Mitogenomic analysis of Rüppell's fox (*Vulpes rueppellii*) confirms phylogenetic placement within the Palaearctic clade shared with its sister species, the red fox (*Vulpes vulpes*)

Ali E. Basuony[1,2,*], Mostafa Saleh[2], and Frank Hailer[1.3,*]

[1] Organisms and Environment, School of Biosciences, Sir Martin Evans Building, Museum Avenue, Cardiff University, CF10 3AX, Cardiff, Wales, UK.

[2] Department of Zoology, Faculty of Science, Al-Azhar University, 11751, Cairo, Egypt.

[3] Cardiff University-Institute of Zoology Joint Laboratory for Biocomplexity Research, Beijing, China

[*] Correspondence to Ali Basuony (basuonyae@cardiff.ac.uk, alibasuony@azhar.edu.eg), Frank Hailer (HailerF@cardiff.ac.uk)

Ali Basuony: ORCiD, 0000-0003-4919-3379

Frank Hailer: ORCiD 0000-0002-2340-1726

Running title: Mitogenomic analysis of Rüppell's fox (*Vulpes rueppellii*).

Abstract

The Rüppell's fox (*Vulpes rueppellii*) inhabits desert regions across North Africa, the Arabian Peninsula and southwestern Asia. Its phylogenetic relationship with other fox species, especially within the phylogeographic context of its sister species, *V. vulpes*, remain unclear. We here report the sequencing and de-novo assembly of the first annotated mitogenome of *V. rueppellii*, analysed with data from other foxes (tribe Vulpini, subfamily Caninae). We used four bioinformatic approaches to reconstruct the *V. rueppellii* mitogenome, obtaining identical sequences except for the incompletely assembled tandem-repeat region within the D-loop. The mitogenome displayed an identical organization, number and length of genes as *V. vulpes*. We found high support for clustering of both known subclades of *V. rueppellii* within the Palearctic clade of *V. vulpes*, rendering the latter species paraphyletic, consistent with previous analyses of shorter mtDNA fragments. More work is needed for a full understanding of the evolutionary drivers and consequences of hybridization in foxes.

Keywords

De novo assembly, mtDNA, Paraphyly, Reference-mapping, *Vulpes rueppellii/Vulpes vulpes*

Introduction

The Rüppell's fox (*Vulpes rueppellii*; Carnivora: Canidae; IUCN category of 'Least Concern'; Mallon et al. 2015) is widespread in desert regions of North Africa and Asia. Its geographical range comprises deserts of North Africa from Mauritania to Somalia and of Asia from the Arabian Peninsula to Iran and Pakistan, with up to six described subspecies (Mallon et al., 2015; Rosevear, 1974; Sillero-Zubiri et al., 2004; Williams et al., 2002). Typical habitats of *V. rueppellii* include sand dunes, sand sheets, gravel plains (Murdoch et al., 2007), stony habitats with few grass species that receive little rainfall (Lenain, 2000), and coastal areas with low vegetation cover (Mallon et al., 2015). The species' range partly overlaps with the red fox (*V. vulpes*), which is considered its sister taxon (Geffen et al., 1992; Leite et al., 2015; Lindblad-Toh et al., 2005; Rocha et al., 2023).

Short fragments of various mitochondrial DNA (mtDNA) loci have previously been used to investigate the evolutionary history of the two species (Basuony et al. 2023; Leite et al., 2015), revealing putative clustering of *V. rueppellii* within the variation of *V. vulpes*, rendering the latter species paraphyletic and potentially questioning the distinctiveness of the former. Due to the higher phylogenetic resolution provided by longer nucleotide sequences (Anijalg et al., 2018; du Plessis et al. 2023a,b; Keis et al., 2013), complete mitogenomes can provide more robust and detailed insights into phylogenetic relationships than short mtDNA fragments (Finstermeier et al., 2013). While the mitogenome has been fully sequenced (Arnason et al., 2006) and characterized (Zhong et al., 2010) for *V. vulpes*, this resource is to date relatively limited for *V. rueppellii.* Recently, Rocha et al (2023) sequenced mitogenomes of *V. rueppellii*, but without a detailed characterization or investigation of its phylogenetic relationship with the known clades and phylogeographic history of *V. vulpes*.

Various bioinformatic approaches have been developed for retrieval of mitogenome sequences from whole-genome sequencing data, e.g. mapping of sequencing reads directly against a typically closely related reference genome (reviewed in Briscoe et al. 2016), or approaches involving de-novo assembly of the reads, reducing or removing reliance on a reference genome, e.g. NOVOPlasty **(**Dierckxsens et al. 2017) and MITObim (Hahn et al., 2013). These approaches now allow efficient reconstruction of mitogenomes from whole-genome sequencing data.

The aim of this study was to obtain and characterize the complete mitogenome of *V. rueppellii* to 1) better understand its phylogenetic relationship with its sister species, *V. vulpes* and other fox species, and 2) compare the performance of four different mitogenome assembly approaches (two de novo and two different reference-based approaches) for obtaining mitogenome sequence data from Illumina whole-genome shotgun sequencing data.

Materials and Methods

Sampling and data generation

We extracted DNA from a male *V. rueppellii* tissue sample (ID M00375 in Basuony et al., 2023) collected from Wadi Umm Gheig, Eastern Desert, Egypt (25°36'55.01" N, 34°23'58.99" E), using a salting-out protocol modified from Rivero et al. (2006). We assessed the quality and concentration of the genomic DNA by electrophoresis in 1% agarose gels and a Qubit fluorometer v.3.0, respectively. DNA was subsequently sent to Neogen (Ayr, Scotland, UK) for library preparation and whole-genome sequencing. DNA was randomly sheared into short fragments, size-selected to ca. 350 base pairs (bp), A-tailed, ligated to Illumina adapters (5'-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT-3' and 5'-GATCGGAAGAGCACACGTCTGAACTCCAGTCACGGATGACTATCTCGTATGCCGTCTTCTGCTTG-3'), PCR-amplified, and purified. After subsequent quantification and checks for fragment size distribution using Qubit, real-time PCR and a bioanalyzer, the library was sequenced on an Illumina Novaseq instrument using paired-end reads (2x 151 bp).

