

# New insights from RADseq data on differentiation in the Hottentot golden mole species complex from South Africa

Samantha Mynhardt<sup>a,b</sup>, Nigel C. Bennett<sup>b,c</sup> and Paulette Bloomer<sup>a,c</sup>

<sup>a</sup> Molecular Ecology and Evolution Programme, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

<sup>b</sup> Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

<sup>c</sup> Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

\*Corresponding author at: Molecular Ecology and Evolution Programme, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa.

Email; samantha.mynhardt@up.ac.za

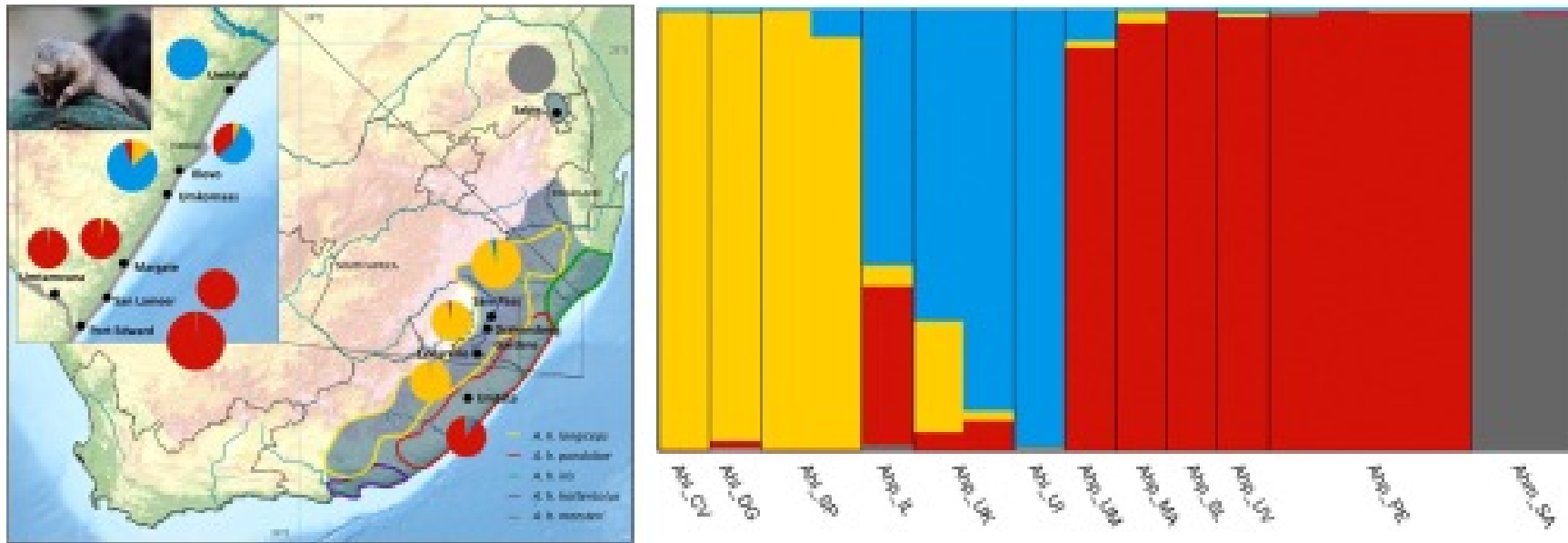
## Highlights

- Our results support *Amblysomus hottentotus meesteri* as a distinct divergent lineage.
- Central coastal *A. h. pondoliae* is divergent from southern *A. h. pondoliae*.
- Central coastal *A. h. pondoliae* may be worthy of specific or subspecific status.
- Mito-nuclear discordance found in a cryptic mtDNA lineage from Umtata.
- These cryptic lineages may be in need of conservation attention.

## Abstract

Golden moles (Family Chrysochloridae) are small subterranean mammals, endemic to sub-Saharan Africa, and many of the 21 species are listed as threatened on the IUCN Red List. Most species have highly restricted ranges; however two species, the Hottentot golden mole (*Amblysomus hottentotus*) and the Cape golden mole (*Chrysochloris asiatica*) have relatively wide ranges. We recently uncovered cryptic diversity within *A. hottentotus*, through a phylogeographic analysis of this taxon using two mitochondrial gene regions and a nuclear intron. To further investigate this cryptic diversity, we generated nuclear SNP data from across the genome of *A. hottentotus*, by means of double-digest restriction-site associated DNA sequencing (ddRADSeq), and mapped reads to the Cape golden mole genome. We conducted a phylogenetic analysis and investigated population differentiation. Our results support the distinctiveness of *A. h. meesteri*. Furthermore, we provide evidence from nuclear SNPs in support of our previous finding that Central coastal samples represent a unique cryptic lineage that is highly divergent from *A. h. pondoliae* farther south. Although mtDNA suggests that Umtata may represent a unique lineage sister to *A. h. longiceps*, mito-nuclear discordance from our RADseq data indicate that these samples may instead be closer to *A. h. pondoliae*, and therefore may not represent a distinct lineage. We stress the importance of recognizing that understudied populations, such as that of Umtata, may represent populations or ESUs under threat and in need of conservation attention. We present a high-quality filtered SNP dataset, comprising thousands of SNPs, which may serve as a useful resource for future golden mole studies. We have thus added to the growing body of research demonstrating the power and utility of RADseq to investigate population differentiation.

## Graphical abstract



**Keywords:** *Amblysomus hottentotu*; Cryptic diversity; ddRADseq; Population differentiation; Small mammal; Southern Africa

## Abbreviations

CV: Cedarville  
DG: Drakensberg Gardens  
IL: Illovo  
MA: Margate  
PE: Port Edward  
SA: Sabie  
SL: San Lameer  
SP: Sani Pass  
UK: Umkomaas  
UI: Umhlali  
UM: Umtata  
UV: Umtamvuna

## 1. Introduction

Golden moles (Family Chrysochloridae, Order Afrotheria) are small, fossorial mammals endemic to sub-Saharan Africa. Phenotypically they may resemble other subterranean mammals, such as lipotyphlan moles (Talpidae), burrowing rodents (Bathyergidae), certain armadillos (*Chlamyphorus*), and marsupial moles (Notoryctidae), but do not share a close phylogenetic relationship with any of these groups. Instead, chrysochlorids belong to the Afrotheria, a radiation of endemic African mammals, also including hyracoids (hyraxes), proboscideans (elephants), sirenians (sea cows), macroscelidids ( sengis or elephant shrews), tubulidentates (the aardvark), and tenrecids (tenrecs) (Stanhope et al., 1998, Seiffert, 2007, Asher et al., 2009). Golden moles are elusive and largely understudied, mostly due to the challenges in collecting them and in observing their subterranean behaviour in the wild. Rainfall and soil characteristics are major factors that affect their distributions (Bronner, 1996b, Jackson and Robertson, 2011). These animals are thought to feed mainly on earthworms and crickets, and have very specific soil type requirements, being often restricted to deep sandy or soft loam soils, where there is also an abundance of invertebrate prey (Kuyper, 1985). This restriction, along with other anthropogenic factors, such as urbanisation, mining and agricultural developments, have resulted in most species having significantly restricted ranges and, consequently, highly fragmented and/or isolated distributions, a feature which is common to subterranean species (Busch et al., 2000).

