The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2011, **104**, 192–206. With 3 figures

Cryptic speciation in the southern African vlei rat *Otomys irroratus* complex: evidence derived from mitochondrial cyt *b* and niche modelling

ADRIAAN ENGELBRECHT[1], PETER JOHN TAYLOR[2], SAVEL REGAN DANIELS[1] and RAMUGONDO VICTOR RAMBAU[1,3]*

[1]*Evolutionary Genomics Group, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa*
[2]*Department of Ecology & Resource Management, School of Environmental Sciences, University of Venda, P/Bag X5050, Thohoyandou 0950, South Africa*
[3]*DST-NRF Centre of Excellence for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602 South Africa*

Received 26 October 2010; revised 23 March 2011; accepted for publication 23 March 2011

The vlei rat *Otomys irroratus* has a wide distribution in southern Africa with several datasets indicating the presence of two putative species (*O. irroratus* and *O. auratus*). In the present study we use mitochrondrial cyt *b* data (~950 bp) from 98 specimens (including museum material) collected throughout the range of the species to determine the geographical limits of the two recognized species, and we link this to niche modelling to validate these species. Phylogenetic analysis of the DNA sequence data, using maximum parsimony, neighbour joining and Bayesian inference, retrieved two divergent statistically well-supported clades. Clade A occurs in the Western and Eastern Cape while Clade B occurs in the Free State, KwaZulu-Natal, Northern Cape and Mpumalanga provinces of South Africa and Zimbabwe. Mean sequence divergence between the two clades (A and B) was 7.0% and between sub-clades comprising clade B it was 4.8%; the two clades diverged during the Pleistocene. Within Clade A the mean sequence divergence among specimens was 1.91%. Niche modelling revealed that the incipient species occupy distinct bioclimatic niches associated with seasonality of precipitation. Our data allow insightful analysis into the factors that could have led to cladogenesis within this rodent. More significantly, the new data enable us to pinpoint the Eastern Cape province as a contact zone for the divergent species. © 2011 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2011, **104**, 192–206.

ADDITIONAL KEYWORDS: contact zone – mtDNA – nested clade analysis – phylogeography.

INTRODUCTION

Phylogeographical approaches have been used successfully to uncover population genetic structure of numerous small mammal species, including rodents (Avise, 2000; Michaux *et al*., 2004; Vega *et al*., 2007; Yuasa *et al*., 2007; Mitsainas *et al*., 2009). These studies have revealed that species characterized by wide geographical distribution ranges often harbour multiple evolutionary units (Rambau, Robinson & Stanyon, 2003; Michaux *et al*., 2004; Willows-Munro & Matthee, 2009; Taylor *et al*., 2009a). The vlei rat *Otomys irroratus* is a mesic adapted grassland species that is widely distributed throughout southern Africa. In southern Africa it occurs along the southern and eastern coastal belt and adjacent interior into Lesotho, eastern Zimbabwe, Swaziland and western Mozambique (Skinner & Chimimba, 2005). Several studies have revealed a complex evolutionary history for this endemic southern African species (Contrafatto *et al*., 1992a; Rambau, Elder & Robinson, 2001; Taylor *et al*., 2009b).

In the first instance, *O. irroratus* has a variable chromosome number, $2n = 23–32$, comprising at least

*Corresponding author. E-mail: rvr2@sun.ac.za

four cytotypes which are defined by the presence or absence of a compound chromosome, centric fusions, pericentric inversions, B chromosomes, and the addition/deletion of heterochromatic short arms (Meester *et al.*, 1992; Contrafatto *et al.*, 1992a, c; Rambau *et al.*, 2001; Taylor *et al.*, 2009b). The various *O. irroratus* cytotypes appear to be distributed in different bioclimatic regions in South Africa (Taylor, Contrafatto & Willan, 1994). However, allozymes and immunochemical data revealed that populations characterized by various cytotypes were genetically invariant ($D = 0.03$; Taylor *et al.*, 1992; Contrafatto, Van Der Berg & Grace, 1997). The low genetic variation detected within *O. irroratus* at the protein level may be due to balancing selection or slow rate of protein evolution which could result in a lack of genetic differentiation among allopatric conspecific populations (Avise, 1998; Crochet, 2000; Piel & Nutt, 2000).

