



The evolutionary history of the Cape hare (*Lepus capensis sensu lato*): insights for systematics and biogeography

Sara Lado^{1,2,6} · Paulo C. Alves^{1,2,3} · M. Zafarul Islam^{4,5} · José C. Brito^{1,2} · José Melo-Ferreira^{1,2}

Received: 17 December 2018 / Revised: 11 April 2019 / Accepted: 19 April 2019 / Published online: 9 May 2019
© The Genetics Society 2019

Abstract

Inferring the phylogeography of species with large distributions helps deciphering major diversification patterns that may occur in parallel across taxa. Here, we infer the evolutionary history of the Cape hare, *Lepus capensis sensu lato*, a species distributed from southern Africa to Asia, by analyzing variation at 18 microsatellites and 9 DNA (1 mitochondrial and 8 nuclear) sequenced loci, from field and museum-collected samples. Using a combination of assignment and coalescent-based methods, we show that the Cape hare is composed of five evolutionary lineages, distributed in distinct biogeographic regions—north-western Africa, eastern Africa, southern Africa, the Near East and the Arabian Peninsula. A deep phylogenetic break possibly dating to the Early Pleistocene was inferred between the African and Asian *L. capensis* groups, and the latter appear more closely related to other Eurasian hare species than to African Cape hares. The inferred phylogeographic structure is shared by numerous taxa distributed across the studied range, suggesting that environmental changes, such as the progressive aridification of the Saharo-Arabian desert and the fluctuations of savannah habitats in Sub-Saharan Africa, had comparable impacts across species. Fine-scale analyses of the western Sahara-Sahel populations showed rich fragmentation patterns for mitochondrial DNA but not for microsatellites, compatible with the environmental heterogeneity of the region and female philopatry. The complex evolutionary history of *L. capensis sensu lato*, which possibly includes interspecific gene flow, is not reflected by taxonomy. Integrating evolutionary inference contributes to an improved characterization of biodiversity, which is fundamental to foster the conservation of relevant evolutionary units.

Supplementary information The online version of this article (<https://doi.org/10.1038/s41437-019-0229-8>) contains supplementary material, which is available to authorized users.

✉ José Melo-Ferreira
jmeloferreira@cibio.up.pt

- ¹ CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Campus de Vairão, 4485-661 Vairão, Portugal
- ² Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal
- ³ Wildlife Biology Program, University of Montana, Missoula, MT 59812, USA
- ⁴ National Wildlife Research Center (NWRC), Post Box 1086, Taif, Saudi Arabia
- ⁵ Ilia State University, Tbilisi 0162, Georgia
- ⁶ Present address: Research Institute of Wildlife Ecology, Vetmeduni Vienna, Vienna, Austria

Introduction

Characterizing the evolutionary history of species and the geographic structure of genetic diversity and lineages (phylogeography) provides valuable and powerful information to understand the processes that influenced organism diversification (Avice 2009). Similar diversification patterns across co-distributed species, in time and space, can be used to explain the general biogeographic processes underlying such patterns (Gutiérrez-García and Vázquez-Domínguez 2011). The analysis of multi-locus sequence datasets in coalescent frameworks, taking into account the variance of lineage sorting, has become an important tool to assist the taxonomic classification of organisms (Fujita et al. 2012), particularly when morphological distinctiveness is unclear or cryptic (Bickford et al. 2007). Even if molecular-based methods primarily detect genetic structure, without distinguishing intra from interspecific processes (Sukumaran and Knowles 2017), the information that molecular inferences generate about the evolutionary processes behind lineage diversification is central to modern integrative taxonomy

(Fujita et al. 2012), with direct implications for the definition of appropriate conservation units (Scheffers et al. 2012; Adams et al. 2014).

The taxonomy of the Cape hare (*Lepus capensis*) is historically controversial. Current classifications consider that *L. capensis sensu lato* exhibits a disjunct distribution in Africa, with no gene flow between southern and northern range margins, and that its distribution extends to the Middle East (Schai-Brown and Hackländer 2018) and possibly to Iran and Pakistan (Drew et al. 2008). Although earlier assessments indicated that the distribution of *L. capensis* could extend to China, Mongolia and Russia, the formal recognition of *Lepus tolai* and *Lepus tibetanus* as valid species removes *L. capensis* from these regions, and possibly also from Iran and Pakistan (Schai-Brown and Hackländer 2018). The wide phenotypic dissimilarity of *L. capensis sensu lato* over its range, such as in fur color, body size or ear length (Ben Slimen et al. 2007; Schai-Brown and Hackländer 2018), resulted in a variable number of classified subspecies over time (see Flux and Angermann 1990; Schai-Brown and Hackländer 2018). However, it remains uncertain whether this morphological diversity reflects gradients of local adaptation or deep evolutionary divergence. Hoffman and Smith (2005) informally divided the species into four major geographic partitions—southern Africa, eastern Africa, Arabia and north-western Africa—that could merit specific classification, but the validation of these partitions awaits additional data. Moreover, the relationships of the Cape hare with neighboring hare species is historically controversial (Ben Slimen et al. 2008b; Schai-Brown and Hackländer 2018 and references therein). In addition to *L. capensis sensu lato*, five other species are currently classified in Africa (*Lepus saxatilis*, *Lepus victoriae*, *Lepus habessinicus*, *Lepus fagani* and *Lepus starcki*), distinguished by morphological characters (Kingdon 2013). However, their evolutionary history, genetic differentiation and distribution remain poorly understood, despite recent progresses in the *habessinicus*–*fagani*–*starcki* group (Tolesa et al. 2017). Another difficulty is posed by the frequent sharing of mitochondrial DNA (mtDNA) variation and discordant differentiation patterns with nuclear markers (Ben Slimen et al. 2006, 2008a), which may result from a tendency of hare species to hybridize and to exchange genetic variation over the contact zones (Alves et al. 2006; Melo-Ferreira et al. 2012, 2014b). Therefore, a formal revision of the taxonomy of *L. capensis sensu lato* awaits a more integrative and precise understanding of population structure and evolutionary history of the species over its range.

The Cape hare, as currently classified, is distributed across a wide geographical area, including the savannah and dry desert regions of Southern Africa, the dry savannah regions of Central, West and North Africa, and parts of the

Saharo-Arabian region (Hoffmann and Smith 2005; Drew et al. 2008; Schai-Brown and Hackländer 2018) (Fig. 1). Comparative studies have identified common phylogeographic patterns of organisms associated with these habitats. In sub-Saharan Africa, a north–south divide has been shown to occur across a diverse array of savannah-dwelling taxa, which has been explained by the periodic population confinement on either side of Central Africa tropical forests (e.g. Lorenzen et al. 2012; Bertola et al. 2016). East Africa has also been shown to harbor divergent evolutionary lineages across taxa, which may be related with regional spatial and temporal heterogeneity and rifting promoting isolation and divergence (Lorenzen et al. 2012; Aghová et al. 2017). In North Africa, phylogeographic patterns are usually associated with barriers and dispersal corridors created by the cyclic expansions and contractions of the Saharo-Arabian desert and changes in hydrologic networks (Brito et al. 2014; Mairal et al. 2017; Stewart et al. 2017).

In this work, we use a multi-locus genetics approach and the widest sampling of the species to date to (i) infer population structure of *L. capensis sensu lato* and relate genetic differentiation to geography, (ii) evaluate relationships and divergence among evolutionary lineages, and (iii) discuss the inferred evolutionary history according to known biogeographic patterns across taxa. We test the validity of the putative subdivision of *L. capensis sensu lato* in four geographic explicit groups (Hoffmann and Smith 2005), and assess whether major bioclimatic and physiographic factors that affected phylogeographic patterns across savannah-associated species also influenced the population history of the Cape hare.

Materials and methods

Sampling and DNA extraction

In total, 162 *L. capensis sensu lato* samples from Africa and the Near East were collected (Tables S1 and S2). Part of the used samples (72) resulted from road-kill specimens found during fieldwork along Northwest Africa. The remaining samples were kindly provided by other researchers (75) and museums (15) (see complete museum sampling information in Table S3). Although samples were classified as *L. capensis* based on geographic location and morphology, no a priori assignment to a given population or lineage was considered for the analyses, given the controversies on the ranges of distinct African hares (e.g. Moores et al. 2012).

