DNA barcoding of the mesic adapted striped mouse, *Rhabdomys dilectus* in the Eastern Cape and KwaZulu-Natal provinces of South Africa

Emmanuel Matamba¹, Leigh R. Richards², Michael I. Cherry¹, Ramugondo V. Rambau¹

¹ Stellenbosch University, Department of Botany and Zoology, Private Bag X1, Matieland 7602, South Africa [mic@sun.ac.za, rvr2@sun.ac.za]
² Durban Natural Science Museum, P O Box 4085, Durban, South Africa, 4001 [Leigh.Richards@durban.gov.za]

http://zoobank.org/6D34D59C-A94A-4ABE-8408-1B458E86CD46

Corresponding author: Emmanuel Matamba (ematamba@sun.ac.za)

**Abstract**

South African small mammals are under-represented in DNA barcoding efforts, particularly from the eastern forested regions of the country. This study reports DNA barcoding of *Rhabdomys* taxa from previously unsampled parts of the Eastern Cape and KwaZulu-Natal provinces of South Africa. The complete mitochondrial DNA cytochrome oxidase I (COI) gene was sequenced for 101 *Rhabdomys* sp. individuals from 16 localities from all three main forest groups (coastal, mistbelt, and scarp forests). Molecular data were supplemented with external morphological measurements, including those deemed potential taxonomically diagnostic characters. Findings indicate the area to be inhabited solely by *Rhabdomys dilectus chakae*. Haplotypes distributed across the three forest groups were separated by shallow sequence divergences ranging from 0.001–0.015 (Kimura 2-parameter model) and displayed very little population genetic structure (*F<sub>ST</sub> = 0.071787*). Morphological data revealed some regional metric differences in external morphology, but all the head-and-body to tail (HB: tail) ratios match that of *R. d. chakae*, and consequently, molecular and morphological data are congruent. These data confirm a range extension of *R. d. chakae*, supporting the utility of COI barcodes in the identification of small mammalian species.

**Keywords**

COI barcodes, forest fragmentation, morphology, *Rhabdomys dilectus*

**Introduction**

The four-striped grass mouse genus *Rhabdomys* (Thomason, 1916), is distributed throughout southern Africa, occurring in a variety of habitats with ample grass cover (Skinner and Chimimba 2005; Monadjem et al. 2015; Du Toit et al. 2016). Its distribution includes forest edges, where it is a pioneer species following forest clearances (Castley 1997). *Rhabdomys* is considered a species complex (Rambau et al. 2003; du Toit et al. 2012; Engelbrecht et al. 2016), with seven subspecies defined using morphological criteria (Meester et al. 1986). At least four subspecies which are: *Rhabdomys pumilio* (Sparrman, 1784), *R. bechuanae* (Thomas, 1893), *R. dilectus* (De Winton, 1897), and *R. intermedius* (Wroughton, 1905) have been elevated to species level based on genetic and chromo-
of chromosomal data (Rambau et al. 2003; du Toit et al. 2012; Ganem et al. 2020). Of these, *R. pumilio* and *R. dilectus* were the first to be recognized based on chromosomal work combined with molecular data (Rambau et al. 2003), with the latter further divided into two subspecies *R. d. chakae* and *R. d. dilectus*. Subsequently, further molecular analyses retrieved three paraphyletic lineages of *R. pumilio* and two monophyletic lineages of *R. dilectus*, each with distinct geographical distributions (du Toit et al. 2012). Two of these paraphyletic lineages of *R. pumilio* have been referred to as *R. bechuanae* and *R. intermedius* (Ganem et al. 2020). There are contact zones between the *R. pumilio*, *R. bechuanae* and *R. intermedius* that have been identified in the Free State and Sandveld (Ganem et al. 2012; du Toit et al. 2012; Ganem et al. 2020); and at Fort Beaufort in the Eastern Cape (du Toit et al. 2012; Ganem et al. 2020). Other contact zones for *R. d. chakae* and *R. d. dilectus* are described from the northern Gauteng Province (Le Grange et al. 2015; Ganem et al. 2020). Altogether, fifteen contact zones for *Rhabdomys* have been documented in South Africa (Ganem et al. 2020), and while one of these has been identified in the vicinity of the Eastern Cape Province in Fort Beaufort, this has been based on limited sampling despite the species being known to occur there. It remains unclear whether the Eastern Cape population is an extension of *R. pumilio*, which typically occurs in the southern parts of Eastern Cape Province, or *R. dilectus* (specifically *R. d. chakae*), which is associated with most sampled parts of the Eastern Cape and the KwaZulu-Natal provinces.