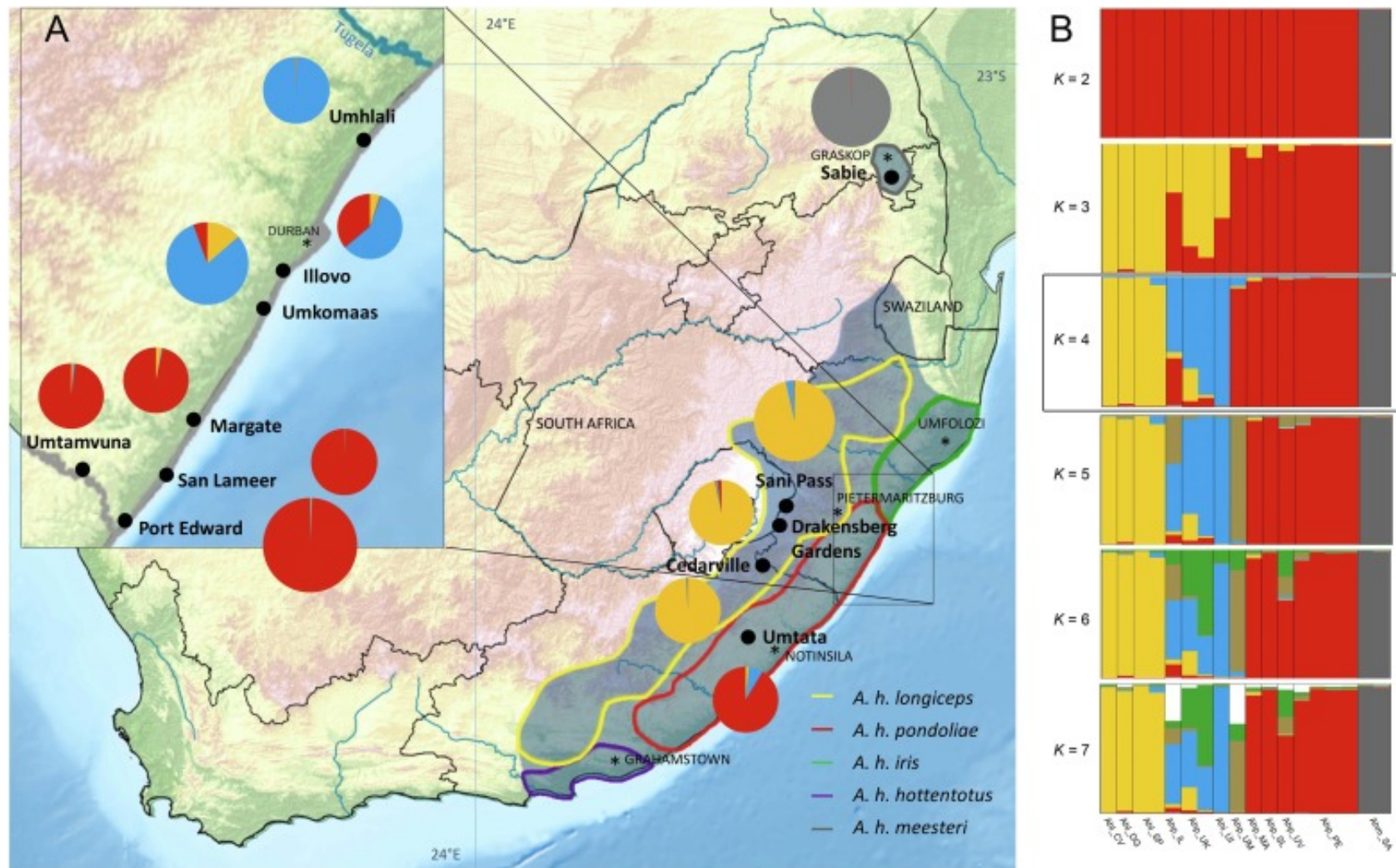
The Chrysochloridae comprises 21 species, of which ten are listed as threatened on the IUCN Red List (The IUCN Red, 2017). Establishing a sound taxonomy for golden moles is paramount to informing conservation efforts for the various taxa under threat. Early taxonomic revisions of chrysochlorids (Forcart, 1942, Simonetta, 1968, Meester, 1974, Petter, 1981; and others) have been mostly intuitive, or based on elementary statistical evaluation of relatively few specimens. These focused mainly on inter-specific and generic relationships, without thoroughly assessing intra-specific variation. Later revisions have focused on hyoid shape (Bronner, 1991), chromosome morphology (Bronner, 1995), and craniodental anatomy (Bronner, 1996b). Most recently, partial sequences of the nuclear GHR gene were combined with morphological characters to derive a broad phylogeny for the family (Asher et al., 2010). However, the relationships and intra-specific variation of many taxa within the family remain unresolved.

The relatively widespread and abundant Hottentot golden mole, *Amblysomus hottentotus* (A. Smith, 1829; IUCN Least Concern), endemic to the Greater Maputaland-Pondoland-Albany biodiversity hotspot in South Africa (GMPA; Perera et al., 2011, Perera et al., 2018), is distinguished from other golden mole species by morphological characteristics as well as cytogenetic distinctions (Bronner, 1995). It has a diploid chromosome number of 30, as opposed to those of 34 and 36 found in some other golden mole species. Considerable geographic variation exists in the size and colouration of *A. hottentotus* (Taylor et al., 2018), possibly as a result of stabilising selection acting on localised populations of individuals with low vagility, which is common in subterranean mammals that consequently often display marked geographic differences in morphology and karyotypic properties (Nevo, 1979). Metabolic rate and other physiological properties of *A. hottentotus* also vary geographically in relation to body size (Kuyper, 1985).

The subdivision of *A. hottentotus* is generally in line with the distribution of localized populations in South Africa (Fig. 1). *Amblysomus h. hottentotus* and *A. h. pondoliae* (type localities Grahamstown and Notinsila respectively; Thomas and Schwann, 1905) are intermediate-sized and were recognised as distinct subspecies on the basis of craniometric differences, and *A. h. iris* (type locality Umfolozi station; Thomas and Schwann, 1905), the smallest subspecies, on the basis of colourimetric and craniometric comparisons. *Amblysomus h. longiceps* (type locality Pietermaritzburg; Broom, 1907) does not differ evidently from coastal populations in cranial shape, and was afforded subspecific rank based on its markedly larger body size. *Amblysomus h. meesteri* (type locality Graskop; Bronner, 2000) closely resembles the coastal populations in craniometric properties, but was recognised as a subspecies on the basis of colourimetric differences, with a mid-dorsal band of reddish-black fur distinguishing it from all other forms. This taxon is also distinguished from the geographically proximate *A. septentrionalis* and *A. robustus* on the basis of differences in dentition, craniometrics and cytogenetics (Bronner, 1996a). Thus, the current classification is primarily based on subtle morphological distinctions, but many of these characteristics have proven to be ambiguous and inconclusive (Bronner, 1996b). For instance, dentition has been found to be an unreliable character in chrysochlorid taxonomy since variation occurs due to an unusual sequence of tooth placement resulting in deciduous and permanent teeth occurring in the same specimen (Bronner, 1996b). These concerns, along with non-geographic variation in morphology and geographic isolation of some of the subspecific populations, have raised uncertainty over the current intraspecific taxonomic classification of *A. hottentotus*.

A comprehensive karyotypic comparison among six species/subspecies of golden moles based on G-banding and chromosome painting revealed that *A. h. meesteri* groups separately to a clade containing *A. h. hottentotus*, *A. h. longiceps*, *A. h. pondoliae* and *A. robustus* (Bronner, 2000), based on a shared intrachromosomal rearrangement and the detection of telomeric sequences in the centromeres of all chromosomes of the three *A. hottentotus* subspecies and *A. robustus*, but not in those of *A. h. meesteri* (Gilbert et al., 2008). These findings indicate an absence of gene flow suggesting that *A. h. meesteri* might represent a distinct species.

In a previous phylogeographic study we uncovered substantial cryptic diversity within *A. hottentotus*, suggesting that this taxon represents a species complex with most recognized subspecies being worthy of species status (Mynhardt et al., 2015). However, this hypothesis requires further support from nuclear DNA and morphological data, since the study was based predominantly on mitochondrial data and included only one nuclear intron. In addition,



**Fig. 1.** (a) Map of southern Africa indicating the distribution of *Amblysomus hottentotus* (shading) and its five subspecies, each subspecies type locality (asterisks) and the 12 sampling sites (black circles). Proportionate assignment to four putative populations is indicated in the associated pie charts. (b) STRUCTURE analysis for the full dataset, based on 805 SNPs, indicating the probability of each individual belonging to each of  $K$  clusters (Structure's  $Q$ ). Results are shown for  $K=2-7$ , with  $K=4$  being the most likely, based on delta  $K$  estimation. Vertical bars represent the 18 individuals, with the 12 sampling localities indicated below.

there have been no studies addressing population structure in golden moles, and as a result very little is known about the connectivity among populations of these animals.

We therefore conducted an investigation into the genomic diversity and differentiation in *A. hottentotus*, by means of restriction-site associated DNA sequencing (RADseq; Miller et al., 2007, Baird et al., 2008), a reduced-representation approach which enables SNP discovery and simultaneous genotyping in non-model species such as golden moles. This method has become increasingly popular in ecological, evolutionary and phylogenomics (Davey et al., 2011, Andrews et al., 2016, Catchen et al., 2017, Lowry et al., 2017, McKinney et al., 2017), and has frequently been used successfully for population genetic analyses in mammalian systems, including assessment of population structure, hybridization, demographic history, phylogeography and migration (Demos et al., 2015, Lanier et al., 2015, Knowles et al., 2016, Lah et al., 2016, Sovic et al., 2016, Svengren et al., 2017, Wang et al., 2017, Humble et al., 2018, Krohn et al., 2018).