We used FASTQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to assess the quality of the reads, and TRIMMOMATIC v0.39 (Bolger et al. 2014) to remove adapters and to trim off low-quality bases (settings: minimum length 50 bp, sliding window 10:15).

Mitogenome assembly

We used four parallel approaches to obtain the *V. rueppellii* mitogenome. Approach 1: De novo assembly with the program NOVOPlasty v4.3.1 (Dierckxsens et al., 2017), using the quality-controlled sequencing reads (see above) as an input, and using default parameter settings, except setting insert size to 350, K-mer size to 33 and leaving Max memory blank (to capture 100% of the input reads). As a seed to initiate the assembly, we used the GenBank-curated mtDNA reference genome of the sister taxon, *V. vulpes* (GenBank accession: NC_008434) (Arnason et al., 2006). Approach 2: De novo assembly using baiting and iterative mapping approaches implemented in MIRA v4.0.2 (Chevreux et al., 1999) and MITObim v1.9.1 (Hahn et al., 2013). Based on default parameter settings, we used MIRA to build an initial reference by mapping the trimmed reads to the mitochondrial reference genome of *V. vulpes* (Arnason et al., 2006). Next, the MITObim.pl script (https://github.com/chrishah/MITObim) was used to iteratively retrieve additional reads from the shotgun sequencing data and to map them against the reference obtained from the previous iteration. This was repeated until gaps were closed and a stationary number of mapped reads was reached for the mitogenome. The approach returns a single-padded consensus sequence, with any sequence fragments connected by 'N' to indicate that the fragments are not connected by reads and therefore not contiguous in the MITObim assembly (Machado et al., 2016). Then, the script circules.py (using a *k*-mer approach; distributed with MITObim) was used to for circularize the sequence. Approaches 3 and 4: Reference-based read mapping was performed using two different parameter settings (see below), by aligning the trimmed sequencing reads against the *V. vulpes* reference genome (assembly version: GCF_003160815.1_VulVul2.2; Kukekova et al., 2018) using BWA-MEM v0.7.17 (Li and Durbin, 2009) with default parameters. We then used SAMtools v1.10 (Li et al., 2009) to obtain sorted bam files, followed by using GATK v4.2.2.0 (https://gatk.broadinstitute.org/hc/en-us) to remove PCR duplicates

using MARKDUPLICATESSPARK and to filter out bad read mates, reads with mapping quality zero and reads which mapped ambiguously (Nater et al., 2017), and again SAMtools to extract the mitochondrial reads that mapped to the mtDNA scaffold (NC_008434, Arnason et al. 2006) of the reference genome. HAPLOTYPECALLER in GATK was used to call variants, using two different parameter settings, using as values for the flag *--sample-ploidy:* 1 for haploid (ploidy 1; Approach 3), and 2 for diploid (ploidy 2; Approach 4), each yielding a separate VCF file. Finally, FastaAlternateReferenceMaker from GATK was used to convert the two VCF files from Approaches 3 and 4 to FASTA format.

Geneious Prime v2022.2.2 (http://www.geneious.com) was used to align and annotate the genes of all obtained mitogenome sequences to the mitogenome of *V. vulpes* (NC_008434, Arnason et al. 2006) and for trimming the poorly-aligned and incompletely assembled tandem repeat region within the D-loop (see Results). We verified the annotation and visualisation of the circular mitogenome using the web-based programme GeSeq – Annotation of Organellar Genomes (Tillich et al. 2017, hosted at Chlorobox 2020, https://chlorobox.mpimp-golm.mpg.de/index.html).

Phylogenetic analysis

To determine the phylogenetic relationship of *V. rueppellii* with *V. vulpes* and other foxes, we generated an alignment of various fox (subfamily Vulpini) outgroup sequences, along with a comprehensive coverage of all previously described lineages/clades of both focal species. We downloaded from GenBank representative haplotypes from Statham et al. (2014), sequences from Leite et al. (2015), reconstructed mitogenomes from two *V. rueppellii* individuals from Morocco, based on whole genome sequencing data by Rocha et al. (2023) using NOVOPlasty (Dierckxsens et al., 2017), available *V. vulpes* complete mitogenome sequences, along with other available mitogenomes from Vulpini, plus raccoon dog (*Nyctereutes procyonoides*) as an outgroup (table 1). Then, MUSCLE v3.8 (Edgar, 2004) as implemented in Geneious Prime was used for aligning the sequences and to generate a FASTA file. We used W-IQ-TREE (Trifinopoulos et al., 2016) to construct a phylogenetic tree using a maximum likelihood approach based on the TIM2 model of sequence evolution, including a discrete Gamma model with 4 rate categories (TIM2+F+G4), which had been determined as the optimal model by Modelfinder (Kalyaanamoorthy et al., 2017) implemented in IQ-TREE. We used 1000 ultrafast bootstrap replications (Minh et al., 2013) to determine branch support. Missing data in the short sequences was assigned as N, allowing IQ-TREE to use any non-gap characters (GCAT) at that site for site likelihood calculation. This ensures full utilization of the available data in the alignment, while also accurately treating missing data (see http://www.iqtree.org/doc/Frequently-Asked-Questions). FIGTREE 1.4.4 (https://github.com/rambaut/figtree/releases) was used to visualize the obtained phylogeny.

