Paraphyly of the widespread generalist red fox (*Vulpes vulpes*): introgression rather than recent divergence of the arid-adapted Rüppell’s fox (*Vulpes rueppellii*)?

ALI E. BASUONY1,2,*, MOSTAFA SALEH2, MOUSTAFA SARHAN3,4, MAHMOUD YOUNES2, FOUAD ABDEL-HAMID2, CARLOS RODRIGUEZ FERNANDES5,6, PAUL VERCAMMEN7, FARAJ ABOSHAALA8, FARID BOUNACEUR9, ELIZABETH A. CHADWICK1 and FRANK HAILER1,*,

1School of Biosciences, Sir Martin Evans Building, Museum Avenue, Cardiff University, Cardiff, CF10 3AX, Wales, UK
2Department of Zoology, Faculty of Science, Al-Azhar University, 11751, Cairo, Egypt
3Department of Biomedical Sciences, College of Clinical Pharmacy, King Faisal University, 31982, Saudi Arabia
4Department of Zoology, Faculty of Science, Al-Azhar University, 71524 Assuit, Egypt
5C3E3c - Centre for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute, Departamento de Biologia Animal Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
6Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal
7Breeding Centre for Endangered Arabian Wildlife, 29922, Sharjah, United Arab Emirates
8Department of Zoology, Faculty of Science, Misurata University, 2478, Misurata, Libya
9Agronomy and Environment Laboratory, Department of Natural and Life Sciences, Tissemsilt University, 38000, Algeria

Received 8 August 2022; revised 29 November 2022; accepted for publication 19 December 2022

Understanding of the evolutionary history of two closely related canid sister taxa, the geographically restricted, arid-adapted Rüppell’s fox (*Vulpes rueppellii*) and the widespread generalist red fox (*Vulpes vulpes*), has been hampered by limited sampling in the biogeographically complex region of North Africa and the Middle East. We sequenced mitochondrial DNA (mtDNA) cytochrome *b* and D-loop fragments from 116 samples for both species and combined these data with previously published sequences, resulting in 459 haplotypes. Obtained phylogenies showed high support for most branches, including for a newly described ‘Palearctic clade’ that includes North African and Asian individuals from both species. All *V. rueppellii* individuals fell within the Palearctic clade, forming two previously undescribed subclades that were intermingled with, but not shared with *V. vulpes*. Our robust placement of *V. rueppellii* within *V. vulpes* renders the latter paraphyletic. We propose three scenarios that could explain these observations: (1) rapid, recent speciation of *V. rueppellii* from *V. vulpes*, (2) incomplete lineage sorting, or (3) ancient divergence followed by introgression and secondary mtDNA similarity. The third scenario is in best agreement with evidence from the fossil record, and morphometric and ecological distinctiveness between the two taxa, and therefore seems most likely.


INTRODUCTION

Except for unusual cases such as hybrid speciation (Lavrenchenko, 2014; Lamichhaney *et al.*, 2018; Masello *et al.*, 2019), the evolution of distinct species is typically considered a slow process that, given enough time of reproductive isolation, will lead to reciprocally monophyletic lineages. During the Pleistocene, populations of many mammalian species were separated into distinct refugia and evolved pronounced phylogeographic structuring (Avise *et al.*,...
In contrast, the much less extensively studied *V. rueppellii* is a species of xeric conditions, occupying arid habitats from North Africa to Pakistan, with up to six described subspecies (Rosevear, 1974; Sillero-Zubiri et al., 2004). Mitochondrial and microsatellite analysis of *V. rueppellii* from north-west Africa and one sample from north-east Africa (Egypt) did not reveal any clear genetic structuring (Leite et al., 2015), although this finding could have resulted from limited geographic coverage and small sample size (Leite et al., 2015). Based on mtDNA analysis, Leite et al. (2015) revealed paraphyly of *V. vulpes* and clustering of *V. rueppellii* within *V. vulpes*, with *V. rueppellii* being most closely related to two *V. vulpes* clades found in Morocco. The authors therefore proposed that *V. rueppellii* could represent an ecotype of *V. vulpes*, or that past introgression from *V. vulpes* into *V. rueppellii* could have occurred.

Although *V. vulpes* is a well-studied taxon in Eurasia and North America (e.g. Frati et al., 1998; Inoue et al., 2007; Perrine et al., 2007; Aubry et al., 2009; Teacher et al., 2011; Edwards et al., 2012; Yu et al., 2012a; Kutschera et al., 2013; Ibi et al., 2014), the authors of the most comprehensive phylogeographic study of *V. vulpes* to date (Statham et al., 2014) emphasized that the North African range remains only relatively sparsely characterized to date. Indeed, several previous studies of *V. vulpes* phylogeography highlighted that sampling gaps in biogeographically important regions still remain (Frati et al., 1998; Inoue et al., 2007; Perrine et al., 2007; Aubry et al., 2009; Teacher et al., 2011; Edwards et al., 2012; Yu et al., 2012a; Kutschera et al., 2013). Hence, previous work in North Africa and the Middle East lacked a comprehensive representation of ecoregions that are occupied by the two species. Cryptic or shared lineages within either species might therefore have remained undetected in previous studies.