In the Eastern Cape, there are broad forest group types that extend to the neighboring KwaZulu-Natal and Western Cape provinces: Southern mistbelt, Sarp, and Southern coastal forest groups (Von Maltitz et al. 2003; Cowling et al. 2004; Mucina et al. 2006). South African forests cover a small land surface area (Low and Rebelo 1996), and their distribution is severely fragmented (Geldenhuys 2004; Harrison et al. 1997; Skinner and Chimimba 2005; Matthee et al. 2008; Losos, 2011; Jacobs et al. 2013), making taxonomic identifications challenging is convergence in external morphology and pelage coloration (Lai et al. 2008; Losos, 2011; Jacobs et al. 2013), making validation using molecular data crucial. Common errors associated with field identifications often result from the absence of referral or comparative specimens; ambiguity due to undescribed species complexes; and verification of the age of the specimen (Borisenko et al. 2008). The benefits of correct taxonomic identification of rodent taxa include recognition of species that are carriers of diseases (Ramatla et al. 2019); an understanding of their role in parasitic transmission cycles and distribution patterns (Engelbrecht et al. 2016); and knowing which rodents are agricultural pests (Nicolas et al. 2012; Engelbrecht et al. 2016; Ramatla et al. 2019). To date, several studies have demonstrated the utility of COI in the differentiation of geographic forms of *Rhabdomys* (du Toit et al. 2012; Grange et al. 2015; Engelbrecht et al. 2016; Coetzer and Grobler 2018; Matthee et al. 2018; Ganem et al. 2020). In the first instance, COI has been used successfully to differentiate museum specimens into two groups - *R. dilectus* and *R. pumilio* (Coetzer and Grobler 2018), and more recently *R. pumilio* group has been divided further to *R. pumilio*, *R. bechuanae* and *R. intermedius* (Ganem et al. 2020). While these studies cover most of the range of the genus in South Africa, there is a noticeable absence of data from parts of the Eastern Cape, particularly from Alexandria to the southern parts of KwaZulu-Natal.
Materials and methods

Specimen collection and morphological analyses

Animals were collected from three forest types in the Eastern Cape and southern parts of the KwaZulu-Natal Province of South Africa (Figure 1, Table 1). Small and large folding Sherman traps were placed approximately 2 m apart along transects in the forest, forest edge, and the grasslands in the surrounding locality. Standard measurements including total length, tail length, hindfoot length (with claw) and ear length, to the nearest 1 mm were taken, as juveniles or sub-adults. Data were log-transformed and pooled according to forest type and not locality, in order to bolster sample sizes. Descriptive statistics were calculated according to the forest type and locality, in order to bolster sample sizes. Descriptive statistics were calculated, and data were screened for normality, skewness, and kurtosis. Owing to uneven sample sizes, and some data sets that did not conform to normal distribution, a Kruskal-Wallis test was used to test for statistical difference between individuals of the three forest types. Statistical procedures were conducted using SPSS 27.0 (IBM Corp. 2020). For molecular analysis, tissue samples were collected from released animals, and voucher specimens, where applicable, were deposited in the Durban Natural Science Museum, South Africa. Supplementary material from the same museum was also used for this study.

Molecular analysis

DNA was extracted using the NucleoSpin (R) Tissue Kit technique (Macherey-Nagel) following the manufacturer’s manual. The COI gene fragment was isolated using L1490 (5’-GGT CAA CAA ATC ATA AAG ATA TTG G-3’) and H2198 (5’-TAA CTT CAG GGT GAC CAA AAA ATC A-3’) primers (Folmer et al. 1994). The 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 MgCl₂, and distilled water to make up 25 μl reaction volumes. Cycling was performed using a GeneAmp PCR system (Applied Biosystems). PCR conditions involved an initial denaturing step at 94°C for 4 min followed by 35 cycles of 30 s denaturation at 94°C, 30 s, annealing at 42.5°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 then sequenced at Central Analytical Facilities of Stellenbosch University.

Sequence data analysis and phylogenetic reconstruction

Sequences generated were edited and assembled manually in BioEdit v7.0.9 (Hall 1999) and aligned using ClustalW (Thompson et al. 2002). 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) for similarity check with published COI sequences of Rhabdomys on GenBank. Sequences were then deposited on the Barcode of Life Data System (BOLD, Ratnasingham and Hebert 2007) through the online interface at www.barcodinglife.org under the Foundational Biodiversity Information Programme (FBIP) in the Foundational Biodiversity Eastern Cape project.