Results and Discussion

After trimming and quality control steps (see Methods), a total of 216,237,628 read pairs (out of 218,591,573) remained for the sequenced *V. rueppellii* individual. For NOVOPlasty, the number of assembled mitogenome read pairs was 792,000 (0.36% of total raw read pairs), yielding an average mitogenome coverage of 16,225x. For MITObim, the number of assembled read pairs was 353,396 (16 % of trimmed read pairs; average coverage not determined), and for reference mapping 427,955 (0.20%) mapped, trimmed read pairs gave an average coverage of 7,372x. The resulting total mitogenome length for *V. rueppellii* from the two from reference-mapping approaches was 16,813 bp, as expected matching the length and genomic organisation (Fig. 1) of the *V. vulpes* mtDNA reference genome (Arnason et al. 2006) that the sequencing reads had been mapped against. Obtained de novo assembly lengths were 16,517 bp (NOVOPlasty) and 17,842 bp (MITObim).

After aligning sequences from all four approaches, at the 5' end there was 219 bp overhang from MITObim and a 369 bp from NOVOPlasty. These regions matched the end of the D-loop, and after moving these sequences to the ends of the MITObim and NOVOPlasty assemblies, respectively, the four approaches aligned perfectly, except for a region towards the end of the D-loop: a 712 bp portion of the D-loop (positions 16,102 to 16,813 in the *V. vulpes* mtDNA reference genome; Arnason et al. 2006) contained a repetitive region known to contain tandemly repeated variations of a ca. 8-14 bp repeat, and showed unreliable alignment characteristics (indels, uneven read coverage and apparent heterozygous sites in the read data, despite mtDNA being a haploid genome). We attributed this to the failure of short-read based sequencing methods to properly assemble the complete D-loop (du Plessis et al., 2023b; Formenti et al., 2021) especially around the tandem repeat region, based on the 151 bp read length used here. Following the trimming of this 712 bp region of the D-loop from the four approaches and the 3' flanking region of the tandem repeat region, we retained positions 1-16,101 bp of the alignment for phylogenetic analysis. Across this remaining alignment, the four mitogenome sequences obtained from different bioinformatic approaches yielded identical sequences. The sequences from the four approaches matched the *V. vulpes* mitochondrial reference genome (Arnason et al., 2006) in terms of organisation and length of all coding sections (13 protein-coding, two rRNA and 22 tRNA genes), and also in the placement of the non-coding region (D-loop) within the mitogenome (Fig. 1).

No previous studies have specifically compared the performance of the four assembly approaches used here to extract the whole mitogenome of *V. rueppellii*. One study by Machado et al. (2016) on frogs compared de novo (not including NOVOPlasty) and reference-based mapping using different software and pipelines than we used here, and found the baiting and iterative mapping approach by MIRA/MITObim to be the best approach to extract the mitogenome, even from a low number of reads. It is possible that this better performance of MIRA/MITObim could be specific for frog mitogenomes (Machado et al. 2016). In our study, the MITObim assembly included numerous tandem repeat insertions (which were lacking in the reference genome and the NOVOPlasty assembly) within the tandem repeat region (positions 16113 and 16522 of the reference genome from Arnason *et al.* (2006)). The number of assembled mitogenome reads and coverage from NOVOPlasty exceeds those from both MITObim and reference-mapping approaches. In Dierckxsens et al. (2017), NOVOPlasty outperformed MITObim slightly in terms of accuracy and memory usage, although its benefits may be especially prominent for AT-rich genomic regions. Similarly, Termignoni-Garcia et al (2023) found that NOVOPlasty gave more accurate mitogenome assemblies compared to MITObim for Corvidae (Aves),

where the latter produced occasional insertions. Our analyses suggest that, at least when sufficient coverage is obtained and the tandem repeat region is excluded, the investigated approaches can yield identical results. However, we caution that different de novo assembly and reference mapping approaches could yield different mitogenome sequences (Ip et al., 2022; Termignoni-Garcia et al., 2023), so a careful evaluation of any methodological should be conducted before joint evolutionary analysis of sequences obtained from different approaches.

Maximum likelihood analysis yielded a phylogenetic tree (Fig. 2) in which two subclades *V. rueppellii* clustered inside the previously identified Palearctic clade (Basuony et al. 2023; Statham et al. 2014) of *V. vulpes* with a bootstrap value, BV = 95, rendering the latter paraphyletic. This paraphyly is in accordance with previous findings by Basuony et al. (2023) and Leite et al. (2015), who based on shorter mtDNA fragments (partial cytochrome b and D-loop sequences) demonstrated clustering of *V. rueppellii* within *V. vulpes.* This topology was also recovered by Rocha et al. (2023) based on mitogenome data. The support for the main clades in our obtained tree was high (ranging from BV>80 to 100), highlighting the usefulness of complete mitogenome in phylogenetic inference. The phylogenetic relationships of both *V. vulpes* and *V. rueppellii* with other fox species was also in accordance with previous genetic studies by Rocha et al. (2023) and Leite et al. (2015) (Fig 2).

Although tentative due to the low sample size for mitogenome sequences analysed here, there appears to be a phylogeographic pattern within *V. rueppellii:* the two SRR samples and VruRMo1 (all from south Morocco) represent subclade 1, which occurs from west of the Nile in Egypt across north Africa to the Atlantic Ocean (Basuony et al., 2023), while the newly sequenced mitogenome of sample 375 represents subclade 2, which is found in the Egyptian eastern desert and Arabia (Basuony et al., 2023). This pattern tentatively suggests some level of longitudinal phylogeographic structuring within *V. rueppellii*.

In contrast to the mitochondrial paraphyly signal, Rocha et al (2023) demonstrated that nuclear DNA (nuDNA) shows *V. vulpes* and *V. rueppellii* as distinct lineages, where individuals (according to their average genomic distances) cluster as reciprocally monophyletic sister species. Such conflict between mtDNA and nuDNA (mito-nuclear discordance) is widespread in mammals and other animals (Toews and Brelsford., 2012). Introgression and incomplete lineage sorting (ILS) have been proposed as the main factors leading to mito-nuclear discordance (Scornavacca and Galtier, 2017; Tamashiro et al., 2019). Rocha *et al*. (2023) examined these two scenarios for *V. vulpes* and *V. rueppellii* and found clear evidence of introgression from *V. rueppellii* into North African populations of *V. vulpes*, intriguingly involving genomic regions which are enriched for genes putatively conveying adaptations to persistence in arid environments. Our high support for the mitochondrial paraphyly suggest that also this portion of the genome has likely passed the species boundary between *V. rueppellii* and *V. vulpes* due to introgressive hybridization.