The reported paraphyly of *V. vulpes* and hence the absence of reciprocally monophyletic mtDNA of *V. rueppellii* could result from various mechanisms. These include (1) ILS, (2) introgressive hybridization, (3) insufficient spatial sampling and low sample size in key biogeographic areas, and (4) analysis of short mtDNA sequences. First, ILS can contribute to non-monophyly when within-species polymorphism persists longer than the time between two successive speciation events (Funk & Omland, 2003; Lopes et al., 2021). Second, introgressive hybridization during a secondary contact of the two species, possibly during periods of fluctuating climate (Barton & Hewitt, 1985; Melo-Ferreira et al., 2005; Rieseberg et al., 2007) might have contributed to that paraphyly. Indeed, prominent cases of mammalian hybridization occur in scenarios of secondary contact of previously allopatric species (Colella et al., 2018). Third, increased sampling can affect the inference of phylogenetic relationships
(Nabhan & Sarkar, 2012; Figueroa et al., 2016). Since V. rueppellii has so far mainly been sampled from north-west Africa, a small part of its range (Fig. 1), mtDNA lineages distinct from those in V. vulpes might have remained undetected in previous works. Fourth, analysis of relatively short mtDNA sequences in previous work resulted in phylogenetic trees with partly low branch support, possibly masking true phylogenetic relationships between the two species. Analysis of longer sequences could hence help identify accurate phylogenetic and phylogeographic structuring (Keis et al., 2013).

Here, we present novel mtDNA data (cytochrome b and D-loop) for V. vulpes and V. rueppellii from North Africa and the Middle East. Our goals were to: (1) investigate the phylogeographic relationship between disjunct populations of V. vulpes and V. rueppellii in North Africa and the Middle East within the context of previously published data; (2) assess the validity of the reported paraphyly of V. vulpes based on longer DNA sequence alignments and improved sampling in key biogeographic regions in the sympatric range of both species.

MATERIAL AND METHODS

SAMPLE COLLECTION

A total of 128 fox samples were newly obtained for this study (Fig. 1). Our sampling included 88 samples from Egypt (65 V. vulpes and 23 V. rueppellii); seven from road-killed animals from Libya (five V. vulpes and two V. rueppellii); four road-killed V. vulpes from Algeria; 24 from road-killed animals from the Middle East (seven V. vulpes tissue samples, 11 V. vulpes hair samples and six V. rueppellii hair samples); and five road-killed V. vulpes obtained from the Vale of Glamorgan Council and Cardiff Council (Wales, UK) (Supporting Information, File S2).

LABORATORY PROCEDURES

DNA extraction

Genomic DNA was extracted from tissue samples using a salting-out protocol modified from Rivero et al. (2006), which in turn was based on the Puregene DNA Extraction Kit (Qiagen, Hilden, Germany). DNA extractions from hair samples were conducted using

Figure 1. Sampling distribution of V. vulpes and V. rueppellii from North Africa, the Middle East and southern Europe. Additional samples from outside this region are not shown here, but were included in some analyses, e.g., the Bayesian tree. *unpublished GenBank sequences, precise coordinates for these samples are unknown. Not all samples are discernible, due to spatial overlap of symbols (for details see Supporting Information, File S2). Prepared using QGIS 3.8.3 (http://www.qgis.org).
DNeasy Blood & Tissue Kits (Qiagen), following the manufacturer’s recommendations, and quality was assessed by electrophoresis in 1% agarose gels.

**PRIMER DESIGN**

Among the previous studies of the two *Vulpes* species that included more than one locus, most sequenced fragments spanned various and often non-overlapping regions of cytochrome *b* and the D-loop (Supporting Information, File S1: Fig. S1). To include as many as possible of the previously published sequences for the geographical regions of interest, especially those of Statham et al. (2014) and Leite et al. (2015) for both cytochrome *b* and the D-loop, we designed new primers for both loci using primer3 v.4.1.0 (http://primer3.ut.ee/) (Table 1). For cytochrome *b*, three primer pairs were initially designed. All of them produced a strong band with a PCR reaction, but only one pair (Vv.CY14144AF and Vv.CY15117AR) consistently produced clear and reliable Sanger sequences. For the D-loop, we designed a primer pair (Vv.CR2AF and Vv.CR2AR) which produced a strong band in PCRs and consistently high-quality Sanger sequences. For hair samples, the designed cytochrome *b* primers did not amplify, likely due to DNA degradation, so we used the primer pair L14724 and H15149 (Kocher et al., 1989; Irwin et al., 1991) that targets a 464-bp amplicon of cytochrome *b*. Locations of the sequenced fragments are shown in Supporting Information, File S1: Fig. S1.

**PCR AMPLIFICATION AND SEQUENCING**

We amplified a 615 bp fragment from the 5’ end of the mitochondrial D-loop (for both tissue and hair samples), and for cytochrome *b*, 974 and 464 bp fragments, respectively, for tissue and hair samples (Table 1). PCR amplification for tissue samples for both markers was performed in 15 µL reaction mixtures containing: 1× GoTaq Flexi buffer (Promega), 167 µM of each dNTP, 0.017 U GoTaq G2 polymerase (Promega), 2 mM MgCl₂, 200 µM of each primer for cytochrome *b*, 400 µM of each D-loop primer and 1 µL DNA extract. PCR cycling conditions were 3 min at 94 °C, followed by 30 cycles of 1 min at 94 °C, 1 min at 50 °C, and 1.5 min at 72 °C, followed by a 7 min step at 72 °C. For hair samples, PCRs for both the D-loop and cytochrome *b* were performed in 20 µL reaction mixtures containing 1× GoTaq Flexi buffer (Promega), 163 µM of each dNTP, 0.023 U GoTaq G2 polymerase, 4.0 mM MgCl₂, 300 µM of each primer and 3 µL DNA extract. Cycling conditions were 3 min at 94 °C, 40 cycles of 1 min at 94 °C, 1 min at 50 °C and 1.5 min at 72 °C, followed by a final 10 min step at 72 °C. The quality of PCR products was verified by electrophoresis in 2% agarose gels. Sanger sequencing of PCR products was performed by Eurofins Genomics (Wolverhampton, UK) on an ABI 3100 Genetic Analyzer.