A haplotype dataset was created using DnaSP v6 (Rozas et al. 2017) from aligned sequences and it was subsequently used to generate a network tree using TCS networks (Clement et al. 2000) in PopART (http://popart.otago.ac.nz). Haplotype diversity, population size changes and population demographic expansions (Tajima’s D) were also generated using DnaSP v6 (Rozas et al. 2017). Individual sequences were grouped according to the forest type they were collected from, and the p-distances among and within the forest types were calculated using the Kimura 2-parameter model (Kimura 1980); these evolutionary analyses were conducted in MEGA version X (Kumar et al. 2018). To test the significance of population structure, three-way analyses of molecular variance (AMOVA) were done. The genetic differentiation among populations (as defined by forest types) was calculated by pairwise FST tests. AMOVA and FST tests were calculated in Arlequin 3.1 (Excoffier et al. 2005) with 1 000 permutations of AMOVA.
All sequences named *R. dilectus* and *R. pumilio* retrieved from GenBank were aligned with sequences generated from this study. Phylogenetic 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 using MEGA version X (Kumar et al. 2018); support for the nodes was obtained by bootstrap resampling (1000 iterations). GenBank derived COI sequences for *Apodemus flavicollis* Myomyscus sp., *Praomys* sp., *Rattus rattus*, and the sequence of *Otomys irratus* (collected from this study) were used as outgroups.

The Barcoding Gap is defined as the difference between the greatest intraspecific variation and the smallest interspecific variation in a group of organisms and is used as the basis for species delimitation (Meyer and Paulay 2005). The mean sequence divergences among specimens were generated using the Kimura 2-parameter (Kimura, 1980) commonly used for barcoding analysis. For estimating the barcoding gap and testing the species hypothesis we used the Automatic Barcoding Gap Discovery (ABGD, Puillandre et al. 2012). 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 selecting the Kimura (K80) TS/TV (2.0). Pairwise differences between individual sequences were calculated and their distributions were grouped to form the histogram using the ABGD online interface. In the same way, the number of preliminary species hypotheses (PSHs) obtained for each prior intraspecific divergence were recorded.

### Results

#### Sequence data and BLAST search results

The COI gene fragment (658 bp) was successfully amplified for all specimens (N=101). The nucleotide content of the gene fragment comprised 30.3% of thymine, 25.6% adenine, 27.2% cytosine and 16.9% guanine. Of the 658 bp, only 13.98% was variable (92/658 bp) while 8.5% of sites were parsimony informative and retrieved 56 haplotypes (Table 1). However, when these sequences were combined with all downloaded sequences, the alignment had a few unaligned bases that had to be trimmed. The new alignment contained 280 sequences that were 582 bp long, excluding outgroups. BLAST analysis (Altschul et al. 1990) of our COI sequences attained 99% similarity with *R. dilectus*, with KC296586.1 top of the list.

Modeltest selected the HKY + G + I (Hasegawa-Kashino-Yano with Gamma and invariant) for the COI data set. However, because of barcoding analysis preference, phylogenetic reconstructions were undertaken using the Kimura 2-parameter model (Kimura 1980) for NJ and MP. Maximum Likelihood (ML) trees were constructed with both models, and the resulting trees were similar. The sequences downloaded from GenBank include *R. bechuanae*, *R. dilectus* (inclusive of *R. d. dilectus* and *R. d. chakae*), *R. intermedius* and *R. pumilio*. The phylogenetic tree retrieved six clades which were robustly supported with high bootstrap values for ML, MP, and NJ. The me-
sic adapted individuals of *R. d. chakae* and *R. d. dilectus* shared a recent common ancestor, while the arid-adapted individuals separated into two paraphyletic clades. One clade contained the individuals of *R. pumilio* coastal A (individuals ranging from the Eastern Cape Province coast downwards the coastline to Western Cape Province and upwards through the coast of Northern Cape Province in South Africa to Namibian coastal regions) and *R. pumilio* coastal B (restricted to Eastern Cape around Fort Beaufort). The other clade contained the sister species *R. bechuanae* (with individuals distributed from Free State Province, South Africa to Namibia) and *R. intermedius* (individuals from Eastern Cape Province towards the Western Cape Province along with the central parts of these provinces). The insert in Figure 2 indicates all specimens of *R. d. chakae*.

