VERTEBRATE ZOOLOGY

70 (4): 667–678
2020

SENCKENBERG

DNA barcoding and molecular taxonomy of dark-footed forest shrew *Myosorex cafer* in the Eastern Cape and KwaZulu-Natal, South Africa

Emmanuel Matamba[1], Leigh R. Richards[2], Michael I. Cherry[1] & Ramugondo V. Rambau[1]

[1] Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa; ematamba@sun.ac.za — [2] Department of Mammalogy, Durban Natural Science Museum, P O Box 4085, Durban, 4001, South Africa

Submitted June 2, 2020.
Accepted October 20, 2020.
Published online at www.senckenberg.de/vertebrate-zoology on November 23, 2020.
Published in print Q4/2020.

Editor in charge: Clara Stefen

Abstract

There is a paucity of molecular DNA barcoding informatics on the South African fauna, particularly on terrestrial small mammals. This study tested the utility of DNA barcoding in the dark-footed forest shrew (*Myosorex cafer*) from forested regions of the Eastern Cape and KwaZulu-Natal provinces of South Africa. Sampled forests included coastal scarp, dune forests and inland Afromontane mistbelt forests. Sequences of mtDNA cytochrome oxidase subunit I (COI, 623 bp), were generated for a total of 78 specimens representing *Myosorex cafer* ($n=72$), *Myosorex varius* ($n=2$), *Crocidura cyanea* ($n=2$) and *C. mariquensis* ($n=2$). Due to the fragmented nature of these forests, we also investigated the cranial morphology of *Myosorex cafer,* which is strictly confined to forests. Analyses of sequence data produced phylogenetic trees that were consistent with morphological identifications. Genetic data suggest that the movement of these animals between other forest types and the Amatole mistbelt forests has been restricted, as they are too far west of scarp forests to have been recolonized by them. This is the first study that supplies COI sequences of a South African *Myosorex* species, thus increasing the availability of DNA barcodes of South African small mammals on BOLD.

Key words

DNA barcoding, forest, forest fragmentation, shrews.

Introduction

Although mammals are ranked as the best-studied animal group (Wilson & Reeder, 2005), small mammal taxonomic identification is limited without the collection of voucher specimens for dental and cranial analysis (Baker & Bradley, 2006). Morphological convergence and similarities in pelage may hamper species identification (Lai *et al*., 2008; Losos, 2011; Jacobs *et al*., 2013). Historically, the taxonomy of African shrews was mostly based on morphological characters (Meester, 1986). However, the identification of small mammals using morphology may be ambiguous, and molecular work has demonstrated that morphology is not always reliable in species identification. For example, the morphologically conservative genus *Sylvisorex* was found to be polyphyletic using molecular markers (Quérouil *et al*., 2001). More recently *Myosorex tenuis* (previously classified as *Myosorex cafer*) was separated from *Myosorex cafer* based on DNA differences (Taylor *et al*., 2013; Taylor *et al*., 2017). To solve the challenges associated with morphological identifications a DNA barcoding system for animal life has been proposed. This barcoding system is based on cytochrome c oxidase subunit I (COI) sequences to classify animals (Hebert *et al*., 2003a). The main purpose of DNA barcoding is to provide a quick and easy method of animal identification without the need for taxonomic expertise (Smith *et al*., 2005; Hajibabaei *et al*.,

2006). DNA barcoding has been used as an important tool in species inventories, based on its precision in differentiating species (Ward *et al*., 2005; Smith *et al*., 2005; Clare *et al*., 2007). Further, it also helps the researcher to conduct rapid taxonomic assignments and it can detect cryptic species (Hajibabaei *et al*., 2007; Borisenko *et al*., 2008). Consequently, DNA barcoding has been successful in assigning individuals to species level, particularly among small mammals (Hebert *et al*., 2003a, b; Hajibabaei *et al*., 2005; Hajibabaei *et al*., 2006; Ivanova *et al*., 2006; Hajibabaei *et al*., 2007; Borisenko *et al*., 2008; Lu *et al*., 2012; Nicolas *et al*., 2012). DNA barcoding involves tree-based analysis such as neighbour joining which is a distance-based method for reconstructing phylogenetic tree, as well as tree-independent approaches such as Automatic Barcode Gap Discovery (ABGD), Clustering 16S rRNA for OTU Prediction (CROP), the General Mixed Yule Coalescent (GMYC) and the Bayesian Species Delineation by BPP (Pons *et al*., 2006; Hao *et al*., 2011; Puillandre *et al*., 2012; Yang, 2015; DE Salle & Goldstein, 2019). The process of species delimitation in DNA barcoding can be considered accurate if more than one method is used to allow cross-validation (Jörger *et al*., 2012). In this study morphology, tree-based and tree independent, ABGD methods have been used. The ABGD is an automatic procedure that groups sequences based on their barcoding gap (nucleotide difference between intraspecific and interspecific difference) and it infers a model-based one-sided confidence limit for intraspecific divergence using a wide range of prior intraspecific divergence (Puillandre *et al*., 2012).

There is a need to barcode South African animal species because the country has more than 65500 known animal species (Hamer, 2013), and as of 2013, only 2.3% of South African species had representation on the Barcoding of Life Database (BOLD; Ratnasingham & Hebert, 2007). Barcodes represented were dominated by fish (approximately 36%), birds (5.4%) and mammals (4.89%); for other taxonomic groups, the species representation is lower than 2% (Da Silva and Willows-Munro, 2016). African shrews represent one of the underrepresented taxa and no *Myosorex* species are represented on BOLD. *Myosorex cafer* is a forest habitat specialist which has similar habitat preferences as the closely related *M. tenuis* (Thomas & Schwann, 1905) which shares a nearly identical phenotype consistent with the low sequence divergence (Taylor *et al*., 2013); the sister taxa to these species is the savanna dwelling *M. varius* (Skinner & Chimimba, 2005). The three congeneric species have close phylogenetic associations and while their morphological differences are minor, they can be discerned by multivariate analyses of craniodental measurements (Meester, 1958; Kearney, 1993).

South African forests cover a small land surface area (Low & Rebelo, 1996), and their distribution is severely fragmented (Geldenhuys, 1989; Lawes, 1990; Eeley *et al*., 1999). This fragmentation is attributed to repeated historical climatic fluctuations dating up to 100,000 years ago, including the Last Glacial Maxima approximately 18,000 years before the present (Lawes, 1990). More recently, the forest fragmentation has been exacerbated by human settlements and human-mediated activities such as overharvesting of medicinal plants, fuelwood, and alien timber plantations (Dold & Cocks, 1998; von Maltitz *et al*., 2003). The impact of the climate oscillations is particularly evident in the Eastern Cape where at least three forest groups co-occur namely: i) Mistbelt forests, ii) Scarp forests, and iii) Coastal forests, these forests extend to the nearby province of KwaZulu-Natal (von Maltitz *et al*., 2003; Mucina *et al*., 2006). Altogether this forest biome encompasses the Maputaland-Pondoland-Albany (MPA) biodiversity hotspot (Mittermeier *et al*., 2004), which is one of three biodiversity hotspots in South Africa. Currently, nearly 200 mammal species occur in this biome of which six are endemic to the region (Perera *et al*., 2011). By separating *M. tenuis* from *M. cafer,* Taylor *et al*., (2017) made the latter the seventh endemic mammal species to this region.

This study focusses on *M. cafer* whose distribution is restricted to the moist and densely vegetated forest biome (Baxter, 2005; Churchill, 2007). The life history of *M. cafer* suggests that the species is highly vulnerable to forest fragmentation as the animal is small, has low dispersal ability and a high metabolic rate (Taylor *et al*., 2017). These characteristics have previously been shown to have a profound impact on shrew distributions, particularly concerning habitat fragmentation (Fedorov *et al*., 2008; Chavel *et al*., 2017; Bannikova *et al*., 2010; Jacquet *et al*., 2014, 2015; Nicolas *et al*., 2008; Bani *et al*., 2017). These characters could possibly also explain the patchy distribution of *M. cafer* in South Africa. Although the species is classified as 'Least Concern' by International Union for Conservation of Nature (IUCN) Red List assessment (Cassola, 2016), it is categorized as 'Vulnerable' according to regional Red Data Listing (Willows-Munro *et al*., 2016). Here we present DNA barcodes of *M. cafer* in order to test the utility of COI DNA barcoding in this species. Given the fragmented nature of the forest biome, we hypothesize that the genetic structure of *M. cafer* should be influenced by habitat fragmentation.