**DATA ANALYSIS**

Electropherograms were checked manually, and sequences were aligned using Geneious Prime 2020.1.1 (https://www.geneious.com). Previously published DNA sequences from *V. vulpes* and *V. rueppellii* were downloaded from GenBank, including 257 *V. vulpes* haplotypes from Statham et al. (2014), nine haplotypes from ten *V. rueppellii* individuals and 24 haplotypes from 31 *V. vulpes* individuals from Leite et al. (2015), six *V. rueppellii* (accession numbers, cytochrome *b*: KU378368–KU378373, D-loop: KU378374–KU378379) and 90 *V. vulpes* (accession numbers, cytochrome *b*: KU378491–KU378580, D-loop: KU378398–KU378486) haplotypes (Harsn et al., unpublished data), five complete mitogenomes (accession numbers: KF387633 (Zhang et al., 2015), AM181037 (Arnason et al., 2006), GQ374180 (Zhong et al., 2010), KP342452 (Sun et al., 2016a), JN711443 (Yu et al., 2012b)) and 25 *V. vulpes* haplotypes from Inoue et al. (2007) (Supporting Information, File S2). We used *Vulpes lagopus* (Linnaeus, 1758) (accession numbers, cytochrome *b*: KU378607–KU378611, D-loop: KU378612–KU378616) as outgroup species for both loci using primer3 v.4.1.0 (http://primer3.ut.ee/) (Table 1). For cytochrome *b*, three primer pairs were initially designed. All of them produced a strong band with a PCR reaction, but only one pair (Vv.CY14144AF and Vv.CY15117AR) consistently produced clear and reliable Sanger sequences. For hair samples, the designed cytochrome *b* primers did not amplify, likely due to DNA degradation, so we used the primer pair L14724 and H15149 (Kocher et al., 1989; Irwin et al., 1991) that targets a 464-bp amplicon of cytochrome *b*. Locations of the sequenced fragments are shown in Supporting Information, File S1: Fig. S1.

**Table 1.** Mitochondrial primers utilized in this study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer length (bp)</th>
<th>Sequence (5’–3’</th>
<th>Fragment length (bp) including primers</th>
<th>Locus</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vv.CR2AF</td>
<td>25</td>
<td>GCCAACCATTAGCATATTGAAAAAC</td>
<td>615</td>
<td>D-loop</td>
<td>This study</td>
</tr>
<tr>
<td>Vv.CR2AR</td>
<td>21</td>
<td>ACCAAATGATGACACCACAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vv.CY14144AF</td>
<td>26</td>
<td>GACATGAAAAATCATGTTGATTTTC</td>
<td>974</td>
<td>Cytochrome <em>b</em></td>
<td>This study</td>
</tr>
<tr>
<td>Vv.CY15117AR</td>
<td>20</td>
<td>TTTGAGGTGTGAGGTGRRG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L14724</td>
<td>20</td>
<td>GATATGAAAAAATCATGTTG</td>
<td>464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H15149</td>
<td>20</td>
<td>CAGAATGATATATTGTCCCA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2023 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2023, **XX**, 1–17
no. KP342451) as an outgroup (Sun et al., 2016b). Genuaneous Prime was used to generate alignments using MUSCLE v.3.8 (Edgar, 2004), and to concatenate cytochrome b and D-loop sequences.

Bayesian phylogenetic analysis was conducted using BEAST v.2.6.0 (Bouckaert et al., 2019). We partitioned the data set into four regions: 1st, 2nd, and 3rd codon positions of the cytochrome b gene, and the D-loop, and determined the most appropriate models of DNA substitution using the Akaike Information Criterion (AIC) in jModelTest 2.1.10 (Darriba et al., 2012). For the cytochrome b partitions of the data set, the GTR+G model was used, and GTR+I+G for the D-loop partition. In BEAST, we used the coalescence constant size model as a tree prior, with default values for other parameters. We conducted and combined five independent BEAST runs for 50 million generations each, sampling every 1000 generations, and subsequently combined these for further analyses. Trace plots were verified using TRACER v.1.7 (Rambaut et al., 2018), confirming good mixing of chains. A burn-in of 10% was found to be suitable, and an effective sample size (ESS) above 200 indicated convergence for all posterior parameter estimates. A maximum clade credibility tree with posterior probabilities for each node was obtained using TREANNO TATOR v.2.6.0 (Bouckaert et al., 2019), and visualized using FIGTREE 1.4.4 (https://github.com/rambaut/figtree/releases).

We reconstructed statistical parsimony haplotype networks using the TCS algorithm (Clement et al., 2000) as implemented in PopArt v.1.7 (https://popart.maths.otago.ac.nz/), using a 95% minimum connection probability limit, and excluded gaps and missing data. Haplotype frequencies, haplotype and nucleotide diversity, Fu’s Fs (Fu, 1997), Tajima’s D (Tajima, 1989) and the average number of nucleotide substitutions per site between groups (Dxy) were calculated using DnaSP v.6.12.03 (Rozas et al., 2017).