Golden moles are thought to have an exceptionally large genome of approximately 4.21 Gb (*Chrysochloris asiatica* genome; GenBank WGS: AMDV01000001-AMDV01391343, 4.21 Gb; Broad Institute 2012). RADseq is an ideal method for SNP discovery in organisms of medium to large genome size, where sequencing the entire genome would be too expensive or too elaborate to address the questions at hand. Double-digest RAD (ddRAD; Peterson et al., 2012), which uses a double restriction enzyme (RE) digest, enables size selection by eliminating the random shearing involved in traditional RADseq. This size selection can further reduce complexity, and thereby increase the depth of sequencing, which is especially useful for organisms of medium to large genome size.

The aim of this study was thus to gain a more thorough understanding of the genetic differentiation among populations of the Hottentot golden mole, *A. hottentotus*, in South Africa. We follow the taxon names, linked to formerly proposed distributions of the subspecies in order to compare the nuclear DNA differentiation to our previous study (Mynhardt et al., 2015). We however acknowledge that future species delimitation based on multi-gene and morphological analyses and inclusion of material from all type localities could change the current interpretation. Furthermore, we present a high-quality filtered SNP dataset, which may serve as a useful resource for future golden mole studies.

## 2. Material and methods

### 2.1. Sample collection

A total of 123 Hottentot golden mole (*Amblysomus hottentotus*) samples were collected between 2002 and 2011 (Permit numbers: MPB5304, CPB6003769, 1731/2005, 232/2007, WRO 23/05WR, WRO 77/07WR) across the entire known species distribution range, of which 18 high-quality representative samples were selected for this study (Fig. 1). The sampling scheme was designed to minimize the cost of sequencing many individuals, while maximizing the potential for SNP discovery in *A. hottentotus*, and therefore included two *A. h. meesteri* from Sabie, four *A. h. longiceps* from the Drakensberg region (Sani Pass, Drakensberg Gardens and Cedarville), ten *A. h. pondoliae* from six localities within the subspecies' distribution (Illovo, Umkomaas, Margate, Port Edward, San Lameer and Umtamvuna), one *A. h. iris* from Umhlali and one from a potentially unique cryptic lineage at Umtata. Individuals were captured with Hickman live-traps (Hickman, 1979), which were baited with worms or crickets from the native habitat. This study was conducted in

accordance with the UK Home Office Animals (Scientific Procedures) Act 1986 and with the regulations of the University of Pretoria's Animal Ethics Committee (ethics clearance no. EC100-13). Golden moles are highly elusive, and intensive efforts are required to obtain samples. These animals do not respond well to stress and even non-fatal sampling methods often result in fatal stress. Other non-invasive sampling methods, such as faecal sampling, were ruled out by the challenging nature of collecting such samples from a subterranean habitat. Therefore, the animals used in this study were euthanized with halothane (Safe Pharmaceuticals Pvt. Ltd, Florida, South Africa), and all efforts were made to minimize suffering. Specimens were stored frozen at  $-20^{\circ}\text{C}$  and later dissected to obtain tissue samples (heart, liver, kidney, pectoral muscle) that were stored in 70% ethanol or at  $-20^{\circ}\text{C}$ .

## 2.2. DNA extraction, library preparation and sequencing

Extraction of genomic DNA and subsequent quantity and quality assessment were conducted for all 123 samples, as described in Mynhardt et al. (in submission). A total of 18 high-quality samples were selected for library preparation (Table 1, Fig. 1). These high-quality samples were sent to the Beijing Genomics Institute (BGI, Shenzhen, China) to undergo the ddRAD sequencing protocol as per Peterson et al. (2012). Library preparation, restriction enzyme digestion, sequencing and initial quality filtering were conducted as described in Mynhardt et al. (in submission).

## 2.3. Bioinformatics and SNP discovery

CLC GENOMICS WORKBENCH 11.0 was used to evaluate read quality and to map reads to the closest reference genome, the Cape golden mole (*Chrysochloris asiatica*, GCA\_000296735.1). Reads were also separately mapped to the *C. asiatica* complete (GCA\_000296735.1) and mitochondrial (NC\_004920.1) genomes using Geneious v. 11.1.4 (Kearse et al., 2012). See Mynhardt et al. (in submission) for more details.

The STACKS v. 1.44 (Catchen et al., 2013) pipeline was used, along with custom bioinformatics expressions, to clean, cluster and catalogue reads, and also for assembly of paired reads and SNP calling. The STACKS script *denovo\_map.pl* was used for SNP identification (Catchen et al., 2013), and optimal parameters were determined by running the script multiple times and recording the SNP yield whilst varying the core parameters. See Mynhardt et al. (in submission) for more details on parameter testing.

The optimised parameters ( $-m\ 3\ -M\ 3\ -n\ 1$  and 50% sample representation cutoff) were used to build a catalogue. We manually filtered out singleton allele SNPs in Microsoft Excel (v. 16.27), after finding that inclusion of singletons obscured the underlying signal of population structure (Mynhardt et al., in submission). A recent study investigating the effects of minor allele frequency (MAF) thresholds on inference of populations structure also reported the same finding (Linck and Battey, 2019). We similarly excluded SNPs with heterozygosity values of 0 or 1, in order to filter out both mitochondrial loci ( $H_E = 0$ ) and loci with heterozygote excess ( $H_E = 1$ ) that could mask population structure (Roesti et al., 2012, Rodríguez-Ezpeleta et al., 2016). We retained only the first SNP at each locus, since some downstream analyses require only one SNP per locus.

**Table 1**

Sample information for all Hottentot golden mole RAD sequences generated in this study.

Subspecies	Locality (name of town/reserve, province)	GPS co-ordinates	Sample names
<i>A. h. meesteri</i>	Sabie, MP <sup>3</sup>	25°06' S 30°47' E	AHM_SA1, 2
<i>A. h. longiceps</i>	Cedarville, EC <sup>1</sup>	30°19' S 29°02' E	AHL_CV1
	Drakensberg Gardens, KZN <sup>2</sup>	29°45' S 29°13' E	AHL_DG1
	Sani Pass, KZN	29°39' S 29°26' E	AHL_SP1, 3
<i>A. h. pondoliae</i>	Margate, KZN	30°50' S 30°21' E	AHP_MA1
	Port Edward, KZN	31°03' S 30°13' E	AHP_PE1, 2, 3, 6
	San Lameer, KZN	30°56' S 30°18' E	AHP_SL26
	Umtamvuna, KZN	31°00' S 30°09' E	AHP_UV1
	Umtata, EC	31°35' S 28°46' E	AHP_UM3
<i>A. h. pondoliae</i> (Central Coast)	Umkomaas, KZN	30°12' S 30°47' E	AHP_UK1, 2
	Illovo, KZN	30°07' S 30°50' E	AHP_IL1
<i>A. h. iris</i> (Central Coast)	Umhlali, KZN	29°29' S 31°14' E	AHI_UI1

<sup>1</sup> EC - Eastern Cape.<sup>2</sup> KZN – KwaZulu-Natal.<sup>3</sup> MP – Mpumalanga.



## 2.4. Analysis of taxon and population differentiation

Phylogenetic relationships were estimated using the SNAPP template implemented in BEAST v2.5.1 (Bouckaert et al., 2014). We used the default model parameters in BEAUti for U and V equal to one. We assigned a Gamma distribution to our Lambda prior, with an Alpha of 2 and a Beta of 200. On the Snap prior we assigned an Alpha of 1, a Beta of 250, and a Lambda of 0.01. We ran the analyses for 1 000 000 MCMC generations, sampling every 1000 generations.

We visualized the complete tree sets in DENSITREE v2.2.6 (Bouckaert and Heled, 2014), removing the first 10% of trees as burn-in. We used TREEANNOTATOR v2.5.1 to generate a maximum clade credibility tree, also with 10% burn-in, and visualized the resulting tree in FIGTREE v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree>).