Materials and methods

Specimens

The handling of animals was approved by the ethics committee of the University of Stellenbosch (protocol number 1285) using guidelines from the Animal Care and Use Committee of the American Society of Mammalogists (Sikes and the Animal Care and Use Committee of the American Society of Mammalogists, 2016). Collection of tissues and specimens were done according to the sampling permits which were granted by the sampling locality authorities, including the Department

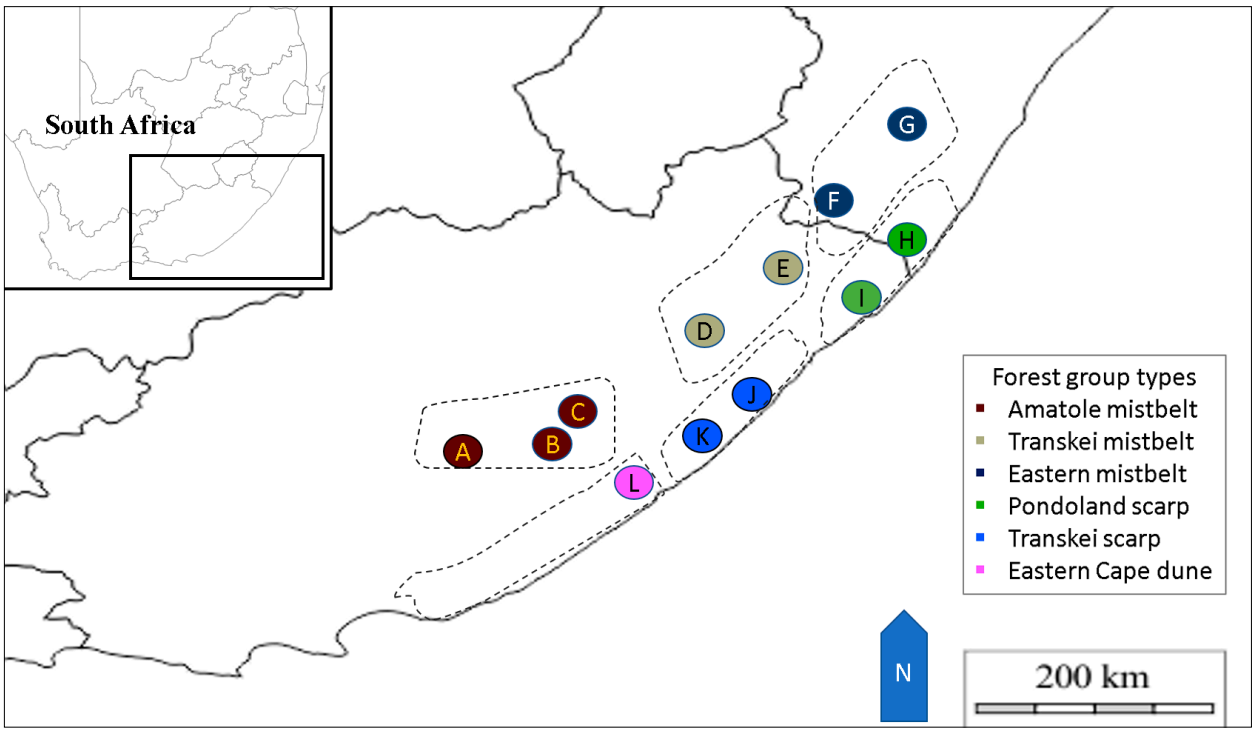

Fig. 1. The study localities in the Eastern Cape and Kwa-Zulu provinces, South Africa. Colour indicates the forest types and the letters represent the study locations as shown in Table 1. The relative position of the study area within South Africa is indicated by the rectangle in the insert.

of Environment, Forestry and Fisheries (DEFF), The Department of Agriculture, Forestry and Fisheries (DAFF), The Eastern Cape Parks and Tourism Agency (ECPTA), Ezemvelo Kwa-Zulu Natal Wildlife (EKZNW), South African National Parks (SANParks) and private landowners.

Samples were collected from three forest groups namely, mistbelt forest, scarp forest and coastal forest in the Eastern Cape and southern parts of the KwaZulu-Natal province of South Africa (Fig. 1, Table 1). For each of these forests, sampling was undertaken in at least two localities within each forest group. Trapping was achieved by randomly placed traplines in the forest, forest edge and the grasslands in the surrounding area, with traps approximately two meters apart. Animals were initially identified in the field based on a suite of diagnostic characters that included dorsal and ventral pelage, hindfoot colour, tail morphology including presence/absence of guard hairs, ear length and colour, and tail to head-and-body ratio. Standard measurements including total length, tail length, hindfoot length (with claw) and ear length, to the nearest 1 mm, were taken in the field. Tissue samples were collected from released animals, while voucher specimens were deposited at the Durban Natural Science Museum, South Africa. Specimens were prepared as wet specimens with crania extracted for morphological examination and verification of field-based taxonomic identification. The preliminary field and final species-level identification of individuals were based on diagnostic characters and taxonomic keys listed in Skinner & Chimimba (2005).

Molecular analysis

DNA was extracted using the NucleoSpin (R) Tissue Kit technique (Macherey-Nagel) following the manufacturer's manual. The COI fragment was isolated using COI primers L1490 and H2198 (Folmer *et al*., 1994). PCR reactions consisted of 0.8 µM of each primer, 1 unit of *Taq* DNA Polymerase, 1 X Buffer, 0.2 mM of each dNTP, 2.5 mM of $MgCl_2$ and distilled water to make up 25 ul reaction volumes. Cycling was performed using a GeneAmp PCR system (Applied Biosystems). PCR conditions involved initial denaturing step at 94 °C for 4 min followed by 35 cycles of 30 s denaturation at 94°C, 30 s annealing at 42°C to 47 °C, 30 s extension at 72°C, and then a final extension step at 72°C for 12 min. PCR products were visualized with one percent agarose gel and sequenced using big dye chemistry and sequencing was done at Central Analytical Facilities at Stellenbosch University.

Sequence data analysis and phylogenetics

Sequences generated were edited and assembled manually in bioedit v7.0.9 (Hall, 1999) and aligned using clustalw (Thompson *et al*., 2003). Sequences were run on the Basic Local Alignment Search Tool (BLAST, Altschul *et al*., 1990) at the NCBI website (http://www.ncbi.nlm.nih.gov/blast) to check for highly similar *Myosorex* taxa already in GenBank. The sequences were then deposited on the Barcode of Life Data System (BOLD, Ratnasingham & Hebert 2007) through the online interface at www.barcodinglife.org.

Table 1. Summary of sample details including study localities at which samples were collected. Letters match those in the map of Fig. 1.

Map Labels	Locality, geographic coordinates	Forest type	COI (623 bp) N (haplotype)	Voucher numbers
A	Fort Fordyce, – 31.68237, 26.482653	Amatole mistbelt	2 (Hap 6)	DM15227, DM15230
B	Isidenge, – 33.69975, 26.359898	Amatole mistbelt	13 (Hap 10, 11, 12, 13)	DM15008, DM15009, DM15010, DM15011, DM15021, DM15242, DM15243, DM15244 DM15245, DM15246
C	Kologha, – 32.535330, 27.36258	Amatole mistbelt	2 (Hap 10, 12)	DM15276, DM15277
D	Baziya, – 31.547255, 28.437503	Transkei mistbelt	10 (Hap 1, 2, 3, 4, 5)	**LRR180214BSF_MC1, LRR180211BSF_MC1, LRR180212BSF_MC1, LRR180213BSF_MC1***
E	Gomo, – 31.013194, 29.344585	Transkei mistbelt	16 (Hap 2, 4, 7, 8, 9)	EM78041217, EM128021217, EM89031217, EM04021217, EM124041217, EM40051217, EM140051217 #
F	Ngeli, – 30.542827, 29.680979	Eastern mistbelt	15 (Hap 1, 7, 15, 16, 17)	**LRR180314NSF_MC1, LRR180314NSF_MC2, LRR180314NSF_MC3, LRR180314NSF_MC4, LRR180314NSF_MC5***
G	Nxumeni, – 29.927042, 29.844179	Eastern mistbelt	3 (Hap 7)	**LRR181210NXSF_MCA1, LRR181212NXSF_MCA2, LRR181213NXSF_MCA1***
H	Umtamvuna, – 31.007467, 30.094205	Pondoland scarp	1 (Hap 7)	DM1121,
I	Mbotyi, – 31.418887, 29.723978	Pondoland scarp	3 (Hap 2, 15)	DM15167, DM15173
J	Dwesa, – 32.445594, 28.60766	Transkei scarp	1 (Hap 1)	DM15188
K	Manubi, – 33.990792, 25.360209	Transkei scarp	2 (Hap 1, 14)	DM15211
L	Morgans Bay, – 32.701243, 28.353646	Eastern Cape dune	4 (Hap 2)	EM240818ECNR_MCA1, EM240818ECNR_MCA2, EM240818ECNR_MCA3, EM240818ECNR_MCA4 #
Total	12 sites	6 forest types	**72(17)**	**Haplotype diversity, Hd: 0.8505 Nucleotide diversity (per site), Pi: 0.00454**
				* = not yet processed, # in Stellenbosch University (released animals not indicated)