<table>
<thead>
<tr>
<th>#</th>
<th>Locality, geographic co-ordinates</th>
<th>Forest type</th>
<th>Sample sizes (Haplotype number)</th>
<th>Voucher number only for collected specimens (all released animals not shown)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Fort Fordyce, -31.68237, 26.482653</td>
<td>Amatole mistbelt</td>
<td>8 (Hap 8, 9, 17, 18, 19, 20, 21, 22)</td>
<td>DM15041, DM15240, EM10070118, EM29070118</td>
</tr>
<tr>
<td>B</td>
<td>Hogsback, -32.613906, 26.972192</td>
<td>Amatole mistbelt</td>
<td>3 (Hap 28, 29)</td>
<td>DM15266, DM15267</td>
</tr>
<tr>
<td>C</td>
<td>Isidenge, -33.69975, 26.359898</td>
<td>Amatole mistbelt</td>
<td>11 (Hap 4, 8, 21, 26 28,30, 31,33,34)</td>
<td>DM4253, DM4249, DM15016, DM15024, DM15024, DM15252, DM15253, DM15254, DM15255</td>
</tr>
<tr>
<td>D</td>
<td>Kolqoga, -32.535330, 27.36258</td>
<td>Amatole mistbelt</td>
<td>6 (Hap 11, 25, 26, 27,32)</td>
<td>DM15273, DM15274, DM15275</td>
</tr>
<tr>
<td>E</td>
<td>Baziy, -31.547255, 28.437503</td>
<td>Transkei mistbelt</td>
<td>12 (Hap 7, 8, 9, 10, 11, 45, 46)</td>
<td>LRR180212BSF_RP1, LRR180212BSF_RP2, LRR180212BSF_RP3, LRR180212BSF_RP4, LRR180212BSF_RP5, LRR180213BSF_RP7</td>
</tr>
<tr>
<td>F</td>
<td>Gomo, -31.013194, 29.344585</td>
<td>Transkei mistbelt</td>
<td>5 (Hap 9, 11, 23, 24)</td>
<td>EM73031217, EM84041217, EM9021217</td>
</tr>
<tr>
<td>G</td>
<td>Ngeli, -30.542827, 29.680979</td>
<td>Eastern mistbelt</td>
<td>6 (Hap 11, 28, 42, 43, 44)</td>
<td>LRR180314NSF_RD1, LRR180315NSF_RD1, LRR180315NSF_RD2, LRR180317NSF_RD1, LRR18001NSF_RD1, LRR181001NSF_RD2</td>
</tr>
<tr>
<td>H</td>
<td>Nxumeni, -29.927042, 29.844179</td>
<td>Eastern mistbelt</td>
<td>14 (Hap 10, 11, 12, 13, 44, 47, 48)</td>
<td>LRR180321BF_RD1, LRR180321BF_RD2, LRR180321BF_RD3, LRR180321BF_RD6, LRR180321BF_RD8, LRR180324BF_RD2, LRR180324BF_RD5, LRR180325BF_RD1</td>
</tr>
<tr>
<td>I</td>
<td>Umntemvuna, -31.007467, 30.094205</td>
<td>Pondoland scarp</td>
<td>4 (Hap 49, 54, 55, 56)</td>
<td>DM1174, DM1177, DM1178, DM1182</td>
</tr>
<tr>
<td>J</td>
<td>Mkambati, -31.247618, 29.900551</td>
<td>Pondoland scarp</td>
<td>4 (Hap 10, 28, 39)</td>
<td>DM151163, DM15164, DM15165, DM15166</td>
</tr>
<tr>
<td>K</td>
<td>Mbotyi, -31.418887, 29.723978</td>
<td>Pondoland scarp</td>
<td>5 (Hap 11, 28, 37, 38)</td>
<td>DM15176, DM15177</td>
</tr>
<tr>
<td>M</td>
<td>Dwesa, -32.445594, 28.60766</td>
<td>Transkei scarp</td>
<td>3 (Hap 1, 14, 15)</td>
<td>DM15190, DM15191, DM15192</td>
</tr>
<tr>
<td>N</td>
<td>Manubi, -33.900792, 25.360209</td>
<td>Transkei scarp</td>
<td>6 (Hap 9, 10, 16, 35, 36)</td>
<td>DM15214, DM15215, DM15216</td>
</tr>
<tr>
<td>O</td>
<td>Morgans Bay, -32.701243, 28.353646</td>
<td>Eastern Cape dune</td>
<td>3 (Hap 40, 41)</td>
<td>EM201, EM207</td>
</tr>
<tr>
<td>P</td>
<td>Alexandria, -33.699756, 26.359898</td>
<td>Albany</td>
<td>6 (Hap 1, 2, 3, 4, 5, 6)</td>
<td>No vouchers.</td>
</tr>
<tr>
<td>16 study localities</td>
<td>7 Forest types</td>
<td>101(56 haplotypes)</td>
<td></td>
<td>LRR=Not catalogued, EM=Stellenbosch university vouchers, DM=Durban Natural Science Museum catalogued specimens (includes historic specimen material)</td>
</tr>
</tbody>
</table>

**Genetic diversity and networks of *Rhabdomys dilectus***

The mean $p$-distance among and within forest types for *R. dilectus* varied from 0.001±0.001 to 0.017±0.003 (Table 3a). The lowest divergence was within the Eastern Cape dune forest grouping (0.001±0.001), while the highest divergence was found amongst Pondoland Coastal scarp forest animals (0.017±0.003). Most observed divergences within the forest types were higher than the divergences among these provinces.
between forest types. The significant value of Tajima’s D of \(-2.54872\), \(P<0.001\) indicated that the population of *R. d. dilectus* from the sampled area had undergone recent expansion. Tajima’s D remained significant (\(-2.60563\), \(P<0.001\)) when GenBank derived sequences of *R. d. chakae* were added. The AMOVA revealed that variation within the population was greater (92.81%) than variation among the population (7.19%) with the fixation index \(F_{ST}\) of 0.071787, \(P<0.001\) (Table 3a). Variation within the population remained high even when the genetic structure was examined between broader forest groups (i.e., Southern coastal forest Southern mistbelt, and Scarp groups).

