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

Aude Lalis, Raphaël Leblois, Emmanuelle Stoetzel, Touria Benazzou, Karim Souttou, Christiane Denys, Violaine Nicolas

The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2016, **118**, 262–279. With 4 figures.

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

AUDE LALIS[1]*, RAPHAEL LEBLOIS[2], EMMANUELLE STOETZEL[1,3], TOURIA BENAZZOU[4], KARIM SOUTTOU[5], CHRISTIANE DENYS[1] and VIOLAINE NICOLAS[1]

[1]*Institut de Systématique, Evolution, Biodiversité, ISYEB-UMR 7205 CNRS, MNHN, UPMC, EPHE, Muséum national d'Histoire Naturelle, Sorbonne Universités, 57 rue Cuvier, CP 51, 75005, Paris, France*
[2]*INRA-UMR1062 CBGP, F-34988 Montferrier-sur-Lez, France*
[3]*Histoire 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*
[4]*Département de Biologie, Faculté des Sciences, BP1014, Rabat, Morocco*
[5]*Laboratoire 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 optimums, 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. *Biological Journal of the Linnean Society*, 2016, **118**, 262–279.

KEYWORDS: biogeographical patterns – fossil calibration – population genetics – rodent.

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

(Harris & Perera, 2009; Ben Faleh *et al.*, 2012; Boratynski, Brito & Mappes, 2012; Husemann *et al.*, 2013; Nicolas *et al.*, 2014). North-western Africa has yielded a rich palaeontological record allowing a better understanding of the evolution of small mammal species and communities over time in relation with Quaternary climatic changes (Stoetzel, 2013). Moreover, this region is now recognized as a major area for the emergence and dispersal of anatomically modern humans from at least 160 kya (Mcbrearty & Brooks, 2000; Manica *et al.*, 2007; Smith *et al.*, 2007; Garcea, 2012). North-western Africa thus constitutes a region of interest for investigating how the arrival of early representatives of *Homo sapiens* and their recent impact on the environment influenced faunal genetic diversity and distribution. Understanding how fluctuations in climate and human-mediated habitat change influenced the demography and genetic structure of historic species is not only interesting for fundamental research, but also can be employed to predict future demographic and genetic parameters of species. This is particularly true in the Mediterranean region of North Africa where species and landscape diversity is currently threatened by global climate change and increasing human activities, resulting in the progressive disappearance of certain biotopes (e.g. Mediterranean dry woodlands and steppe ecoregions; Olson *et al.*, 2001).

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 species in this group: *M. shawii* and *Meriones grandis* (Cabrera, 1907; Pavlinov, 2000) but with a large biometric overlapping and without clear geographical distribution. Other studies consider them only as 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 Témara 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, palaeoenvironmmental, 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/CFF) 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 NucleoSpinR 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 β-fibrinogen

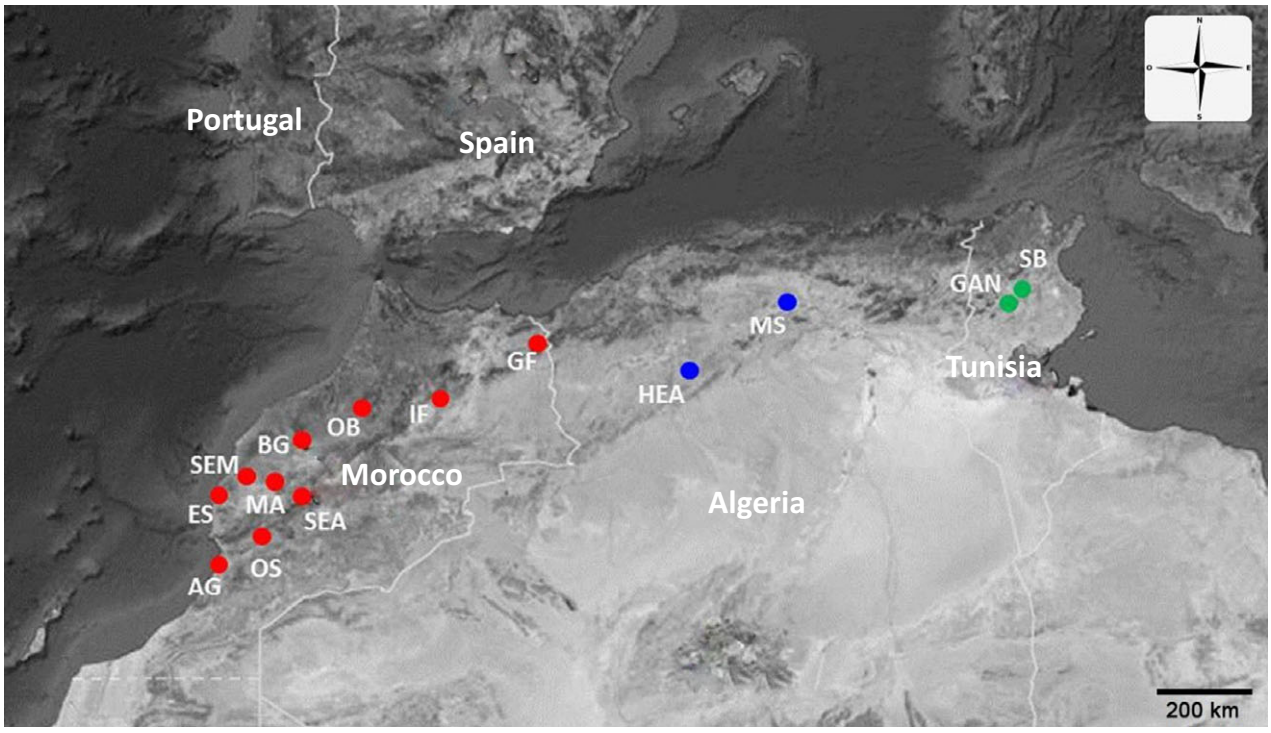

Figure 1. Map of the sampled geographical localities.