Breeding experiments in which specimens with a diploid number of $2n = 24$ (with a complex tandem fusion) were crossbred with specimens with $2n = 30$ (without the tandem fusion) resulted in high levels of antagonism or reduced viability in the progeny (Pillay, Willan & Meester, 1992; Pillay, Willan & Cooke, 1995). These results suggest both pre- and post-zygotic barriers may preclude gene flow between some populations, and allude to the presence of two evolutionary lineages within the taxon. This was confirmed by a recent analysis of several *O. irroratus* populations, using mtDNA cytochrome *b* (cyt *b*), skull morphometrics, and cytogenetics which retrieved two lineages (Taylor *et al.*, 2009b). The two evolutionary lineages (designated as *O. irroratus* and *O. auratus*) were demarcated by 6.4% sequence divergence, the presence or absence of pericentric inversions, and significant differentiation in cranial morphology. However, the detection of two lineages (one in the Western Cape and the other in the Eastern Cape and KwaZulu-Natal provinces) was inferred from limited samples ($N = 19$) and geographical coverage (17 localities) and consequently did not allow for a thorough description of the phylogeography and demographic history of the species. In this study, we build on these initial findings and we significantly increase the sample size (using fresh tissue samples and museum material) to 98, of which the majority were newly sequenced. We also improved the geographical coverage, which now allows insightful analysis into component species limits, demography and mode of speciation. Using the almost complete cyt *b* gene (950 bp) and niche modelling approach (MaxEnt program: Phillips, Anderson & Schapire, 2006), our aim was to determine the exact geographical limits of the putative species (lineages) and also to identify the contact zone of these species. Although our data retrieved two divergent mtDNA lineages separated by elevated sequence divergence (as in Taylor *et al.*, 2009b), the extensive sampling regime allowed us to identify further subdivision within the previously recognized species. The genetic data are supplemented with niche modelling data, both indicating that following the divergence of the two species, speciation followed different evolutionary trajectories which converge in the Eastern Cape province in the vicinity of the Alice area.

MATERIAL AND METHODS

Sampling

Tail biopsies were taken from 55 live *O. irroratus* which were collected in the Western Cape, Eastern Cape, Free State, KwaZulu-Natal and Mpumalanga provinces of South Africa (Fig. 1; Table 1). The sampling was supplemented with skin samples from 43 museum specimens obtained from the Durban Natural Science Museum (DM) (museum accession numbers provided in Table 1). Furthermore, an additional four sequences from specimens that were analysed in previous investigations (GenBank accession numbers FJ619554.1, FJ619555.1, FJ619556.1, and FJ619562.1) were included in the analysis.

DNA extraction, PCR and DNA sequencing

Total genomic DNA was extracted using a Qiagen DNEasy kit following the protocol of the manufacturer. Cross contamination is commonly associated with museum material (see Stuart *et al.*, 2006 for use of museum material); hence a blank DNA extraction was set up each time DNA was extracted from museum material. The mitochondrial cyt *b* gene was amplified using the universal primers L 14724 (5′-AAAAAGCTTCCATCCAACATCTCAGCATGATGAAA-3′) (Pääbo & Wilson, 1988; Kocher *et al.*, 1989) and H 15915 (5′-CTGCAGTCATCTCCGGTTTACAAGAC-3′) (Irwin, Kocher & Wilson, 1991) to amplify a 953-bp fragment from the fresh material. Museum material was amplified using primers targeting the first 400 base pairs of the 5′ end of the cyt *b* gene (tRNA-GluA: 5′-TGACTTGAARAACCAYCGTTG-3′ and tRNA-GluJ: 5′-CCCTCAGAATGATATTTGTCCTCA-3′) (Palumbi *et al.*, 1991). Each PCR reaction contained 14.9 μL of millipore water, 3.5 μL of 25 mM $MgCl_2$, 3 μL of 10× Mg^{2+}-free buffer, 0.5 μL of a 10 mM dNTP solution and 0.5 μL (10 mM) of the respective primer pairs, 0.1 μL of *Taq* polymerase and 2–4 μL of template DNA.

All PCR reactions had the following temperature cycle: 94 °C for 4 min, 94 °C for 30 s, 48 °C for 45 s and 72 °C for 35 s. The last three steps were repeated for 40 cycles followed by a final extension of 15 min at