The median-joining networks for COI (Figure 3) showed the majority of haplotypes differ by one mutation (i.e., singletons); and those haplotypes were not restricted to forest types or groups. A high number of haplotypes were from the mistbelt and scarp forests, with common haplotypes dominated by individuals from these localities. Of the six common haplotypes, four comprised individuals from both mistbelt and scarp forests, while only two of these haplotypes comprised individuals from coastal forests and either mistbelt or scarp forests. Three of the common haplotypes with individuals from the mistbelt and scarp forests were ancestral haplotypes, with most other haplotypes joined by one
mutation from these haplotypes. Although there were no clear patterns of forest segregation indicated by the haplotypes, there were few haplotypes that are individually divergent from the others. A total of 56 haplotypes (from n=101) resulted in a high haplotype diversity of 0.972 (Table 1).

Barcoding Gap among *Rhabdomys* species

The groupings as retrieved above by the phylogenetic tree were used for calculating divergences between and within *Rhabdomys* taxa (Table 3b). The results revealed that the nucleotide difference was always smaller within all the groups (between 0.004±0.001 and 0.023±0.004) than the lowest nucleotide difference between groups (which is between *R. pumilio* coastal A and coastal B groups 0.050±0.002). The highest nucleotide difference within a group was for *R. d. dilectus* (0.023±0.004); this nucleotide difference was still lower than the lowest nucleotide difference between groups: between *R. pumilio* coastal A and coastal B groups (0.050±0.002). There was a notable lower nucleotide difference within *R. d. chakae* (0.008±0.002), compared to the nucleotide difference between *R. d. chakae* and the other groups. The low nucleotide differences between groups were between *R. pumilio* coastal A with *R. pumilio* coastal B (0.050±0.002); between *R. d. chakae* and *R. d. dilectus* (0.053±0.001); and between *R. bechuanae* and *R. intermedius* (0.079±0.001). The ABGD analyses (Figure 4) retrieved seven groups up to prior maximal distance P= 0.002783. Following the grouping used above, individuals of *R. d. dilectus* were separated into two groups, and all other individuals have grouped accordingly as *R. bechuanae*, *R. d. chakae*, *R. pumilio* coastal A, *R. pumilio* coastal B and *R. intermedius*. From the prior maximal distance of P= 0.004642 to the prior maximal distance of P=0.012915 each of the six groups was distinct, although at prior maximal distance P=0.021544 the analysis combined individuals of *R. pumilio* coastal A and *R. pumilio* coastal B. The ABGD analysis did not combine *R. d. dilectus* and *R. d. chakae*.

---

**Figure 3.** Median-joining haplotype network generated from COI of 101 specimens of *R. dilectus* from forests in the Eastern Cape and southern KwaZulu-Natal provinces of South Africa. The haplotype colors represent the forest groups from which the individuals were collected. Forest groups are indicated in the insert (also in Figure 1). The number of mutations separating haplotypes is indicated using hatch marks. The network was drawn in PopArt.
Table 2. Morphological analyses of 78 Rhabdomys dilectus specimens across different forest types. Standard measurements including total length (TL), tail length (T), hindfoot length (with the claw, HF) and ear length (E), to the nearest 1 mm were taken in the field; the head and body length (HB) was calculated as the difference of tail and the total length of the animal. Sample sizes are presented in parentheses. Descriptive statistics (mean, standard deviation) and Kruskal-Wallis tests were conducted in SPSS 27.0 (IBM Corp., 2020).