Genomic DNA from field-collected samples was extracted from preserved tissues (liver or ear) using the JETQUICK Tissue DNA Kit (Genomed). The museum samples (dry skins) were extracted following the protocol described by Bi et al. (2013), in an isolated and autonomous

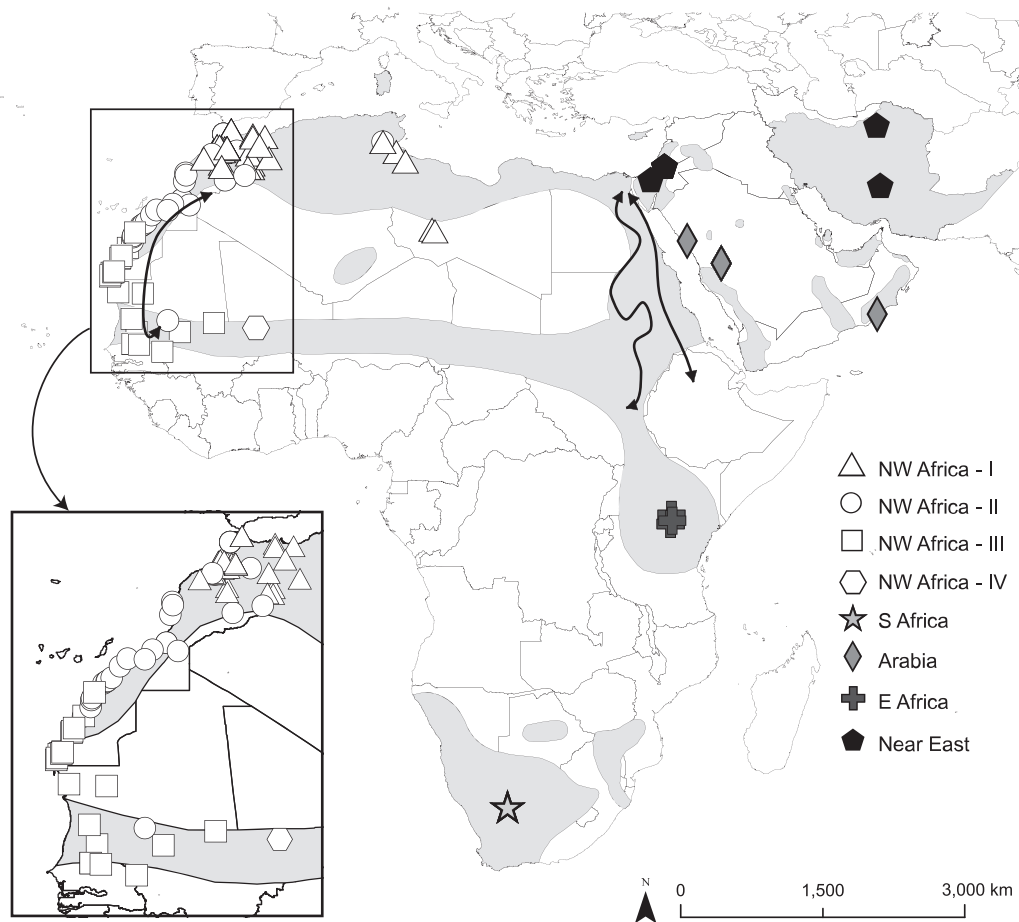


Fig. 1 Geographic distribution of *Lepus capensis* and samples unambiguously attributed to a microsatellite cluster. Gray background indicates the distribution of *L. capensis* according to IUCN Red List (www.redlist.org). Symbols indicate sampling localities: shapes indicate the eight mtDNA lineages (according to the BAPS and network analyses, which agrees with the phylogenetic inference shown in Fig. 2) and their

gray shades depict the five evolutionary groups inferred from microsatellites (see Fig. S1). Symbols are not proportional to the number of individuals samples in the locality and more than one specimen may be represented in each geographical point. Arrows indicate hypothesized dispersal corridors in the Sahara-Sahel—Atlantic Sahara, Nile River and Red Sea mountains (Brito et al. 2014)

room with devoted sterilized equipment, in order to prevent contamination with modern DNA.

Genotyping of microsatellites and DNA sequencing

Microsatellite loci for African *L. capensis sensu lato* were newly developed in this work by Genoscreen (<http://www.genoscreen.fr/>), using pooled high-quality DNA of 12 samples from north-west Africa (see Table S1). Genomic DNA libraries were enriched for microsatellites with microprobes, and sequenced via high throughput Titanium pyrosequencing on a 454-GsFLX® sequencer (Roche Diagnostics). AUTODIMER (Vallone and Butler 2004) was used to predict and avoid hairpin structures and dimers within primer multiplexes. Initially, 38 microsatellite loci were selected to optimize polymerase chain reactions, evaluate genotyping success (in a ABI3130xl Genetic Analyzer; Applied Biosystems), and test their compatibility in

multiplexed reactions. Variability in a selected set of 19 loci was then assessed in 50 specimens from north-west Africa, southern Africa and the Near East (Table S1). Eighteen variable microsatellite loci were finally selected, organized in three multiplexes with fluorescent marked primers (M13 tails) and genotyped (see details in Table S4). GENE-MAPPER v4.0 (Applied Biosystems) was used to read and score the genotyping results, followed by visual inspection. Museum samples were amplified four times and 30% of field-collected samples were re-amplified to ensure consistency of genotype determination. The final dataset included the 162 *L. capensis sensu lato* specimens.

Partial mitochondrial cytochrome *b* was sequenced in 139 *L. capensis sensu lato* samples (three from Melo-Ferreira et al. 2012). This dataset was complemented with 41 sequences from putative *L. capensis* samples from China and four additional species (*L. saxatilis*, *Lepus timidus*, *Lepus granatensis* and *Lepus europaeus*), obtained from

previous studies (Halanych et al. 1999; Matthee et al. 2004; Melo-Ferreira et al. 2007, 2011, 2012; Ramírez-Silva et al. 2010; Liu et al. 2011b). Eight nuclear DNA loci (DARC—Duffy blood group, chemokine receptor; HPX—hemopexin; SPTBN1—spectrin, beta, non-erythrocytic 1; TF—transferrin; OXA1L—oxidase assembly 1-Like; TG—thyroglobulin; TSHB—thyroid stimulating hormone beta; UCP2—uncoupling protein 2) were sequenced in 61 *L. capensis sensu lato* specimens (58 newly sequenced, and 3 from Melo-Ferreira et al. 2012). These samples represented the population clusters identified with microsatellites (except eastern Africa for which only museum samples were obtained, and which could not be amplified). This dataset was also complemented with available sequences from putative *L. capensis* from China and three other hare species (*L. granatensis*, *L. europaeus* and *L. timidus*) retrieved from GenBank (Alves et al. 2003, 2006; Matthee et al. 2004; Melo-Ferreira et al. 2009, 2011, 2012; Liu et al. 2011a, unpublished). Table S2 provides a detailed description of the final sequence datasets, including missing information.

Fragments for newly sequenced specimens were amplified using described primers (Wallner et al. 2001; Matthee et al. 2004; Melo-Ferreira et al. 2009). For 15 museum samples, only a smaller cytochrome *b* portion could be amplified, using a second primer pair designed in this work (two replicates were done to ensure the absence of contamination from exogenous DNA). Detailed primer information and PCR conditions are shown in Table S5. Purified PCR products were sequenced using the standard Sanger sequencing protocol at Macrogen Inc. (Netherlands), using both forward and reverse primers.

Population genetics analyses

Genepop (Raymond and Rousset 1995) was used to test for deviations to Hardy–Weinberg and linkage equilibria, which could be indicative of allele dropout, non-independence among loci and other biases. Markov chain parameters for exact tests were set at 10,000 dememorizations and 100 batches, with 5000 iterations per batch, and the Bonferroni correction for multiple tests was applied. Given the a priori uncertainties about population structure, which could cause spurious deviations from equilibrium, we arbitrarily divided the dataset according to the countries of origin of the samples, and performed the tests in these subsamples. For low sample sizes, samples from neighboring countries were pooled for a minimum sample size of 7 (Table S6).

The Bayesian assignment method implemented in STRUCTURE v.2.3.3 (Pritchard et al. 2000; Falush et al. 2003) was used to infer population structure, without prior assignment of specimens to populations. Given that the

model considering correlated allele frequencies performs better in separate analyses of divergent populations (Evanno et al. 2005 and STRUCTURE documentation), we performed a hierarchical structure analysis (Coulon et al. 2008; Cheng et al. 2014; Pisa et al. 2015). At each round of the analysis, the dataset was split according to the best K number of groups and re-run independently until no structure was found (i.e. no sorting of specimens per K clusters with high probability). Individuals with probability of assignment lower than 80%, i.e. potentially admixed specimens, could not be unequivocally attributed to a cluster and were discarded for the following rounds. For each K number of clusters ($K = 1$ to $K = 10$), three independent replicates of 1,000,000 generations of burn-in followed by 1,000,000 MCMC generations were run using the admixture model with correlated allele frequencies. The best number of K populations for each dataset was determined using the ΔK method (Evanno et al. 2005), as implemented in STRUCTURE HARVESTER v.0.6.94 (Earl and vonHoldt 2012).