The sequences were collapsed into a haplotype dataset using DNASP v5.10 and subsequently used to generate network tree using TCS networks (CLEMENT *et al*., 2002) in POPART (http://popart.otago.ac.nz). Haplotype diversity and population demographic expansions (Tajima's D) were analysed using DNSAP v5.10 (LIBRADO *et al*., 2009). To test for the significance of population structure three-way analyses of molecular variance (AMOVA) were undertaken. The genetic differentiation among populations (as defined by forest types) was calculated by pairwise F_{ST} tests. AMOVA and F_{ST} tests were calculated in ARLEQUIN 3.1 (EXCOFFIER *et al*., 2005) with 1000 permutations of AMOVA. Trees were reconstructed using Neighbour-joining (NJ), Maximum Parsimony (MP), and Maximum Likelihood (ML) based on the best substitution model which was identified using MODELTEST as implemented in MEGA 7 (KUMAR *et al*., 2016); support for the nodes was obtained by bootstrap resampling (1000 iterations).

Barcoding Gap: Our sequences were combined with sequences that were retrieved from BLAST analysis where sequences showing more than 80% are considered to be in the same family, 90–98% considered to be the same genus, and 99–100% correct species. The mean sequence divergences among specimens were generated using the Kimura 2-parameter (KIMURA, 1980) as it is commonly used for barcoding analysis. For estimating the barcoding gap and testing the species hypothesis Automatic Barcoding Gap Discovery (ABGD, PUILLANDRE *et al*., 2012) was used. Sequences were uploaded on the online interface of ABGD on https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html and analyses were carried out using default settings except that Kimura (K80)TS/TV (2.0) was selected instead of the default distance method. The pairwise differences between individual sequences were calculated and their distributions were grouped to form the histogram using the ABGD online interface. The same way the number of preliminary species hypothesis (PSHs) obtained for each prior intraspecific divergence were recorded.

Cranial morphological analyses

Crania of 30 vouchered specimens were used in the morphological analyses. Thirteen craniodental variables were recorded from *Crocidura cyanea* ($n=6$), *C. mariquensis* ($n=4$), *Myosorex cafer* ($n=15$) and *M. varius* ($n=5$). Nine cranial measurements were recorded from specimens using a Mitutoyo digital calipers to the nearest 0.01 mm and included: condylo-incisive length (CIL) – from the occipital condyles to the anterior margin of the incisors; braincase breadth

(BCB) – maximum width of braincase; braincase height (BCH) – maximum height as measured from the basic cranium to the dorsum of the braincase; post-glenoid width (PGW) – width at the outer margins of the glenoid fossa; greatest maxillary width (GMW) – width is taken across the outer margins of M2; interorbital width (IOW) – the greatest interorbital constriction of the cranium; post-palatal length (PL) – from the end of the hard palate to the anterior margin of the incisors; mandibular length (ML) – from the mandibular condyle to the anterior-most point of the incisors; and mandibular coronoid height (CORH) – from the uppermost point of the coronoid process to the arch of the mandible. The four dental measurements were: maxillary toothrow (UTR) – taken from the posterior margin of M^3 to the anterior margin of the incisors; maxillary molar toothrow (M^1M^3) – from the posterior margin of M^3 to the anterior margin of M^1; maxillary premolar and molar toothrow (PM^3) – taken from the anterior alveolus of the premolar to the posterior-most margin of M^3; and mandibular toothrow (LTR) – recorded from the posterior-most margin of M_3 to the most anterior point of the incisor. Cranial and dental measurements were recorded by a single observer to minimize measurement error.

Data were missing for nine variables for six individuals, for which the mean value for the craniometric variable for the particular taxon, was used. Owing to small sample sizes, no attempts were made to investigate sexual-size dimorphism (SSD) or any potential SSD heterogeneity between taxa. Kearney (1993) and Taylor *et al*., (2013) reported no SSD in southern African *Myosorex* species; based on this evidence and our small samples, we pooled data for males and females of all four species. Data were log_{10}-transformed, and taxon-specific tests for normality, skewness, kurtosis and descriptive statistics (mean ± stdev, minimum value, maximum value) were executed. Although data sets were parametric, owing to small sample sizes and unequal sample variance for certain variables, a Kruskal-Wallis test was performed, using taxon as the predictor variable. Log_{10}-transformed data were subjected to a principal component analysis of the covariance matrix with the extraction of the first three principal components. All descriptive and statistical analyses were carried in IBM SPSS version 21 (IBM Corp).

Results

Sequence data and BLAST search results

The COI gene was successfully amplified for all specimens ($n=78$), and of the 658 bp comprising the gene, 623 bp were used for further analysis. The nucleotide content of the gene fragment comprises 33.8% of Thymine, 26.4% Adenine, 24.8% Cytosine and 15% Guanine. Of the 623 bp, only 3.2% is variable (20/623 bp) and 2.1% sites are parsimony informative. The sequences revealed 17 haplotypes for COI. As far as we have been able to ascertain, GenBank does not have COI sequences of *M. cafer*.

Phylogenetic reconstructions were undertaken using our data ($n=78$) and 47 sequences that were kept from the BLAST analysis (Altschul *et al*., 1990) of our COI sequences with at least 83% similarity to *M. cafer*. These sequences were downloaded from GenBank and represent the species *Crocidura attenuata, C. brunnea, C. flavescens, C.* cf. *tanakae*, *C. russula, C. olivieri, C. nimbasilvanus, C. somalica, C. suaveolens* and *C. wuchihensis* resulting in a total of 116 sequences including the European shrew *Sorex araneus* used to root the trees (Fig. 2). Modeltest selected the Kimura-2 parameter with the gamma substitution model (Kimura, 1980) for the COI dataset The phylogenetic analyses were conducted using the haplotypes generated from 72 sequences of *M. cafer*, two sequences of *M. varius*, two sequences of *C. mariquensis,* and two sequences of *C. cyanea*; the other sequences were downloaded from the GenBank database. All other *Crocidura* species used for this analysis formed a well-supported clade with no samples from this study. Further, all 17 haplotypes of *M. cafer* constitute a monophyletic clade that is closely related to *M. varius.*

Barcoding gap among shrews (*Crocidura* and *Myosorex*)

The barcoding gap analysis for COI was performed using our data and sequences retrieved with BLAST analysis at a cutoff of 80%, resulting in $n=155$, these include *M. cafer* (72), *M. varius* (2), *C. mariquensis* (2), *C.* cf. *tanakae* (21), *C. wuchihensis* (4), *C. olivieri* (4), and *C. obscurior* (50). The cutoff of 80% similarity has resulted in more sequences than the 83% similarity from BLAST, however, all species with only one representative sequence in the cutoff were not included for barcoding gap analysis. These analyses revealed that the interspecific divergence for *M. cafer* and *M. varius* (0.027) is the lowest interspecific divergence; however, it is higher than all intraspecific divergences. The ABGD analyses retrieved nine groups with a prior maximal distance of $p=0.100000$ (Fig. 4). This data reveals there is a barcoding gap which only occurs for six of the nine groups. These six groups combined two samples of *M. varius* from this study with *M. cafer* samples while all other species formed independent groups.