<table>
<thead>
<tr>
<th></th>
<th>Mistbelt</th>
<th>Scarp</th>
<th>Coastal</th>
<th>The combined average for all groups</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total length (mm)</td>
<td>196.0±5.09 (42)</td>
<td>188.6±7.47 (25)</td>
<td>186.9±7.75 (11)</td>
<td>192.4±5.09 (78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Head-and-body (mm)</td>
<td>103.1±3.40 (42)</td>
<td>99.0±4.83 (25)</td>
<td>97.6±4.52 (11)</td>
<td>101.1±3.40 (78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Tail length (mm)</td>
<td>92.9±2.33 (42)</td>
<td>89.6±3.07 (25)</td>
<td>89.3±3.47 (11)</td>
<td>91.3±2.33 (78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ear (mm)</td>
<td>12.0±0.14 (42)</td>
<td>12.1±0.44 (25)</td>
<td>12.5±0.52 (11)</td>
<td>12.1±0.38 (78)</td>
<td>0.010</td>
</tr>
<tr>
<td>Hindfoot (mm)</td>
<td>20.4±0.92 (42)</td>
<td>18.9±1.20 (25)</td>
<td>18.5±1.63 (11)</td>
<td>19.5±1.41 (78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HB:Tail ratio</td>
<td>1.11±0.03 (42)</td>
<td>1.10±0.03 (25)</td>
<td>1.09±0.02 (11)</td>
<td>1.11±0.03 (78)</td>
<td>0.123</td>
</tr>
</tbody>
</table>

Figure 4. Barcode gap analysis of Rhabdomys taxa generated by Automatic Barcode Discovery Gap Discovery (Puillandre et al. 2012) using 280 sequences in the COI gene (582 bp). The gap was calculated with the default setting using Kimura (K80) TS/TV distances. (a) Histograms showing the frequency of pairwise nucleotide differences, and (b) shows a plot with the number of PSHs (preliminary species hypothesis) obtained for each prior intraspecific divergence (0.01–0.1000).

Morphology

The average total length of specimens was 186.9±7.75 mm for coastal forest, 188.6±7.47 mm in the scarp forest and 196.0±5.09 mm for mistbelt forests; mean head-and-body length was 97.6±4.52 mm for coastal forest, 99.0±4.83 in the scarp forest and 103.1±3.40 mm for mistbelt forests; and mean tail length was 89.3±3.47 mm for coastal forest, 89.6±3.07 in the scarp forest and 92.9±2.33 mm in the mistbelt forests (Table 2). No significant difference was found for mean values of head-and-body to tail length ratio (p=0.123): HB: tail was 1.09±0.02 in the coastal
Table 3a. The mean p-distances among and within (in bold) forest types in the Eastern Cape and KwaZulu-Natal provinces of South Africa for COI (658bp) for Rhabdomys dilectus. The number of base substitutions per site, derived from the mean of all sequence pairs between groups, is shown with the standard error estimate(s) next to each. Analyses were conducted using the Kimura 2-parameter model (Kimura 1980). Evolutionary analyses were conducted in MEGA version X (Kumar et al. 2018).

<table>
<thead>
<tr>
<th></th>
<th>Albany</th>
<th>Transkei mistbelt</th>
<th>Eastern mistbelt</th>
<th>Transkei scarp</th>
<th>Amatole mistbelt</th>
<th>Pondoland scarp</th>
<th>Eastern Cape dune</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albany</td>
<td>0.011±0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transkei Mistbelt</td>
<td>0.008±0.002</td>
<td>0.004±0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern Mistbelt</td>
<td>0.012±0.002</td>
<td>0.008±0.001</td>
<td>0.01±0.002</td>
<td>0.012±0.002</td>
<td>0.011±0.002</td>
<td>0.017±0.003</td>
<td>0.001±0.001</td>
</tr>
<tr>
<td>Transkei Scarp</td>
<td>0.011±0.002</td>
<td>0.008±0.002</td>
<td>0.012±0.002</td>
<td>0.012±0.002</td>
<td>0.011±0.002</td>
<td>0.017±0.003</td>
<td>0.001±0.001</td>
</tr>
<tr>
<td>Amatole Mistbelt</td>
<td>0.011±0.002</td>
<td>0.008±0.001</td>
<td>0.012±0.002</td>
<td>0.012±0.002</td>
<td>0.011±0.002</td>
<td>0.017±0.003</td>
<td>0.001±0.001</td>
</tr>
<tr>
<td>Pondoland Scarp</td>
<td>0.014±0.002</td>
<td>0.012±0.002</td>
<td>0.015±0.002</td>
<td>0.015±0.002</td>
<td>0.015±0.002</td>
<td>0.015±0.002</td>
<td>0.015±0.002</td>
</tr>
<tr>
<td>Eastern Cape Dune</td>
<td>0.009±0.003</td>
<td>0.008±0.003</td>
<td>0.012±0.003</td>
<td>0.010±0.003</td>
<td>0.011±0.003</td>
<td>0.013±0.003</td>
<td>0.001±0.001</td>
</tr>
<tr>
<td>Tajima’s D: –2.60563</td>
<td>P&lt;0.001</td>
<td>Va = 92.81</td>
<td>Vb = 7.19</td>
<td>Fixation Index</td>
<td>FST: 0.071787</td>
<td>P&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

Table 3b. The number of base substitutions per site in Rhabdomys taxa. These were derived from the mean of all sequence pairs between and within species, with standard error estimate(s) shown next to each. Analyses were conducted using the Kimura 2-parameter model (Kimura 1980). The analysis involved 280 nucleotides in the COI dataset which had a total of 582 bp. Evolutionary analyses were conducted in MEGA version X (Kumar et al. 2018).