An additional Bayesian clustering analysis was performed using TESS v2.3.1 (Chen et al. 2007; Durand et al. 2009), which takes into account the spatial distribution of samples. An admixture analysis using the conditional autoregressive (CAR) Gaussian model was performed with a linear trend degree. Ten independent replicates of 130,000 sweeps from $K = 2$ to 10 were performed, discarding the first 30,000 sweeps as burn-in. The best number of K clusters was determined based on the stabilization of average DIC (deviance information criterion) with K_{\max} . An additional 140 replicates (total 150) were performed for the best K_{\max} . Results were summarized from the 10% of replicates with the lowest DIC values, using CLUMPP (Jakobsson and Rosenberg 2007). F_{ST} between pairs of clusters estimated with Structure was calculated using FSTAT (Goudet 1995) and significance was determined through genotype randomization (correcting the $p = 0.05$ significance threshold using the Bonferroni correction for multiple tests).

Genealogical relationships among mtDNA haplotypes were determined using the Median–Joining (MJ) algorithm with software POPART v.1.7 (Leigh and Bryant 2015). In addition, spatial clustering of individuals was performed for the mtDNA dataset using the Bayesian approach implemented in BAPS 6 (Corander et al. 2008; Cheng et al. 2013). Prior upper values for the number of clusters was specified with a maximum of $K = 10$ and five independent runs.

Phylogenetic analyses

The cytochrome *b* phylogeny was determined using both Bayesian (BI) and Maximum Likelihood (ML) inference

methods, with BEAST v1.8.1 (Drummond and Rambaut 2007) and Garli v1.0 (Zwickl 2006), respectively. The European wild rabbit, *Oryctolagus cuniculus* was used as an outgroup. The best-fit model of evolution was chosen using jModeltest among 88 possible models and the AICc criterion (Darriba et al. 2012). In Garli v1.0, five independent search replicate runs were performed, specifying the optimal mutation model but not fixing the model parameters. No starting topology was defined and the program was set to run until no significantly better scoring topology was found after 50,000,000 generations. Tree support was estimated using 500 bootstrap replicates. BI was performed using BEAST v1.8.1, setting as prior the determined mutation model, or the next-most parameterized model available in BEAST when the best-fit model was not implemented. Posterior probabilities were determined using the Yule tree prior and the uncorrelated lognormal relaxed clock (Drummond et al. 2006), with three replicate runs of 100,000,000 generations, sampling trees and parameter estimates every 10,000 generations. Convergence was assessed in Tracer v1.7 (Rambaut et al. 2018).

Phylogenetic relationships among species and evolutionary groups (inferred by microsatellite analyses) were inferred from nuclear DNA sequences using the multilocus coalescent-based method *BEAST (Heled and Drummond 2010), implemented in BEAST v1.8.1. First, PHASE v2.1.2 (Stephens et al. 2001) was used to determine the phase of alleles, running 1000 generations after 1000 generations of burn-in, with a thinning interval of 1. *BEAST assumes no gene flow among the determined groups, and we minimized violation of this assumption using specimens that were determined as non-admixed by the microsatellite analyses (thus used as proxy of genomic admixture levels). Also, mtDNA sequences were not included given the suspicion of secondary mtDNA introgression (data not shown), and the widespread occurrence of this phenomenon in hares (Thulin et al. 2006; Liu et al. 2011b; Melo-Ferreira et al. 2012, 2014b; Levänen et al. 2018). Finally, we opted to maintain the full alignments, as it has been shown that recombination has minimal impacts in species-tree determination (Lanier and Knowles 2012). Specimens were assigned to distinct populations considering (i) the determined microsatellite population clusters, (ii) the putative *L. capensis* sequences from China, and (iii) three other species: *L. granatensis*, *L. europaeus* and *L. timidus*. Outgroup sequences were not included, as the method combines estimates of the root of each gene tree using the multi-species coalescent (Heled and Drummond 2010). Phylogenies were inferred from the full sequence dataset, including a string of Ns with the same length of the alignment in few cases where a locus was not available for a population (southern Africa and China for 2 and 5 loci, respectively). A separate analysis that excluded the Chinese

and southern Africa populations was performed to assess the robustness of the topology of the remaining populations/species.

Nucleotide substitution and relaxed clock models were set for each locus as described above. A Yule tree prior was used, considering each terminal branch of the tree as an independent evolutionary unit. Three independent runs of 500,000,000 generations sampling every 80,000 generations were performed. Node dating was performed using a rate of 3.17×10^{-3} substitutions per site per million years for the reference locus TF (Melo-Ferreira et al. 2012), estimated from a calibrated divergence between the European rabbit and hares of 11.8 Ma (Matthee et al. 2004).

Isolation-with-migration and coalescent simulations

The divergence history among distinct evolutionary entities was further explored using the isolation-with-migration (IM) framework (Hey 2010) and the multilocus nuclear DNA sequence dataset. This analysis was applied in three instances. First, to re-assess the divergence parameters between entities representing the deepest phylogeographic divide inferred in our phylogenetic analysis—Arabian Peninsula and north-western Africa populations. Second, to re-estimate the divergence between southern and northern African *L. capensis*. Third, to clarify the apparent discordance between nuclear and mtDNA divergence between the Near East *L. capensis* population and *L. europaeus*—close mtDNA relationship but deep nuclear divergence. IMA2 (Hey 2010) was used to infer effective population sizes (populations 1, 2 and ancestral), divergence time and gene flow rates (Hey and Nielsen 2004) in the three pairwise analyses. Datasets were reduced to the largest non-recombining blocks using IMgc (Woerner et al. 2007), which has been shown to reduce biases in the final IM estimates (Strasburg and Rieseberg 2010). Three independent runs were performed, varying the parameters' upper bound priors and the starting seeds and using the HKY mutation model (Hasegawa et al. 1985). A likelihood ratio test was applied to assess whether migration was significantly different from zero (Nosil et al. 2009). The locus-specific mutation rates estimated by Melo-Ferreira et al. (2012) were used, except for the TG locus, for which it was newly determined using the same methodology.

To investigate whether the cyto-nuclear discordance inferred for the Near East *L. capensis* population may have resulted from incomplete lineage sorting alone, coalescent simulations were done with SIMCOAL V2.1.2 (Laval and Excoffier 2004). 10,000 cytochrome *b* datasets were simulated under a coalescent model with no gene flow, using the divergence time (t) and effective population sizes (N_e) modeled with IMA2 for Near East and *L. europaeus*, and the mutation rate determined by Melo-Ferreira et al. (2012) for

cytochrome *b*. An ancestral haploid population of size $Ne_A/2$ was simulated to split into two descendant populations of sizes $Ne_1/2$ and $Ne_2/2$, t generations ago, with no gene flow occurring after the split. For each simulation, the minimum uncorrected sequence divergence (D_{xy}) between descendant populations was calculated and used to build the distribution of minimum expected distances.

Results

Population structure

The 18 newly discovered microsatellite loci were successfully genotyped in 162 *L. capensis sensu lato* samples. The number of alleles per locus varied between 8 and 28 (see Table S4). Considering nine data partitions based on geographic proximity, 3.7% and 0.6% of the tests rejected conformation to Hardy–Weinberg and linkage equilibria, respectively. Given that significant deviations were sporadic across loci and not consistent across partitions, all loci were retained for subsequent analyses.

The STRUCTURE clustering analyses supported the subdivision in two clusters ($K = 2$) at the uppermost hierarchical level: a predominantly north-western population (hereafter NW Africa), and a second cluster with the remaining individuals (Figs. 1, S1 and S2). The following round divided the second group in 4 sub-clusters ($K = 4$)—East Africa (E Africa), the Near East, Arabia and South Africa (S Africa) (Figs. 1, S1 and S2). In the NW Africa group, vast mixed assignment of individuals to K clusters with no geographic partition was found, and thus no further subdivision was considered (Figs. 1, S1 and S2). After both rounds of hierarchical analysis, 140 specimens were attributed to a single cluster, while 22 showed mixed assignment probability to distinct clusters (i.e. <0.8 probability) (Table S2).