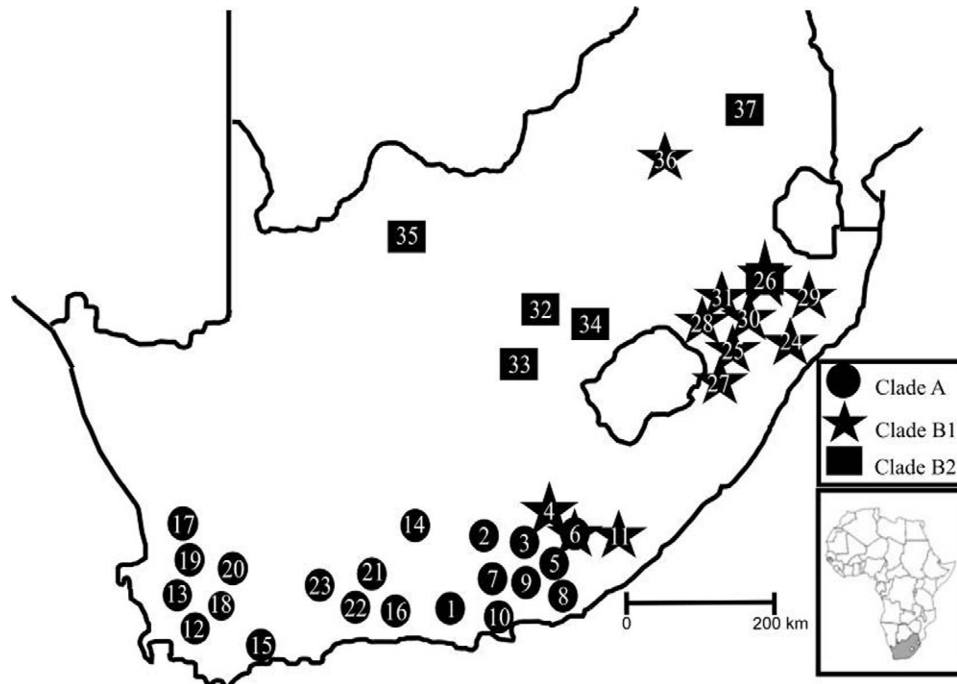

Figure 1. The geographical distribution of *Otomys irroratus* samples used in the current study. The numbers in each circle correspond to the locality in Table 1: 1, Baviaanskloof; 2, Somerset East; 3, Kroomie; 4, Hogsback; 5, Thomas Baines Nature Reserve; 6, Alice; 7, Groendal Nature Reserve; 8, Grahamstown; 9, Sam Knott Nature Reserve; 10, Port Elizabeth; 11, Stutterheim; 12, Stellenbosch; 13, Porterville; 14, Beaufort West; 15, De Hoop Nature Reserve; 16, Oudtshoorn; 17, Van Rhynsdorp; 18, Bainskloof; 19, Algeria; 20, Tweede Tol; 21, Gamkaskloof; 22, Swartberg; 23, Prince Albert; 24, Lemonwood; 25, Kamberg; 26, Karkloof; 27, Loteni Nature Reserve; 28, Bergville; 29, Umgeni valley; 30, Fort Nottingham; 31, Mgeni; 32, Theunissen; 33, Bloemfontein; 34, Ficksburg; 35, Kuruman; 36, Rietvlei; 37, Lydenburg. The localities of sequences obtained from GenBank are listed in Table 1 (Tygerkloof, KwaZulu Natal; Springs, Gauteng provinces in South Africa; Chingamwe, in Zimbabwe).

72 °C. PCR products were separated in 1% agarose gels, and excised with the aid of a sterile surgical blade. An Illustra™ (GE Healthcare) commercial kit was then used to purify the gene fragments from the agarose gel. An automated sequencer (ABI 3100, Applied Biosystems) was used to run cycle sequencing products. To authenticate the sequences obtained from the museum material all the sequences were blasted and compared with sequences on GenBank and aligned with sequences obtained from fresh material. To ascertain whether a pseudo-gene was amplified, sequences were checked for stop codons using EMBROSS/Transec (http://www.ebi.ac.uk/Tools/emboss/transeq/index.html).

Phylogenetic analysis

Sequences were edited in Sequence Navigator ver. 1.0.1 (Applied Biosystems, 1994) and MODELTEST ver. 3.06 was used to estimate the mode of nucleotide substitution (Posada & Crandall, 1998). Using the Akaike information criterion (AIC), which optimizes the number of parameters that describe the data, the best-fit maximum-likelihood (ML) score was chosen (Akaike, 1973; Nylander *et al*., 2004). Subsequently, phylogenetic relationships were inferred using maximum parsimony (MP) and neighbour joining (NJ) as implemented in PAUP*4 ver. beta 10 (Swofford, 2002). The heuristic search option was selected with tree bisection and reconnection branch swapping using 1000 random taxon stepwise additions and gaps were treated as missing characters. Confidence nodes on the MP and NJ trees were determined with the aid of bootstrapping (Felsenstein, 1985); bootstrap values above 75% were considered well supported. Further analysis was done using Bayesian inference (BI; MrBayes 3.0b4 for ML; Huelsenbeck & Ronquist, 2001). Ten Monte Carlo Markov chains were run for five million generations of which 10% were burn-in. The chains were sampled randomly for a tree every 1000th generation. A consensus tree was then constructed from the generated trees using the 50% majority rule and support for the nodes was estimated by posterior probabilities using the percentage