Table 1. *Meriones shawii* sampling locality details and genetic diversity showing the number of individuals analyzed (N), averaged values of observed heterozygosity (H_O) and expected heterozygosity (H_E), mean number of alleles (N_A), allelic richness (A) within-population coefficient of inbreeding (FIS), nucleotide diversity (Pi), and haplotypic diversity (H_D)

				Microsatellite diversity						mtDNA diversity: *cytb*			nDNA diversity: *BFIBR*		
Locality	Code	Latitude	Longitude	N	H_O	H_E	N_A	A	F_{IS}	N	Pi	H_D	N	Pi	H_D
Morocco															
BenGuerir	BG	32.19	−7.97	29	0.488	0.462	4.20	3.30	−0.038	28	0.01397	0.918	8	0.00404	0.883
Guenfouda	GF	MM	−2.01	18	0.528	0.556	4.20	3.59	0.080	20	0.00866	0.632	7	0.01061	0.835
Ifrane	IF	33.42	−5.04	11	0.518	0.493	3.80	3.54	−0.004	12	0.00484	0.318	11	0.00393	0.827
Sour El Az	$EA	31.84	−7.01	33	0.509	0.464	4.00	3.03	−0.081	33	0.00289	0.845	5	0.00216	0.844
Ouled Boughadi	OB	33.07	−6.72	33	0.355	0.411	4.60	3.20	0.048	33	0.00406	0.748	6	0.00222	0.530
Aglou	AG	29.83	−9.77	–						2			–		
2okm $ Essaoulra	E$	31.32	−9.71	–						2			–		
20 km *N* Manakech	MA	31.83	−7.97	–						1			–		
Qued Souss	OS	3040	−8.92	–						3			–		
Sidl El Moctar	$EM	31.53	−9.00	–						2			1		
Algeria															
MSila	MS	35.7	4.5	13	0.43	0.558	4.10	3.74	0.009	14	0.00275	0.912	4	0.01001	0.964
Hadjeb El Djmel	HEA	34.3	2.97	–						25	0.00446	0.923	–		
Tunisia															
Garat An Njila	GAN	35.07	9.6	–						2			1		
Sidi Bouzid	SB	35.03	9.5	–						1			2		
Total/mean				137						178			45		

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 CODONCODE ALIGNER, version 3.5.6. Sequences of 1062 and 732 bp long were retained 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 & Lambourdiere (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 (TMRCA). Divergence times and their credibility intervals were estimated using a relaxed clock model with branch rates drawn from an uncorrelated log-normal 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 Kazakstan 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 airensis* (AJ012021), *Acomys cahirinus* (AJ233953), *Acomys chudeaui* (FJ415538), *Acomys cilicicus* (AJ233957), *Acomys dimidiatus* (AJ233959), *Acomys ignitus* (Z96064), *Acomys johannis* (HM635823), *Acomys minous* (GU046553), *Acomys nesiotes* (AJ233952), *Acomys percivali* (EF187818), *Acomys russatus* (FJ415485), *Acomys spinosissimus* (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), *Gerbilliscus tytonis* (AJ430559), *Sekeetamys calurus* (AJ851276), *Gerbillus campestris* (AJ851271), *Gerbilliscus gerbillus* (AJ851269), *Gerbilliscus henleyi* (JQ753050), *Gerbilliscus nanus* (JQ753051), *Gerbilliscus poecilops* (JQ753064), *Gerbilliscus simoni* (GU356577), *Gerbilliscus tarabuli* (GU356573), *Desmodilliscus braueri* (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.

Genetic 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 DISTRUCT, 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 μ 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 θ_{anc}) into effective population sizes (N and N_{anc}), and scaled times (D) into times in years (T), we used: (1) a generation time of 1 year; (2) mutation rates of 5×10^{-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 phylogenetic (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 jacknife 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.