The spatial Bayesian clustering analysis suggested $K = 5$ as the best number of clusters (Fig. S2), which however corresponded to four partitions with proportions of individual assignment >0.8 (the fifth cluster had no assigned individuals). Results were consistent across replicate runs, from $K_{max} = 4$ to 10. Spatial clustering of African samples was compatible with STRUCTURE results, suggesting clusters in NW Africa, E Africa and S Africa, but placing the Near East and Arabia specimens in a single group. A separate analysis with these specimens and $K_{max} = 2$, confirmed the STRUCTURE separation in Near East and Arabia (Fig. S1). No signs of error-driven batch effects for museum samples were found, as these were assigned logically according to geography.

Pairwise F_{ST} among the five inferred clusters varied between 0.11 (Near East vs. NW Africa) and 0.27 (S Africa

vs. E Africa) (Table S7). Differentiation was generally significant, except between Arabia and East Africa (note that the significance of F_{ST} involving southern Africa could not be assessed due to low sample size).

The BAPS analysis of mtDNA structure suggested six evolutionary groups for *L. capensis sensu lato*: three groups in NW Africa, S Africa, Near East/East Africa and Arabia (Fig. S3). The mtDNA haplotype network (Fig. S4) nevertheless showed genealogical separation of Near East and East Africa haplotypes and of the haplotype from Mali (which represents the fourth lineage from NW Africa). We therefore mapped the distribution of eight maternal haplogroups—NWA-I to NWA-IV, S Africa, E Africa, Near East, Arabia (Fig. 1). These eight mtDNA haplogroups coincide with the evolutionary units identified using microsatellites, with the exception of the NW Africa, for which mtDNA suggested four different, yet closely related, sublineages (NW Africa I–IV; Figs. 1 and S4).

Phylogenetic analyses

Eight nuclear and one mitochondrial DNA loci were sequenced (GenBank accession numbers MK775981–MK776533) and aligned together with sequences retrieved from GenBank (Table S2), for a total of 5164 bp (see Table S5). The inclusion of sequences from other *Lepus* species in the cytochrome *b* phylogeny showed that the eight lineages sampled in *L. capensis sensu lato* (Fig. 1) do not form a monophyletic group (Figs. 2 and S5): the Near East and E Africa lineages are closely related to *L. europaeus*, and the putative *L. capensis* from China to *L. timidus*. The mtDNA phylogeny excluding museum samples and using the longer DNA fragment showed a similar topology (but removed the East Africa haplotypes) (Fig. S5).

In the nuclear loci, 86% of the sequence length corresponded to introns. The multi-locus coalescent-based phylogeny inferred with the nuclear DNA dataset suggested a major phylogenetic divide that splits African from Asian and European *Lepus* lineages around 1 million years ago (Ma) (Fig. 3). The Near East and Arabian populations were found to share a very recent common ancestor, appearing more closely related to putative *L. capensis* from China and the Eurasian hare species (*L. europaeus*, *L. granatensis* and *L. timidus*), than to the remaining African *L. capensis* lineages (NW Africa and S Africa; note that E Africa specimens could not be included in this analysis). Even if posterior probabilities are not high for some nodes, the grouping of non-African lineages of *L. capensis sensu lato* with other hare species was inferred with high probability (0.98; Fig. 3a). This phylogenetic pattern remains when removing the lineages for which not all loci were sequenced (*L. capensis* populations from China and S Africa) (Fig. 3b).

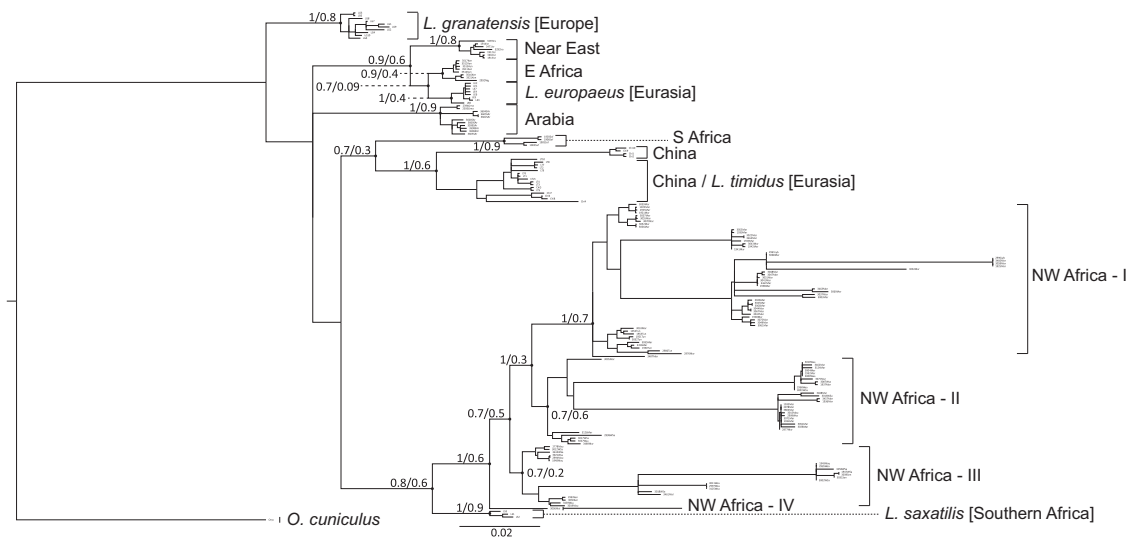


Fig. 2 Bayesian inference mtDNA phylogeny (cytochrome *b*) of *L. capensis*, *L. saxatilis* and Eurasian hare species (*L. timidus*, *L. europaeus* and *L. granatensis*), rooted by a rabbit (*Oryctolagus cuniculus*) haplotype. Bayesian posterior probabilities and maximum likelihood

bootstrap supports (scaled from 0 to 1) are shown next to nodes when the first is above 0.50 (Bayesian/Maximum likelihood inferences). *L. capensis* clades agree with those depicted in Fig. 1

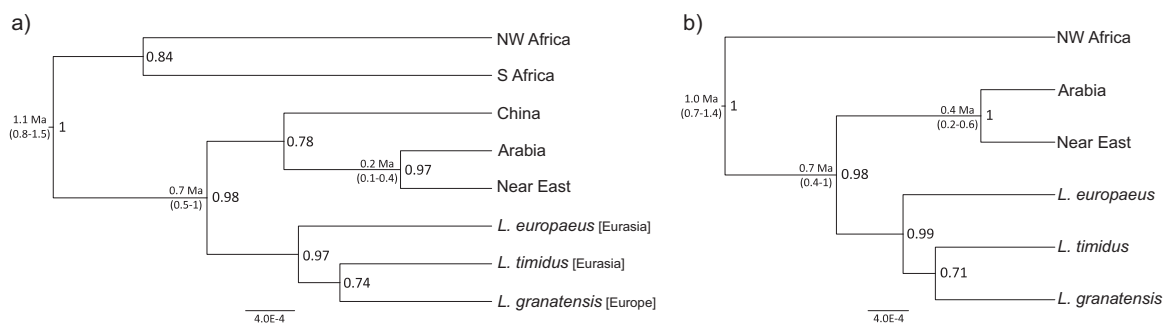


Fig. 3 Nuclear DNA species tree inferred from eight nuclear loci (posterior probabilities are shown on the right of each node), including **a** African and non-African populations of *L. capensis* and Eurasian *Lepus* species and **b** excluding South African and Chinese populations,

for which not all loci were sequenced. Estimates of split times in units of million years (Ma) are indicated for *L. capensis sensu lato* nodes supported by posterior probabilities above 0.95 (95% confidence intervals shown in brackets)

Isolation-with-migration and coalescent simulations

Estimates retrieved from the isolation-with-migration analyses were consistent across replicate runs (Table 1). Keeping the non-recombining blocks reduced the dataset to 8–18% of the alignment length. Among the analyzed pairs of populations, nuclear gene flow was not significantly different from zero. Even if in some instances the posterior density curves did not allow to estimate the 95% confidence intervals of the divergence times, point estimates were consistent with the inferences based on the multi-locus species tree (Fig. 3).

In order to determine whether the low divergence between Near East *L. capensis* and *L. europaeus* inferred for mtDNA was compatible with divergence history inferred from the nuclear DNA, we used these parameters to simulate the expected Cytb divergence under a model with

no gene flow. The empirical pairwise mtDNA divergence was found to be smaller than the 5th percentile of the simulated minimum distances in all instances (Fig. 4), suggesting that incomplete lineage sorting does not explain the close mtDNA relationship between European *L. europaeus* and Near East *L. capensis* in our dataset (see Melo-Ferreira et al. 2012, 2014b).