Putative loci under selection were identified by detection of  $F_{ST}$  outliers using BAYESCAN 2.1 (Foll and Gaggiotti, 2008). Three *a priori* populations were assigned: (1) all *A. h. longiceps* individuals (CV1, DG1, SP1 and SP3), (2) all *A. h. pondoliae* and *A. h. iris* (including all Central Coastal samples) and the one sample from Umtata (IL1, MA1, PE1, PE2, PE3, PE6, SL26, UK1, UK2, UI3, UV1 and UM3), and (3) the two *A. meesteri* individuals from Sabie (SA1 and SA2). Using the SNPs retained after bioinformatics filtering, the analysis was based on 20 pilot runs each consisting of 5000 iterations, followed by 100 000 iterations with a burn-in of 50 000 iterations. An R function, provided along with the Bayescan software package, was used to plot and identify outliers using different criteria from the Bayescan output file.

Population differentiation was assessed using STRUCTURE v. 2.3 (Pritchard et al., 2000). All 805 SNP loci that passed the filtering process were used to assess structuring of the 18 individuals. For all STRUCTURE runs, the admixture model was used with allele frequencies correlated among populations, 500 000 MCMC steps after a burnin of 50 000 steps for  $K = 1$  to  $K = 12$ , with 20 iterations at each  $K$ . The results were run through STRUCTURE HARVESTER (Earl and vonHoldt, 2012) in order to reformat files for use in downstream programs. The program CLUMPP (CLUster Matching and Permutation Program) v.1.1.2 (Jakobsson and Rosenberg, 2007) was used to permute and optimise the results from all replicate analyses, and the cluster visualisation program DISTRUCT (Rosenberg, 2004) was subsequently used for the graphical display of the output.

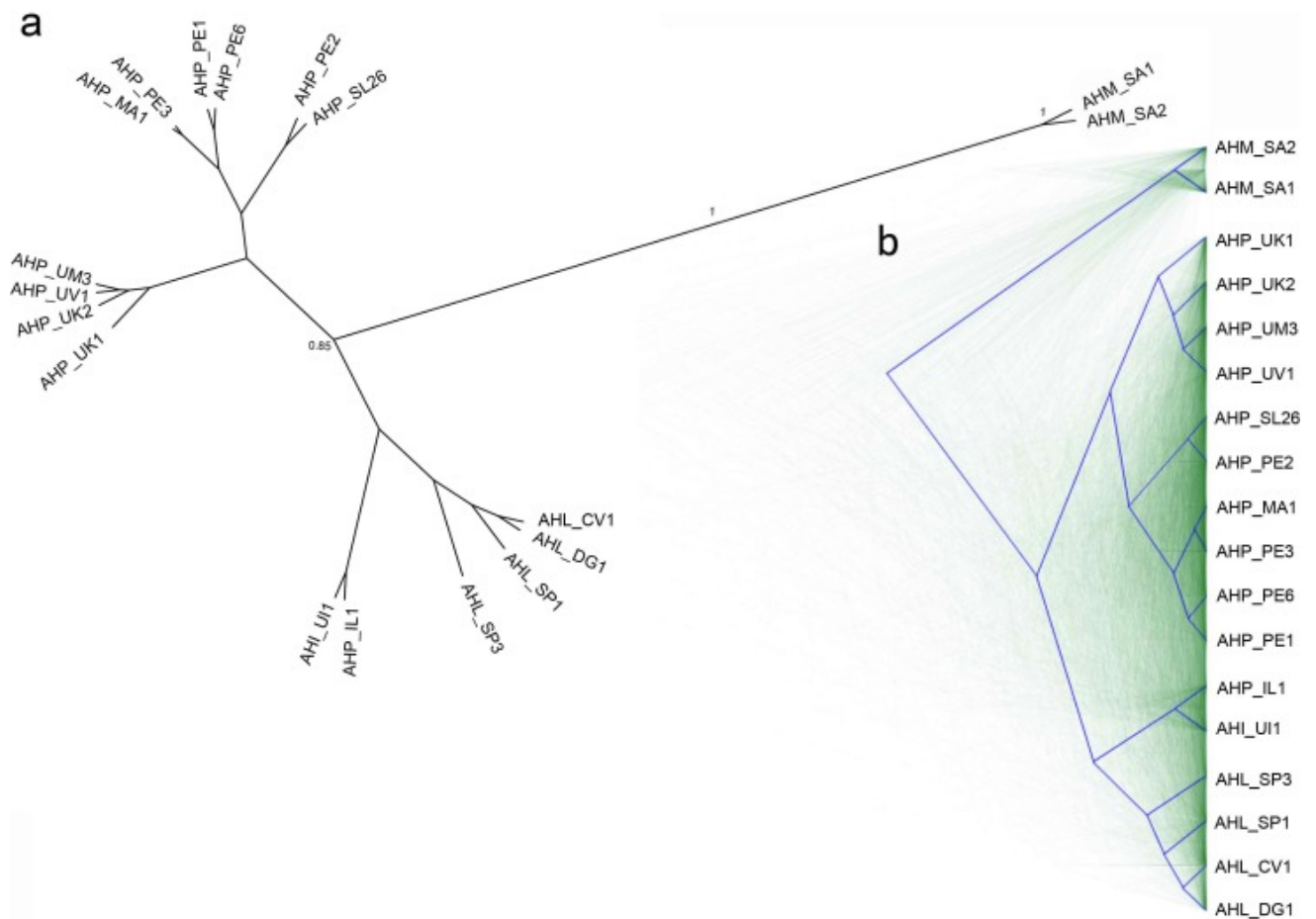
As an independent assessment of population differentiation we used the FINER-ADSTRUCTURE software package (Malinsky et al., 2018) to construct a co-ancestry matrix from the RADseq data. We used a 1 000 000 burn-in followed by 1 000 000 MCMC steps, sampling every 10 000 steps, and the tree was constructed with 10 000 hill-climbing iterations. FINESTRUCTURE GUI (Lawson et al., 2012) was used to visualize the results and conduct principal components analysis (PCA), by eigenanalysis of a normalised version of the co-ancestry matrix.

## 3. Results

### 3.1. Bioinformatics and SNP discovery

Millions of high quality reads were produced for the 18 individuals sequenced (Mynhardt et al. in submission). Mapping reads against the Cape golden mole (*Chrysochloris asiatica*)

reference genome resulted in an average genome coverage of 11.28% per individual and a total genome coverage of 51%. On average 58.94% of all reads were mapped to the *C. asiatica* nuclear genome, and 0.25% to the mitogenome, with a reference mitogenome coverage of 72.88% per individual and 97.6% overall. Initial processing of the data (involving cleaning, discarding unpaired reads, and filtering out reads with adapter contamination) identified over 118 million high-quality paired reads, with an average of 6.6 million reads per individual. The optimised STACKS parameters ( $-m\ 3\ -M\ 3\ -n\ 1$ ) resulted in a total of 46 479 SNPs and when the additional stringent parameters were applied (50% sample representation cutoff, filtering out singletons, and retaining only the first SNP at any locus), 805 SNPs remained. See Mynhardt et al. (in submission) and associated tables and figures for more details.



**Fig. 2.** Phylogenetic estimate for the full dataset, based on 805 SNPs. (a) Maximum clade credibility tree indicating posterior probabilities for well supported nodes only. (b) Densitree indicating all 1000 trees (green) and maximum clade credibility tree (blue).

