3) the length in pb of the two sequenced loci; and (4) a correction factor for the loss of polymorphic sites that do not fit the ISM for the sequenced loci.

RESULTS

mtDNA and nDNA Phylogeography: Sequence Variation and Population Structure

According to our phylogetic (GTR + I + G model of evolution) and network analyses, three mtDNA clades can be identified within the M. shawii complex (Figs 2, 3): clade A is present in all sampled Moroccan localities, except Oued Souss (OS) and Aglou (AG); clade B is present in five of the 10 sampled Moroccan localities: Ben Guerir (BG), Ouled Boughadi (OB), Sidi El Moctar (SEM), Oued Souss (OS), and Aglou (AG); and clade C is present in Algeria (HEA, MS) and Tunisia (GAN, SB), and contains one specimen from eastern Morocco (MA234, locality of Guenfouda, represented by a red star in Fig. 2). The median TMRCA of each clade varies from 81–120 kya, with large CIs (Fig. 2). The two Moroccan clades (A and B) form a monophyletic group, with a TMRCA of 0.414 Mya (0.240–0.617 Mya). The TMRCA of M. shawii is estimated at 1.361 Mya (0.878–1.873 Mya). Clades A and B differ by 2.8% of sequence divergence (K2P distance). Clade C differs by 7.7% and 7.8% of sequence divergence from clades A and B, respectively. Within-clade haplotype diversity varies between 0.910 ± 0.956 for clade B and 0.950 ± 0.016 for clade C. Within-clade nucleotide diversity varies between 0.00385 ± 0.00031 for clade C and 0.00683 ± 0.00103 for clade B. The F-statistic analysis yields a highly significant level of population structure among all localities ($F_{ST} = 0.71, P < 0.001$). The $F_{ST}$ value between Morocco and Algeria is high and significantly different from 0 ($F_{ST} = 0.90, P < 0.001$). Pairwise $F_{ST}$ values between Moroccan and Algerian populations are also high ($0.96 > F_{ST} > 0.87$) and significantly different from 0 (Table 2). Pairwise $F_{ST}$ values between Moroccan populations are lower ($0.49 < F_{ST} < 0.17$) but still significantly different from 0. Lastly, the $F_{ST}$ value between the two Algerian populations is low and nonsignificant ($F_{ST} = 0.04, P > 0.05$). The SASHA analysis shows that M. shawii complex haplotypes are significantly aggregated (OM = 13 km, expected = 508 km; $P = 0.001$). This is also true when only Moroccan specimens of either clade A + B (OM = 7 km, expected = 236 km; $P = 0.001$) or clade A (OM = 7 km, expected = 225 km; $P = 0.001$) are considered. For the three analyses, the haplotype-by-haplotype analysis shows that all common haplotypes are significantly aggregated and the jackknife analyses indicate the robustness of the overall results, which remain qualitatively the same when any single haplotype is removed. When all specimens are considered, a significant positive relationship between geographical and genetic distances is found (Mantel test, $P < 0.001$, slope of 0.736). However, no significant correlation between geographical and genetic distances is observed within clade A ($P = 0.071$, slope of 0.451), which is the clade with the highest sample size and the best geographical coverage.
Figure 2. Time calibrated phylogeny inferred using BEAST, depicting phylogenetic relationships within the species *Meriones shawii*. Numbers at nodes represent clade posterior probabilities. To improve clarity, outgroup taxa are not shown.
Concerning the BFIBR, the alignment of the 90 phased haplotypes required the addition of 11 gaps. According to our network analyses, two main groups of haplotypes differing by 19 mutations can be identified (Fig. 3): clade 1 groups all but two Moroccan individuals; clade 2 groups all Algerian and Tunisian individuals, as well as two individuals from the Moroccan locality of Guenfouda (MA205 and MA210). The mean percentage of nucleotide differences between clades 1 and 2 is 1.9%. For a given specimen the two phased haplotypes always cluster within the same clade. There is congruence between the cytb and BFIBR data: the cytb monophyletic group A + B corresponds to BFIBR clade 1, and cytb clade C corresponds to BFIBR clade 2. The only exception concerns specimens from Guenfouda: two specimens of clade A cluster within clade 2 and one specimen of clade C clusters within clade 1.