In conclusion, we here report the first annotated mitogenome of *V. rueppellii*, which is 'near-complete' due to incomplete characterisation of the tandem repeats in the D-loop. This genome will be useful for future phylogenetic and other evolutionary studies of the little-studied *V. rueppellii* and its relatives. Our results showed consistency of the de novo and reference-based approaches in extracting near-complete mitogenomes, at least when excluding the tandem repeat region. Assembling highly repetitive regions such as this will likely require read lengths that span across the entire repeat region. We also confirmed the paraphyly of *V. vulpes*, with clustering of *V. rueppellii* inside the Palearctic clade of the former, adding to previous suggestions (Leite et al. 2015, Basuony et

al. 2023, Rocha et al. 2023) of introgressive hybridization between the two species. Future combined analyses of mitogenomes and nuDNA for additional individuals across the range of the two species may shed further light on their evolutionary history, including how such introgression has shaped their adaptations and population structuring.

Acknowledgements

AEB was supported by the Cardiff University Vice Chancellor's International Scholarship of Excellence. We thank the Cardiff University School of Biosciences for funding and support, Elizabeth Chadwick for support, Sarah du Plessis for discussions and advice, and the following organizations and individuals for assistance with sample collection: Zoology Department, Faculty of Science, Al-Azhar University, Egypt; Egyptian Environmental Affairs Agency (EEAA), Nature Conservation Sector, Egypt; Dr. Fouad Abdel-Hamid; Dr. Ahmed Badry; Mr. Mohamed Anwar late Dr. Ali Al-Korey; Dr. Magdy Saleh; Mr. Naser Nabil and Dr. Mahmoud Younes. Sampling was conducted in accordance with national wildlife conservation and veterinary regulations. Sample import was under the Animal and Plant Health Agency (APHA) permit ITIMP21.0069. Bioinformatic analyses were performed on the Cardiff School of Biosciences' Biocomputing Hub HPC/Cloud infrastructure, and we thank Dr Ian Merrick and Andy Ells for their support. We thank the anonymous reviewers and editor for their helpful suggestions.

Declaration of competing interests

There is no conflict of interest.

Data availability

We have deposited the obtained mitogenome sequence from NOVOPlasty in GenBank (accession number: OQ274912) and the raw reads in SRA (BioProject accession number: PRJNA1083499).

References

Anijalg, P., Ho, S.Y.W., Davison, J., Keis, M., Tammeleht, E., Bobowik, K., Tumanov, I.L., Saveljev, A.P., Lyapunova, E.A., Vorobiev, A.A., Markov, N.I., Kryukov, A.P., Kojola, I., Swenson, J.E., Hagen, S.B., Eiken, H.G., Paule, L., Saarma, U., 2018. Large-scale migrations of brown bears in Eurasia and to North America during the Late Pleistocene. J Biogeogr 45, 394–405. doi.org/10.1111/jbi.13126

Arnason, U., Gullberg, A., Janke, A., Kullberg, M., Lehman, N., Petrov, E.A., Väinölä, R., 2006. Pinniped phylogeny and a new hypothesis for their origin and dispersal. Mol Phylogenet Evol 41, 345–354. doi.org/10.1016/j.ympev.2006.05.022

Basuony, A.E., Saleh, M., Sarhan, M., Younes, M., Abdel-Hamid, F., Rodriguez Fernandes, C., Vercammen, P., Aboshaala, F., Bounaceur, F., Chadwick, E.A., Hailer, F., 2023. Paraphyly of the widespread generalist red fox (*Vulpes vulpes*): introgression rather than recent divergence of the arid-adapted Rüppell's fox (*Vulpes rueppellii*)? Biological Journal of the Linnean Society 138, 453–469. doi.org/10.1093/biolinnean/blad001

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. doi.org/10.1093/bioinformatics/btu170

Briscoe, A.G., Hopkins, K.P., Waeschenbach, A., 2016. High-throughput sequencing of complete mitochondrial genomes. Marine Genomics: Methods and Protocols 45–64.

Chevreux, B., Wetter, T., Suhai, S., 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) '99, GCB, Hannover, Germany. 45–56.

Dierckxsens, N., Mardulyn, P., Smits, G., 2017. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. Nucleic Acids Res 45. doi.org/10.1093/nar/gkw955

du Plessis, S. J., Hong, S., Lee, B., Koepfli, K. P., Chadwick, E. A., & Hailer, F., 2023a. Mitochondrial genome-based synthesis and timeline of Eurasian otter (*Lutra lutra*) phylogeography. *Animal Cells and Systems*, 27(1), 366-377.

du Plessis, S.J., Blaxter, M., Koepfli, K.-P., Chadwick, E.A., Hailer, F., 2023b. Genomics reveals complex population history and unexpected diversity of Eurasian otters (*Lutra lutra*) in Britain relative to genetic methods. *Mol Biol Evol* 40(11):msad207. doi.org/10.1093/molbev/msad207

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792–1797. doi.org/10.1093/nar/gkh340

Finstermeier, K., Zinner, D., Brameier, M., Meyer, M., Kreuz, E., Hofreiter, M., Roos, C., 2013. A Mitogenomic Phylogeny of Living Primates. PLoS One 8, 1–10. doi.org/10.1371/journal.pone.0069504