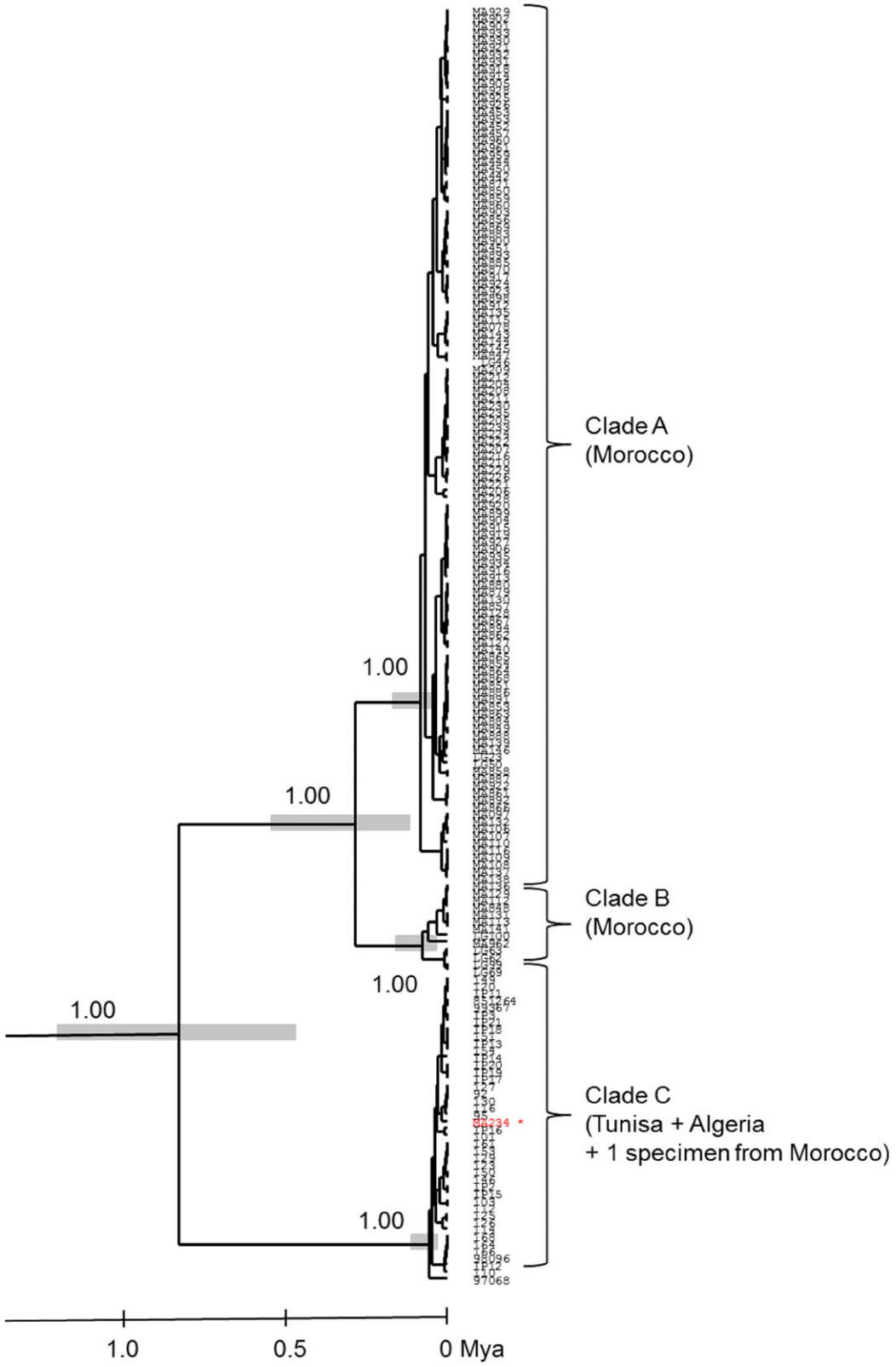

Figure 2. Time calibrated phylogeny inferred using BEAST, depicting phylogenetic relationships within the species *Meriones shawii*. Numbers at nodes represent clade posterior probabilities. To improve clarity, outgroup taxa are not shown.