Genetic diversity and networks of *Myosorex cafer*

The mean *p*-distance among and within forest groups types for *M. cafer* revealed very low variation in COI sequences per forest group (from 0.000 ± 0.000 to 0.006 ± 0.002; Table 2). Generally, the low genetic variation where divergences between forest groups and within forest groups range from 0.002 ± 0.001 to 0.044 ± 0.003. The significant negative values of the Tajima's D at −2.63069,

Table 2. Mean p-distances for *Myosorex cafer* among and within forest types. All codon positions were included and the total positions in the final dataset are also shown. Values were obtained in MEGA7 (TAMURA *et al*., 2016). The analyses of molecular variance (AMOVA) were calculated on ARLEQUIN 3.1 (EXCOFFIER *et al*., 2005), the variation among groups (Va), the variation among populations within groups (Vb), and the variation within populations are presented with their indices.

	Transkei mistbelt	**Transkei Scarp**	**Amatole mistbelt**	**Pondoland Scarp**	**Eastern Cape Dune**	**Eastern Mistbelt**
Transkei Mistbelt	0.004 ± 0.002					
Transkei Scarp	0.005 ± 0.002	0.002 ± 0.001				
Amatole mistbelt	0.006 ± 0.002	0.005 ± 0.002	0.002 ± 0.001			
Pondoland Scarp	0.003 ± 0.001	0.004 ± 0.002	0.006 ± 0.002	0.004 ± 0.002		
Eastern Cape Dune	0.003 ± 0.001	0.002 ± 0.001	0.003 ± 0.002	0.003 ± 0.001	0.000 ± 0.000	
Eastern mistbelt	0.004 ± 0.001	0.006 ± 0.003	0.008 ± 0.003	0.003 ± 0.001	0.005 ± 0.002	0.002 ± 0.001
Fixation Indices	**FSC: 0.27425**	**FST: 0.58620**	**FCT: 0.42983**	**Va = 42.98**	**Vb = 15.64**	**Vc = 41.38**

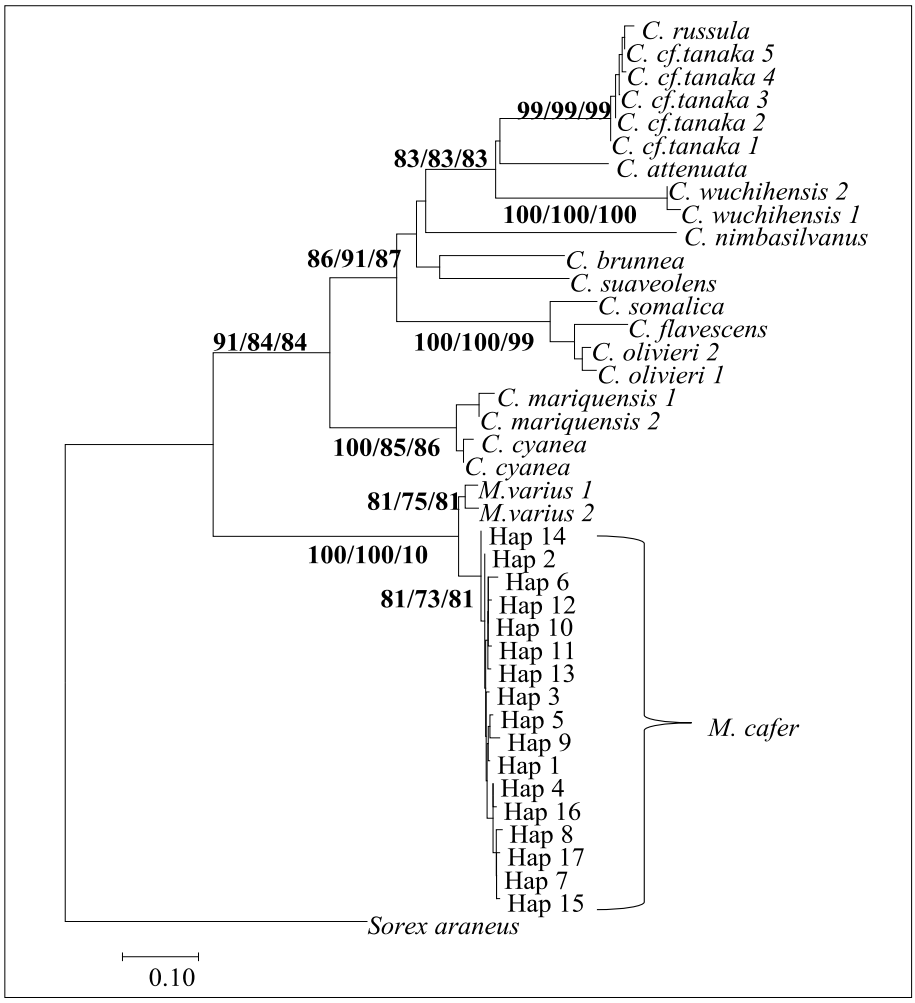

Fig. 2. A Neighbour-joining (NJ) phylogram illustrating the phylogenetic relationships among haplotypes generated from a total of 116 COI sequences. The sequences include 72 specimens of *M. cafer* (Hap 1–17), two *Myosorex varius*, two *Crocidura mariquensis*, two *C. cyanea* and a further 38 sequences of *Crocidura* species that were retrieved from the BLAST search (ALTSCHUL *et al*., 1990), while *Sorex araneus* was used as outgroup. Maximum Likelihood (ML), Maximum Parsimony (MP) and NJ analysis were drawn using the Kimura-2-Parameter model (KIMURA, 1980). Bootstrap resampling support (1000 iterations) are only shown for well supported clades (i.e. > 75%; NJ/MP/ML). Evolutionary analyses were conducted in MEGA7 (KUMAR *et al*., 2016).

($P < 0.001$) indicate that this population of *M. cafer* has undergone recent expansion. The pairwise F_{ST} and analyses of molecular variance (AMOVA) were done with individuals grouped according to their localities, and individuals grouped according to the two phylogroups. The pairwise F_{ST} is 0.58620, and the variation is 42.98% among groups, 15% within groups, and 41.38% within populations.

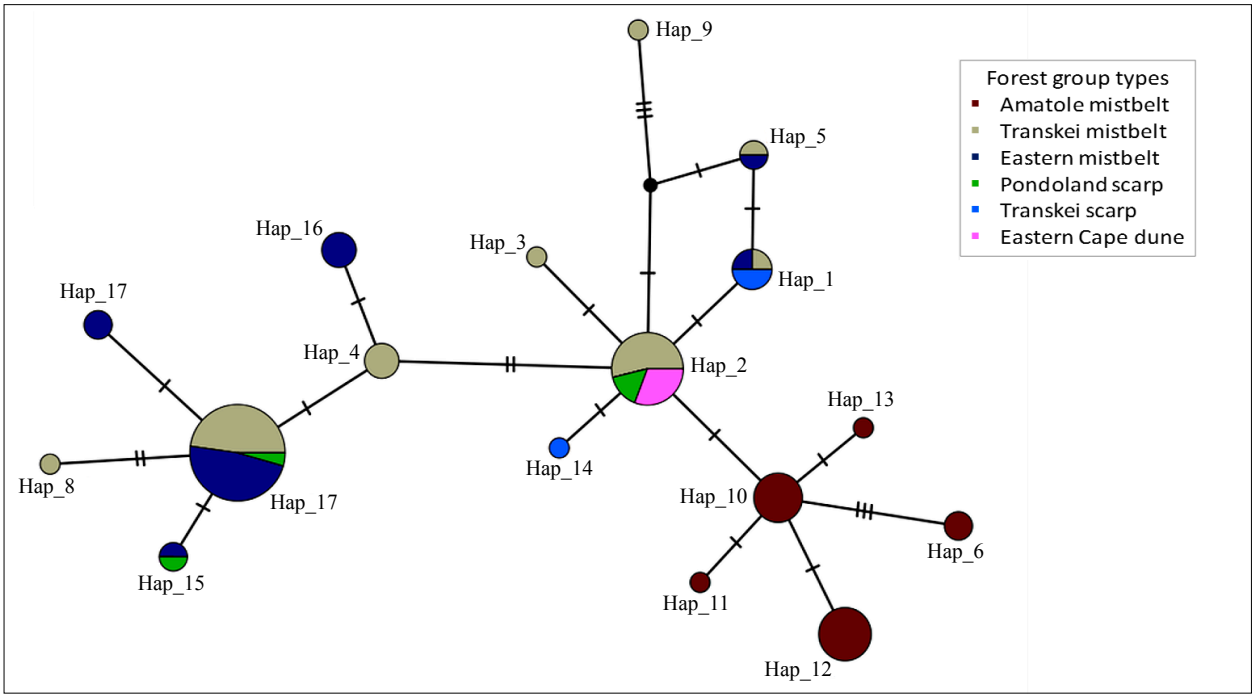

Fig. 3. Median-joining haplotype network generated from 72 specimens of *Myosorex cafer* with the COI dataset. The haplotype colours represent the forest group type from which the individuals where collected from the colours are indicated in the insert (see Table1, Figure 1). The number of mutations separating haplotypes is indicated using hatch marks. The network was drawn on POPART.