Discussion

Population structure with deep divergence in *Lepus capensis sensu lato*

Our microsatellite analyses suggest that *L. capensis sensu lato* is composed of five evolutionary units with geographic structure—north-western Africa, southern Africa, eastern

Table 1 Isolation-with-migration maximum likelihood estimates (95% posterior density intervals in parentheses, when estimated) of demographic parameters obtained with IMA2 between three pairs of populations

Pop.1	Pop.2	$Ne1^a$	$Ne2^a$	NeA^a	t^b	$2Nm1 \rightarrow 2^c$	$2Nm2 \rightarrow 1^c$
NW Africa	Arabia	1,080,674 (830,014; 1,411,733)	177,353 (96,953; 190,859)	63,847 –	716,508 (489,495; 1,000,746)	0.02 (n.s.) (0.00; 0.81)	0.00 (n.s.) –
S Africa	NW Africa	218,983 –	1,107,427 (798,766; 1,548,868)	68,823 –	767,482 (433,794; 1,190,153)	0.01 (n.s.) –	0.15 (n.s.) –
Near East	<i>L. europaeus</i>	352,933 (207,503; 686,358)	161,392 (83,356; 278,445)	221,691 –	860,755 –	0.01 (n.s.) –	0.06 (n.s.) –

^aEffective population size of population 1 ($Ne1$), 2 ($Ne2$) and the ancestral population (NeA)

^bTime in years since species 1 and 2 split

^cPopulation migration rate into population 1 ($2Nm2 \rightarrow 1$) and population 2 ($2Nm1 \rightarrow 2$), n.s.—not significant

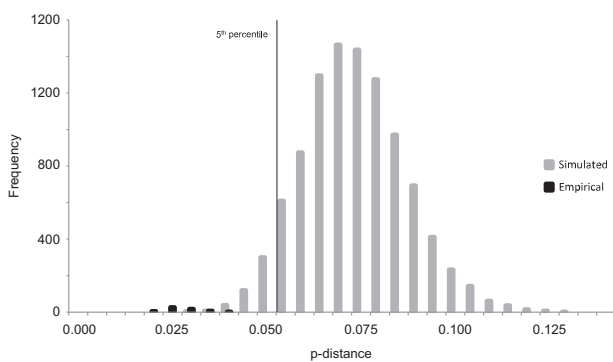


Fig. 4 Distribution of empirical and simulated uncorrected p -distances between European *L. europaeus* and the Near East *L. capensis* population. Simulations were conducted based on population parameters estimated with the isolation-with-migration multi-locus analysis. Gray bars show the distribution of minimum pairwise uncorrected p -distances per simulation between *L. europaeus* and the Near East *L. capensis* population (the vertical line indicates the 5th percentile) and black bars depict the empirical pairwise p -distances between the same populations

Africa, Arabia and the Near East (Fig. 1). These results are concordant with the insights of Hoffmann and Smith (2005), who restricted *L. capensis* to Southern Africa, and informally suggested three species-level division for the remaining *capensis*-type hares—Northwest Africa, East Africa and Arabia-Near East. Given the gaps in our sampling scheme, we cannot completely exclude that smoother transitions of genetic structure could occur between evolutionary units. However, our multilocus coalescent-based nuclear DNA phylogeny supports two major *L. capensis sensu lato* clades, and does not retrieve the species as monophyletic when including other *Lepus* lineages, as the Near East and Arabian Peninsula specimens appear more closely related to Eurasian hare species (Fig. 3). This major phylogenetic/phylogeographic divide could not be created by spurious structure from spatial correlation of our gapped

opportunistic sampling scheme. These results therefore show that the Asian *L. capensis* groups have an independent evolutionary origin from their African counterparts (Fig. 3). The relevant uncertainties of molecular dating calibrations (Ho et al. 2015) indicate caution is needed when interpreting the absolute inferred ages of the divergence events. Still, our analyses suggest that the African–Asian split is the oldest divergence event, and possibly occurred during the Early Pleistocene (Fig. 3). We could not, however, include East African samples in our nuclear DNA phylogeny. The microsatellite analysis and mtDNA phylogeny suggest that this is a separate group, and potentially more closely related with the Near East haplotypes (Figs. 1 and 2). If that is the case, the major phylogenetic divide in *L. capensis sensu lato* may therefore run across East Africa. Regardless, our results confirm that the classification of the Asian evolutionary units as *L. capensis* is controversial (Hoffmann and Smith 2005), and indicate that a reassessment of the relationships of these populations with other currently classified Asian hare species, such as *L. tolai* or *L. tibetanus* (to which the sequences obtained from specimens from China recovered here may belong to) is needed. Our work further suggests the distinction of the Near East and Arabian evolutionary units, in line with the reported morphological uniqueness of the Arabian hares (Drew 2000; Zubair et al. 2011). Whether this would merit taxonomic recognition requires further investigation of the biology, ecology and genetics of these groups.

In Africa, the Cape hare has a disjunct northern and southern distribution and our analyses recover these as separate evolutionary units (Figs. 1 and 3). This is supported by the absence of admixture in all instances of our analyses—microsatellites (Fig. S1), mtDNA (Fig. S4) and IM model applied to the multilocus dataset (Table 1). The split between northern and southern Africa *L. capensis* may have occurred during the Middle Pleistocene, in a period of strong climate oscillations (Brown et al. 2007). However,

given the small sample size of the Southern population, further analyses with larger sample sets are needed to clarify its degree of differentiation and divergence.

In northern Africa, the microsatellite analyses of our comprehensive sampling did not show geographic substructure (Figs. 1 and S1). This result is surprising given the environmental and topographic heterogeneity of the range occupied by the species in the region. Indeed, the analyses of mtDNA variation suggest subdivision into several geographically structured haplogroups (Figs. 1 and 2). While the haploid nature and uniparental transmission of mtDNA can contribute to faster lineage sorting at local geographic scales, mtDNA structure could have been reinforced by female philopatry and male-mediated dispersal. The female-linked marker can thus mark the regional historical structure of the species, which was homogenized at biparentally transmitted loci by recent male-mediated gene flow (Fahey et al. 2014). Female philopatry has been suggested in several hare species (Fickel et al. 2005; Hamill et al. 2007; Melo-Ferreira et al. 2014a) including in southern Africa *L. capensis* and *L. saxatilis* (Kryger et al. 2002). In addition, we cannot exclude that some of these clades may have introgressed from an unsampled neighboring species.

Secondary mitochondrial DNA introgression

Our results suggest that the mtDNA haplotypes sampled in the Near East and East Africa regions were more closely related to the European brown hare (*L. europaeus*) than expected considering the inferred model of divergence of the species (Figs. 2–4). The Cape hare specimens from the Near East are close to the possible contact area between *L. capensis* and *L. europaeus*, and our simulations suggest that the mtDNA similarity may result from introgression (Fig. 4). This suspicion adds to numerous works showing that mtDNA introgression is a pervasive phenomenon among hare species, both in current and historical contacts (e.g. Melo-Ferreira et al. 2005, 2014b). However, understanding whether these results reflect a smoother local transition with gene flow between the currently classified species in the region (see Ben Slimen et al. 2008b) demands a detailed analysis of regional genetic variation, with precise estimates of differentiation and divergence at the nuclear and mtDNA levels. Another instance of cyto-nuclear discordance of phylogenetic patterns concerns the relatedness of putative *L. capensis* from China to *L. timidus* for mtDNA. Studies based on mtDNA have reported this similarity (Yu 2004; Wu et al. 2005) which could result from mtDNA introgression (Alves et al. 2006, 2008), regardless of the most appropriate taxonomic classification of the specimens from China (Hoffmann and Smith 2005; Liu et al. 2011b; Cheng et al. 2012). Even though we could not formally explore this, our analyses show a discordance of mtDNA and

nuclear phylogenies (Figs. 2 and 3), which may result from mtDNA introgression.