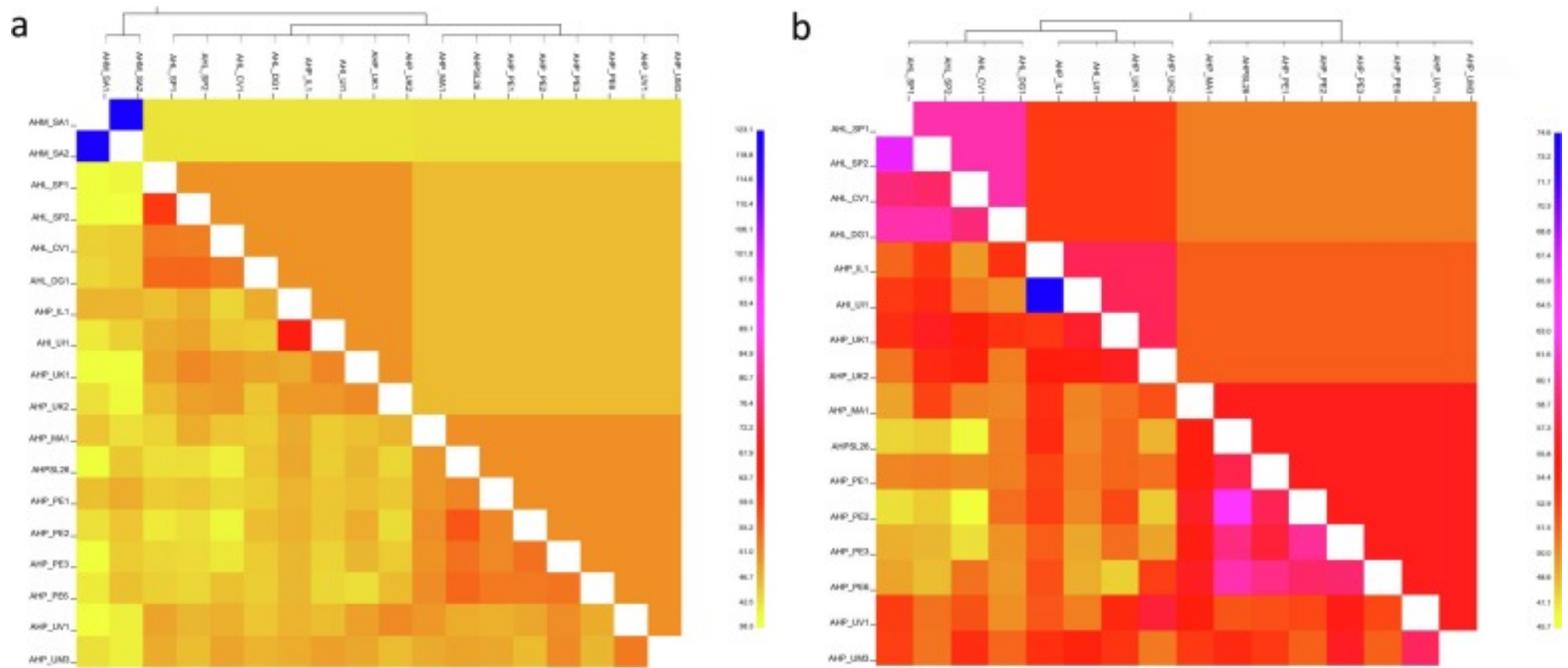
### 3.2. Analysis of taxon and population differentiation

Phylogenetic analyses indicate that two highly divergent lineages are present, representing (1) *A. h. meesteri* and (2) all other *A. hottentotus* (Fig. 2). Further differentiation exists within *A. hottentotus*, essentially splitting *A. h. pondoliae* (Port Edward, Margate, San Lameer,

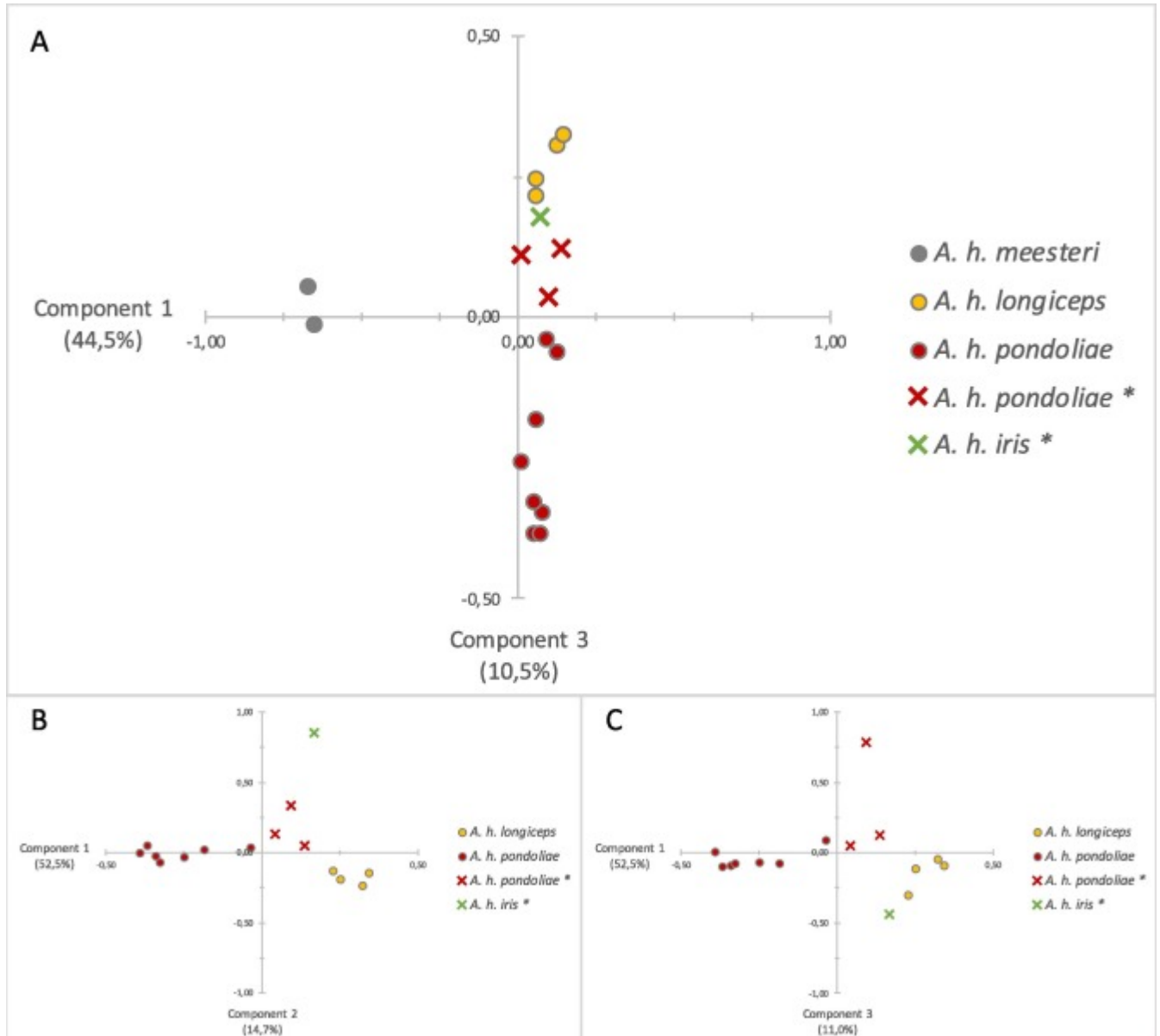
Umtamvuna and Umtata) and *A. h. longiceps* (Cedarville, Drakensberg Gardens, Sani Pass). Two samples from the Central coastal clade (*A. h. pondoliae* from Illovo and *A. h. iris* from Umhlali) cluster with *A. h. longiceps*, while the other two Central coastal samples (*A. h. pondoliae* from Umkomaas) cluster with *A. h. pondoliae*.

STRUCTURE analyses corroborate the two highly divergent lineages at  $K=2$  (mean assignment across 20 replicates:  $0.998 \pm 0.003$ ; Fig. 1, Table S1). According to delta  $K$  estimation, the number of clusters that best fit the data is  $K=4$  (Fig. S1): (1) *A. h. meesteri* (Sabie), (2) *A. h. longiceps* (Cedarville, Drakensberg Gardens, Sani Pass), (3) *A. h. pondoliae* (Margate, Port Edward, San Lameer, Umtamvuna, Umtata), and (4) Central coastal clade (*A. h. pondoliae* from Umkomaas and Illovo, and *A. h. iris* from Umhlali). The mean assignment of all samples at  $K=4$  across 20 replicates is 0.938 (Table S1). It is also worth noting that at  $K=5$  an additional fifth cluster is evident, with high assignment for Umtata (mean assignment across 20 replicates: 0.971) and also partial assignment for Illovo (mean assignment: 0.372), although Illovo is still more strongly assigned to the Central coastal clade (mean assignment: 0.523; Table S1).