RESULTS

Out of the 128 novel samples, ten hair samples failed to amplify, and two (one tissue and one hair) were excluded due to signals of heteroplasy and/or nuclear mitochondrial copies (see the text in Supporting Information, File S1), leaving 116 newly obtained sequences (Supporting Information, File S2).

Most new sequences represented novel haplotypes, except three V. vulpes sequences from Egypt that were identical to the Egyptian haplotype from Leite et al. (2015). The concatenated sequences comprised 109 longer sequences (1400 bp: 864 bp cytochrome b + 536 bp D-loop), and seven shorter sequences from lower-quality samples (939 bp: 403 bp cytochrome b + 536 bp D-loop) (Supporting Information, File S2). The alignment of the longer (1400 bp) sequences contained 129 segregating sites that formed 37 haplotypes (26 for V. vulpes and 11 for V. rueppellii). In addition, we encountered five haplotypes (two for V. vulpes and three for V. rueppellii) for the seven short sequences, across 39 polymorphic sites (Supporting Information, File S2). Tajima’s D deviated non-significantly from zero (P > 0.5) for a total dataset of 148 individuals comprising 664 bp of concatenated sequences (cytochrome b: 360 bp; D-loop: 304 bp) and for each species separately, being -0.104 for 34 individuals of V. rueppellii, and 0.133 for 114 V. vulpes individuals, consistent with neutral evolution of the sequences.

**MAIN PHYLOGENETIC CLADES OF V. VULPES AND V. RUEPPELLII**

A Bayesian phylogenetic tree of 459 mtDNA haplotype sequences grouped V. rueppellii inside the diversity of V. vulpes with high support (Bayesian Posterior Probability; BPP > 0.99), showing paraphyly of V. vulpes (Fig. 2A; see Supporting Information, File S3 for the complete tree file). Figure 2C shows the distribution of V. vulpes and V. rueppellii clades in North Africa and Middle East and their sample frequencies. We obtained high support (BPP > 0.99) for the ‘Holarctic’ and ‘Nearctic’ clades described by Statham et al. (2014), and also obtained such high support (BPP > 0.99) for a clade containing newly obtained sequences along with previously published ‘Palearctic basal haplotypes’ from Statham et al. (2014). This clade, henceforth referred to as ‘Palearctic clade’, contains sequences from V. vulpes from North Africa and Asia, along with all sequences from V. rueppellii that have been generated to date—from across North Africa, Saudi Arabia, United Arab Emirates and Iran. Further, we obtained high support (BPP > 0.99) for two African clades (Africa 1 and Africa 2), which in turn clustered together with high support (BPP > 0.99). These two African clades correspond to Maghreb 1 and Maghreb 2 described by Leite et al. (2015) for north-west Africa. The support for the two African clades to cluster with the joint Holarctic/Nearctic clades was moderate (BPP: 0.82) and did not increase when we restricted the analysis to long sequences only, nor when cytochrome b and the D-loop were analysed separately (details not shown). Haplotype networks showed groupings consistent with these main clades, both for shorter (Fig. 2B) and longer (Supporting Information, File S1; Fig. S2) alignment lengths.

All analysed V. rueppellii sequences clustered into two main sub-clades within the Palearctic
Figure 2. Phylogenetic and phylogeographic results. A, maximum clade credibility tree from concatenated cytochrome b and D-loop sequences (459 haplotypes, 430 V. vulpes and 29 V. rueppellii). Bayesian posterior support values ≥ 80% are
clade, each receiving high support (BPP > 0.99). The average number of nucleotide substitutions per site between the two subclades was \( D_{XY} = 2.1\% \). Subclade 1 was restricted to North Africa, and subclade 2 was found in Iran, Arabia and east of the Nile (Egypt) (Fig. 2A, C). The two subclades were sympatric only in one region, east of the Nile in Egypt.

*Vulpes vulpes* sequences were found within all major clades. The Palearctic clade was of particular interest, since it contains both *V. vulpes* and *V. rueppellii*, so it is presented in greater detail. Palearctic-clade *V. vulpes* comprised eight haplotypes from North Africa, the Middle East and East Asia (Japan) (Fig. 2B, C). Two haplotypes (PS12 and PS18) were widely distributed along the Nile and western desert oases in Egypt (27 and ten samples, respectively), one (PS30) was found in six samples from the United Arab Emirates, one (PS50) in four samples from Japan, and four additional haplotypes were rare and geographically restricted (three in Egypt; one in Japan; see Supporting Information, File S2). Table S1, Supporting Information, File S1 shows the divergence between the main clades of short (Fig. 2B) and long (Supporting Information, File S1: Fig. S2) sequences. The haplotype network for a subset of longer sequences (Supporting Information, File S1: Fig. S2) showed the same overall topology, but with increased divergence between the main clades.

The Holarctic clade contained the greatest number of haplotypes and individuals, and was also the geographically most widely distributed, occurring in North Africa, Europe, Asia and North America. Most newly obtained haplotypes within the Holarctic clade were from Europe, West Asia and the Sinai Peninsula, along with a few from North Africa (Supporting Information, File S2). The Nearctic clade only contained samples from North America, as found previously (Kutschera et al., 2013; Statham et al., 2014). The Africa 1 clade was restricted to central and north-west Africa (Libya, Tunisia, Algeria, and Morocco). The Africa 2 clade was found in samples from the Mediterranean coastal desert in Egypt, Libya and the western Atlas, comprising two newly obtained Egyptian haplotypes, two Libyan haplotypes and two previously described haplotypes from Morocco ['Maghreb 2' subclade of Leite et al. (2015)].