Formenti, G., Rhie, A., Balacco, J., Haase, B., Mountcastle, J., Fedrigo, O., Brown, S., Capodiferro, M.R., Al-Ajli, F.O., Ambrosini, R., Houde, P., Koren, S., Oliver, K., Smith, M., Skelton, J., Betteridge, E., Dolucan, J., Corton, C., Bista, I., Torrance, J., Tracey, A., Wood, J., Uliano-Silva, M., Howe, K., McCarthy, S., Winkler, S., Kwak, W., Korlach, J., Fungtammasan, A., Fordham, D., Costa, V., Mayes, S., Chiara, M., Horner, D.S., Myers, E., Durbin, R., Achilli, A., Braun, E.L., Phillippy, A.M., Jarvis, E.D., Kirschel, A.N.G., Digby, A., Veale, A., Bronikowski, A., Murphy, B., Robertson, B., Baker, C., Mazzoni, C., Balakrishnan, C., Lee, C., Mead, D., Teeling, E., Aiden, E.L., Todd, E., Eichler, E., Naylor, G.J.P., Zhang, G., Smith, J., Wolf, J., Touchon, J., Delmore, K., Jakobsen, K., Komoroske, L., Wilkinson, M., Genner, M., Pšenička, M., Fuxjager, M., Stratton, M., Liedvogel, M., Gemmell, N., Minias, P., Dunn, P.O., Sudmant, P., Morin, P., Ayub, Q., Kraus, R., Vernes, S., Smith, S., Lama, T., Edwards, T., Smith, T., Gilbert, T., Marques-Bonet, T., Einfeldt, T., Venkatesh, B., Johnson, W., Warren, W., Bukhman, Y., 2021. Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. Genome Biol 22, 1–22. doi.org/10.1186/s13059-021-02336-9

Geffen, E., Mercure, A., Girman, D.J., Macdonald, D.W., Wayne, R.K., 1992. Phylogenetic relationships of the fox-like canids: mitochondria1 DNA restriction fragment, site and cytochrome b sequence analyses. J Zool 228, 27–39.

Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. Nucleic Acids Res 41. doi.org/10.1093/nar/gkt371

Ip, E.K., Troup, M., Xu, C., Winlaw, D.S., Dunwoodie, S.L. and Giannoulatou, E., 2022. Benchmarking the effectiveness and accuracy of multiple mitochondrial DNA variant callers: practical implications for clinical application. *Frontiers in Genetics*, *13*, p.692257.Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat Methods 14, 587–589. doi.org/10.1038/nmeth.4285

Keis, M., Remm, J., W Ho, S.Y., Davison, J., Tammeleht, E., Tumanov, I.L., Saveljev, A.P., Männil, P., Kojola, I., Abramov, A. V, Margus, T., Saarma, U., 2013. Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. J Biogeogr 5, 916–927. doi.org/10.1111/jbi.12043

Koepfli, K.P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., Thalmann, O., Silva, P., Fan, Z., Yurchenko, A.A., Dobrynin, P., Makunin, A., Cahill, J.A., Shapiro, B., Álvares, F., Brito, J.C., Geffen, E., Leonard, J.A., Helgen, K.M., Johnson, W.E., O'Brien, S.J., Van Valkenburgh, B., Wayne, R.K., 2015. Genome-wide evidence reveals that African and Eurasian golden jackals are distinct species. Current Biology 25, 2158–2165. doi.org/10.1016/j.cub.2015.06.060

Kukekova, A. v., Johnson, J.L., Xiang, X., Feng, S., Liu, S., Rando, H.M., Kharlamova, A. v., Herbeck, Y., Serdyukova, N.A., Xiong, Z., Beklemischeva, V., Koepfli, K.P., Gulevich, R.G., Vladimirova, A. v., Hekman, J.P., Perelman, P.L., Graphodatsky, A.S., O'Brien, S.J., Wang, X., Clark, A.G., Acland, G.M., Trut, L.N., Zhang, G., 2018. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. Nat Ecol Evol 2, 1479–1491. doi.org/10.1038/s41559-018-0611-6

Leite, J.V., Álvares, F., Velo-Antón, G., Brito, J.C., Godinho, R., 2015. Differentiation of North African foxes and population genetic dynamics in the desert—insights into the evolutionary history of two sister taxa, *Vulpes rueppellii* and *Vulpes vulpes*. Org Divers Evol 15, 731–745. doi.org/10.1007/s13127-015-0232-8

Lenain, D.M., 2000. Fox populations of a protected area in Saudi Arabia. Master's thesis. University of Hertfordshire.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. doi.org/10.1093/bioinformatics/btp352

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E.A., Ponting, C.P., Galibert, F., Smith, D.R., DeJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.P., Parker, H.G., Pollinger, J.P., Searle, S.M.J., Sutter, N.B., Thomas, R., Webber, C., Lander, E.S., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803–819. doi.org/10.1038/nature04338

Machado, D.J., Lyra, M.L., Grant, T., 2016. Mitogenome assembly from genomic multiplex libraries: Comparison of strategies and novel mitogenomes for five species of frogs. Mol Ecol Resour 16, 686–693. doi.org/10.1111/1755-0998.12492

Mallon, D., Murdoch, J.D., Wacher, T., 2015. Vulpes rueppellii. The IUCN Red List of Threatened Species 2015.

Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol 30, 1188–1195. doi.org/10.1093/molbev/mst024

Murdoch, J.D., Drew, C., Llanes, I.B., Tourenq, C., 2007. Rüppell's foxes in Al Dhafra, United Arab Emirates. Canid News 10, 1–6.