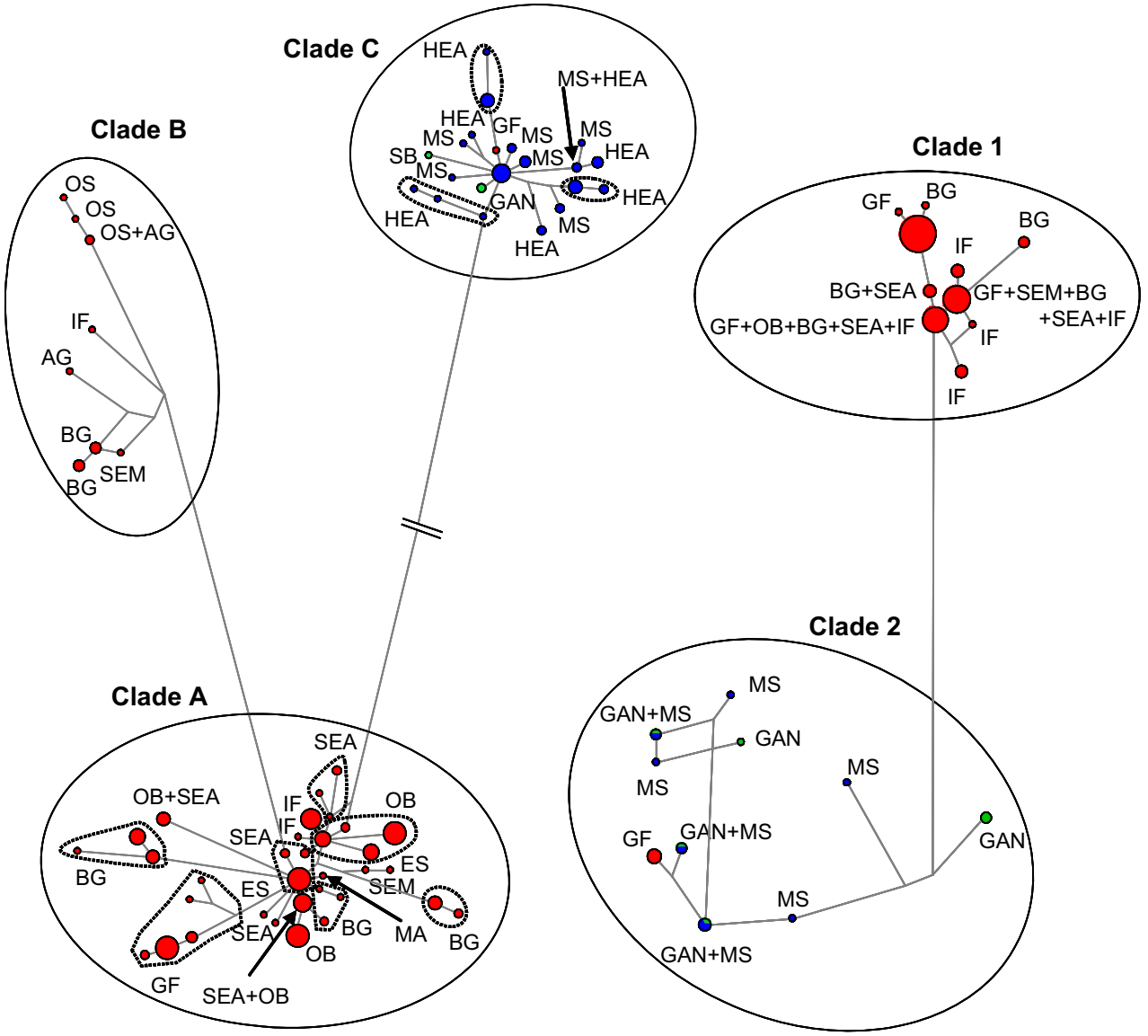

Figure 3. Minimum spanning network of *Meriones shawii cytb* [mitochondrial (mt)DNA, left] and *BFIBR* (nuclear DNA, right) haplotypes. Circle sizes are proportional to the number of similar haplotypes observed in the data set. Branch lengths are proportional to the number of mutations between haplotypes (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).

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.

Table 2. Mitochondrial DNA sequence (above diagonal) and microsatellite (below diagonal) pairwise F_{ST} statistics between *Meriones shawii* populations. ($P < 0.001$)

Population	BG	GF	IF	SEA	OB	MS	HEA
BG		0.3309*	0.2153*	0.2233*	0.2634*	0.8713*	0.8803*
GF	0.0549		0.4651*	0.4932*	0.4711*	0.9139*	0.9131*
IF	0.0966	0.1135*		0.2983*	0.2344*	0.9509*	0.9399*
SEA	0.0389	0.067	0.1178*		0.1678*	0.9625*	0.9535*
OB	0.0675	0.1113*	0.1542*	0.038		0.9514*	0.9445*
MS	0.1877*	0.1767*	0.2420*	0.1976*	0.2050*		0.0401

*Significant values.

Haplotype diversity is 0.814 ± 0.027 for clade 1 and 0.922 ± 0.039 for clade 2. Within-clade nucleotide diversity is 0.00345 ± 0.00022 for clade 1 and 0.00961 ± 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 heterozygosity ranges from 0.355 (OB) to 0.528 (GF), whereas the expected heterozygosity ranges 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.

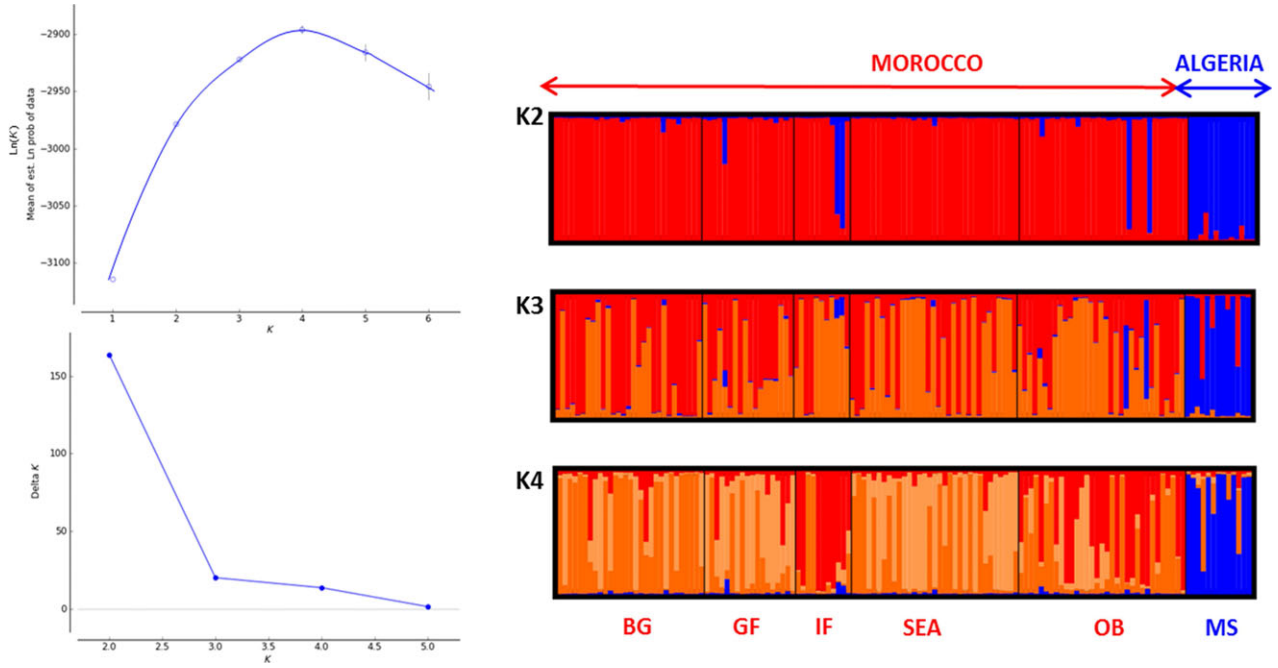

Figure 4. *Meriones shawii* complex populations clustering based on STRUCTURE Bayesian inference (K = 2–4); the graph illustrates the log posterior probabilities of the microsatellite data [Ln P(K)] for each number of genetic groups (K) and the number of contributing populations was tested using the ad-hoc Evanno statistic (Delta K) for K = 1–6. Each individual is represented by a thin vertical line, which is partitioned into K-coloured segments that indicate an individual's estimated membership fraction in K clusters.