Biogeographic insights from comparative phylogeography

Our multilocus nuclear DNA phylogeny of *L. capensis sensu lato* suggested that the most ancient phylogenetic divide separates the African (north-western and southern) from the Asian (Near East, Arabian) populations (Fig. 3). The transition from Africa to Asia has been shown to be an important phylogeographic barrier for several animals, with estimates of segregation time varying from the Pliocene to the late Pleistocene, for example in jackals (Koepfli et al. 2015), cheetahs (Charruau et al. 2011), African desert jerboas (Ben Faleh et al. 2012) or geckos (Metallinou et al. 2012). The progressive aridification of the Saharo-Arabian region following the desert onset after the Late Miocene–Pliocene likely imposed strong dispersal barriers (Pokorný et al. 2015; Mairal et al. 2017). The alternation between arid and humid conditions during the Pleistocene possibly allowed dispersal events between north-eastern Africa and the Arabian Peninsula (Stewart et al. 2017), which may explain the time-frame of the divergence inferred here for *L. capensis* (~1 Ma). For some groups, the Asian clade is closely related to the eastern Africa one, such as for the helmeted terrapin (Wong et al. 2010), or Dorcas gazelles (Lerp et al. 2011). Our lack of sampling from the East Africa group for multilocus phylogenetic analyses does not allow clarifying the possible link between East African and the Near East, but the mtDNA data seems to point in that direction (Fig. 2), notwithstanding the possibility of mtDNA introgression from *L. europaeus* (see above). The maintenance of dispersal corridors along the Nile and the Red Sea mountains may have maintained North–South ecological corridors (Metallinou et al. 2012; Brito et al. 2014) (Fig. 1). The inferred admixture of our Niger sample in the microsatellite analysis and its inclusion in the eastern Africa mtDNA clade may support East–West dispersal corridors of savannah found immediately south of the desert, as suggested by vicariance-expansion cycles inferred for murid rodents (Brouat et al. 2009; Dobigny et al. 2013).

In Africa, we found three major evolutionary units for the Cape hare—North, East and South (Fig. 1). These phylogeographic divides are remarkably similar to those inferred for species associated with African savannah, such as ungulates (Lorenzen et al. 2012), lions (Bertola et al. 2016), giraffes (Brown et al. 2007; Fennessy et al. 2016) or rodents (Granjon et al. 2012; Bryja et al. 2014). This suggests species persistence in refugia created by the expansion and contraction of tropical forests and consequent savannah fragmentation, mostly associated with changes in precipitation regimes along the glacial cycles (Dupont 2011).

Interestingly, despite similar geographic structures, estimated times of diversification vary among taxa, from the Pliocene to Pleistocene. This suggests repeated contraction–expansions along Plio-Pleistocene, possibly resulting in deeper divergence in less mobile species (such as rodents) and more recent divergence times for more mobile mammals (such as hares or ungulates).

The marked phylogeographic structure we found for mtDNA in North Africa (Fig. 1) provides important insights on the complexity of Sahara-Sahel biogeographical patterns, even though this structure appeared diluted at biparentally inherited markers (Figs. 1 and S1). The largest mtDNA haplogroup, NWA-I (Figs. 1 and 2), covers most of the Maghreb, and expands throughout Tunisia and Libya. This suggests that the Atlas Mountains are permeable to gene flow in hares, which contrasts with animals with low dispersal abilities (Brown et al. 2002). Mitochondrial DNA clade NWA-II (Figs. 1 and 2) appears endemic to the Atlantic Coastal Sahara, from the southern slopes of the Atlas Mountains to the Tagant Mountain of Mauritania. Though delimited by the desert, this region benefits from humid winds, and is prone to harbor endemic forms for several species, such as lizards (Velo-Anton et al. 2018) and snakes (Goncalves et al. 2018). The transition between clades II and III does not coincide with apparent landscape barriers to gene flow (terrain is mostly flat and the river present in the area, Lakra, is dry most of the year) and may represent a secondary contact from populations temporarily restricted to refugia. A similar pattern was found in the Schokari sand racer (Goncalves et al. 2018). Clade NW-IV is represented by a single haplotype in our work (Figs. 1 and 3), sampled in Central Sahel, Mali, and may represent a distinct lineage delimited by the Niger river (e.g. Dobigny et al. 2013), as suggested in other river systems (e.g. Brouat et al. 2009). Further sampling is needed to clarify this distinction. To summarize, the complex physiography of the Sahara-Sahel with its mountain chains, river systems and Atlantic and Mediterranean influence, seems to have promoted differentiation, generating a heterogeneous gene pool that may have been homogenized at the nuclear DNA level by male-mediated dispersal, likely along the Atlantic Sahara corridor (Fig. 1).

Conclusions

The inferred evolutionary history of *L. capensis sensu lato* over its range is complex, with differentiation, deep divergence of African and Asian populations, and instances of interspecific gene flow, possibly more pronounced at mtDNA. The phylogeographic structure of the Cape hare finds parallel in other taxa, particularly those associated with African open savannah habitats, with distinct lineages found in North, East and South Africa. Furthermore, it

suggests that the complex environmental heterogeneity in coastal areas of North Africa and of the Sahara-Sahel region promoted differentiation, which appears to have been eroded at the nuclear DNA possibly by male-biased dispersal. In Northeast Africa, the deep phylogeographic break may be associated with cyclic barriers created by the aridification of the Saharo-Arabian desert. Our results suggest that *L. capensis sensu lato* is not monophyletic and current taxonomy does not reflect the complexity of its evolutionary history. This study provides valuable information to guide future taxonomic revisions, which will be important to foster the conservation of the evolutionary groups. Filling population sampling gaps in future studies, including of neighboring species, will provide a deeper understanding of the transition between population groups, and allow a powerful quantification of the magnitude of genetic isolation barriers that may persist among lineages. This study shows that assessing the evolutionary history of species using multilocus approaches allows precise inferences of population divergence, taking into account and measuring gene flow. Such inferences contribute to an improved quantification and characterization of biodiversity, and to the definition of appropriate conservation units.

Data archiving

Newly obtained DNA sequences were deposited in GenBank with accession numbers MK775981–MK776533. The microsatellite genotypes and full DNA alignments were deposited in Dryad: <https://doi.org/10.5061/dryad.th71ss0>.

Acknowledgements This work was financially supported by Portuguese national funds through the Fundação para a Ciência e a Tecnologia (FCT; project HybridAdapt, FCT-ANR/BIA-EVF/0250/2012). JCB and JM-F were supported by Programa Operacional Potencial Humano-Quadro de Referência Estratégico Nacional (POPH-QREN) funds from the European Social Fund and Portuguese Ministério da Ciência, Tecnologia e Ensino Superior through FCT (IF/00459/2013 and IF/00033/2014 Investigador FCT research contracts, respectively). All samples used in this work were collected before the Nagoya Protocol entered into force. Museum samples from the Natural History Museum in Berlin were collected under the financial support provided by SYNTHESYS grant DE-TAF-4131 (funded by the European Union under FP7 grant agreement 226506). North African field samples were collected with funds from National Geographic Society (CRE-7629-04, CRE-8412-08), Mohammed bin Zayed Species Conservation Fund (11052709, 11052707, 11052499, 13257467), Fundação para a Ciência e Tecnologia (PTDC/BIA-BEC/099934/2008, PTDC/BIA-BIC/2903/2012), and FEDER through COMPETE-Operational Programme for Competitiveness Factors (FCOMP-01-0124-FEDER-008917, -028276) to JCB. Additional support was obtained from project NORTE-01-0145-FEDER-000007 supported by the Norte Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). The authors thank MS Shah for sampling in Saudi Arabia. Franz Suchentrunk, Conrad Matthee, João Maia, Raquel Vasconcelos, Teresa Luísa Silva, Lahoussine Ouragh, the Natural History Museum in Berlin, and the Natural History

Museum in Vienna kindly contributed with samples for this study (voucher codes for the museum specimens are in Table S3). The authors thank the curators of these Natural History Museums for support. The authors also thank James Harris for his kind revision of the manuscript.