Nater, A., Mattle-Greminger, M.P., Nurcahyo, A., Nowak, M.G., de Manuel, M., Desai, T., Groves, C., Pybus, M., Sonay, T.B., Roos, C., Lameira, A.R., Wich, S.A., Askew, J., Davila-Ross, M., Fredriksson, G., de Valles, G., Casals, F., Prado-Martinez, J., Goossens, B., Verschoor, E.J., Warren, K.S., Singleton, I., Marques, D.A., Pamungkas, J., Perwitasari-Farajallah, D., Rianti, P., Tuuga, A., Gut, I.G., Gut, M., Orozco-terWengel, P., van Schaik, C.P., Bertranpetit, J., Anisimova, M., Scally, A., Marques-Bonet, T., Meijaard, E., Krützen, M., 2017. Morphometric, Behavioral, and Genomic Evidence for a New Orangutan Species. Current Biology 27, 3487-3498.e10. doi.org/10.1016/j.cub.2017.09.047

Rivero, E.R.C., Neves, A.C., Silva-Valenzuela, M.G., Sousa, S.O.M., Nunes, F.D., 2006. Simple salting-out method for DNA extraction from formalin-fixed, paraffin-embedded tissues. Pathol Res Pract 202, 523–529. doi.org/10.1016/j.prp.2006.02.007

Rocha, J.L., Silva, P., Santos, N., Nakamura, M., Afonso, S., Qninba, A., Boratynski, Z., Sudmant, P.H., Brito, J.C., 2023. North African fox genomes show signatures of repeated introgression and adaptation to life in deserts. https://doi.org/10.1038/s41559-023-02094-w

Rosevear, D.R., 1974. The carnivores of West Africa. Trustees of the British Museum (Natural History), London, United Kingdom.

Scornavacca, C., Galtier, N., 2017. Incomplete lineage sorting in mammalian phylogenomics. Syst Biol 66, 112–120. https://doi.org/10.1093/sysbio/syw082

Sillero-Zubiri, C., Hoffmann, M., Macdonald, D.W., 2004. Canids: Foxes, Wolves, Jackals and Dogs: Status survey and conservation action plan. Gland, Switzerland: IUCN/SSC Canid Specialist Group.

Statham, M.J., Murdoch, J., Janecka, J., Aubry, K.B., Edwards, C.J., Soulsbury, C.D., Berry, O., Wang, Z., Harrison, D., Pearch, M., Tomsett, L., Chupasko, J., Sacks, B.N., 2014. Range-wide multilocus phylogeography of the red fox reveals ancient continental divergence, minimal genomic exchange and distinct demographic histories. Mol Ecol 23, 4813–4830. doi.org/10.1111/mec.12898

Sun, L.W., Yang, Y., Li, G.Y., 2019. The complete mitochondrial genome of the raccoon dogs (Canidae: *Nyctereutes ussurienusis*) and intraspecific comparison of three Asian raccoon dogs. Mitochondrial DNA B Resour 4, 670–671. https://doi.org/10.1080/23802359.2017.1419081

Sun, W.-L., Zhong, W., Bao, K., Liu, H.-L., Ya-Han, Y., Wang, Z., Li, G.-Y., 2016. The complete mitochondrial genome of silver fox (Caniformia: Canidae). Mitochondrial DNA A DNA Mapp Seq Anal 27, 3348–3350. doi.org/10.3109/19401736.2015.1018216

Tamashiro, R.A., White, N.D., Braun, M.J., Faircloth, B.C., Braun, E.L., Kimball, R.T., 2019. What are the roles of taxon sampling and model fit in tests of cyto-nuclear discordance using avian mitogenomic data? Mol Phylogenet Evol 130, 132–142. doi.org/10.1016/j.ympev.2018.10.008

Termignoni Garcia, F., Bougiouri, K. and Edwards, S.V., 2023. Heteroplasmy and tandem repeats reveal adaptation to elevation in the New World Jays (Aves: Corvidae). *bioRxiv*, doi.org/10.1101/2023.12.05.570157

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., Greiner, S., 2017. GeSeq - Versatile and accurate annotation of organelle genomes. Nucleic Acids Res 45, W6–W11. doi.org/10.1093/nar/gkx391

Timbó, R.V., Togawa, R.C., Costa, M.M.C., Andow, D.A., Paula, D.P., 2017. Mitogenome sequence accuracy using different elucidation methods. PLoS One 12, 1–15. doi.org/10.1371/journal.pone.0179971

Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol 21, 3907–3930. doi.org/10.1111/j.1365-294X.2012.05664.x

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44, W232–W235. doi.org/10.1093/NAR/GKW256

Williams, J.B., Lenain, D., Ostrowski, S., Tieleman, B.I., Seddon, P.J., 2002. Energy expenditure and water flux of Rüppell's foxes in Saudi Arabia. Physiological and Biochemical Zoology 75, 479–488. doi.org/10.1086/344490

Yan, S.Q., Guo, P.C., Yue, Y., Li, W.H., Bai, C.Y., Li, Y.M., Sun, J.H., Zhao, Z.H., 2016. The complete sequence of the mitochondrial genome of Arctic fox (*Alopex lagopus*). Mitochondrial DNA Part A 27, 4095–4096. doi.org/10.3109/19401736.2014.1003860

Yang, X., Zhao, C., Zhang, H., Zhang, J., Chen, L., Sha, W., Liu, G., 2016. Complete mitochondrial genome of the fennec fox (*Vulpes zerda*). Mitochondrial DNA 27, 637–638. doi.org/10.3109/19401736.2014.908467

Yu, J.N., Kim, S., Oh, K., Kwak, M., 2012. Complete mitochondrial genome of the Korean red fox *Vulpes vulpes* (Carnivora, Canidae). Mitochondrial DNA 23, 118–119. doi.org/10.3109/19401736.2011.653800

Zhang, J., Zhang, H., Zhao, C., Chen, L., Sha, W., Liu, G., 2015. The complete mitochondrial genome sequence of the Tibetan red fox (*Vulpes vulpes montana*). Mitochondrial DNA 26, 739–741. doi.org/10.3109/19401736.2013.845766

Zhao, C., Zhang, H., Liu, G., Yang, X., Zhang, J., 2016. The complete mitochondrial genome of the Tibetan fox (*Vulpes ferrilata*) and implications for the phylogeny of Canidae. C R Biol 339, 68–77. doi.org/10.1016/j.crvi.2015.11.005

Zhong, H.M., Zhang, H.H., Sha, W.L., Zhang, C. De, Chen, Y.C., 2010. Complete mitochondrial genome of the red fox (*Vulpes vulpes*) and phylogenetic analysis with other canid species. Zool Res 31, 122–130. doi.org/10.3724/SP.J.1141.2010.02122