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 [i.e. $D_{95\%\ \mathrm{CI}}$ = (0.01–1.4)], whereas the mtDNA *cytb* sequence analysis shows very wide CIs for this estimation [$D_{95\%\ \mathrm{CI}}$ = (0.008–∞)]. Second, point estimates and the CI of the scaled mutation rates (θ and θ_{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 leads 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

		Nuclear DNA sequences *BFIBR*		Mitochondrial DNA sequences *cytb*	
	Microsatellites	Deleted problematic sites (11 haplotypes and 24 SS left)	Deleted problematic individuals (8 haplotypes and 23 SS left)	Deleted problematic sites (27 haplotypes and 111 SS left)	Deleted problematic individuals (23 haplotypes and 45 SS left)
N	124	76	49	136	77
pGSM	0.37 (0.15–0.55)	NA	NA	NA	NA
θ	1.9 (1.4–2.7)	2.3 (0–99)	0.09 (0–3.6)	5.6 (0.4–21)	15 (0–225)
D	6.9 (2.3–10.5)	0.152 (0–∞)	0.2 (0.01–1.4)	0.057 (0.008–∞)	0.94 (0–∞)
θ*anc*	566 (43–876)	6 (0–∞)	7.1 (3.0–19)	24 (14–73)	0.5 (0–∞)
Population size ratio equivalent θ/θ_{anc}	0.0034 (0.0021–0.042)	0.37 (0.0000024–13 000)	0.013 (0.000001–0.7)	0.24 (0.0022–0.7)	29 (0.0000001–50 000)
Past variation in population size	si: significant contraction	No significant variation	Significant contraction	Significant contraction	No significant variation
N	950 (700–1400)		3500 (1–140 000)	54 000 (3,800–200 000)	
T	26 000 (6.400–59.000)		2800 (1–800.0001	3000 (30–∞)	
Nanc	280 000 (22 000–440 000)		275 000 (115 000–475 000)	230 000 (130 000–700 000)	

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) *p*GSM, the parameter of the geometric distribution of the generalized stepwise mutation model for the microsatellite loci; (2) $\theta = 2N\mu$ and $\theta_{anc} = 2N_{anc}\mu$, the scaled current and ancestral population sizes; (3) $D = T_{\text{in generation}}/2N$, 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) N_{anc}, 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 physiographical 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 Harhoura 2 is attributed to the Neolithic culture, few faunal 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 Témara, the impact of human activities on the landscape and the faunal communities should have been moderate. The difference in species proportions 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, development of Mediterranean vegetation in North Africa, and reduction of the desert areas) and extant *Meriones* species avoid semi-humid and humid regions (Aulagnier & 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 Harhoura 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 phylogenetic 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 isolation 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 palaeoenvironmmental data) points towards a climatic rather than anthropogenic influence with respect to explaining *M. shawii* complex population size reduction in the Holocene. The ability of the *M. shawii* complex to thrive in cultivated crops and irrigation structures (to the point where it is considered to be a crop pest) (Aulagnier *et al.*, 2008 in IUCN 2012) must have appeared very recently. The present study demonstrates the advantage of using integrative and multidisciplinary approaches for investigating species abundance and demography. We also provide relevant information on the genetic structure of *M. shawii* complex, structured geographically in two major clades, which can be of importance to those concerned with pest control and human disease because this species is considered to represent an economically significant pest (Aulagnier *et al.*, 2008 in IUCN 2012).

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 Evolution' (CD) and the project CMEP TASSILI MDU 09MDU755. RL has been partially funded by two ANR projects (EMILE 09-blan-0145-01 and IM-Model@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, MOHMIE and the LabEx BCDiv. Fieldwork was made possible through collaboration with the 'Institut Scientifique 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 collectors, particularly Abderrahmane Mataame, Hicham El Brini, Arnaud Delapré, Léa Bourg, Loubna Tifarouine, and Laurent Granjon. Molecular analyses were financially supported by the 'ATM MNHN: Taxonomie moléculaire, DNA Barcode & Gestion Durable 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 GENOTOUL, and MIGALE bioinformatics platforms and the computing grid of the CBGP laboratory. We thank K. Gavrilchuk for helpful comments on an earlier version of the manuscript and three anonymous reviewers for their helpful comments.

REFERENCES

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

molecular and morphological evidence. *Herpetological Journal* **8:** 57–64.

Aulagnier S 1992. Zoogéographie des Mammifères du Maroc: de l'analyse spécifique à la typologie de peuplement à l'échelle régionale. State Thesis, University Montpellier 2.

Aulagnier S, Thevenot M. 1986. Catalogue des mammifères sauvages du maroc. *Travaux De l'institut Scientifique Série Zoologie* **41:** 1–163.

Aulagnier S, Granjon L, Amori G, Huttere R, Krystufek B, Yigit N, Mitsain G 2008. Meriones shawi. In: IUCN 2012. IUCN Red List Of Threatened Species, Version 2012.2. http://www/iucnredlist.org

Baker RJ, Bradley RD. 2006. Speciation in mammals and the genetic species concept. *Journal of Mammalogy* **87:** 643–662.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Barata M, Harris DJ, Castilho R. 2008. Comparative phylogeography of northwest African *Natrix maura* (Serpentes: Colubridae) inferred from mtDNA sequences. *African Zoology* **43:** 1–7.