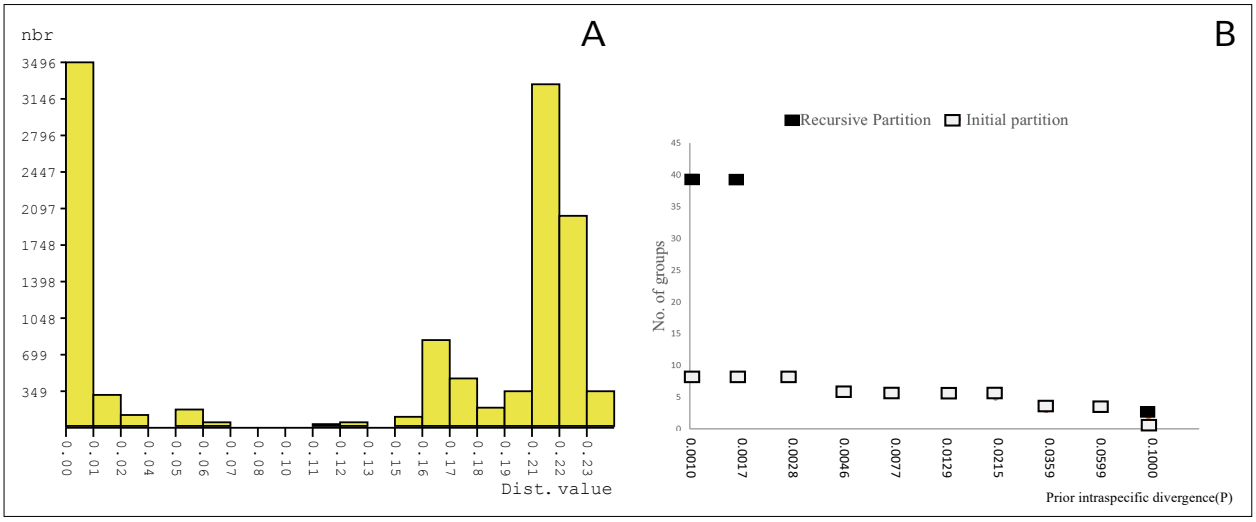

Fig. 4. Barcode gap analysis of shrews generated by Automatic Barcode Discovery Gap Discovery (Puillandre *et al*., 2012). The gap was calculated with the default setting using Kimura (K80) TS/TV distances in the COI gene. (a) Histograms showing the frequency of pairwise nucleotide differences, and (b) shows a plot diagram with the number of PSHs (preliminary species hypothesis) obtained for each prior intraspecific divergence (0.01 – 0.1000).

The median-joining networks for COI (Fig. 3) revealed the two phylogroups, with haplotypes from Amatole forest clustering together to the exclusion of haplotypes from all other localities. The haplotypes from the Amatole forest group are separated by only a single mutation to the rest, with no further grouping. The haplotypes from other localities have unique haplotypes from one forest group. However, it is also important that there are haplotypes composed of individuals from more than one forest group. There is a haplotype with individuals from Transkei mistbelt and Eastern mistbelt forests; a haplotype with individuals from Transkei mistbelt, Transkei scarp and Eastern mistbelt forests; a haplotype with individuals from Transkei mistbelt, Pondoland scarp and Eastern Cape dune forests; a haplotype with individuals from Transkei mistbelt and Pondoland scarp; and a haplotype with individuals from Transkei mistbelt, Pondoland scarp and Eastern mistbelt forests. Of all these six shared haplotypes, the Transkei mistbelt forest appears on five of them, while the Eastern mistbelt forest appears on four.

Table 3. Summary of cranial and dental measurements of the four shrew taxa sampled from Eastern Cape and KwaZulu-Natal forested habitats in this study.

Craniometric variable		***Myosorex cafer*** (*n* = 14)	***Myosorex varius*** (*n* = 5)	***Crocidura cyanea*** (*n* = 6)	***Crocidura mariquensis*** (*n* = 4)
CIL	Mean ± stdev	22.2 ± 0.22	22.4 ± 0.65	19.5 ± 0.57	21.4 ± 0.93
	Range	21.61 – 22.63	21.43 – 22.88	18.64 – 20.37	20.15 – 22.35
BMW	Mean ± stdev	6.6 ± 0.14	6.8 ± 0.18	5.9 ± 0.11	6.5 ± 0.22
	Range	6.34 – 6.88	6.49 – 6.98	5.76 – 6.05	6.21 – 6.69
BCB	Mean ± stdev	10.8 ± 0.21	10.5 ± 0.19	8.6 ± 0.13	9.1 ± 0.42
	Range	10.47 – 11.12	10.20 – 10.65	8.45 – 8.81	8.53 – 9.50
BCH	Mean ± stdev	6.1 ± 0.16	6.2 ± 0.26	4.6 ± 0.17	5.5 ± 0.17
	Range	5.73 – 6.28	5.93 – 6.42	4.36 – 4.83	5.27 – 5.66
IOC	Mean ± stdev	4.2 ± 0.13	4.2 ± 0.19	4.0 ± 0.07	4.0 ± 0.17
	Range	4.09 – 4.46	3.92 – 4.43	3.92 – 4.10	3.77 – 4.14
MTR	Mean ± stdev	9.6 ± 0.26	9.7 ± 0.29	8.3 ± 0.27	9.3 ± 0.20
	Range	9.13 – 10.13	9.33 – 10.05	7.89 – 8.62	9.08 – 9.54
M^1M^3	Mean ± stdev	4.1 ± 0.17	4.1 ± 0.10	3.4 ± 0.16	3.8 ± 0.11
	Range	3.66 – 4.32	4.00 – 4.27	3.14 – 3.55	3.71 – 3.97
PM^3	Mean ± stdev	5.6 ± 0.18	5.7 ± 0.17	4.7 ± 0.21	5.0 ± 0.13
	Range	5.31 – 5.91	5.48 – 5.93	4.44 – 4.94	4.88 – 5.17
PGW	Mean ± stdev	7.2 ± 0.16	7.2 ± 0.24	6.0 ± 0.13	6.00 ± 0.18
	Range	6.96 – 7.53	6.98 – 7.48	5.81 – 6.15	5.72 – 6.13
PPL	Mean ± stdev	10.0 ± 0.19	10.2 ± 0.41	8.6 ± 0.24	9.7 ± 0.23
	Range	9.69 – 10.26	9.74 – 10.70	8.32 – 8.99	9.42 – 9.94
LTR	Mean ± stdev	8.8 ± 0.23	8.9 ± 0.35	7.3 ± 0.47	8.5 ± 0.17
	Range	8.36 – 9.21	8.41 – 9.29	6.65 – 7.69	8.34 – 8.74
ML	Mean ± stdev	13.9 ± 0.28	13.9 ± 0.48	11.6 ± 0.27	13.0 ± 0.42
	Range	13.42 – 14.36	13.28 – 14.55	11.31 – 11.97	12.75 – 13.77
CORH	Mean ± stdev	5.4 ± 0.17	5.55 ± 0.19	4.2 ± 0.10	5.0 ± 0.28
	Range	5.21 – 5.81	5.37 – 5.73	4.13 – 4.39	4.62 – 5.26

Craniometric variation

The mean, standard deviation and range of 13 craniometric variables for the four shrew taxa sampled during this study are presented in Table 3. Based on craniometrics measurements, *Crocidura cyanea* is the smallest-sized taxon. There was a substantial overlap in variable mean and ranges of the two *Myosorex* species. Kruskal-Wallis H test revealed significant variation ($p < 0.01$) in all 13 variables across the four taxa. A Kaiser-Meyer Olkin (KMO) test score of 0.920 indicated that sampling adequacy was sufficient to permit principal component analysis (PCA). A PCA revealed considerable separation along with principal component 1 (PC1) between *C. cyanea*, *C. mariquensis* and the cluster of *Myosorex* taxa (Fig. 5). Based on factor loadings associated with PC1 (considered a general size vector as all variables exhibited high loadings), and PC2 shown in Table 4, *C. cyanea* was distinguished by its overall smaller cranial geometry and relatively broader braincase breadth (BCB) and postglenoid width (PGW) in relation to other craniometrics variables. *Crocidura mariquensis* was discerned from *C. cyanea* by its larger size and proportionately narrow BCB and PGW. There was some overlap between *M. cafer* and *M. varius*, undoubtedly compounded by an aberrantly large-sized male *M. varius* recorded from Mkambati Nature Reserve, which accounted for the greater standard deviation values presented in Table 3. In general, *M. varius* could be distinguished by a proportionately narrower BCB in comparison to *M. cafer* (see Tables 3 and 4).