© 2011 The Linnean Society of London,

Table 1. List of all the specimens used in this study

Locality number and locations	Province	Locality coordinates	N	$2n$	GenBank accession numbers/Museum accession numbers
1 Baviaanskloof	Eastern Cape	33°39′38″S, 24°37′57″E	5	28	HM363664-HM363668; Fresh
2 Somerset East	Eastern Cape	32°41′62″S, 25°37′80″E	4	28,29,32	HM363689-HM363692; Fresh
3 Kroomie	Eastern Cape	27°08′07″S, 20°32′11″E	5	28	HM363673-HM363677; Fresh
4 Hogsback	Eastern Cape	32°36′19″S, 27°01′30″E	8	24	HM363657-HM363663; HM363708; Fresh; DM2934;
5 Thomas Bains NR	Eastern Cape	33°23′32″S, 26°29′41″E	1		HM363696; DM3080
6 Alice	Eastern Cape	32°47′00″S, 26°50′12″E	6	28	HM363698; HM363730; HM363744; HM363746; Fresh; DM2009; DM3581; DM2018
7 Groendal NR	Eastern Cape	33°42′07″S, 25°19′20″E	4		HM363699; HM363700; HM363703; HM363720; DM8391; DM8400; PT25
8 Grahamstown	Eastern Cape	33°18′15″S, 26°31′08″E	7		HM363701; HM363702; HM363714-HM363717; HM363736; DM3085; DM3578; PT33; PT34; PT35; PT1 PT2
9 Sam Knott NR	Eastern Cape	33°05′27″S, 26°42′58″E	3		HM363718; HM363719; HM363728; DM8504; DM8388; DM8397
10 Port Elizabeth	Eastern Cape	33°57′29″S, 25°36′00″E	1		HM363733; DM3005
11 Stutterheim	Eastern Cape	32°32′06″S, 27°22′57″E	3		HM363747-HM363748; HM363740; DM2994; DM2917; DM2919
12 Stellenbosch	Western Cape	33°55′54″S, 18°49′47″E	5	28	HM363669-HM363672; Fresh
13 Porterville	Western Cape	32°59′15″S, 19°01′28″E	5	28	HM363678-HM363682; Fresh
14 Beaufort West	Western Cape	32°15′19″S, 22°34′25″E	4	28	HM363685-HM363688; Fresh
15 De Hoop NR	Western Cape	34°26′03″S, 20°32′52″E	1		HM363693; Fresh
16 Oudtshoorn	Western Cape	33°39′56″S, 22°07′39″E	1	28	HM363694; Fresh
17 Van Rhynsdorp	Western Cape	31°44′46″S, 18°46′26″E	1	28	HM363695; Fresh
18 Bainskloof	Western Cape	33°34′50″S, 19°09′13″E	1		HM363697; DM4883
19 Algeria	Western Cape	32°26′19″S, 19°05′05″E	2		HM363710; HM363712; DM4196; DM4301
20 Tweede Tol	Western Cape	33°34′28″S, 19°09′32″E	1		HM363711
21 Gamkaskloof	Western Cape	33°21′50″S, 22°03′41″E	2		HM363713; HM363725; DM7130
22 Swartberg	Western Cape	33°21′36″S, 22°03′14″E	1		HM363363731; DM5096
23 Prince Albert	Western Cape	33°13′12″S, 22°01′33″E	1		HM363732; DM4210
24 Lemonwood	KwaZulu-Natal	29°28′03″S, 30°06′16″E	3		HM363704-HM363706; Fresh
25 Kamberg	KwaZulu-Natal	29°24′54″S, 29°40′21″E	2		HM363707; HM363739; DM3908; DM3107
26 Karkloof	KwaZulu-Natal	29°18′18″S, 30°13′30″E	2		HM363734; HM363709; DM3094; DM3093
27 Loteni	KwaZulu-Natal	29°27′01″S, 29°32′08″E	2		HM363721; HM363737; DM3112; DM3113
28 Bergville	KwaZulu-Natal	28°44′00″S, 29°22′00″E	1		HM363723; DM2374
29 Umgeni valley	KwaZulu-Natal	29°29′43″S, 29°48′55″E	2		HM363724; HM363741; DM3482; DM3898
30 Fort Nottingham	KwaZulu-Natal	29°24′46″S, 29°54′00″E	4		HM363749; HM363750; HM363726; HM363735; DM3575; DM3493; DM3576; DM3497
31 Mgeni	KwaZulu-Natal	29°48′36″S, 31°02′08″E	2		HM363727; HM363738; DM3489; DM3490
32 Theunissen	Free State	28°30′06″S, 26°48′07″E	2	28,30	HM363683; HM363684; Fresh
33 Bloemfontein	Free State	29°07′02″S, 26°14′01″E	1	28	HM363656; Fresh
34 Ficksburg	Free State	28°47′33″S, 27°53′35″E	2		HM363742; HM363743; Fresh
35 Kuruman	Northern Cape	27°27′12″S, 23°26′37″E	1		HM363729; DM3931
36 Rietvlei	Gauteng	25°53′49″S, 28°17′38″E	1		HM363745; DM3567
37 Lydenburg	Mpumalanga	25°05′46″S, 30°26′49″E	1	28	HM363751; Fresh
38 Tygerkloof	KwaZulu-Natal	25°58′05″S, 31°34′01″E	2		FJ619554.1; FJ619555.1
39 Springs	Gauteng	26°21′07″S, 28°45′00″E	1		FJ619556.1
40 Zimbabwe	Chingamwe	18°45′00″S, 32°57′00″E	1		FJ619562.1