Batista V, Carranza S, Carretero MA, Harris DJ. 2006. Genetic variation within *Bufo viridis*: evidence from mitochondrial 12S and 16S RRNA DNA sequences. *Butlletí de la Societat Catalana d'Herpetologia* **17:** 24–33.

Beaumont MA. 1999. Detecting population expansion and decline using microsatellites. *Genetics* **153:** 2013–2029.

Ben Faleh A, Granjon L, Tatard C, Boratynski Z, Cosson JF, Said K. 2012. Phylogeography of two cryptic species of african desert jerboas (*Dipodidae: Jaculus*). *Biological Journal of the Linnean Society* **107:** 27–38.

Beukema W, de Pous P, Donaire D, Escoriza D, Bogaerts S, Toxopeus AG, de Bie CAJM, Roca J, Carranza S. 2010. Biogeography and contemporary climatic differentation among Moroccan *Salamandra algira*. *Biological Journal of the Linnean Society* **101:** 626–641.

Boratynski Z, Brito JC, Mappes T. 2012. The origin of two cryptic species of African desert jerboas (*Dipodidae: Jaculus*). *Biological Journal of the Linnean Society* **105:** 435–445.

Boratynski Z, Melo-Ferreira J, Alves PC, Berto S, Koskela E, Pentikäinen OT, Tarroso P, Ylilauri M, Mappes T. 2014. Molecular and ecological signs of mitochondrial adaptation: consequences for introgression? *Heredity* **113:** 277–286.

Brun A. 1989. Micoflores et paléovégétations en Afrique du Nord depuis 30 000 ans. *Bulletin de la Société Géologique de France* **1:** 25–33.

Brun A 1991. Réflexions sur les Pluviaux et Arides au Pléistocène supérieur et à l'Holocène en Tunisie. *Palaeoecology of Africa*. **22:** 157–183.

Bryja J, Granjon L, Dobigny G, Patzenhauerova H, Konecny A, Duplantier JM, Gauthier P, Colyn M, Durnez L, Lalis A, Nicolas V. 2010. Plio-pleistocene history of west african sudanian savanna and the phylogeography of the *Praomys daltoni* complex (rodentia): the environment/geography/genetic interplay. *Molecular Ecology* **19:** 4783–4799.

Cabrera A. 1907. Meriones grandis. IN: Algunos roedores nuevos de Marruecos. *Boletin la Real Sociedad Española de Historia Natural* **7:** 175–177.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J. 1996. A comprehensive map of the human genome based on 5,264 microsatellites. *Nature* **380:** 152–154.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214–222.

Ducroz JF, Volobouev V, Granjon L. 2001. An assessment of the systematics of Arvicanthine rodents using mitochondrial DNA sequences: evolutionary and biogeographical implications. *Journal of Mammalian Evolution* **8:** 173–206.

Ellegren H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics* **16:** 552–558.

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

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

Falush D. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164:** 1567–1587.

Fedorov VB, Goropashnaya AV, Boeskorov GG, Cook JA. 2008. Comparative phylogeography and demographic history of the wood lemming (*Myopus schisticolor*): implications for late quaternary history of the taiga species in Eurasia. *Molecular Ecology* **17:** 598–610.

Fritz U, Harris DJ, Fahd S, Rouag R, Gracia Martinez E, Gimenez Casalduero A, Siroky P, Kalboussi M, Jdeidi TB, Hundsdörfer AK. 2009. Mitochondrial phylogeography of *Testudo graeca* in the Western Mediterranean: old complex divergence in North Africa and recent arrival in Europe. *Amphibia-Reptilia* **30:** 63–80.

Garcea E. 2012. Successes and failures of human dispersals from North Africa. *Quaternary International* **270:** 119–128.

Goudet J. 1995. FSTAT: a computer program to calculate f-statistics. *Journal of Heredity* **86:** 485–486.

Goudet J, Raymond M, De Meeüs T, Rousset F. 1996. Testing differenciation in diploid populations. *Genetics* **144:** 1933–1940.

Gündüz I, Rambau RV, Tez C, Searle JB. 2005. Mitochondrial DNA variation in the western house mouse (*Mus musculus domesticus*) close to its site of origin: studies in Turkey. *Biological Journal of the Linnean Society* **84:** 473–485.

Harris D, Perera A. 2009. Phylogeography and genetic relationships of North African *Bufo mauritanicus* Schlegel, 1841 estimated from mitochondrial DNA sequences. *Biologia* **64:** 356–360.

Hooghiemstra H, Stalling H, Agwu COC, Dupont L. 1992. Vegetational and climatic changes at the northern fringe of the Sahara 250,000-5000 years BP: evidence from 4 marine pollen records located between Portugal and the Canary Islands. *Review of Palaeobotany and Palynology* **74:** 1–53.

Hudson RR, Kaplan NL. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111:** 147–164.

Husemann M, Schmitt T, Zachos FE, Ulrich W, Habel JC. 2013. Palaearctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaearctic biota. *Journal of Biogeography* **41:** 81–94.

de Iorio M, Griffiths RC. 2004a. Importance sampling on coalescent histories. *Advances in Applied Probability* **36:** 417–433.

de Iorio M, Griffiths RC. 2004b. Importance sampling on coalescent histories II. Subdivided population models. *Advances in Applied Probability* **36:** 434–454.

de Iorio M, Griffiths RC, Leblois R, Rousset F. 2005. Stepwise mutation likelihood computation by sequential importance sampling in subdivided population models. *Theoretical Population Biology* **68:** 41–53.