Discussion

Barcoding efficiency in *Myosorex cafer*

The fundamental feature in the success of COI DNA barcoding as a tool of species identification is the gap between intraspecific and interspecific variation (HEBERT *et al*., 2004; BORISENKO *et al*., 2008; LU *et al*., 2012). In the present study, the intraspecific variation ranges from 0.003 ± 0.001 (*C.* cf. *tanakae*) to 0.017 ± 0.002 (*C. obscurior*), and for *M. cafer* the intraspecific variation is 0.005 ± 0.002, which is the range within these species (TAYLOR *et al*., 2013). The interspecific variation ranges from 0.027 ± 0.006 (between *M. cafer* and *M. sclateri*) to 0.245 ± 0.027 (between *M. varius* and *C. wuchihensis*). The lowest interspecific variation (0.027 ± 0.006) is at least more than three times the intraspecific divergence

Table 4. Factor loadings for 13 craniometric variables derived from principal component analyses. The first three principal components accounted for 93.43% of sample variance.

Craniometric variable	Principal component 1	Principal component 2	Principal component 3
CIL	0.967	0.142	−0.021
BMW	0.917	0.227	−0.055
BCB	0.932	−0.333	−0.045
BCH	0.980	0.036	−0.126
IOC	0.668	−0.219	0.400
MTR	0.955	0.223	0.138
M^1M^3	0.934	0.120	0.275
PM^3	0.948	−0.089	0.186
PGW	0.884	−0.437	0.015
PPL	0.961	0.223	0.026
LTR	0.927	0.233	−0.027
ML	0.974	0.157	−0.017
CORH	0.966	0.034	−0.155

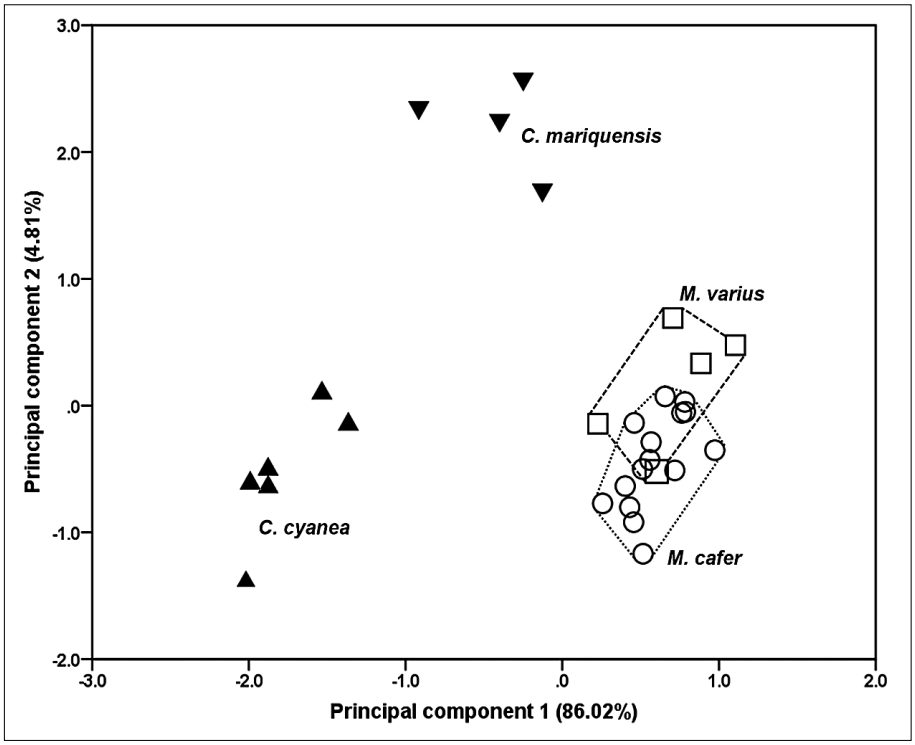

Fig. 5. A biplot derived from principal components analysis of log10-transformed craniometrics variables, depicting sample variation amongst the 30 shrew individuals along with the first two principal components

value of *M. cafer*. This difference between the intraspecific and interspecific variation (commonly known as barcoding gap), is consistent with a variation for *Crocidura* species where the interspecific variation was also very high (from 0.131 ± 0.017 to 0.200 ± 0.023) compared with intraspecific variation. The barcoding gap is thus the most important tool for discrimination between species, identification and assignation of new species (Janzen *et al*., 2005; Smith *et al*., 2005; Puillandre *et al*., 2012).

The method of preliminary species hypothesis (PSH) in ABGD (Puillandre *et al*., 2012) has been applied. This method can be used even when the divergences can overlap to set the dataset into candidate species (Puillandre *et al*., 2012). In the present study, the ABGD revealed consistently that the data contain six candidate species with the initial partition, and in all instances, the six groups have combined our samples of *M. cafer* and *M. varius* in one group. With the recursive partition, there are instances where *M. cafer* and *M. varius* are not in one group as the first four partitions with a prior maximal distance of $p = 0.004642$, where nine groups are recognized, with *C. obscurior* subdivided into two groups.

Cranial morphology of *Myosorex*

Although limited by sample size, cranial morphological data did allow for partial separation of the four shrew taxa recorded during this study. Descriptive statistics and explorative multivariate analysis of craniometrics variables confirmed the difficulty in separating *M. cafer* and

M. varius using morphological criteria (see KEARNEY, 1993). Unlike TAYLOR *et al*., (2013), we did not find differences in cranial size between *M. cafer* and *M. varius*. This is all likelihood an artifact of insufficient sampling; and skewing of data in our *M. varius* sample owing to an unusually large-sized male, confirmed by genetic data and recorded from Mkambati Nature Reserve in the Eastern Cape. Our findings do, however, indicate that braincase breadth (BCB, Table 3) might be a means of discerning *M. cafer* from *M. varius*.

The impact of forest fragmentation on population structure

Our genetic analysis revealed shallow phylogeographic structure among the populations that were analysed from the Amatole mistbelt, Transkei mistbelt, Eastern mistbelt, Pondoland scarp, Transkei scarp and Eastern Cape dune forests. Our data are consistent with the previous phylogeographic analysis of *M. cafer* using the partial control region of mitochondrial DNA, which revealed an absence of isolation by distance (i.e. poor population genetic structure) (WILLOWS-MUNRO, 2008). Although the current study lacks samples from some of the previous locations analysed by WILLOWS-MUNRO, (2008), our study utilized a larger sample size from all forest types in the Eastern Cape and indicates for the first time that the Amatole mistbelt population may be diverging from the more eastern populations (at least based on the network analysis shown below).

Extinction filtering event on *M. cafer*

The nucleotide divergences support the fact that individuals from the Amatole mistbelt forest are isolated from the rest of the population, and that sub-populations from the Transkei mistbelt forest are more diverse (Fig. 3). This is based on the Transkei mistbelt having more haplotypes, and also having shared haplotypes with all other forest group types except the Amatole mistbelt forest. These data thus do not support the forest refugia theory which postulates that following the last glacial maxima (LGM) the scarp forests have acted as refugia for small mammals that replenished adjacent forests (LAWES *et al*., 2007). Our findings are consistent with the findings tested genetically with chameleons in Kwazulu-Natal which did support the theory of forest refugia in the scarp forests (DA SILVA & TOLLEY, 2017). Our study suggests that the Amatole mistbelt is more ancient and was separated from the other forests earlier and was too far west of scarp forests to have been re-colonized after the LGM.

Conclusions

This study reinforces the utility of COI barcodes in the identification and discrimination of species. These data are consistent with morphology data, although there is no clear cranial differentiation between the two *Myosorex* species. Still, this could be attributed to low sample size. There is a clear morphological differentiation between *Crocidura* and *Myosorex* species, irrespective of low sample sizes for *Crocidura* species. Altogether, the morphology and COI barcoding can be used in the identification of these shrews, and possibly with larger datasets, there will be clear differentiation of *M. cafer* and *M. varius*. Although TAYLOR *et al*., (2013) showed a clear distinction between *M. cafer* and *M. varius* based on confirmatory multivariate analyses, our morphological data fall within the same range as their data and support the assertion of KEARNEY, (1993) that it is difficult to separate *M. cafer* and *M. varius* on cranial morphology.