<table>
<thead>
<tr>
<th></th>
<th>R. bechuane</th>
<th>R. d. dilectus</th>
<th>R. d. chakae</th>
<th>R. intermedius</th>
<th>R. pumilio coastal A</th>
<th>R. pumilio Coastal B</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. bechuane</td>
<td>0.009±0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. d. dilectus</td>
<td>0.115±0.014</td>
<td>0.023±0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. d. chakae</td>
<td>0.123±0.015</td>
<td>0.053±0.001</td>
<td>0.008±0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. intermedius</td>
<td>0.079±0.011</td>
<td>0.105±0.002</td>
<td>0.113±0.002</td>
<td>0.004±0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. pumilio coastal A</td>
<td>0.131±0.016</td>
<td>0.086±0.001</td>
<td>0.099±0.002</td>
<td>0.115±0.002</td>
<td>0.007±0.002</td>
<td></td>
</tr>
<tr>
<td>R. pumilio Coastal B</td>
<td>0.118±0.015</td>
<td>0.094±0.002</td>
<td>0.098±0.002</td>
<td>0.114±0.002</td>
<td>0.050±0.002</td>
<td>0.005±0.002</td>
</tr>
</tbody>
</table>

forest; 1.10±0.03 in the scarp forest; and 1.11±0.03 in the mistbelt forests. A Kruskal-Wallis test revealed significant differences in mean values for all other external measurements evaluated; mean values for mistbelt forest individuals were, in general, greater than those originating from scarp or coastal forests.

Discussion

Barcoding efficiency in delineating Rhabdomys species

Based on DNA barcoding, all samples in this study were attributed to R. d. chakae. The two sub-species of R. dilectus (R. d. chakae and R. d. dilectus) have long since been recognized: R. d. chakae is endemic to the country, while the geographic distribution of R. d. dilectus extends beyond South Africa (Ramabu et al. 2003; du Toit et al. 2012; Ganem et al. 2020). With the exception of R. d. dilectus (with a variation slightly higher 0.023±0.004), there was very low genetic variation within populations of each of the five groups (between 0.005±0.002 and 0.009±0.002), suggesting that R. d. chakae could be an independent species. However, the low variation between R. pumilio coastal A and R. pumilio coastal B (0.050±0.002); and between R. d. chakae and R. d. dilectus (0.053±0.001), which comprise the lowest variation between groups, suggest that these groups are subspecies or lineages (R. dilectus and R. pumilio) respectively. This is supported by the finding that divergences for small mammals with more than one distinct lineage can have more than 3 % intraspecific variation (only up to a maximum of 5.4 %), and only a few species without known lineages show more than 2 % intraspecific variation (Borisenko et al. 2008). To support this the ABGD (Puillandre et al. 2012), preliminary species hypothesis (PSH) suggested six distinct groups should be considered as species, and the distance phylogenetic tree revealed six groups and each group had at least 99% bootstrap support.

Our combined (sequence divergence, ABGD and phylogenetic analyses) findings are largely congruent, and suggest six groups may be assigned to two sub-species of R. dilectus; two lineages of R. pumilio; and two species R. bechuanae and R. intermedius. These findings are consistent with the previous findings relating to this genus (du Toit et al. 2012; Ganem et al. 2020). But our study provides much-needed barcoding data for Rhabdomys from the Eastern Cape Province which was not provided by previous studies (Castiglia et al. 2011; Meynard et al. 2012; Ganem et al. 2012; Ganem et al. 2020). Our study documents the distribution of this taxa in the unsampled parts of the Eastern Cape and southern KwaZulu-Natal.
provinces and found this area to be inhabited solely by *R. d. chakae*.

**Morphology**

Ganem et al. (2020) suggested that morphological comparison within *Rhabdomys* is unreliable, because whenever there is more than one collector, their measuring criteria may be different. But despite this reservation, the head-and-body to tail ratio has been used to separate the mesic group from the arid group (Ganem et al. 2020). The analysis involved controlling for mass and data collector effect and found that tails of the mesic species *R. dilectus* (*R. d. dilectus* and *R. d. chakae*) were, in general, shorter than those of their arid-adapted counterparts. Our study revealed possible clinal variation in *R. d. chakae* external morphology, with individuals from more temperate climates (mistbelt) slightly larger or longer than their coastal conspecifics. Animals recorded from coastal forests appeared to be the shortest of the three groups investigated. However, we found no significant difference in the ratio of the head-and-body to tail across the forest types which averaged 1.11±0.03 mm, matching the ratio of *R. d. chakae* from Ganem et al. (2020). This confirms that all specimens in this study are indeed *R. d. chakae*, supporting the taxonomic assignments based on DNA barcoding in this group.