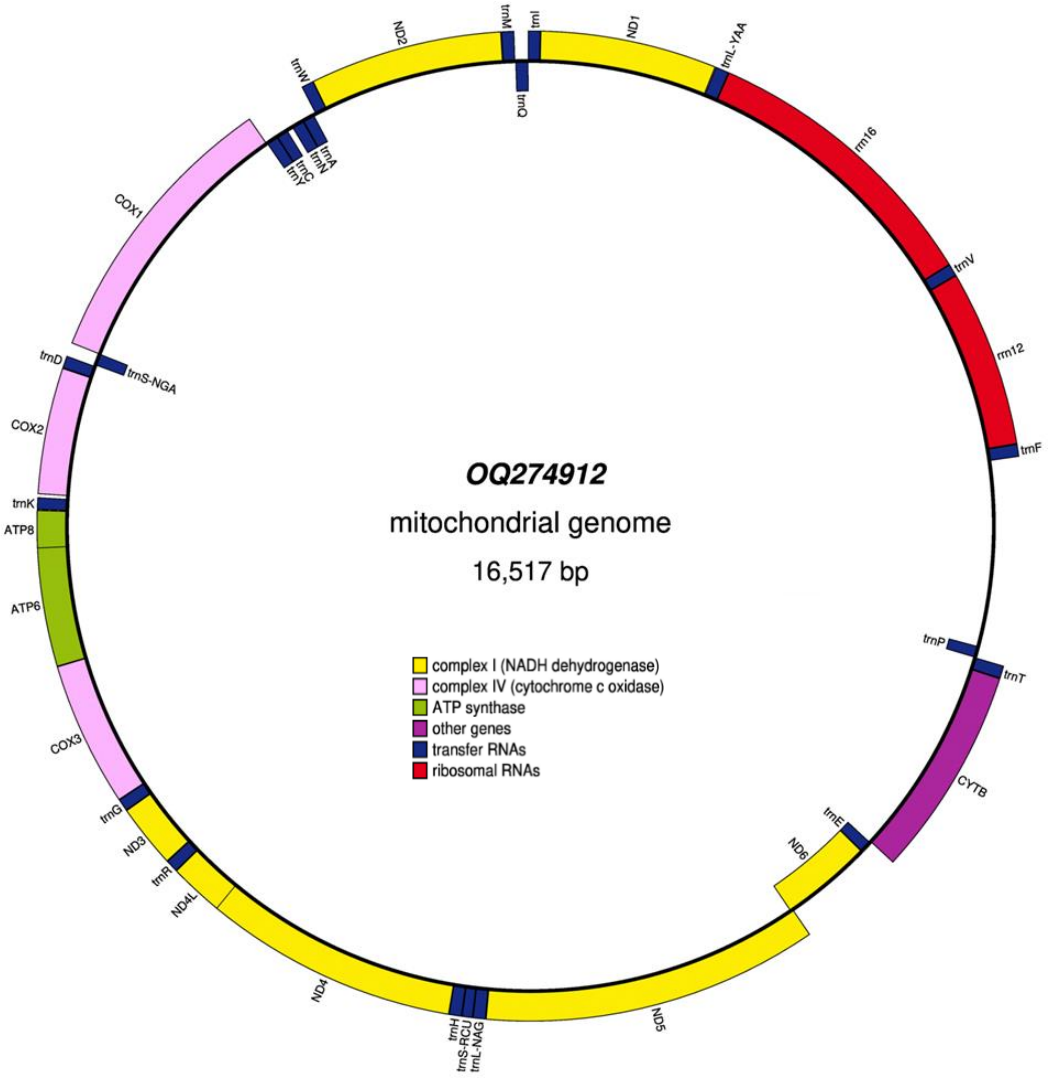
Tables

Table 1: Sequences included in the phylogenetic analysis. Names of mitochondrial clades and subclades follow the terminology of Statham et al. (2014) and Basuony et al. (2023).

Sample ID/Haplotype/ Accession number/ SRA ID	Mitochondrial		Reference for sequence
	Clade	Subclade	
V. rueppellii			
375 (MITObim), NOVOPlasty, and reference-based: ploidy 1&2	Palearctic	2	This study
SRR24040959	Palearctic	1	Rocha et al. 2023
SRR24040997	Palearctic	1	Rocha et al. 2023
* V.ruRMo1 (KJ597994 and KJ597968)	Palearctic	1	Leite et al. 2015
V. vulpes			
* Oo24	Nearctic	I	Statham et al. 2014
* Fo12	Nearctic	II	Statham et al. 2014
* Ao63	Nearctic	III	Statham et al. 2014
* B2o106	Holarctic	I	Statham et al. 2014
* Uo211	Holarctic	II	Statham et al. 2014
* Go78	Holarctic	III	Statham et al. 2014
* Wo156	Holarctic	IV	Statham et al. 2014
* W4o175	Holarctic	V	Statham et al. 2014
* U35o98	Holarctic	VI	Statham et al. 2014
* U32o107	Holarctic	VII	Statham et al. 2014
* U12o115	Holarctic	VIII	Statham et al. 2014
* U8o118	Holarctic	IX	Statham et al. 2014
* Xo244	Africa 1	n/d	Statham et al. 2014
* X2o252	Africa 2	n/d	Statham et al. 2014
* X3o262	Africa 2	n/d	Statham et al. 2014
* V.vuMO4 (KJ598014, KJ597980)	Africa 1	n/d	Leite et al. 2015
* V.vuMO1 (KJ597977, KJ598009)	Africa 2	n/d	Leite et al. 2015
* Y2o197	Palearctic	I	Statham et al. 2014
* Yo202	Palearctic	II	Statham et al. 2014
* Yo155	Palearctic	III	Statham et al. 2014
* Y9o117	Palearctic	IV	Statham et al. 2014
KP342452	Nearctic	n/d	Sun et al. 2016b
GQ374180	Holarctic	n/d	Zhong et al. 2010
KF387633	Holarctic	n/d	Zhang et al. 2015
JN711443	Holarctic	n/d	Yu et al. 2012
AM181037	Holarctic	n/d	Arnason et al. 2006
MN122913	Holarctic	n/d	DNAmark project, unpublished
KT448287	Holarctic	n/d	Koepfli et al., 2015
V. corsac, NC_023958	NA	n/d	Unpublished
V. ferrilata, NC_027935	NA	n/d	Zhao et al., 2016
V. lagopus, NC_026529	NA	n/d	Yan et al. 2016
V. zerda, KJ603240	NA	n/d	Yang et al. 2016
V. pallida, SRR24040935	NA	n/d	Rocha et al. 2023
V. chama, ON756054.3	NA	n/d	Unpublished
Nyctereutes procyonoides, MG256392	NA	n/d	Sun et al. 2019