**GENETIC DIVERSITY**

To infer the genetic diversity within and among *V. vulpes* and *V. rueppellii* populations, we trimmed our data according to Leite et al. (2015), a dataset of particular interest since it includes *V. vulpes* and *V. rueppellii* from Africa, and *V. vulpes* from Europe and the Middle East. This combined data set contained 148 individuals [109 from this study, 39 from Leite et al. (2015)], comprising 664 bp of concatenated sequences (cytochrome b: 360 bp; D-loop: 304 bp; Table 2).

Consistent with the deeply divergent clades in *V. vulpes*, this species showed higher nucleotide diversity and numbers of variable sites than *V. rueppellii*, although the latter showed slightly higher haplotype diversity (Table 2). The high nucleotide diversity among *V. vulpes* populations along with west of the Nile coincides with clade admixture in these populations (west of the Nile: Africa 2, Holarctic and Palearctic clades; along the Nile: Holarctic and Palearctic clades). In contrast, *V. vulpes* populations from north-west Africa, Europe and east of the Nile contained only one clade—the African clade for north-west Africa, and Holarctic clade for both Europe and east of the Nile—yielding lower nucleotide variability estimates. Fu’s \( F_s \) was non-significant for all investigated geographic groupings except north-west African *V. rueppellii*, for which a significantly negative value was observed (Table 2).

**DISCUSSION**

We here provide a comprehensive phylogenetic and phylogeographic analysis of *V. vulpes* and *V. rueppellii*, allowing us to evaluate their matrilineal evolutionary history. Our study incorporates newly obtained sequences from both species, along with previously published homologous mtDNA data from across their geographic ranges. Based on longer sequence alignments than most previous studies (Supporting Information, File S1: Fig. S1), our obtained phylogeeny demonstrates that the 'Palearctic basal haplotypes' by Statham et al. (2014) form a distinct Palearctic clade that is shared between *V. vulpes* and *V. rueppellii*. Importantly, we show that all analysed *V. rueppellii*, sampled across North Africa and the Middle East, are nested within this Palearctic clade, rendering *V. vulpes* paraphyletic. These findings are consistent
with previous work by Leite et al. (2015), who found V. rueppellii to cluster with two African clades (Maghreb 1 and 2) of V. vulpes. Our results link this paraphyly to Palearctic-clade sharing with V. vulpes populations across North Africa and Asia.

### Evolutionary History of V. rueppellii and Paraphyly of V. vulpes

Our results lead us to propose three evolutionary scenarios for the phylogenetic relationships of the two species (Fig. 3). Edwards et al. (2011) proposed similar scenarios to explain the paraphyly of brown bears.

**Scenario 1: ‘Ecotype scenario’ – the rapid evolution of V. rueppellii from Palearctic-clade V. vulpes (Fig 3A).**

A parsimonious explanation for V. vulpes paraphyly and the low divergence of V. rueppellii from Palearctic clade V. vulpes sequences would be a recent and rapid evolution of V. rueppellii. This scenario could support the classification of V. rueppellii as a desert ecotype of V. vulpes (see Leite et al., 2015). The term ecotype is typically used to describe genetically distinct forms within a species that are highly adapted to a specific environment (Begon et al., 2005). Indeed, other species of canids have previously been suggested to contain distinct ecotypes, such as wolves (Carmichael et al., 2007; Leonard et al., 2007; Musiani et al., 2007; Muñoz-Fuentes et al., 2009; Hendricks et al., 2019) and Arctic foxes (Dalén et al., 2005; Norén et al., 2011). However, we consider this scenario to be unlikely for V. rueppellii, for several reasons:

(a) The fossil record suggests that V. rueppellii as a species is much older than suggested by nesting of mtDNA within V. vulpes diversity. Geraads (2011) recorded two V. rueppellii fossils from Tighenif, Algeria (north-west Africa): one of them dating to about 0.5 Mya and showing a similar morphotype to V. rueppellii today, and the other form from 0.8 Mya was interpreted as a fossil precursor species to V. rueppellii, suggesting an even earlier divergence from V. vulpes.

(b) The morphological and physiological differentiation between the two species is considerable, and well supported: V. vulpes is overall larger, with longer hind legs, a longer tail and proportionally shorter ears than the sympatric V. rueppellii (Lariviere & Seddon, 2001). Ecologically,
Figure 3. Three hypothetical scenarios for the evolution of V. rueppellii and current paraphyly of V. vulpes: (A) ‘Ecotype scenario’: rapid evolution of V. rueppellii from Palearctic-clade V. vulpes; (B/C) Old divergence and recent introgression of mtDNA between the two species. B, introgression of the V. rueppellii mitogenome into V. vulpes. C, introgression of the V. vulpes mitogenome into V. rueppellii. Divergence times within V. vulpes are based on Statham et al. (2014). Interspecific
behaveurally and physiologically, *V. rueppellii* is adapted to xeric conditions (Rosevear, 1974; Williams *et al.*, 2002; Siller-Zubiri *et al.*, 2004), whereas *V. vulpes* avoids such habitats, is distributed throughout the Holarctic and shows a wide plasticity in terms of habitat requirements (Siller-Zubiri *et al.*, 2004; Soulsbury *et al.*, 2010). An analysis of external measurements (head and body length, tail length, ear length, shoulder height and weight) showed a large difference between the two species (Siller-Zubiri *et al.*, 2004). That dataset included *V. rueppellii* from Arabia (Lenain, 2000) and Egypt (Osborn & Helmy, 1980), and *V. vulpes* from across its distribution except North Africa (UK, Hattingh, 1956; Australia, McIntosh, 1963; Canada, Voigt, 1987; Japan, Zhan *et al.*, 1991 and several studies from Cavallini, 1995). These results appear comparable to those from other mammalian sister species pairs, which according to a meta-analysis by Avise *et al.* (1998) typically diverged more than a million years ago. Hence, the significant physical differentiation between *V. rueppellii* and *V. vulpes* tentatively suggests a longer time since speciation than suggested by mtDNA.