Figure 3. Minimum spanning network of Meriones shawii cytb [mitochondrial (mt)DNA, left] and BFIBR (nuclear DNA, right) haplotypes. Circle sizes are proportional to the number of similar haplotypes observed in the data set. Branch lengths are proportional to the number of mutations between haplotypes (except between clades A and C for the mtDNA dataset). Colours represent the country of sampling (red = Morocco, blue = Algeria, green = Tunisia). Two- or three-letter codes refer to the locality of sampling (for localities codes, see Table 1).
Table 2. Mitochondrial DNA sequence (above diagonal) and microsatellite (below diagonal) pairwise $F_{ST}$ statistics between *Meriones shawii* populations. ($P < 0.001$)

<table>
<thead>
<tr>
<th>Population</th>
<th>BG</th>
<th>GF</th>
<th>IF</th>
<th>SEA</th>
<th>OB</th>
<th>MS</th>
<th>HEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG</td>
<td>0.3309*</td>
<td>0.2153*</td>
<td>0.2233*</td>
<td>0.2634*</td>
<td>0.8713*</td>
<td>0.8803*</td>
<td></td>
</tr>
<tr>
<td>GF</td>
<td>0.0549</td>
<td>0.1135*</td>
<td>0.4932*</td>
<td>0.4711*</td>
<td>0.9139*</td>
<td>0.9131*</td>
<td></td>
</tr>
<tr>
<td>IF</td>
<td>0.0966</td>
<td>0.1178*</td>
<td>0.2983*</td>
<td>0.2344*</td>
<td>0.9509*</td>
<td>0.9399*</td>
<td></td>
</tr>
<tr>
<td>SEA</td>
<td>0.0389</td>
<td>0.0670</td>
<td>0.1678*</td>
<td>0.9625*</td>
<td>0.9535*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OB</td>
<td>0.0675</td>
<td>0.1113*</td>
<td>0.1542*</td>
<td>0.038</td>
<td>0.9514*</td>
<td>0.9445*</td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>0.1877*</td>
<td>0.1767*</td>
<td>0.2420*</td>
<td>0.1976*</td>
<td>0.2050*</td>
<td>0.0401*</td>
<td></td>
</tr>
</tbody>
</table>

*Significant values.

Haplotype diversity is $0.814 \pm 0.027$ for clade 1 and $0.922 \pm 0.039$ for clade 2. Within-clade nucleotide diversity is $0.00345 \pm 0.00022$ for clade 1 and $0.00961 \pm 0.00113$ for clade 2. The SASHA analysis show that *M. shawii* complex *BFIBR* haplotypes are significantly aggregated (OM = 227 km, expected = 494 km; $P = 0.001$). However, when only Moroccan haplotypes of clade 1 are considered, a random haplotype distribution is observed (OM = 228 km, expected = 231 km; $P = 0.857$). As a result of low sample sizes, $F_{ST}$ calculations and Mantel tests were not performed on the *BFIBR* data set.

**MICROSATELLITES: GENETIC VARIATION AND POPULATION STRUCTURE**

The analysis of 137 *M. shawii* complex genotypes reveals a variable degree of polymorphism across the six locations studied. A total of 249 alleles are detected across all loci, ranging from a maximum of nine alleles detected in population OB for locus MS-10 to a minimum of one allele detected in populations GF and IF for locus MS-3, MS and OB for locus MS-2, and OB for locus MS-4 (see Supporting information, Table S2). Considering the 10 loci together, the population from OB shows the highest allele mean number (4.6), whereas the IF population has the lowest allele mean number (3.8). Observed heterozygosities range from 0.355 (OB) to 0.528 (GF), whereas the expected heterozygosities range from 0.411 (OB) to 0.558 (MS) (Table 1). No significant linkage disequilibrium is detected after adjusting for multiple comparisons. Significant deviations from HWE are observed in 18 out of 60 locus/population combinations (see Supporting information, Table S2). Performing the HWE global test, heterozygote deficiencies are observed in four out of six populations. The presence of non-amplifying, null alleles at microsatellites loci is usually cited to explain deviation from HWE. For this reason, allele frequencies are analysed by MICRO-CHECKER, version 2.2.1 (Van Oosterhout *et al.*, 2004) and null alleles are inferred for two loci (MS-7 and MS-9). The data set is corrected accordingly and HWE is tested again. The result based on this new data set does not differ from the results of the original one (not shown) because HWE and heterozygous deficiencies remained identical, suggesting that other factors (e.g. capture of sibling individuals) may also play a role in reducing heterozygosity.