* Fragments of cytochrome b and D-loop (each <400 bp), included in the present phylogenetic analysis to anchor analysed mitogenomes to existing *Vulpes* clade terminology. GenBank accession numbers for Statham et al. (2014) haplotypes are provided in their supplementary information. n/d: not determined. NA: not applicable.

Figures



425

Figure 1: Annotated mitogenome of *V. rueppellii* as obtained from GeSeq. OQ274912 is the GenBank accession number.

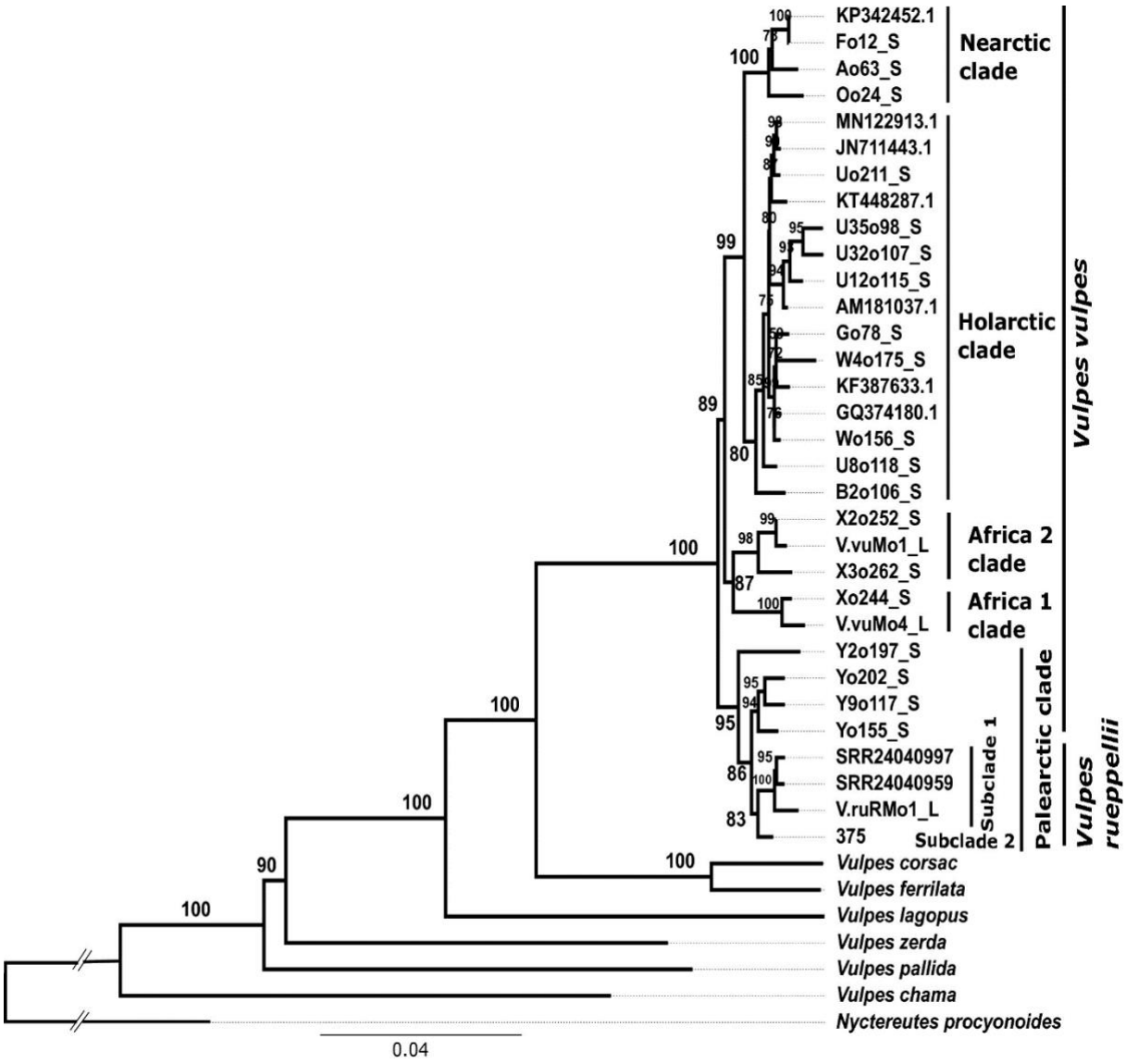


Figure 2: Maximum likelihood mitogenome tree obtained from IQ-TREE based on an alignment of 16,163 bp for 38 sequences (28 *V. vulpes*, 4 *V. rueppellii* and 6 other fox species) and *Nyctereutes procyonoides* as an outgroup. Numbers on branches denote ultrafast bootstrap support from 1000 replicates. Sample names are followed by **S** for sequences from Statham *et al.* (2014) and **L** for those from Leite *et al*. (2015), IDs that are not followed by a letter are GenBank and SRA accession numbers and the newly sequenced *V. rueppellii* is 375 (see table 1). Scale bar shows nucleotide substitutions per site.