**Phylogeography**

The reticulate tree (Figure 3) indicates that this species has haplotypes comprising individuals from all over the study area. It should be noted that most of the shared haplotypes include individuals from the mistbelt and scarp forests, including almost all single haplotypes. The distribution and assemblage of forest animals in South Africa have been heavily influenced by climatic fluctuations and species recolonization following these fluctuations. However, forest-associated species such as *R. d. chakae* were less affected by such events, as they were able to survive in adjacent habitats (Lawes et al. 2007). The high number of single haplotypes and high haplotype diversity within *R. d. chakae* may be attributed to very high reproductive rates: an average of two litters per season producing up to 11 offspring per litter (Taylor 1998; do Toit et al. 2016). There was no specific structure detected with the AMO-VAs, and the phylogenetic tree (insert in Figure 2) has robust bootstrap support. This indicates that our sampled individuals are indeed one species, with no indications of lineages in this species, and exhibit a low variation of less than 1% within this group (0.008±0.002, Table 3b). These data show that there is a continuous distribution of the species ranging from Alexandria in the Eastern Cape to KwaZulu-Natal. There are also known records of *R. d. chakae* in Eswatini and Lesotho (Castiglia et al. 2012; du Toit et al. 2012). These findings confirm the continuous distribution occupied by this species (du Toit et al. 2012; du Toit et al. 2016; Ganem et al. 2020), reflecting its generalist nature as it utilizes a wide range of resources (Brown 1984).

Although a wide range of rodents are good candidate taxa for phylogeographical analysis (Bryja et al. 2014), it is important to note that rodents have different habitat responses and sensitivity to environmental change (Dondina et al. 2017). Generalist species, such as the subspecies *R. d. chakae* are hypothesized to conform to the predictions of meta-population theory rather than island biogeography (MacNally and Bennet 1998). Indeed, the genetic analysis of *R. d. chakae* supports this postulation, as no phylogeographic structure was detected. These data indicate that there is connectivity between the populations in different forests and reflects the ability of the species to survive in disturbed, semi-urban and agricultural landscapes (Ganem et al. 2012). The distribution of *R. d. chakae* may have been limited by the Bedford gap (Lawes 1990; Lawes et al. 2000; Lawes et al. 2007), which has been shown to limit the distribution of other small mammal species (Willows-Munro 2008, Abrahams 2018). There are no records of the subspecies beyond this gap: south-west of the Bedford area there is a population of *R. pumilio* extending to the Western Cape Province (du Toit et al. 2012). The recorded contact zones in the Eastern Cape Province are in the Bedford area (du Toit et al. 2012; Ganem et al. 2020) and may be attributed to recent colonization (as Tajima’s D indicates) that the population of *R. dilectus* from the sampled area has undergone recent expansion.

**Conclusions**

The utility of DNA barcoding in identifying *Rhabdomys* taxa has been supported. This study focused on supplying information on the distribution of the taxa in the Eastern Cape and southern parts of KwaZulu-Natal provinces and confirmed that all samples included in this study were of *R. d. chakae*. These findings are consistent with previous revisions of this genus (du Toit et al. 2012, Ganem et al. 2020), suggesting four species with two subspecies of *R. dilectus* and two geographic and genetic distinct lineages of *R. pumilio*. Our molecular and morphological data support the delineation of *R. d. chakae* and *R. d. dilectus*. There may be intraspecific morphological clinal variation within the populations studied here, which merits further investigation. Our findings support the notion that although the ecological requirements of species may differ geographically, generalist species tend to show connectivity between subpopulations within fragmented habitats.

**Acknowledgments**

The project was funded by the Foundational Biodiversity Information Programme (FBIP 98871) of the Department of Science and Technology of South Africa through the National Research Foundation. The
student extension funding was provided by the National Research Foundation and the Stellenbosch University. We express gratitude to the Durban Natural Science Museum and eThekwini Municipality for co-funding field trips and provisioning equipment. We express gratitude to private land-owners, the former Department of Environment, the former Department of Agriculture, Forestry and Fisheries, Eastern Cape Province: Department of Economic Development Environmental Affairs and Tourism, The Eastern Cape Parks and Tourism Agency, Ezemvelo KZN Wildlife, and South African National Parks and private landowners for permits and facilitating sample collection. We also acknowledge the staff of the Durban Natural Science Museum and fellow research students for assistance during sample collection; as well as the collection of samples by Yusi Martins and Dale Wright. The capturing and handling of animals was approved by the ethics committee of the University of Stellenbosch (protocol number 1285) and followed the 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).

References