(c) Nuclear microsatellite data show a strong differentiation between *V. rueppellii* and *V. vulpes* \( F_{ST} = 0.14 \) (Leite *et al.*, 2015) with larger interspecific differences than mtDNA. Such mito-nuclear discordance has been found in other paraphyletic mammals and their sibling species, where paraphyly for mtDNA is accompanied by significant differentiation at nuclear loci (Good *et al.*, 2008; Hailer *et al.*, 2012). However, we caution that this pattern for microsatellites in *V. vulpes* and *V. rueppellii* could hypothetically result from strong/rapid genetic drift, rather than long evolutionary time. Under such a scenario one would predict decreased intrapopulation variability. However, when compared to their North African and Eurasian counterparts of *V. vulpes*, unbiased expected heterozygosity and allelic richness in *V. rueppellii* are c. 105\% and 102\% for allelic richness and 90\% and 87\% of expected heterozygosity, respectively (Leite *et al.*, 2015). These findings do not reveal clear evidence of the strong and recent genetic drift, but are consistent with the long time frames indicated by the fossil record of *V. rueppellii* (Geraads, 2011).

(d) For red foxes, Statham *et al.* (2014) estimated the time to most recent common ancestor \( T_{MRCA} \) of the Palearctic group at c. 70–98 kya. Hence, *V. rueppellii* would have evolved from a lineage within the Palearctic *V. vulpes* clade, subsequently adapting rapidly to arid habitats. If the *V. rueppellii* lineage indeed were this young, the vast current geographic range (Fig. 1) would predict clear signals of demographic growth. However, our analyses only revealed signals of population growth for north-west African *V. rueppellii* sequences, but not for any other regions studies (or all sequences combined) (Table 2).

### Scenario 2: ILS explains intermingled lineages
The oldest fossil remains of *V. rueppellii* are from north-west Africa, dating back to c. 0.8 Mya (Geraads, 2011). The divergence between *V. vulpes* and *V. rueppellii* therefore likely occurred in or before the Middle Pleistocene. ILS can cause species-level non-monophyly, if divergence between the species was too recent for ancestral polymorphisms to have sorted into reciprocally monophyletic lineages (Funk & Omland, 2003; McKay & Zink, 2010). ILS has previously been suggested to cause non-monophyly in European bison (*Bison bonasus*) (Wang *et al.*, 2018). Structuring within Eurasian and Nearctic *V. vulpes* populations has so far been interpreted as the result of biogeographic barriers, or isolation-by-distance (Kutschera *et al.*, 2013; Statham *et al.*, 2014). Therefore, if ILS explains lineage branching patterns between *V. vulpes* and *V. rueppellii*, then perhaps the intraspecific phylogeographic patterns of *V. vulpes* would need to be re-evaluated as well.

Lineage sorting for mtDNA requires on average \( 1 \times N_e \) generations (where \( N_e \) is the effective female population size; Nichols, 2001). Indeed, in *V. vulpes*, this corresponds to only c. 100–200 kya – based on an ancestral \( N_e \) of 91 000 (Statham *et al.*, 2014) and a generation time of 2 years (Statham *et al.*, 2018). ILS therefore appears unlikely to impact red foxes mtDNA beyond a few 100 kyr, a time frame younger than the divergence time suggested by the fossil record (Geraads, 2011).

### Scenario 3: Old divergence and recent introgression of mtDNA between the two species (Fig. 3B, C)
There are numerous examples of introgressive hybridization in the genus *Canis*, e.g. between Ethiopian wolves (*Canis simensis*) and domestic dogs (*Canis familiaris*) (Gottelli *et al.*, 1994), and between red wolves (*Canis rufus*) and coyotes (*Canis latrans*; Adams *et al.*, 2003; Hailer & Leonard, 2013).
Even hybridization between taxa with differing chromosome numbers has been described for mammals (Horn et al., 2012; Giménez et al., 2016). Intraspecific hybridization in Vulpes has been described for V. vulpes and the kit fox (Vulpes macrotis) (Creeel & Thornton, 1974), and between V. macrotis and the swift fox (Vulpes velox) (Say, 1823) (Dragoo & Wayne, 2003). The previous two cases suggest that hybridization between V. vulpes and V. rueppellii should not be excluded, despite the differences in their chromosome number: 2n = 34 plus 0–8 B chromosomes for the former (Graphodatsky et al., 2000) and 2n = 40 for the latter (Ewer, 1973).

Behaviourally, V. vulpes typically dominates other fox species, especially smaller species such as V. lagopus (Tannerfeldt et al., 2002), Vulpes corsac (Linnaeus, 1768) and V. macrotis Merriam, 1888 (Sillerro-Zubiri et al., 2004). However, prezygotic interspecific barriers can break down under e.g. Allee effects (Courchamp et al., 1999) acting at low population densities or other population pressures (Adams et al., 2003; Hailer & Leonard, 2008; Seehausen et al., 2008). Hybridization between V. vulpes and V. rueppellii therefore remains a reasonable scenario, although its occurrence has not been described to our knowledge.