The highest posterior probability of the model given by STRUCTURE is obtained for $K = 4$, and the highest posterior probability of Delta $K$ is obtained for $K = 2$ (Fig. 4). One cluster is comprised of the Algerian population, whereas all of the Moroccan populations are part of the second cluster. Some Moroccan individuals show a high proportion of their genome belonging to the Algerian cluster; however, they appear to belong to the Moroccan clusters for $K = 3$ and for $K = 4$ (Fig. 4). The $F_{ST}$ analysis between Moroccan and Algerian populations is highly significant ($F_{ST} = 0.19, P < 0.001$). Indeed, pairwise $F_{ST}$ values between Moroccan and Algerian populations are non-negligable ($0.17 < F_{ST} < 0.24$) and the differentiation is also always significant (Table 2). By contrast, pairwise $F_{ST}$ values between Moroccan populations only are low ($0.03 < F_{ST} < 0.15$) and differentiation is not always significant. We did not find any significant correlation between $F_{ST}/(1 - F_{ST})$ values and the logarithm of the geographical distance in separating Moroccan populations ($P = 0.582$, slope of 0.065).

**DEMOGRAPHIC HISTORY**

All MIGRAINE results are presented in Table 3. Analyses of modified data sets (mtDNA and nDNA sequence) fit to the ISM give nonhomogeneous results, suggesting that modifications to fit the ISM may have been too sharp to retain most of the information within the original sequences. This could be a result of multiple recurrent mutations or recombination events that occurred on those two loci. However, for both genes, a significant degree of past contraction is detected by one of the two analyses, although parameter estimation differs slightly for the two markers.
First, the nuclear BFIBR locus provides more information on the time when the past contraction began and suggests that it was a relatively recent contraction \([D_{95\% CI} = (0.01–1.4)]\), whereas the mtDNA cytb sequence analysis shows very wide CIs for this estimation \([D_{95\% CI} = (0.008–\infty)]\). Second, point estimates and the CI of the scaled mutation rates \((\theta \text{ and } \theta_{\text{anc}})\) show larger values for the mtDNA data set than for the BFIBR one. However, given that the effective population size in the number of genes is four times higher for the nuclear locus but the mutation rate for the mtDNA cytb locus may be much higher than that for the nuclear BFIBR locus, those differences may not be unrealistic.

MIGRAINE results from the microsatellite data show a highly significant signal of past contraction with parameter estimations that are not always concordant with those obtained with the mtDNA and nDNA data sets. As expected, because of the greater number of loci, estimation precision is much better than for the DNA sequence analyses. The magnitude of past contraction is better inferred with a point estimate for the ratio of current over ancestral population sizes of 0.034 (0.0021–0.042, 95% CI).

The conversion of our estimates of scaled parameters into unscaled demographic parameters is shown in Table 3. Population contraction was substantial (from several hundreds of thousands individuals to a few hundred or thousands) and likely occurred a few thousand years ago. Interestingly, all analyses lead to very similar results in terms of ancestral population size, the parameter for which there is the more information in the data.

**DISCUSSION**

**SPATIAL GENETIC STRUCTURE**

With both mtDNA and nDNA, we recover two main genetic clades: (1) an eastern clade, which includes all individuals from Tunisia and Algeria, as well as some individuals from the most eastern Moroccan locality (Guenfouda), and (2) a western clade, comprising most specimens from Morocco. The cytb K2P genetic difference between these two clades is 7.8%, which often corresponds to the genetic distance observed between sister species in rodents (Baker & Bradley, 2006; Boratynski et al., 2012; Ndiaye et al., 2012; ). It is interesting to note that the specimens from Guenfouda, which group with either the eastern or the western clades, differ in the two datasets. For cytb, MA234 specimen clusters with the eastern clade, whereas all other specimens from this locality cluster with the western clade (Figs 2 and 3), but,
Table 3. Inferences on demographic history using MIGRAINE on the pooled Moroccan data set