Jacobs Z, Roberts RG, Nespoulet R, El Hajraoui MA, Debenath A. 2012. Singlegrain OSL chronologies for Middle Palaeolithic deposits at El Mnasra and El Harhoura 2, Morocco: implications for Late Pleistocene human–environment interactions along the Atlantic coast of northwest Africa. *Journal of Human Evolution* **62:** 377–394.

Jaeger JJ, Tong H, Denys C. 1986. The age of the *Mus-Rattus* divergence: paleontological data compared with the molecular clock. *Comptes Rendus de l'Académie des Sciences Serie II Paris* **302:** 917–922.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23:** 1801–1806.

Jolly D, Harrisson SP, Damnati B, Bonnefille R. 1998. Simulated climate and biomes of Africa during the late Quaternary: comparison with pollen and lake status data. *Quaternary Science Reviews* **17:** 629–657.

Jombart T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24:** 1403–1405.

Kelly RP, Oliver TA, Sivasundar A, Palumbi SR. 2010. A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity* **101:** 423–436.

Lalis A, Lambourdiere J. 2014. Genetic variation in a North African rodent pest, *Meriones shawi*: microsatellite polymorphism. *African Zoology* **49:** 157–160.

Lawson Handley LJ, Perrin N. 2007. Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology* **16:** 1559–1578.

Leblois R, Pudlo P, Néron J, Bertaux F, Reddy Beeravolu C, Vitalis R, Rousset F. 2014. Maximum likelihood inference of population size contractions from microsatellite data. *Molecular Biology and Evolution* **31**, 2805–2823.

Librado P, Rozas J. 2009. Dnasp V5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451–1452.

Liu W, Wang G, Wan XR, Zhong WQ. 2009. Effects of supplemental food on the social organization of Mongolian gerbils during the breeding season. *Journal of Zoology* **278:** 249–257.

Lopez-Garcia JM, Agusti J, Aouraghe H. 2013. The small mammals from the Holocene site of Guenfouda (Jerada, Eastern Morocco): chronological and paleoecological implications. *Historical Biology* **25:** 51–57.

Manica A, Amos W, Balloux F, Hanihara T. 2007. The effect of ancient population bottlenecks on human phenotypic variation. *Nature* **448:** 346–348.

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* **27:** 209–220.

Marshall CR. 1994. Confidence intervals on stratigraphic ranges: partial relaxation of the assumption of randomly distributed fossil horizons. *Paleobiology* **20:** 459–469.

Mcbrearty S, Brooks AS. 2000. The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution* **39:** 453–563.

Mongelard C, Bentz S, Tirard C, Verbeau O, Catzeflis FM. 2002. Molecular systematics of the Scirugnathi (Rodentia): the mithochondrial cytochrome b and the 12S rRNA genes support the Anomaluroidea (Petetidae and Anomaluridae). *Molecular Phylogenetics and Evolution* **22:** 220–233.

Nabholz B, Glemin S, Galtier N. 2008. Strong variations of mitochondrial mutation rate across mammals, the longevity hypothesis. *Molecular Biology and Evolution* **25:** 120–130.

Ndiaye A, Bâ K, Aniskin V, Benazzou T, Chevret P, Konecny A, Sembene M, Tatars C, Kergoat GJ, Granjon L. 2012. Evolutionary systematics and biogeography of endemic gerbils (Rodentia, Muridae) from Morocco: an integrative approach. *Zoologica Scripta* **41:** 11–28.

Nespoulet R, El Hajraoui MA, Amani F, Ben Ncer A, Debenath A, El Idrissi A, Lacombe JP, Michel P, Oujaa A, Stoetzel E. 2008. Palaeolithic and neolithic occupations in the temara region (Rabat, Morocco): recent data on hominin contexts and behavior. *African Archaeological Review* **25:** 21–39.

Nicolas V, Bryja J, Akpatou B, Konecny A, Lecompte E, Colyn M, Lalis A, Couloux A, Denys C, Granjon L 2008. Comparative phylogeography of two sibling species of forest-dwelling rodent (*Praomys rostratus* and *P. tullbergi*) in west Africa: different reactions to past forest fragmentation. *Molecular Ecology*, **17**, 5118–5134.

Nicolas V, Herbreteau V, Couloux A, Keovichit K, Douangboupha B, Hugot JP. 2012. A remarkable case of micro-endemism in *Laonastes aenigmamus* (Diatomyidae, Rodentia) revealed by nuclear and mitochondrial DNA sequence data. *PLoS ONE* **7:** e48145.

Nicolas V, Ndiaye A, Benazzou T, Souttou K, Delapre A, Denys C. 2014. Phylogeography of the North African Dipodil (Rodentia: Muridae) based on cytochrome b sequences. *Journal of Mammalogy* **95:** 241–253.

Nylander JA. 2004. *MrModeltest v2. Program distributed by the author*. Evolutionary Biology Centre: Uppsala University.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR 2001. Terrestrial ecoregions of the world: a new map of life on earth a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, **51**, 933–938.

Orth A, Belkhir K, Britton-Davidian J, Boursot P, Benazzou T, Bonhomme F. 2002. Hybridation naturelle entre deux espèces sympatriques de souris *Mus musculus domesticus* L. et *Mus spretus* Lataste. *C. R. Biologies* **325:** 89–97.