This data suggest that the movement of *M. cafer* between forests is not restricted, except in the case of the Amatole mistbelt forest. The Amatole mistbelt forest is separated from the rest of the forests, while both other mistbelt forests (Transkei and Eastern) are dominated by shared haplotypes, suggesting past dispersion from these forests to the scarp and coastal forests. Climatic extinction filtering small mammals, as proposed by LAWES *at al*., (2007), does not appear to apply to this taxon, although forest fragmentation has played a role in the current distribution of this species. Studies using a faster-evolving DNA marker are required to further elucidate the genetic consequences of fragmentation in this species.

Acknowledgments

The project was funded by the Foundational Biodiversity Information Programme (FBIP) under the Department of Science and Technology (DST). We express our gratitude to the Durban Natural Science Museum and eThekwini Municipality for funding field trips and providing equipment. We express gratitude to the Eastern Cape Parks and Tourism Agency, Ezemvelo KZN Wildlife, Department of Agriculture, Forestry and Fisheries, and SANParks for permits and facilitating sample collection. We also acknowledge employees and interns of the Durban Natural Science Museum Mammal Department and fellow research students for their assistance during the sample collection, as well as the collection of samples by Vusi Martins and Dale Wright.

References

ALTSCHUL, S.F., GISH, W., MILLER, W., MYERS, E.W. & LIPMAN, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

BAKER, R.J. & BRADLEY, R.D. (2006). Speciation in mammals and the genetic species concept. *Journal of Mammalogy*, **87**, 643–662.

BANI, L., ORIOLI, V., PISA, G., FAGIANI, S., DONDINA, O., FABBRI, E., RANDI, E., SOZIO, G. & MORTELLITI, A. (2017). Population genetic structure and sex-biased dispersal of the hazel dormouse (*Muscardinus avellanarius*) in a continuous and in a fragment-

ed landscape in central Italy. *Conservation Genetics*, **18**, 261–274.

Bannikova, A.A., Dokuchaev, N.E., Yudina, E.V., Bobretzov, A.V., Sheftel, B.I. & Lebedev, V.S. (2010). Holarctic phylogeography of the tundra shrew (*Sorex tundrensis*) based on mitochondrial genes. *Biological Journal of the Linnean Society*, **101**, 721–746.

Borisenko, A.V., Lim, B.K., Ivanova, N.V., Hanner, R.H. & Hebert, P.D. (2008). DNA barcoding in surveys of small mammal communities: a field study in Suriname. *Molecular Ecology Resources*, **8**, 471–479.

Caro, T.I.M. (2013). The colours of extant mammals. *Seminars in Cell & Developmental Biology*; **24**, 542–552.

Chavel, E.E., Imbeau, L., Mazerolle, M.J. & Drapeau, P. (2017). Boreal small mammals show evidence of density-dependent patterns with area-sensitivity. *Forest Ecology and Management*, **400**, 485-501.

Churchfield, S. (2013). Family Soricidae, shrews, pp. 43–50 *in*: Happold, M. & Happold, D.C.D. (eds) *Mammals of Africa, Hedgehogs, Shrews and Bats*, *Vol. 4*. London, Bloomsbury.

Clare, E.L., Lim, B.K., Engstrom, M.D., Eger, J.L. & Hebert, P.D. (2007). DNA barcoding of Neotropical bats: species identification and discovery within Guyana. *Molecular Ecology Notes*, **7**, 184–190.

Clement, M., Posada, D.C.K.A. & Crandall, K.A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Da Silva, J.M. & Willows-Munro, S. (2016). A review of over a decade of DNA barcoding in South Africa: a faunal perspective. *African Zoology*, **51**, 1–12.

Da Silva, J.M. & Tolley, K.A. (2017). Diversification through ecological opportunity in dwarf chameleons. *Journal of Biogeography*, **44**, 834–847.

De Salle, R. & Goldstein, P. (2019). Review and interpretation of trends in DNA barcoding. *Frontiers in Ecology and Evolution*, **7**, 302.

Dold, A.P. & Cocks, M.L. (2002). The trade-in medicinal plants in the Eastern Cape Province, South Africa. *South African Journal of Science*, **98**, 589–597.

Dubey, S., Salamin, N., Ohdachi, S.D., Barrière, P. & Vogel, P. (2007). Molecular phylogenetics of shrews (Mammalia: Soricidae) reveal timing of transcontinental colonizations. *Molecular Phylogenetics and Evolution.* **44,** 126–137.

Eeley, H.A., Lawes, M.J. & Piper, S.E. (1999). The influence of climate change on the distribution of indigenous forest in KwaZulu-Natal, South Africa. *Journal of Biogeography,* **26**, 595–617.

Excoffier, L., Laval, G. & Schneider, S. (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics online*, **1**, 47.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology,* **3**: 294–299.

Geldenhuys, C.J. (1989). *Environmental and biogeographic influences on the distribution and composition of the southern Cape forests* (*Veld type 4*) (Doctoral dissertation, University of Cape Town).

Hajibabaei, M., dewaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., Mackie, P.M. & Hebert, P.D. (2005). Critical factors for assembling a high volume of DNA barcodes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 1959–1967.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W. & Hebert, P.D. (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences*, **103**, 968–971.

Hajibabaei, M., Singer, G.A., Hebert, P.D. & Hickey, D.A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics,* **23**, 167–172.

Hamer, M. (2013). A national strategy for zoological taxonomy (2013–2020). *South African National Biodiversity Institute, Pretoria*, 1–53.

Hall, T. A. (1999). bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41,** 95–98.

Hao, X., Jiang, R. & Chen, T. (2011). Clustering 16S rRNA for OTU prediction: a method of unsupervised Bayesian clustering. *Bioinformatics*, **27**, 611–618.

Hebert, P.D., Cywinska, A., Ball, S.L. & Dewaard, J.R. (2003a). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B*, **270**, 313–321.

Hebert, P.D., Ratnasingham, S. & De Waard, J.R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B*, **270**, 96–99.

Irwin, D.M., Kocher, T.D. & Wilson, A.C. (1991). Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, **32**, 128–144.

Ivanova, N.V., Dewaard, J.R. & Hebert, P.D (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes*, **6**, 998–1002.

Jacobs, D.S., Babiker, H., Bastian, A., Kearney, T., van Eeden, R. & Bishop, J.M. (2013). Phenotypic convergence in genetically distinct lineages of a *Rhinolophus* species complex (Mammalia, Chiroptera). *PloS One*, **8**.

Jacquet, F., Nicolas, V., Bonillo, C., Cruaud, C. & Denys, C. (2012). Barcoding, molecular taxonomy, and exploration of the diversity of shrews (Soricomorpha: Soricidae) on Mount Nimba (Guinea). *Zoological Journal of the Linnean Society*, **166**, 672–687.

Janzen, D.H., Hajibabaei, M., Burns, J.M., Hallwachs, W., Remigio, E. & Hebert, P.D. (2005). Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philosophical Transactions of the Royal Society B*, **360**, 1835–1845.

Jörger, K.M., Norenburg, J.L., Wilson, N.G. & Schrödl, M. (2012). Barcoding against a paradox? Combined molecular species delineations reveal multiple cryptic lineages in elusive meiofaunal sea slugs. *BMC Evolutionary Biology*, **12**, 245.

Kearney, T.C. (1993). *A Craniometric Analysis of Three Taxa of Myosorex from Natal and Transkei*. Pietermaritzburg, Doctoral dissertation, University of Natal.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870–1874.

Lai, Y.C., Shiroishi, T., Moriwaki, K., Motokawa, M. & Yu, H.T. (2008). Variation of coat color in house mice throughout Asia. *Journal of Zoology*, **274**, 270–276.

Lawes, M.J, (1990). The distribution of the Samango monkey (*Cercopithecus mitis erythrarchus* Peters, 1852 and *Cercopithecus mitis labiatus* I. Geoffroy, 1843) and forest history in southern Africa. *Journal of Biogeography*, **17**, 669–680.