<table>
<thead>
<tr>
<th></th>
<th>Nuclear DNA sequences BFIBR</th>
<th>Mitochondrial DNA sequences cytb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deleted problematic sites</td>
<td>Deleted problematic individuals</td>
</tr>
<tr>
<td></td>
<td>(11 haplotypes and 24 SS left)</td>
<td>(8 haplotypes and 23 SS left)</td>
</tr>
<tr>
<td>N</td>
<td>124</td>
<td>136</td>
</tr>
<tr>
<td>pGSM</td>
<td>0.37 (0.15–0.55)</td>
<td>0.152 (0–∞)</td>
</tr>
<tr>
<td>θ</td>
<td>1.9 (1.4–2.7)</td>
<td>0.09 (0–3.6)</td>
</tr>
<tr>
<td>D</td>
<td>6.9 (2.3–10.5)</td>
<td>0.2 (0.01–1.4)</td>
</tr>
<tr>
<td>θanc</td>
<td>566 (43–876)</td>
<td>7.1 (3.0–19)</td>
</tr>
<tr>
<td>Population size ratio</td>
<td>0.0034 (0.0021–0.042)</td>
<td>0.37 (0.0000024–13 000)</td>
</tr>
<tr>
<td>Past variation in population size</td>
<td>si: significant contraction</td>
<td>No significant variation</td>
</tr>
<tr>
<td>N</td>
<td>950 (700–1400)</td>
<td>3500 (1–140 000)</td>
</tr>
<tr>
<td>T</td>
<td>26 000 (6.400–59.000)</td>
<td>2800 (1–800.0001)</td>
</tr>
<tr>
<td>Nanc</td>
<td>280 000 (22 000–440 000)</td>
<td>275 000 (115 000–475 000)</td>
</tr>
</tbody>
</table>

MIGRAINE was only run on the pooled Moroccan sample (i.e. all six Moroccan populations treated as a single population) for all data sets because of (1) the relatively weak population structure observed with all markers and (2) preliminary results being nonsignificant for population-specific analyses on microsatellite data (potentially as a result of low sample sizes and a weak signal of past change in population size). Point estimates and 95% confidence intervals are reported. Inferred parameters are (1) pGSM, the parameter of the geometric distribution of the generalized stepwise mutation model for the microsatellite loci; (2) θ = 2Nθ and θanc = 2Nancθ, the scaled current and ancestral population sizes; (3) D = T_in generations/2N, the scaled time of when the past change in population size started; (4) N, the current population size; (5) T, the time when past change in population size started (years); and (6) Nanc, the ancestral population size. All population sizes are expressed as numbers of genes (i.e. haploid population sizes). N is the sample size for each data set expressed in number of individuals. SS, segregating sites.
for BFIBR, MA205 and MA210 specimens cluster with the eastern clade (Fig. 3), whereas all other specimens cluster with the western clade. For a given specimen, the two phased BFIBR haplotypes always cluster within the same clade. In animal species, nuclear and mitochondrial markers differ in effective size, presence or the absence of recombination and biparental vs. maternal inheritance. Thus, it is not surprising to find different results for mitochondrial and nuclear data. The pattern observed in the locality of Guenfouda could be explained by mtDNA introgression as a result of past hybridization followed by back-crosses with paternal lineages. mtDNA introgression is not rare in nature, particularly for rodents (Bryja et al., 2010; Boratynski et al., 2014). Of the 10 microsatellite loci studied, three loci (MS-6 MS-7 and MS-8), have specific fixed alleles in the Algerian population (not shown) that can be considered as diagnostic alleles of the Algerian populations. Four Moroccan specimens are associated with the Algerian population (specimens MA924 and MA928 from Ouled Boughadi and specimens MA960 and MA961 from Ifrane; Fig. 4) because of the presence of some of these specific alleles (M924 and M928: homozygous on locus MS-7; M960 and M961: heterozygous on locus MS-6). This result is similar to the one found in North African sympatric mice species (Mus musculus domesticus and Mus spretus; Orth et al., 2002) and is the result of a single locus effect of the 10 loci analysis. Our results tend to favour the hypothesis of two sister species in North Africa with possible hybridization between them. Additional analyses with more specimens and more genetic markers are necessary to confirm this hypothesis.