Author contributions JM-F and PCA devised the study with insights from JCB. SL and PCA sampled in the Natural History Museums. JCB led field expeditions to North Africa. MZI led sampling in Saudi Arabia. SL performed all laboratory work and analyses under the supervision of JM-F. SL and JM-F wrote the paper. All authors discussed and interpreted the results, and read, revised and approved the final version of the manuscript. Data archiving Newly obtained DNA sequences were deposited in GenBank with accession numbers MK775981–MK776533. The microsatellite genotypes and full DNA alignments were deposited in Dryad: <https://doi.org/10.5061/dryad.th71ss0>.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Adams M, Raadik TA, BurrIDGE CP, Georges A (2014) Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? *Syst Biol* 63:518–533
- Aghová T, Šumbera R, Piálek L, Mikula O, McDonough MM, Lavrenchenko LA et al. (2017) Multilocus phylogeny of East African gerbils (Rodentia, Gerbilliscus) illuminates the history of the Somali-Masai savanna. *J Biogeogr* 44:2295–2307
- Alves PC, Ferrand N, Suchentrunk F, Harris DJ (2003) Ancient introgression of *Lepus timidus* mtDNA into *L. granatensis* and *L. europaeus* in the Iberian Peninsula. *Mol Phylogenet Evol* 27:70–80
- Alves PC, Harris DJ, Melo-Ferreira J, Branco M, Ferrand N, Suchentrunk F et al. (2006) Hares on thin ice: introgression of mitochondrial DNA in hares and its implications for recent phylogenetic analyses. *Mol Phylogenet Evol* 40:640–641
- Alves PC, Melo-Ferreira J, Freitas H, Boursot P (2008) The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*. *Philos Trans R Soc B* 363:2831–2839
- Avise JC (2009) Phylogeography: retrospect and prospect. *J Biogeogr* 36:3–15
- Ben Faleh A, Granjon L, Tatar C, Boratyński Z, Cosson JF, Said K (2012) Phylogeography of two cryptic species of African desert jerboas (Dipodidae: Jaculus). *Biol J Linn Soc* 107:27–38
- Ben Slimen H, Suchentrunk F, Memmi A, Sert H, Kryger U, Alves PC et al. (2006) Evolutionary relationships among hares from North Africa (*Lepus* sp or *Lepus* spp.), cape hares (*L. capensis*) from South Africa, and brown hares (*L. europaeus*), as inferred from mtDNA PCR-RFLP and allozyme data. *J Zool Syst Evol Res* 44:88–99
- Ben Slimen H, Suchentrunk F, Shahin AB, Ben Ammar Elgaaied A (2007) Phylogenetic analysis of mtCR-1 sequences of Tunisian and Egyptian hares (*Lepus* sp. or spp., Lagomorpha) with different coat colours. *Mamm Biol* 72:224–239
- Ben Slimen H, Suchentrunk F, Ben Ammar Elgaaied A (2008a) On shortcomings of using mtDNA sequence divergence for the systematics of hares (genus *Lepus*): an example from cape hares. *Mamm Biol* 73:25–32
- Ben Slimen H, Suchentrunk F, Stamatis C, Mamuris Z, Sert H, Alves PC et al. (2008b) Population genetics of cape and brown hares (*Lepus capensis* and *L. europaeus*): a test of Petter's hypothesis of conspecificity. *Biochem Syst Ecol* 36:22–39
- Bertola LD, Jongbloed H, van der Gaag KJ, de Knijff P, Yamaguchi N, Hooghiemstra H et al. (2016) Phylogeographic patterns in Africa and high resolution delineation of genetic clades in the lion (*Panthera leo*). *Sci Rep* 6:30807
- Bi K, Linderoth T, Vanderpool D, Good JM, Nielsen R, Moritz C (2013) Unlocking the vault: next-generation museum population genomics. *Mol Ecol* 22:6018–6032
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K et al. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155
- Brito JC, Godinho R, Martinez-Freiria F, Pleguezuelos JM, Rebelo H, Santos X et al. (2014) Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel. *Biol Rev* 89:215–231
- Brouat C, Tatar C, Bâ K, Cosson J-F, Dobigny G, Fichet-Calvet E et al. (2009) Phylogeography of the Guinea multimammate mouse (*Mastomys erythroleucus*): a case study for Sahelian species in West Africa. *J Biogeogr* 36:2237–2250
- Brown DM, Brenneman RA, Koepfli KP, Pollinger JP, Mila B, Georgiadis NJ et al. (2007) Extensive population genetic structure in the giraffe. *BMC Biol* 5:57
- Brown RP, Suárez NM, Pestano J (2002) The Atlas mountains as a biogeographical divide in North-West Africa: evidence from mtDNA evolution in the Agamid lizard *Agama impalearis*. *Mol Phylogenet Evol* 24:324–332
- Bryja J, Mikula O, Šumbera R, Meheretu Y, Aghová T, Lavrenchenko LA et al. (2014) Pan-African phylogeny of *Mus* (subgenus *Nannomys*) reveals one of the most successful mammal radiations in Africa. *BMC Evol Biol* 14:256
- Charruau P, Fernandes C, Orozco-Terwengel P, Peters J, Hunter L, Ziaie H et al. (2011) Phylogeography, genetic structure and population divergence time of cheetahs in Africa and Asia: evidence for long-term geographic isolates. *Mol Ecol* 20:706–724
- Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* 7:747–756
- Cheng C, Ge D, Xia L, Zhou C, Yang Q (2012) Morphometrics study on the so called “Cape hare” (Lagomorpha: Leporidae: *Lepus*) in China. *Acta Theriol Sin* 32:275–286
- Cheng E, Hodges KE, Melo-Ferreira J, Alves PC, Scott Mills L (2014) Conservation implications of the evolutionary history and genetic diversity hotspots of the snowshoe hare. *Mol Ecol* 23:2929–2942
- Cheng YT, Liu J, Yang LQ, Sun C, Kong QP (2013) Mitochondrial DNA content contributes to climate adaptation using Chinese populations as a model. *PLoS ONE* 8:e79536
- Corander J, Marttinen P, Siren J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539
- Coulon A, Fitzpatrick JW, Bowman R, Stith BM, Makarewicz CA, Stenzler LM et al. (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). *Mol Ecol* 17:1685–1701
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772
- Dobigny G, Tatar C, Gauthier P, Ba K, Duplantier JM, Granjon L et al. (2013) Mitochondrial and nuclear genes-based phylogeography of *Arvicanthis niloticus* (Murinae) and sub-Saharan open habitats pleistocene history. *PLoS ONE* 8:e77815

- Drew C (2000) The distribution of the Cape Hare, *Lepus capensis*, in Abu Dhabi Emirate, United Arab Emirates. *Zool Middle East* 20:15–20
- Drew C, O'Donovan D, Simkins G, Al Dosary M, Al Khaldi AM, Mohammed OB et al. (2008) *Lepus capensis*. The IUCN Red List of threatened species. e.T41277A10429185
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4:e88
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
- Dupont LM (2011) Orbital scale vegetation change in Africa. *Quat Sci Rev* 30:3589–3602
- Durand E, Jay F, Gaggiotti OE, Francois O (2009) Spatial inference of admixture proportions and secondary contact zones. *Mol Biol Evol* 26:1963–1973
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Fahey AL, Ricklefs RE, Dewoody JA (2014) DNA-based approaches for evaluating historical demography in terrestrial vertebrates. *Biol J Linn Soc* 112:367–386
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587
- Fennessy J, Bidon T, Reuss F, Kumar V, Elkan P, Nilsson MA et al. (2016) Multi-locus analyses reveal four giraffe species instead of one. *Curr Biol* 26:2543–2549
- Fickel J, Schmidt A, Putze M, Spittler H, Ludwig A, Streich WJ et al. (2005) Genetic structure of populations of European brown hare: implications for management. *J Wildl Manag* 69:760–770
- Flux JEC, Angermann R (1990) The hares and jackrabbits. In: Chapman JA, Flux JEC (eds) Rabbits, hares and pikas. Status survey and conservation action plan. IUCN, Switzerland
- Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol Evol* 27:480–488
- Goncalves DV, Martinez-Freiria F, Crochet PA, Geniez P, Carranza S, Brito JC (2018) The role of climatic cycles and trans-Saharan migration corridors in species diversification: biogeography of *Psammophilus schokari* group in North Africa. *Mol Phylogenet Evol* 118:64–74
- Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* 86:485–486
- Granjon L, Colangelo P, Tatar C, Colyn M, Dobigny G, Nicolas V (2012) Intrageneric relationships within *Gerbilliscus* (Rodentia, Muridae, Gerbillinae), with characterization of an additional West African species. *Zootaxa* 3325:1–25
- Gutiérrez-García TA, Vázquez-Domínguez E (2011) Comparative phylogeography: designing studies while surviving the process. *BioScience* 61:857–868
- Halanych KM, Demboski JR, van Vuuren BJ, Klein DR, Cook JA (1999) Cytochrome b phylogeny of North American hares and jackrabbits (*Lepus*, Lagomorpha) and the effects of saturation in outgroup taxa. *Mol Phylogenet Evol* 11:213–221
- Hamill RM, Doyle D, Duke EJ (2007) Microsatellite analysis of mountain hares (*Lepus timidus hibernicus*): low genetic differentiation and possible sex-bias in dispersal. *J Mammal* 88:784–792
- Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22:160–174
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Mol Biol Evol* 27:570–580
- Hey J (2010) Isolation with migration models for more than two populations. *Mol Biol Evol* 27:905–920
- Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167:747–760
- Ho SYW, Tong KJ, Foster CSP, Ritchie AM, Lo N, Crisp MD (2015) Biogeographic calibrations for the molecular clock. *Biol Lett* 11:20150194
- Hoffmann RS, Smith A (2005) Order Lagomorpha. In: Wilson DE, Reeder DM (eds) Mammal species of the world (third edition). Johns Hopkins University Press, Baltimore
- 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
- Kingdon J (2013) Mammalian evolution in Africa. In: Kingdon J, Butynski TM, Happold DCD, Happold M, Hoffmann M (eds) The mammals of Africa, introductory chapters and Afrotheria. Bloomsbury, Amsterdam
- Koepfli KP, Pollinger J, Godinho R, Robinson J, Lea A, Hendricks S et al. (2015) Genome-wide evidence reveals that African and Eurasian Golden Jackals are distinct species. *Curr Biol* 25:2158–2165
- Kryger U, Robinson TJ, Bloomer P (2002) Isolation and characterization of six polymorphic microsatellite loci in South African hares (*Lepus saxatilis* F. Cuvier, 1823 and *Lepus capensis* Linnaeus, 1758). *Mol Ecol Notes* 2:422–424
- Lanier HC, Knowles LL (2012) Is recombination a problem for species-tree analyses? *Syst Biol* 61:691–701
- Laval G, Excoffier L (2004) SIMCOAL 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. *Bioinformatics* 20:2485–2487
- Leigh JW, Bryant D (2015) popart: full-feature software for haplotype network construction. *Methods Ecol Evol* 6:1110–1116
- Lerp H, Wronski T, Pfenninger M, Plath M (2011) A phylogeographic framework for the conservation of Saharan and Arabian Dorcas gazelles (*Artiodactyla*: Bovidae). *Org Divers Evol* 11:317
- Levänen R, Thulin C-G, Spong G, Pohjoismäki JLO (2018) Wide-spread introgression of mountain hare genes into Fennoscandian brown hare populations. *PLoS ONE* 13:e0191790
- Liu J, Chen P, Yu L, Wu SF, Zhang YP, Jiang XL (2011a) The taxonomic status of *Lepus melainus* (Lagomorpha: Leporidae) based on nuclear DNA and morphological analyses. *Zootaxa* 3010:47–57
- Liu J, Yu L, Arnold ML, Wu CH, Wu SF, Lu X et al. (2011b) Reticulate evolution: frequent introgressive hybridization among Chinese hares (genus *Lepus*) revealed by analyses of multiple mitochondrial and nuclear DNA loci. *BMC Evol Biol* 11:223
- Lorenzen ED, Heller R, Siegmund HR (2012) Comparative phylogeography of African savannah ungulates. *Mol Ecol* 21:3656–3670
- Mairal M, Sanmartín I, Pellissier L (2017) Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora. *J Biogeogr* 44:911–923
- Matthee CA, van Vuuren BJ, Bell D, Robinson TJ (2004) A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. *Syst Biol* 53:433–447
- Melo-Ferreira J, Alves PC, Freitas H, Ferrand N, Boursot P (2009) The genomic legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Mol Ecol* 18:2643–2658
- Melo-Ferreira J, Alves PC, Rocha J, Ferrand N, Boursot P (2011) Interspecific X-chromosome and mitochondrial DNA