Lawes, M.J., Eeley, H.A., Findlay, N.J. & Forbes, D. (2007). Resilient forest faunal communities in South Africa: a legacy of palaeoclimatic change and extinction filtering? *Journal of Biogeography*, **34**, 1246–1264.

Leigh, J.W. & Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, **6**, 1110–1116.

Librado, P. & Rozas, J. (2009). dnsap v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Losos, J.B. (2011). Convergence, adaptation, and constraint. *Evolution: International Journal of Organic Evolution*, **65**, 1827–1840.

Low, A.R. & Rebelo, A.G. (1996). *Vegetation of South Africa, Lesotho and Swaziland. A companion to the vegetation map of*

South Africa, Lesotho and Swaziland. Pretoria, Department of Environmental Affairs and Tourism.

LU, L., CHESTERS, D., ZHANG, W., LI, G., MA, Y., MA, H., SONG, X., WU, H., MENG, F., ZHU, C. & LIU, Q. (2012). Small mammal investigation in spotted fever focus with DNA-barcoding and taxonomic implications on rodent species from Hainan of China. *PLoS One*, **7,** e43479.

LUBKE, R.A., EVERARD, D.A. & JACKSON, S. (1986). The biomes of the eastern Cape with emphasis on their conservation. *Bothalia*, **16**, 251–261.

MACARTHUR, R.H. & WILSON, E.O. (1967). *The Theory of Island Biogeography*. Princeton, NJ, Princeton University Press.

MCDONALD, D.E. & DANIELS, S.R. (2012). Phylogeography of the Cape velvet worm (Onychophora: *Peripatopsis capensis*) reveals the impact of Pliocene/Pleistocene climatic oscillations on Afromontane forest in the Western Cape, South Africa. *Journal of Evolutionary Biology*, **25**, 824–835.

MEESTER, J. (1958). Variation in the shrew genus *Myosorex* in southern Africa. *Journal of Mammalogy*, **39**, 325–339.

MEESTER, J.A.J. (1986). *Classification of Southern African Mammals, Vol. 5*. Pretoria, Transvaal Museum.

MERILAITA, S. & STEVENS, M. (2011). Crypsis through background matching, pp. 17–33 *in*: STEVENS, M. & MERILAITA, S. (eds) *Animal Camouflage: Mechanisms and Function*. Cambridge, Cambridge University Press.

MEYER, C.P. & PAULAY, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, **3**, e422.

MEYER, S., WEISS, G. & VON HAESELER, A. (1999). Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics*, **152**, 1103–1110.

MITTERMEIER, R.A., GIL, P.R., HOFFMAN, M., PILGRIM, J., BROOKS, T., MITTERMEIER, C.G., LAMOREUX, J., DA FONSECA, G.A.B. & SALIGMANN, P.A. (2004). *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Chicago, IL, The University of Chicago Press.

MUCINA, L., GELDENHUYS, C.J. & RUTHERFORD, M.C. (2006). Afrotemperate, subtropical and azonal forests. The vegetation of South Africa, Lesotho and Swaziland. *Strelitzia*, **19**, 584–614.

MUNSHI-SOUTH, J. (2012). Urban landscape genetics: canopy cover predicts gene flow between white-footed mouse (*Peromyscus leucopus*) populations in New York City. *Molecular Ecology*, **21**, 1360–1378.

NEI, M. (1987). *Molecular Evolutionary Genetics*. New York, NY, Columbia University Press.

NICOLAS, V., SCHAEFFER, B., MISSOUP, A.D., KENNIS, J., COLYN, M., DENYS, C., TATARD, C., CRUAUD, C. & LAREDO, C. (2012). Assessment of three mitochondrial genes (16S, Cytb, CO1) for identifying species in the Praomyini tribe (Rodentia: Muridae). *PLoS One*, **7,** e36586.

PERERA, S.J., RATNAYAKE-PERERA, D. & PROCHES, S. (2011). Vertebrate distributions indicate a greater Maputaland-Pondoland-Albany region of endemism. *South African Journal of Science*, **107**, 52–66.

PONS, J., BARRACLOUGH, T.G., GOMEZ-ZURITA, J., CARDOSO, A., DURAN, D.P., HAZELL, S., KAMOUN, S., SUMLIN, W.D. & VOGLER, A.P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595–609.

PUILLANDRE, N., LAMBERT, A., BROUILLET, S. & ACHAZ, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

QUÉROUIL, S., HUTTERER, R., BARRIÈRE, P., COLYN, M., PETERHANS, J.C.K. & VERHEYEN, E. (2001). Phylogeny and evolution of African shrews (Mammalia: Soricidae) inferred from 16s rRNA sequences. *Molecular Phylogenetics and Evolution*, **20**, 185–195.

RAMOS-ONSINS, S.E. & ROZAS, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

RATNASINGHAM, S. & HEBERT, P.D.N. (2007). BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364.

RODRÍGUEZ, S., VISEDO, G. & ZAPATA, C. (2001). Detection of errors in dinucleotide repeat typing by nondenaturing electrophoresis. *Electrophoresis*, **22**, 2656–2664.

SHIELDS, G.F. & KOCHER, T.D. (1991). Phylogenetic relationships of North American ursids based on analysis of mitochondrial DNA. *Evolution*, **45**, 218–221.

SIKES, R. S. & ANIMAL CARE AND USE COMMITTEE OF THE AMERICAN SOCIETY OF MAMMALOGISTS (2016). 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *Journal of mammalogy*, **97**, 663–688.

SKINNER, J.D. & CHIMIMBA, C.T. (2005). *The Mammals of the Southern African Sub-region*. Cambridge, Cambridge University Press.

SMITH, M.A., FISHER, B.L. & HEBERT, P.D. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philosophical Transactions of the Royal Society B*, **360**, 1825–1834.

STEVENS, M. & MERILAITA, S. (2009). Defining disruptive coloration and distinguishing its functions. *Philosophical Transactions of the Royal Society B*, **364**, 481–488.

TAJIMA, F. & NEI, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution*, **1**, 269–285.

TAYLOR, P.J., KEARNEY, T.C., KERBIS PETERHANS, J.C., BAXTER, R.M. & WILLOWS-MUNRO, S. (2013). Cryptic diversity in forest shrews of the genus *Myosorex* from southern Africa, with the description of a new species and comments on *Myosorex tenuis*. *Zoological Journal of the Linnean Society*, **169**, 881–902.

TAYLOR, P.J., OGONY, L., OGOLA, J. & BAXTER, R.M. (2017). South African mouse shrews (*Myosorex*) feel the heat: using species distribution models (SDMs) and IUCN Red List criteria to flag extinction risks due to climate change. *Mammal Research*, **62**, 149–162.

THOMAS, O. & SCHWANN, O. (1905). The Rudd exploration of South Africa. II. List of mammals from the Wakkerstroom District, Southeast Transvaal. *Proceedings of the Zoological Society of London,* **1905**, 129–138

THOMPSON, J.D., GIBSON, T. & HIGGINS, D.G. (2003). Multiple sequence alignment using CLUSTALW and CLUSTALX. *Current Protocols in Bioinformatics*, **2003**, 2.3.1–2.3.22.

VON MALTITZ, G., MUCINA, L., GELDENHUYS, C.J., LAWES, M.J., EELEY, H., ADIE, H., VINK, D., FLEMING, G. & BAILEY, C. (2003). Classification system for South African indigenous forests: an objective classification for the Department of Water Affairs and Forestry. *Environmental Report ENV-PC*, **17**, 1–284.

WARD, R.D., ZEMLAK, T.S., INNES, B.H., LAST, P.R. & HEBERT, P.D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, **360**, 1847–1857.

WILLOWS-MUNRO, S. (2008). *The Molecular Evolution of African Shrews* (*Family Soricidae*). Stellenbosch, Ph.D. Thesis, Stellenbosch University

WILLOWS-MUNRO, S., BAXTER, R. & TAYLOR, P.J. (2016). A conservation assessment of *Myosorex cafer*, pp. 1-5 *in*: CHILD, M.F., ROXBURGH, L., DO LINH SAN, E., RAIMONDO, D. & DAVIES-MOSTERT, H.T. (eds). *The Red List of Mammals of South Africa, Swaziland and Lesotho*. Stellenbosch, South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.

WILSON, D.E. & REEDER, D.M. (eds, 2005). *Mammal Species of the World: a Taxonomic and Geographic Reference, Vol. 1*. Baltimore, MD, Johns Hopkins University Press.

YANG, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, **61**, 854–865.