The suggested two species would have diverged in the early Pleistocene. This type of genetic structuring, with one or several eastern clades (Algeria/Tunisia) and one or several western clade (Morocco), has been found in many other vertebrate species (Arano et al., 1998; Recuero et al., 2007; Barata, Harris & Castilho, 2008; Fritz et al., 2009; Nicolas et al., 2014; Vences et al., 2014). In most of these cases, the Moulouya River and/or the wide arid valley extending along much of the river (except close to the estuary) was referred to as the major geographical and climatic barrier responsible for this structuring. However, Fritz et al. (2009) suggested that this bipartite east–west differentiation may be too simplistic, reflecting incomplete sampling rather than true geographical differentiation. Indeed, most studies lack proper sampling in Algeria.

General phylogeographical patterns in Morocco are diverse. Although several species show a lack of phylogeographical structure (Batista et al., 2006; Harris & Perera, 2009; de Pous et al., 2013), others show the opposite (Pinho, Harris & Ferrand, 2007; Recuero et al., 2007; Stöck et al., 2008; Fritz et al., 2009; Beukema et al., 2010; Nicolas et al., 2014). The strong phylogeographical pattern within Morocco can be largely explained by climate changes during the Plio-Pleistocene, which may have created multiple refugia. In North Africa, Morocco has a high physiological complexity, with several large mountain ranges oriented primarily east–west (Anti-Atlas, High Atlas, Middle Atlas, Rif). These mountains may have aided the survival of populations through altitudinal shifts, allowing them to track suitable microclimates as the general climate fluctuated. Moreover, as a result of its geographical location, Morocco is influenced by both the Atlantic and the Mediterranean sea, resulting in a wide range of climates. These characteristics make it unlikely that Morocco offered a single homogeneous and continuous refuge area throughout the Pleistocene. Instead, the variable distribution and fragmented nature of suitable habitats would have favoured the occurrence of multiple refugia isolated from one another. Our mtDNA data show the existence of two genetic clades within Morocco that likely diverged in the Middle Pleistocene (414 kya; range: 240–617 kya). However, presently, these clades are broadly sympatric, rendering the localization of the two possible refugia impossible to determine. Within Morocco, our mtDNA results show some spatial clustering of haplotypes, although this is less evident in the nuclear (BFIBR and microsatellite) datasets. Isolation-by-distance patterns and geographical aggregation of haplotypes suggest spatially limited dispersal and thus geographically restricted gene flow. Significant mtDNA spatial patterns and nonsignificant nuclear ones may indicate sex-biased dispersal with males dispersing more frequently than females. This pattern is common in mammals (Lawson Handley & Perrin, 2007) and has been observed in M. unguiculatus (Liu et al., 2009).

**DEMOGRAPHIC HISTORY OF THE SPECIES BASED ON PALEONTOLOGICAL AND GENETIC DATA**

According to El Harhoura 2 cave fossil data, Meriones are dominant in relatively stable proportions all along the Late Pleistocene record, although a significant decrease in the proportion of Meriones is observed in the assemblage of level 1 (Holocene approximately 5.8 kya) compared to Pleistocene levels (Stoetzel et al., 2011). Our genetic data confirm a bottleneck in Moroccan M. shawii complex at least during the Holocene, and possibly slightly earlier. However, our dating estimate that marks the beginning of this bottleneck has a large uncertainty. Our data suggest that the pattern observed in the fossil record represents demographic
changes. Two hypotheses can be proposed to explain
M. shawii complex population size changes over time:
(1) environmental changes caused by humans, and (2)
climatic changes unrelated to human activities.
Despite the fact that the Holocene level of El Har-
houra 2 is attributed to the Neolithic culture, few fau-
nal remains represent domestic taxa. Furthermore,
neither commensal species (Mus domesticus, Rattus
rattus, Rattus norvegicus), nor the opening of habitat
under anthropic intervention (agriculture) occurred
(Stoetzel et al., 2014). Thus, at this period in the
region of Temara, the impact of human activities on
the landscape and the faunal communities should
have been moderate. The difference in species propor-
tions observed in recent times could therefore be a
result of environmental factors. Indeed, the Holocene
level of El Harhoura 2 (approximately 5.8 kya) is
characterized by a more humid climate than previous
periods (Stoetzel et al., 2011) corresponding to the last
climatic optimum (warm and humid climate, develop-
ment of Mediterranean vegetation in North Africa,
and reduction of the desert areas) and extant Meriones
species avoid semi-humid and humid regions (Aulag-
nier & Thevenot, 1986; Aulagnier, 1992). Thus, we
could easily argue a climatic context to explain the
M. shawii complex population size reduction in the
Holocene. Despite pronounced climate changes during
the period approximately 20–120 kya (Stoetzel et al.,
2011, 2014), no significant change in the proportion of
M. shawii complex in the small mammal assemblage
of El Harhoura 2 was observed (Stoetzel et al., 2011).
Unfortunately, the previous glacial optimum (Eemian,
approximately 125–130 kya) is not recorded at El Har-
houra 2 (or the corresponding level was not yet
reached), preventing any comparison between these
two interglacials (Eemian and Holocene). Between
them, during the Late Pleistocene glacial period, the
relatively humid interstadials were far less humid and
their impact would have been milder than the last
climatic optimum. According to our phylogeographic
tree analyses, a splitting event occurred within this species
in Morocco at the end of the Middle Pleistocene
(TMRCa of the two clades: 0.292 Mya, with large CI:
0.120–0.559 Mya). This may correspond to the isola-
tion of M. shawii complex populations in two distinct
arid refugia during the 125–130 kya glacial optimum;
however, the large CIs hinder the precision of any
estimates.