FINERADSTRUCTURE and principal components analysis (PCA) showed, once again, that *A. h. meesteri* was highly divergent from the remainder of *A. hottentotus*, and can be distinguished at the highest level of population structure (Figs. 3a; 4a). The co-ancestry matrix (Fig. 3a) shows that the two *A. h. meesteri* samples have a high level of shared co-ancestry with each other, but low co-ancestry with all other samples. The associated tree shows that these samples represent a lineage that is highly divergent from the remaining samples. A further two lineages can be distinguished, whereby the Central coastal samples (Ahp\_IL, Ahi\_UI, Ahp\_UK1 and Ahp\_UK2) cluster along with *A. h. longiceps* to the exclusion of all other *A. h. pondoliae* (Fig. 3a). In the associated PCA, component 1 (accounting for 44.5% of the eigenvalues) separates *A. h. meesteri* from the rest. Component 2 (accounting for 34.4% of the eigenvalues) describes variation within *A. h. meesteri*, and is unable to distinguish clusters due to low variation at the relevant loci among the remaining samples (data not shown). Component 3 (accounting for 10.5% of the eigenvalues) separates *A. h. longiceps*, *A. h. pondoliae*, and the Central coastal clade (*A. h. pondoliae* from Umkomaas and Illovo, and *A. h. iris* from Umhlali) from one another. Umtata clusters closely to Umtamvuna and is thus not clearly separated from *A. h. pondoliae* (Fig. 4a).



**Fig. 3.** FINERADSTRUCTURE results for (a) the full dataset, and (b) the reduced dataset (excluding the two *A. h. meesteri* individuals from Sabie). The co-ancestry matrix indicates pairwise co-ancestry between individuals. Individual pairwise values are shown below the diagonal, and mean pairwise values for populations are shown above the diagonal. High values, indicating high probability of shared co-ancestry are clustered along the diagonal, and these individuals are clustered into populations in the accompanying tree.



**Fig. 4.** FINERADSTRUCTURE results for (a) the full dataset, and (b and c) the reduced dataset (excluding the two *A. h. meesteri* individuals from Sabie). The principal component analyses (PCA) indicate the distribution of individuals into populations according to the first three principal components identified in FINERADSTRUCTURE. \* Individuals belonging to the Central coastal clade are denoted with an "X".

Since it is possible for the highly divergent *A. h. meesteri* to mask the signal for the remaining structure in the dataset, an additional FINERADSTRUCTURE and PCA analysis was performed to the exclusion of these two samples. In this analysis, three major lineages were identified: (1) *A. h. pondoliae*, (2) *A. h. longiceps* and (3) the Central coastal clade (Fig. 3b). In the associated PCA (Fig. 4b and c), component 1 (accounting for 52.5% of the eigenvalues) separates *A. h. pondoliae*, *A. h. longiceps* and the Central coastal clade from one another, while component 2 (accounting for 14.7% of the eigenvalues) further distinguishes *A. h. longiceps* from the rest (Fig. 4b).

## 4. Discussion

Our results conclusively support the distinctiveness of *A. h. meesteri*. Our phylogenetic analysis indicated strong support for the divergence of this lineage from the remainder of *A. hottentotus*. The STRUCTURE analysis identified four major clusters in the data, with the first division (at  $K = 2$ ) separating *A. h. meesteri* from the remaining *A. hottentotus*. The two *A. h. meesteri* individuals from Sabie were consistently assigned to a unique cluster (mean assignment at  $K = 2$  across 20 replicates:  $0.998 \pm 0.003$ ), indicating a high likelihood that they represent a unique population, genetically isolated from other *A. hottentotus* populations. Our FINERADSTRUCTURE analysis supports the distinctiveness of *A. h. meesteri* by identifying it as a unique population at the highest level of structure in the co-ancestry matrix and in the supporting PCA plots (Figs. 3a; 4a). These results support our previous findings, based on mtDNA and limited nuclear intron data, that *A. h. meesteri* is a monophyletic lineage that diverged from the remainder of its extant genus, along with sister species *A. marleyi* from the Lebombo mountains just southeast of Swaziland, approximately 4.42 million years ago (Ma), and later diverged from *A. marleyi*, approximately 1.9 Ma (Mynhardt et al., 2015). Inclusion of RADseq data, or other available nuclear SNP data in future divergence dating analyses, ideally also incorporating *A. marleyi*, which was not included in the present study, would be useful in further testing this hypothesis. *Amblysomus h. meesteri* is a geographically isolated population that also differs morphologically (Bronner, 1996a) and cytogenetically (Gilbert et al., 2008) from other *A. hottentotus*, and may be worthy of specific status, as previously recommended (Mynhardt et al., 2015).

STRUCTURE analyses identified four major clusters corresponding to: (1) *A. h. meesteri* (Sabie), (2) *A. h. longiceps* (Cedarville, Drakensberg Gardens, Sani Pass), (3) *A. h. pondoliae* (Margate, Port Edward, San Lameer, Umtamvuna, Umtata), (4) Central coastal clade (Illovo, Umkomaas and Umhlali). The mean assignment at  $K = 4$  across 20 replicates was 0.938. FINERADSTRUCTURE grouped *A. h. longiceps* and the Central coastal clade together based on the full dataset (Fig. 3a), although substructure between these lineages can be seen in the associated PCA (Fig. 4a). These lineages were separated when analysed to the exclusion of *A. h. meesteri* (Figs. 3b; 4b and c). This is probably due to the strong signal of divergence between *A. h. meesteri* and the remainder of *A. hottentotus* masking the signal in the rest of the full dataset. Two Central coastal samples (*A. h. iris* from Illovo and *A. h. pondoliae* from Umhlali) also cluster alongside *A. h. longiceps* in the phylogenetic analysis (Fig. 2), but without strong support for this relationship. The grouping of the Central coastal clade with *A. h. longiceps* (Fig. 3), supports our previous findings that samples from this region are more closely related to *A. h. longiceps* than they are to the geographically proximate *A. h. pondoliae*. This study showed that the Central coastal mitochondrial haplogroup is separated from *A. h. longiceps* by fewer mutational steps than from *A. h. pondoliae* (Mynhardt et al., 2015).

Individuals from the Central coastal region were fairly confidently assigned to a unique cluster at  $K = 4$  (mean assignment across 20 replicates: 0.796), however there is a possibility of some admixture in this population. Inclusion of true *A. h. iris* representatives in future structure analyses may help to resolve uncertainties, although it is clear that these individuals are different to *A. h. pondoliae* from the southern coastal localities. The FINERADSTRUCTURE analysis revealed that Illovo and Umhlali share a high level of co-ancestry, and in the reduced dataset these individuals cluster together with Umkomaas (Fig. 3b). Taken together, these results support the presence of an additional, although potentially admixed coastal lineage. These individuals represent the recently discovered cryptic Central

coastal clade, which is thought to have diverged from a shared common ancestor with *A. h. longiceps* and *A. h. pondoliae* approximately 2.53 Ma (Mynhardt et al., 2015). The current subspecific classification of *A. h. longiceps* and *A. h. pondoliae* was based on phenetics, however, these two subspecies were only distinguished from one another based on overall body size, and could not be distinguished on other craniometric characters (Bronner, 1996a). *A. h. pondoliae* was distinguished from other coastal samples farther north (*A. h. iris*) and south (*A. h. hottentotus*) on the basis of colourimetric and craniodental variation respectively (Bronner, 1996a, Bronner, 1996b), but samples from the northern and southern extents of the coastal *A. h. pondoliae* distribution were not distinguished morphometrically. Our findings therefore provide evidence from nuclear SNPs that Central coastal samples represent a unique cryptic lineage that is highly divergent from *A. h. pondoliae* farther south. We have already demonstrated that this lineage is divergent from northern coastal *A. h. iris* (Mynhardt et al., 2015). This cryptic lineage may be worthy of specific or subspecific status, as previously recommended (Mynhardt et al., 2015).