If introgression indeed explains the Palearctic clade sharing between the two species, then we might expect to also see clade sharing for the other three clades occurring in sympatry (Holarctic, Africa 1 and Africa 2). Given the extended sample size across sympatric areas in North Africa and the Middle East included in our study, we consider the absence of clade sharing among those three clades to be robust. A more likely scenario therefore involves an ancient divergence between V. vulpes clades (including the Palearctic group that contains current V. rueppellii) at c. 1.15 Mya (Statham et al., 2014), and a secondary contact leading to a gene flow at around 70–98 kya. This introgression is consistent with the estimated time of the diversity of the Palearctic haplotypes (Statham et al., 2014). There are two possible directions of introgression, as follows:

**Scenario 3a: Introgression of V. rueppellii mtDNA into V. vulpes (Fig. 3B)**

The Palearctic clade may originally have evolved in V. rueppellii, having diverged from other V. vulpes clades at c. 1.15 (0.85–1.45) Mya (Statham et al., 2014). Broadly consistent with this timing, Geraads, (2011) recorded two V. rueppellii fossils from Tighenif, Algeria (see above), V. cf. rueppellii (0.8 Mya) and V. rueppellii (0.5 Mya). The latter is closer to the modern V. rueppellii than to any other species (Geraads, 2011). Furthermore, Geraads, (2011) recorded Vulpes hassani Geraads, 2011 (2.5 Mya) as a precursor of V. rueppellii, suggesting an even earlier divergence of V. rueppellii from V. vulpes. Vulpes rueppellii may therefore have evolved from V. cf. rueppellii (Geraads, 2011), and subsequently passed on its mtDNA genome to some V. vulpes populations currently found in parts of North Africa and Eurasia (the Palearctic clade). That may have been related to V. vulpes colonizing arid habitats and/or persisting in low densities, which can favour introgressive hybridization in canids (Hailer & Leonard, 2008).

**Scenario 3b: Introgression of the V. vulpes mtDNA introgression into V. rueppellii (Fig. 3C)**

If we instead assume that the Palearctic clade originally evolved in V. vulpes, clade sharing between the two species today could result from introgression of this clade from V. vulpes into V. rueppellii. The original V. rueppellii mtDNA would thus have been lost (mtDNA replacement), as suggested for e.g., Lepus hare species (Melo-Ferreira et al., 2012) and Ursus bears (Hailer et al., 2012). Such replacement events can be due to a combination of strong genetic drift or potentially driven by selective advantage of introgressed lineages. Alternatively, the original V. rueppellii mtDNA lineage may still persist, undetected despite our increased sampling.

Without additional evidence, we consider scenarios 3a and 3b to be of equal likelihood. Fossil, ancient DNA or modern genomic evidence from biparentally or male-inherited markers may shed further light on these scenarios.

**PHYLOGEOGRAPHY OF V. RUEPELLII**

Only one previous study by Leite et al. (2015) has evaluated the phylogeography of V. rueppellii, finding no clear structuring in mitochondrial and nuclear markers (based on ten samples mainly from north-west Africa: three from Morocco, six from Mauritania and one from Egypt). Our results extend these findings by revealing a second mtDNA clade within the species, and by showing population genetic structuring for these clades across the species’ range (Fig. 2C). Our findings demonstrate that the genetic structuring of V. rueppellii is shallower than that of the V. vulpes, with no deeply divergent lineages present.

Our findings demonstrate, for the first time, the presence of two subclades within the species. These subclades show a predominantly western and eastern distribution, respectively. Populations of V. rueppellii are distributed through three main geographical regions across: (1) North Africa (west of the Nile to the Atlantic Ocean), (2) an intermediate North Africa/Middle East (east of the Nile), and (3) the Middle East
(from the Sinai Peninsula through Arabia to Pakistan). Subclades 1 and 2 correspond to the geographical regions 1 and 3, respectively, while the east Nile populations in Egypt (geographical region 2) share mtDNA haplotypes with both subclade 1 and subclade 2 (Supporting Information, File S1: Fig. S3).

This clear but relatively shallow genetic structuring between populations of V. rueppellii resembles that of the sand cat, Felis margarita (Loche, 1858), which occupies nearly the same habitats and geographic range. Howard-McCombe et al. (2019) investigated the phylogeny of the four established populations (subspecies) of Felis margarita: Felis margarita margarita (North Africa), Felis margarita harrisoni (Arabia), Felis margarita thinobia (west/central Asia) and Felis margarita scheffeli (Pakistan), detecting a significant genetic differentiation between the African subspecies and the other three subspecies, and only low differentiation among the Asian subspecies.