CONCLUSIONS
The combination of three different data sets (fossil
remains, molecular data, and palaeoenvironmental
data) points towards a climatic rather than anthro-
pogenic influence with respect to explaining M. shawii

ACKNOWLEDGEMENTS
The present study was supported by ANR 6ème
extinction ANR-09-PEXT-004, MOHMIE, ‘Modern
Human installation in Morocco, Influence on the
small terrestrial vertebrate biodiversity and Evolu-
tion’ (CD) and the project CMEE TASSILI MDU
09MDU755. RL has been partially funded by two
ANR projects (EMILE 09-blan-0145-01 and IM-Mod-
el@CORAL.FISH 2010-BLAN-1726-01) and by the
INRA (Project INRA Starting Group 'IGGiPop'). ES
was funded by both the ANR-09-PEXT-004, MOHM-
IE and the LabEx BCDiv. Fieldwork was made
possible through collaboration with the 'Institut Sci-
entifique de Rabat' (A. El Hassani, M. Fekhaoui) and
the ‘Haut Commissariat aux Eaux et Forêts et Lutte
contre la Désertification’. We are grateful to all col-
lectors, particularly Abderrahmane Mataame, Hic-
ham El Brini, Arnaud Delapré, Léa Bourg, Loubna
Tifarouine, and Laurent Granjon. Molecular analyses
were financially supported by the ‘ATM MNHN: Tax-
onomie moléculaire, DNA Barcode & Gestion Dur-
able des Collections’, the ‘Service de Systématique
Moléculaire’ of the MNHN (UMS 2700, Paris,
France), and the network ‘Bibliothèque du Vivant’
funded by the CNRS, MNHN, INRA, and CEA
(Genoscope). Part of this work was carried out using
resources from the MNHN UMS 2700, INRA GENO-
TOUL, and MIGALE bioinformatics platforms and the
computing grid of the CBGP laboratory. We thank K.
Gavrilchuk for helpful comments on an ear-
lier version of the manuscript and three anonymous
reviewers for their helpful comments.

REFERENCES
Arano G, Llorente GA, Montori A, Busckley D, Herrera
P. 1998. Diversification in North-west African water frogs:

© 2015 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 262–279


Rambaut A. 2009. FigTree version 1.3.1 [computer program]. Available at: http://tree.bio.ed.ac.uk


© 2015 The Linnean Society of London, Biological Journal of the Linnean Society, 2016, 118, 262–279
SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. List of specimens included in the molecular analyses.
Table S2. Genetic variability of Meriones shawii populations at microsatellite loci. Genetic variability of M. shawii at six locations and 10 loci. Sample size; number of alleles; observed heterozigosity ($H_O$) and expected heterozigosity ($H_E$); and probability of significant deviation from Hardy–Weinberg equilibrium ($P$, Markov chain procedure, $\alpha=0.05$) are reported. Significant $P$ values are shown in bold.