Ouahbi Y, Aberkan M, Serre F. 2003. Recent Quaternary fossil mammals of Chrafate and Ez Zarka. The origin of modern fauna in the Northern Rif (NW Morocco, Northern Africa). *Geologica Acta* **1:** 277–288.

Pavlinov IY. 2000. Contribution to craniometric variation and taxonomy of jirds from the group '*shawi-grandis*' of the genus *Meriones* (Gerbillidae). *Russian Journal of Zoology* **79:** 201–209.

PDB. 2011. The Paleobiology Database. http://paleodb.org/cgi-bin/bridge.pl

Petter F. 1961. Répartition et écologie des rongeurs désertiques (du Sahara occidental à l'Iran oriental). *Mammalia* **25:** 1–222.

Pinho C, Harris DJ, Ferrand N. 2007. Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. *Molecular Ecology* **16:** 1191–11205.

de Pous P, Metallinou M, Donaire-Barroso D, Carranza S, Sanuy D. 2013. Integrating mtDNA analyses and ecological niche modelling to infer the evolutionary history of *Alytes maurus* (Amphibia; Alytidae) from Morocco. *Herpetological Journal* **23:** 153–160.

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

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

Recuero E, Iraola A, Rubio X, Machordom A, Garcia Paris M. 2007. Mitochondrial differentiation and biogeography of *Hyla meridionalis* (Anura: Hylidae): an unusual phylogeographical pattern. *Journal of Biogeography* **34:** 1207–1219.

Reed DN, Barr WA. 2010. A preliminary account of the rodents from Pleistocene levels at Grotte des Contrebandiers (Smuggler's Cave), Morocco. *Historical Biology* **22:** 286–294.

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* **4:** 137–138.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145:** 1219–1228.

Rousset F. 2008. Genepop'007: a complete re-implementation of the genepop software for windows and linux. *Molecular Ecology Resources* **8:** 103–106.

Russo IR, Chimimba C, Bloomer P. 2010. Bioregion heterogeneity correlates with extensive mitochondrial DNA diversity in the namaqua rock mouse, *Micaelamys namaquensis (Rodentia: Muridae)* from southern africaevidence for a species complex. *BMC Evolutionary Biology* **10:** 307–313.

Schenk JJ, Rowe KC, Steppan SJ. 2013. Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. *Systematic Biology* **62:** 837–864.

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

Smith TM, Tafforeau PT, Reid DJ, Grün R, Eggins S, Boutakiout M, Hublin JJ. 2007. Earliest evidence of modern human life history in north african early *Homo sapiens*. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 6128–6133.

Stöck M, Dubey S, Klutsch C, Litvinchuk SN, Scheidt U, Perrin N. 2008. Mitochondrial and nuclear phylogeny of circum-Mediterranean tree frogs from the *Hyla arborea* group. *Molecular Phylogenetics and Evolution* **49:** 1019–1024.

Stoetzel E. 2013. Late Cenozoic micromammal biochronology of northwestern Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology* **392:** 359–381.

Stoetzel E, Bailon S, Nespoulet R, El Hajraoui MA, Denys C. 2010. Pleistocene and Holocene small vertebrates of El Harhoura 2 Cave (Rabat-Temara, Morocco): an annotated preliminary taxonomic list. *Historical Biology* **22:** 303–319.

Stoetzel E, Marion L, Nespoulet R, El Hajraoui MA, Denys C. 2011. Taphonomy and palaeoecology of the Late Pleistocene to Middle Holocene small mammal succession of El Harhoura 2 cave (Rabat-Temara, Morocco). *Journal of Human Evolution* **60:** 1–33.

Stoetzel E, Campmas E, Michel P, Bougariane B, Ouchaou B, Amani F, El Hajraoui MA, Nespoulet R. 2014. Context of modern human occupations in North Africa: contribution of the Témara caves data. *Quaternary International* **320:** 143–161.

Storz JF, Beaumont MA. 2002. Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution* **56:** 154–166.

Sun JX, Helgason A, Masson G, Ebenesersdottir SS, Li H, Mallick S, Gnerre S, Patterson N, Kong A, Reich D, Stefansson K 2012. A direct characterization of human mutation based on microsatellites. *Nature Genetics*, **44**, 1161–1165.

Tolley KA, Burger M, Turner AA, Matthee CA. 2006. Biogeographic patterns and phylogeography of dwarf chameleons (*Bradypodion*) in an african biodiversity hotspot. *Molecular Ecology* **15:** 781–793.

Trauth MH, Larrasoana JC, Mudelsee M. 2009. Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews* **28:** 399–411.

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

Vences M, De Pous P, Nicolas V, Díaz-Rodríguez J, Donaire D, Hugemann K, Hauswaldt JS, Amat F, Barnestein JAM, Bogaerts S, Bouazza A, Carranza S, Galán P, De La Vega JPG, Joger U, Ansari A, El Mouden EH, Ohler A, Sanuy D, Slimani T, Tejedo M. 2014. New insights on phylogeography and distribution of painted frogs (*Discoglossus*) in northern Africa and the Iberian peninsula. *Amphibia Reptilia* **35:** 305–320.

Wang Y, Zhao LM, Fang FJ, Liao JC, Liu NF. 2013. Intraspecific molecular phylogeny and phylogeography of the *Meriones meridianus* (Rodentia: Cricetidae) complex in northern China reflect the processes of desertification and the Tianshan mountains uplift. *Biological Journal of the Linnean Society* **110:** 362–383.

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

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, α=0.05) are reported. Significant P values are shown in bold.