The geological record suggests that arid habitats were widespread and largely contiguous across North Africa and extending into the Middle East at 1.2–0.8 Mya (deMenocal, 2004). Leite et al. (2015) suggested that V. rueppellii might have evolved during the Pleistocene and colonized its existing range while the Sahara was connected to the Arabian and Syrian deserts. Subsequent climatic oscillations introduced more humid and mesic conditions, fragmenting these arid zones. At c. 12 kya, the modern Nile River formed (Said, 1981, 1993), its mesic habitats likely posing a barrier to gene flow for arid-adapted taxa such as V. rueppellii, splitting the populations to the west and east of the Nile. In contrast, these mesic habitats may have allowed more generalist species to colonize, perhaps explaining the arrival of Holarctic clade red foxes to North Africa. Similarly, climatic and sea level fluctuations would have created temporary barriers around the Gulf of Suez. Derricourt (2005) suggested that during drier periods of the Pleistocene, the Gulf of Suez was reduced in area and the Sinai Peninsula was readily accessible from the Eastern Desert, merging these two regions into an arid mountainous zone. Until about 14–15 kya when sea levels rose above about -50 m.a.s.l., the Sinai Peninsula was therefore presumably connected to the Eastern Desert (Derricourt, 2005; Bailey et al., 2007). This could explain the admixture of V. rueppellii subclade 1 and 2 haplotypes east of the Nile. The Eastern desert of Egypt and Sinai Peninsula may therefore represent a transitional region for V. rueppellii. Indeed, the Sinai Peninsula played an important role in the faunal exchange between Africa and Eurasia, linking these regions during periods of low sea level. Such conditions likely occurred frequently throughout the Pliocene and Pleistocene, facilitating multiple dispersion waves (Saleh et al., 2018). Existence of Pleistocene fossils of African mammalian fauna in the Levant dating back to 1.8–1.4 Mya (Tchernov, 1992) suggests the activity of this Afro-Asian route during the Pleistocene.

**CONCLUSION**

This study solidifies our understanding of the phylogeography of both V. rueppellii and V. vulpes, documenting for the first time two subclades and phylogeographic structuring within V. rueppellii. While Holarctic, Nearctic, Palearctic and two African clades had previously been robustly defined for V. vulpes, we here obtained robust statistical support for the previously so-called ‘Palearctic basal haplotypes’ as a ‘Palearctic clade’. We also report the first mtDNA data for V. rueppellii from north-east Africa and the Middle East. Our extended sampling across previously poorly sampled and unsampled regions reinforces that V. rueppellii is matrilineally rooted inside the diversity of the paraphyly V. vulpes. This paraphyly may have resulted from introgressive hybridization rather than recent speciation of V. rueppellii, consistent with evidence from morphometrics and the fossil record. Although our study included V. rueppellii from different ecoregions across its range, additional sampling would be desirable, in particular from the Asian part of the range. The occurrence of the three V. vulpes clades (Holarctic, Palearctic and Africa 2) and both subclades of V. rueppellii in north-east Africa, indicates that this region is a biogeographic diversity hotspot.

As a matrilineal marker that may not reveal genetic differentiation of the rest of the genome (Zhang & Hewitt, 2003; Hailer et al., 2012; Bidon et al., 2014), mtDNA evidence should be revisited with information from independently inherited genetic markers (e.g., autosomal and the Y-chromosome), to shed further light on the possible scenarios for the evolutionary history of the ecologically and morphometrically distinct V. rueppellii and V. vulpes.

**ACKNOWLEDGEMENTS**

A.E.B. was supported by the Cardiff University Vice Chancellor’s International Scholarship of Excellence. We thank the Cardiff University School of Biosciences, the British Ecological Society, the Genetics Society UK and the American Society of Mammalogists for funding. We also thank the following organizations and individuals for help with sample collection: Zoology Department, Faculty of Science, Al-Azhar University, Egypt; Egyptian Environmental Affairs Agency (EEAA), Nature Conservation Sector, Egypt; Breeding Centre for Endangered Arabian Wildlife, Sharjah, United Arab Emirates; Jonathan Greatrex (Vale of

© 2023 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2023, XX, 1–17
Glamorgan Council, UK; the late Dr Ali Al-Korey; Dr Magdy Saleh; Mr Naser Nabil; Mr Mohamed Anwar; Dr Ahmed Ghazy; Dr Saber Riad and Dr Abdullah Nagy. C.R.F. thanks the support of eC3c through an assistant researcher contract (FCiência.ID contract #366), Fundação para a Ciência e a Tecnologia (FCT) for Portuguese National Funds attributed to eC3c within the strategic project UID/BIA/00329/2020, and Faculty of Psychology of the University of Lisbon (FPUL) for an invited assistant professorship. The authors are very grateful to Dr Ahmed Badry who helped significantly with sample collection in Egypt. All sampling was conducted in accordance with national wildlife conservation and veterinary regulations. Samples were imported under Animal and Plant Health Agency (APHA) permit ITIMP21.0069. We thank the editor Prof. John A. Allen and the anonymous reviewers for their helpful comments. The authors have no conflict of interest to declare.

DATA AVAILABILITY

The newly obtained sequence data have been deposited in the EMBL/Genbank archive [accession numbers: MT955790–MT955894, OM001534–OM001544 for the control region (D-loop), and MT941577–MT941681, OL989975–OL989985 for cytochrome b].

REFERENCES


© 2023 The Linnean Society of London, Biological Journal of the Linnean Society, 2023, XX, 1–17
PARAPHYLY OF THE RED FOX (VULPES VULPES) 15


Teacher AG, Thomas JA, Barnes I. 2011. Modern and ancient red fox (Vulpes vulgaris) in Europe show an unusual lack of geographical and temporal structuring, and differing responses within the carnivores to historical climatic change. BMC Evolutionary Biology 11: 214.


SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article on the publisher's website.
File S1. Supplementary text, tables and figures.
File S2. Excel file of data on individuals, sequences and haplotypes.
File S3. Output file from BEAST/Figtree.