We did not find strong support for the uniqueness of Umtata. Although the STRUCTURE analysis provided some support for an additional cluster to which Umtata is strongly assigned (mean assignment at  $K = 5$  across 20 replicates: 0.971), FINERADSTRUCTURE could not provide further support since this population is represented only by a single individual, and FINERADSTRUCTURE clusters samples together on the basis of shared co-ancestry, thereby grouping Umtata with its nearest relatives, the southern *A. h. pondoliae* samples from Margate, Port Edward, San Lameer and Umtamvuna. This is consistent with STRUCTURE clustering at  $K = 4$  (Fig. 1). Taken together, our results support the placement of Umtata within *A. h. pondoliae*, as currently classified based on predicted distribution of the subspecies. Our previous phylogenetic study, based on mitochondrial NADH dehydrogenase 2 (*MT-ND2*), cytochrome *b* (*cyt b*) and a nuclear intron (GHR, growth hormone receptor, intron 9), placed Umtata as a unique lineage sister to *A. h. longiceps*, but with low support (bootstrap 64, posterior probability 0.90; Mynhardt et al., 2015), of which all the support came from *MT-ND2* (bootstrap and posterior probabilities respectively for *MT-ND2*: 82 and 0.81; *cyt b*: 33 and 0.46; GHR: not supported). The underlying reason for this mito-nuclear discordance is yet to be tested, but introgression could be a plausible hypothesis. Mito-nuclear discordance is quite common among mammals, and is often attributed to mitochondrial introgression, for example in voles (Boratyński et al., 2011, Bastos-Silveira et al., 2012), in chipmunks (Sarver et al., 2016) and in hares (Alves et al., 2008, Boratyński et al., 2011, Kinoshita et al., 2019). Sex-biased dispersal, such as that observed in the slit-faced bats, *Nycteris thebaica* (Demos et al., 2019), could also explain this discordance, although there is no evidence for sex-biased dispersal in *A. hottentotus*. Inclusion of *A. h. pondoliae* and *A. h. hottentotus* representatives from farther south along the coast in future analyses, as well as hypothesis testing may help to resolve this. Furthermore, additional sampling from Umtata and the surrounding area will be needed to enable more robust analyses to be conducted for this cryptic lineage, in order to resolve its evolutionary history.

## 5. Conclusion

In this study we investigated differentiation in *A. hottentotus* (*sensu lato*) by means of double digest restriction-site associated DNA sequencing (ddRADseq). We present a high-quality filtered SNP dataset, comprising thousands of SNPs, which may serve as a useful resource for future golden mole studies. We have added to the growing body of research demonstrating the power and utility of RADseq to investigate population differentiation (Benestan et al., 2015, Szulkin et al., 2016, Vendrami et al., 2017). The full set of high-quality SNPs can

easily be re-filtered to obtain different sets of SNPs to the 805 filtered SNPs retained for downstream analysis in this study, in order to address more specific questions pertaining to population structure in *A. hottentotus*. For example, in order to more thoroughly investigate population structure within a particular (sub)species, the data can be filtered specifically for SNPs that are polymorphic within that (sub)species. Mapped sequences can be used to facilitate downstream assay design, potentially also for application in other golden mole species.

This study has provided evidence in support of previous findings, as well as some additional insights. We provide support for the recognition of *A. h. meesteri* as a separate species to *A. hottentotus* (Bronner, 1996a, Gilbert et al., 2008, Mynhardt et al., 2015). We also provide support for the recognition of a fourth coastal *Amblysomus* lineage (Mynhardt et al., 2015). Contrary to our previous finding that Umtata may represent a unique lineage sister to *A. h. longiceps* (Mynhardt et al., 2015), based on whole-genome representative RADseq SNP data, these samples appear to be more closely related to *A. h. pondoliae*, and may not represent a distinct lineage. This mito-nuclear discordance has raised further questions about the evolutionary history of these individuals, which emphasizes the need for additional sampling from the area.

Golden moles represent a family of highly threatened small mammals, with ten of its 21 species listed in threatened IUCN Red List categories (The IUCN Red, 2017). Given the dearth of information available for most species, uncertainties surrounding the taxonomy, and recent discovery of cryptic diversity within *A. hottentotus*, it is important to recognize that understudied populations, such as that of Umtata, may represent populations or ESUs under threat and in need of conservation attention. Therefore, in expanding our dataset for future studies, it will be of utmost importance to consider innovative approaches to sampling these potentially threatened golden mole populations, in order to address some of the more specific questions at hand. We are confident that the genomic resources provided by this study have, and will continue to, assist us and others in gaining a more thorough understanding of the evolutionary history and consequent genetic differentiation among populations of golden moles in South Africa.

## Acknowledgements

We thank S Maree, GN Bronner, I Pelser and JW Wilson for their contributions towards field trips and the acquisition of samples, and S Maree for assistance with sample preparation. We also acknowledge numerous insightful discussions on golden moles with S Maree and GN Bronner over years of collaboration. We thank B Hansson, M Tarka and CS Stobie for their contributions toward the implementation of the Stacks pipeline in our study.

This work was funded in part by the Genomics Research Institute of the University of Pretoria and the National Research Foundation (NRF) South African Research Chair Initiative (SARChI) Chair of Behavioural Ecology and Physiology, grant GUN 64756. The opinions, findings and conclusions expressed in this publication are those of the authors and the NRF accepts no responsibility in this regard.



## References

- Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., Hohenlohe, P.A., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* 17, 81–92.
- Alves, P.C., Melo-Ferreira, J., Freitas, H., Boursot, P., 2008. The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*. *Philosoph. Trans. Royal Soc. B: Biol. Sci.* 363, 2831–2839.
- Asher, R.J., Bennett, N., Lehmann, T., 2009. The new framework for understanding placental mammal evolution. *BioEssays* 31, 853–864.
- Asher, R.J., Maree, S., Bronner, G., Bennett, N.C., Bloomer, P., Czechowski, P., Meyer, M., Hofreiter, M., 2010. A phylogenetic estimate for golden moles (Mammalia, Afrotheria, Chrysochloridae). *BMC Evol. Biol.* 10 (1), 69.
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3, e3376.
- Bastos-Silveira, C., Santos, S.M., Monarca, R., Mathias, M.D.L., Heckel, G., 2012. Deep mitochondrial introgression and hybridization among ecologically divergent vole species. *Mol. Ecol.* 21, 5309–5323.
- Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., Bernatchez, L., 2015. RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Mol. Ecol.* 24, 3299–3315.
- Boratyński, Z., Alves, P.C., Berto, S., Koskela, E., Mappes, T., Melo-Ferreira, J., 2011. Introgression of mitochondrial DNA among *Myodes* voles: consequences for energetics? *BMC Evol. Biol.* 11, 355.
- Bouckaert, R., Heled, J. 2014. DensiTree 2: Seeing trees through the forest. *BioRxiv*, 012401.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10, 1–6.
- Bronner, G.N., 1991. Comparative hyoid morphology of nine chrysochlorid species (Mammalia: Chrysochloridae). *Ann. Transv. Mus.* 35, 295–311.
- Bronner, G.N., 1995. Cytogenetic Properties of Nine Species of Golden Moles (Insectivora: Chrysochloridae). *J. Mammal.* 76, 957–971.
- Bronner, G.N., 1996a. Geographic patterns of morphometric variation in the Hottentot golden mole, *Amblysomus hottentotus* (Insectivora: Chrysochloridae) A multivariate analysis. *Mamm* 60, 729–752.
- Bronner, G.N., 1996b. Non-geographic variation in morphological characteristics of the Hottentot golden mole, *Amblysomus hottentotus* (Insectivora : Chrysochloridae). *Mamm.* 60, 707–728.
- Bronner, G.N., 2000. New species and subspecies of golden mole (Chrysochloridae: *Amblysomus*) from Mpumalanga, South Africa. *Mamm.* 64, 41–54.

- Broom, R., 1907. A Contribution to the knowledge of the cape golden moles. *Trans. S. Afr. Phil. Soc.* 18, 283–311.
- Busch, C., Antinuchi, C., del Valle, J., Kittlein, M., Malizia, A., Vassallo, A., Zenuto, R., 2000. Population Ecology of Subterranean Rodents. In: *Life Underground: The Biology of Subterranean Rodents*. University of Chicago Press, Chicago, Illinois, pp. 183–286.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140.
- Catchen, J.M., Hohenlohe, P.A., Bernatchez, L., Funk, W.C., Andrews, K.R., Allendorf, F.W., 2017. Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Mol. Ecol. Resour.* 17, 362–365.
- Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510.
- Demos, T.C., Kerbis Peterhans, J.C., Joseph, T.A., Robinson, J.D., Agwanda, B., Hickerson, M.J., 2015. Comparative population genomics of African montane forest mammals support population persistence across a climatic gradient and Quaternary climatic cycles. *PLoS ONE* 10, e0131800.
- Demos, T.C., Webala, P.W., Kerbis Peterhans, J.C., Goodman, S.M., Bartonjo, M., Patterson, B.D., 2019. Molecular phylogenetics of slit-faced bats (Chiroptera: Nycteridae) reveal deeply divergent African lineages. *J. Zool. System. Evolut. Res.*
- Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361.
- Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977–993.
- Forcart, L., 1942. Beiträge zur Kenntnis der Insectivoren familie Chrysochloridae. *Rev. Suisse Zool.* 49, 1–7.
- Gilbert, C., Maree, S., Robinson, T.J., 2008. Chromosomal evolution and distribution of telomeric repeats in golden moles (Chrysochloridae, Mammalia). *Cytogenet. Genome Res.* 121, 110–119.
- Hickman, G.C., 1979. A live-trap and trapping technique for fossorial mammals. *S. Afr. J. Zool.* 14, 9–12.
- Humble, E., Dasmahapatra, K.K., Martinez-Barrio, A., Gregório, I., Forcada, J., Polikeit, A.C., Goldsworthy, S.D., Goebel, M.E., Kalinowski, J., Wolf, J.B., Hoffman, J.I., 2018. RAD sequencing and a hybrid Antarctic fur seal genome assembly reveal rapidly decaying linkage disequilibrium, global population structure and evidence for inbreeding. *G3 (Bethesda)* 8, 2709–82722.
- The IUCN Red List of Threatened Species Version 2017-3 [Internet]. 2017 [Available from: <http://www.iucnredlist.org/>].
- Jackson, C.R., Robertson, M.P., 2011. Predicting the potential distribution of an endangered cryptic subterranean mammal from few occurrence records. *J. Nat. Conserv.* 19 (2), 87–94.

- Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649.
- Kinoshita, G., Nunome, M., Kryukov, A.P., Kartavtseva, I.V., Han, S.-H., Yamada, F., Suzuki, H., 2019. Contrasting phylogeographic histories between the continent and islands of East Asia: massive mitochondrial introgression and long-term isolation of hares (*Lagomorpha: Lepus*). *Mol. Phylogenet. Evol.* 136, 65–75.
- Knowles, L.L., Massatti, R., He, Q., Olson, L.E., Lanier, H.C., 2016. Quantifying the similarity between genes and geography across Alaska's alpine small mammals. *J. Biogeogr.* 43, 1464–1476.
- Krohn, A.R., Conroy, C.J., Pesapane, R., Bi, K., Foley, J.E., Rosenblum, E.B., 2018. Conservation genomics of desert dwelling California voles (*Microtus californicus*) and implications for management of endangered Amargosa voles (*Microtus californicus scirpensis*). *Conserv. Genet.* 19, 383–395.
- Kuyper, M.A., 1985. The ecology of the golden mole *Amblysomus hottentotus*. *Mamm. Rev.* 15, 3–11.
- Lah, L., Trense, D., Benke, H., Berggren, P., Gunnlaugsson, P., Lockyer, C., Öztürk, A., Öztürk, B., Pawliczka, I., Roos, A., 2016. Spatially explicit analysis of genome-wide SNPs detects subtle population structure in a mobile marine mammal, the harbour porpoise. *PLoS ONE* 11, e0162792.
- Lanier, H.C., Massatti, R., He, Q., Olson, L.E., Knowles, L.L., 2015. Colonization from divergent ancestors: glaciation signatures on contemporary patterns of genomic variation in Collared Pikas (*Ochotona collaris*). *Mol. Ecol.* 24, 3688–3705.
- Lawson, D.J., Hellenthal, G., Myers, S., Falush, D., 2012. Inference of population structure using dense haplotype data. *PLoS Genet.* 8, e1002453.
- Linck, E., Battey, C.J., 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol. Ecol. Resour.* 19, 639–647.
- Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M.F., Storfer, A., 2017. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Mol. Ecol. Resour.* 17 (2), 142–152.
- Malinsky, M., Trucchi, E., Lawson, D.J., Falush, D., 2018. RADpainter and fineRAD structure: population inference from RADseq data. *Mol. Biol. Evol.* 35, 1284–1290.
- McKinney, G.J., Larson, W.A., Seeb, L.W., Seeb, J.E., 2017. RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al.(2016). *Mol. Ecol. Resour.* 17, 356–361.
- Meester, J.A.J., 1974. Family Chrysochloridae. In: Meester, J., Setzer, H.W. (Eds.), *The mammals of Africa: an identification manual*. Smithsonian Institution Press, Washington D.C., pp. 1–7.

- Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17, 240–248.
- Mynhardt, S., Bennett, N.C., Bloomer, P. In submission. Bioinformatics and SNP discovery in golden mole ddRADseq data.
- Mynhardt, S., Maree, S., Pelser, I., Bennett, N.C., Bronner, G.N., Wilson, J.W., Bloomer, P., 2015. Phylogeography of a morphologically cryptic golden mole assemblage from South- Eastern Africa. *PLoS ONE* 10, e0144995.
- Nevo, E., 1979. Adaptive convergence and divergence of subterranean mammals. *Annu. Rev. Ecol. Syst.* 10, 269–308.
- Perera, S.J., Procheş, Ş., Ratnayake-Perera, D., Ramdhani, S., 2018. Vertebrate endemism in south-eastern Africa numerically redefines a biodiversity hotspot. *Zootaxa* 4382, 56–92.
- Perera, S.J., Ratnayake-Perera, D., Proches, S., 2011. Vertebrate distributions indicate a greater Maputaland-Pondoland-Albany region of endemism. *S. Afr. J. Sci.* 107, 1–15.
- Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7, e37135.
- Petter, F., 1981. Remarques sur le systematique des chrysochlorides. *Mamm.* 49–53.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Rodríguez-Ezpeleta, N., Bradbury, I.R., Mendibil, I., Álvarez, P., Cotano, U., Irigoien, X., 2016. Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection. *Mol. Ecol. Resour.* 16, 991–1001.
- Roesti, M., Salzburger, W., Berner, D., 2012. Uninformative polymorphisms bias genome scans for signatures of selection. *BMC Evol. Biol.* 12, 94.
- Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138.
- Sarver, B.A., Demboski, J.R., Good, J.M., Forshee, N., Hunter, S.S., Sullivan, J., 2016. Comparative phylogenomic assessment of mitochondrial introgression among several species of chipmunks (*Tamias*). *Geno. Boil. Evolut.* 9, 7–19.
- Seiffert, E., 2007. A new estimate of afrotherian phylogeny based on simultaneous analysis of genomic, morphological, and fossil evidence. *BMC Evol. Biol.* 7, 224.
- Simonetta, A., 1968. A new golden mole from Somalia with an appendix on the taxonomy of the family Chrysochloridae (Mammalia, Insectivora). *Monit. Zool. Ital.* 2, 27–55.
- Sovic, M.G., Carstens, B.C., Gibbs, H.L., 2016. Genetic diversity in migratory bats: Results from RADseq data for three tree bat species at an Ohio windfarm. *PeerJ* 4, e1647.

Stanhope, M.J., Waddell, V.G., Madsen, O., de Jong, W., Hedges, S.B., Cleven, G.C., Kao, D., Springer, M.S., 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Natl. Acad. Sci. USA* 95, 9967–9972.

Svengren, H., Prettejohn, M., Bunge, D., Fundi, P., Björklund, M., 2017. Relatedness and genetic variation in wild and captive populations of Mountain Bongo in Kenya obtained from genome-wide single-nucleotide polymorphism (SNP) data. *Glob. Ecol. Conserv.* 11, 196–206.

Szulkin, M., Gagnaire, P.A., Bierne, N., Charmantier, A., 2016. Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. *Mol. Ecol.* 25, 542–558.

Taylor, W.A., Mynhardt, S., Maree, S., 2018. Family Chrysochloridae (Golden Moles). In: Wilson, D.E., Mittermeier, R.A. (Eds.), *Handbook of the Mammals of the World – Volume 8 Insectivores, Sloths and Colugos*. Lynx Edicions, Barcelona, pp. 180–203.

Thomas, O., Schwann, H., 1905. The Rudd exploration of South Africa.—III. List of the mammals obtained by Mr. Grant in Zululand. *Proc. Zool. Soc. Lond.* 254–276.

Vendrami, D.L.J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., Clark, M.S., Leese, F., McMin, C., Moore, H., et al., 2017. RAD sequencing resolves finescale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. *R. Soc. Open Sci.* 4, 160548.

Wang, W., Yan, H., Yu, J., Yi, J., Qu, Y., Fu, M., Chen, A., Tang, H., Niu, L., 2017. Discovery of genome-wide SNPs by RAD-seq and the genetic diversity of captive hog deer (*Axis porcinus*). *PLoS ONE* 12, e0174299.