- introgression in the Iberian hare: selection or allele surfing? *Evolution* 65:1956–1968
- Melo-Ferreira J, Boursot P, Carneiro M, Esteves PJ, Farello L, Alves PC (2012) Recurrent introgression of mitochondrial DNA among hares (*Lepus* spp.) revealed by species-tree inference and coalescent simulations. *Syst Biol* 61:367–381
- Melo-Ferreira J, Boursot P, Randi E, Kryukov A, Suchentrunk F, Ferrand N et al. (2007) The rise and fall of the mountain hare (*Lepus timidus*) during Pleistocene glaciations: expansion and retreat with hybridization in the Iberian Peninsula. *Mol Ecol* 16:605–618
- Melo-Ferreira J, Boursot P, Suchentrunk F, Ferrand N, Alves PC (2005) Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Mol Ecol* 14:2459–2464
- Melo-Ferreira J, Farello L, Freitas H, Suchentrunk F, Boursot P, Alves PC (2014a) Home-loving boreal hare mitochondria survived several invasions in Iberia: the relative roles of recurrent hybridisation and allele surfing. *Heredity* 112:265–273
- Melo-Ferreira J, Seixas FA, Cheng E, Mills LS, Alves PC (2014b) The hidden history of the snowshoe hare, *Lepus americanus*: extensive mitochondrial DNA introgression inferred from multilocus genetic variation. *Mol Ecol* 23:4617–4630
- Metallinou M, Arnold EN, Crochet P-A, Geniez P, Brito JC, Lymberakis P et al. (2012) Conquering the Sahara and Arabian deserts: systematics and biogeography of *Stenodactylus* geckos (Reptilia: Gekkonidae). *BMC Evol Biol* 12:258
- Moores R, Brown D, Martin R, Lees AC (2012) Status and identification of hares *Lepus* sp. in Western Sahara and Southern Morocco. *Go-South Bull* 9:126–130
- Nosil P, Harmon LJ, Seehausen O (2009) Ecological explanations for (incomplete) speciation. *Trends Ecol Evol* 24:145–156
- Pisa G, Orioli V, Spilotros G, Fabbri E, Randi E, Bani L (2015) Detecting a hierarchical genetic population structure: the case study of the Fire Salamander (*Salamandra salamandra*) in Northern Italy. *Ecol Evol* 5:743–758
- Pokorny L, Riina R, Mairal M, Meseguer AS, Culshaw V, Cendoya J et al. (2015) Living on the edge: timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Front Genet* 6:154
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904
- Ramírez-Silva JP, González-Cózatl FX, Vázquez-Domínguez E, Cervantes FA (2010) Phylogenetic position of Mexican jackrabbits within the genus *Lepus* (Mammalia: Lagomorpha): a molecular perspective. *Rev Mex Biodivers* 81:721–731
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
- Schai-Brown S, Hackländer K (2018) Cape hare, *Lepus capensis* Linnaeus 1758. In: Smith AT, Johnston CH, Alves PC, Hackländer K (eds) *Lagomorphs: pikas, rabbits and hares of the world*. John Hopkins University Press, Baltimore, USA
- Scheffers BR, Joppa LN, Pimm SL, Laurance WF (2012) What we know and don't know about Earth's missing biodiversity. *Trends Ecol Evol* 27:501–510
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
- Stewart M, Louys J, Price GJ, Drake NA, Groucutt HS, Petraglia MD (2017) Middle and Late Pleistocene mammal fossils of Arabia and surrounding regions: implications for biogeography and hominin dispersals. *Quat Int.* <https://doi.org/10.1016/j.quaint.2017.11.052>
- Strasburg JL, Rieseberg LH (2010) How robust are “Isolation with Migration” analyses to violations of the IM model? A simulation study. *Mol Biol Evol* 27:297–310
- Sukumaran J, Knowles LL (2017) Multispecies coalescent delimits structure, not species. *Proc Natl Acad Sci USA* 114:1607
- Thulin CG, Fang M, Averianov AO (2006) Introgression from *Lepus europaeus* to *L. timidus* in Russia revealed by mitochondrial single nucleotide polymorphisms and nuclear microsatellites. *Heredity* 143:68–76
- Tolesa Z, Bekele E, Tesfaye K, Ben Slimen H, Valqui J, Getahun A et al. (2017) Mitochondrial and nuclear DNA reveals reticulate evolution in hares (*Lepus* spp., Lagomorpha, Mammalia) from Ethiopia. *PLoS ONE* 12:e0180137
- Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 37:226–231
- Velo-Anton G, Martinez-Freiria F, Pereira P, Crochet PA, Brito JC (2018) Living on the edge: ecological and genetic connectivity of the spiny-footed lizard, *Acanthodactylus aureus*, confirms the Atlantic Sahara desert as a biogeographic corridor and centre of lineage diversification. *J Biogeogr* 45:1031–1042
- Wallner B, Huber S, Achmann R (2001) Non-invasive PCR sexing of rabbits (*Oryctolagus cuniculus*) and hares (*Lepus europaeus*). *Mamm Biol* 66:190–192
- Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics* 23:1851–1853
- Wong RA, Fong JJ, Papenfuss TJ (2010) Phylogeography of the African Helmeted Terrapin, *Pelomedusa subrufa*: genetic structure, dispersal, and human introduction. *Proc Calif Acad Sci* 61:575–585
- Wu C, Wu J, Bunch TD, Li Q, Wang Y, Zhang YP (2005) Molecular phylogenetics and biogeography of *Lepus* in Eastern Asia based on mitochondrial DNA sequences. *Mol Phylogenet Evol* 37:45–61
- Yu X (2004) Molecular systematics of the genus *Lepus* in China. Institute of Zoology, Chinese Academy of Sciences, Beijing
- Zubair M, Shukkur EAA, Azeez PA, Jayson EA (2011) Feeding behaviour of three species of falcons in the wild in united arab emirates. *Millenium Zool* 12:13–19
- Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. University of Texas, Austin