Total number of specimens: 102.

For each specimen the locality coordinates are indicated and the number of samples (N) that was collected and the number of base pairs (bp) sequenced are given. NR, nature reserve. For fresh material a 953-bp fragment was sampled, while a 400-bp fragment was amplified from tissue obtained from the Durban Museum of Natural History (DM); museum accession numbers are indicated below. The diploid numbers of karyotyped specimens and GenBank accession numbers are listed for all the populations.

of time a node was recovered. Nodes with posterior probabilities (pP) < 0.95 were considered poorly supported. Two species of *Otomys*, *O. laminatus* and *O. karoensis,* were used as outgroups as they comprise a monophyletic group which forms a sister group to *O. irroratus* (Taylor, Denys & Mukerjee, 2004).

Molecular clock

To date cladogenic events within *O. irroratus*, we used fossil calibration points in conjunction with our DNA sequence data. Fossil data indicate that the family Otomyinae is approximately five million years old (Pocock, 1976; Avery, 1991; Jansa, Barker & Heaney, 2006), while the oldest known fossil of *Otomys* is dated at approximately 3.1 Ma (Sénégas & Avery, 1998; Sénégas, 2001; Denys, 2003; Taylor *et al*., 2004; Matthews, Denys & Parkington, 2005; Hopley, Latham & Marshall, 2006). Using BEAST (ver. 1.4.8) and three fossil calibration dates (*Mus*/*Rattus*: ~12 Ma; Otomyinae: ~5 Ma, and *Otomys*: ~3.1 Ma) the divergence date between the major clades that were retrieved was estimated using the Bayesian approach (with a length of 20 million generations and a burn-in of 200 000) and the nucleotide evolution model sampled every 10 000 iterations (Drummond & Rambaut, 2006). The Bayesian approach was also used to investigate the confidence at the nodes at 95% pP.

Population genetics analysis

A haplotype network was constructed using the method of Templeton, Crandall & Sing (1992) at the 95% confidence interval, using TCS ver. 1.13 (Clement, Posada & Crandall, 2001) and the network subjected to nested clade analysis (NCA) (Templeton, Routman & Phillips, 1995). NCA was used to determine whether historical events versus current gene flow barriers are operating on the current population genetic structure (Templeton *et al*., 1995; Templeton, 1998, 2004). The haplotype network was nested and analysed using GeoDIS ver. 2.0 (Posada, Crandall & Templeton, 2000).

The population genetic structure was analysed using AMOVA and the clades obtained from the phylogenetic analysis were used for defining groups *a priori* (Excoffier, Smouse & Quattro, 1992). Population subdivision was estimated using F-statistics (Excoffier *et al*., 1992; Frankham, Ballou & Briscoe, 2003; Excoffier, Laval & Schneider, 2005). Population demographics of *O. irroratus* were also evaluated using mismatch analysis (Rogers & Harpending, 1992; Harpending 1994). Tests for goodness-of-fit statistics based on the sum of square deviations for a model of sudden expansion was determined in Arlequin version 3.11 (Excoffier *et al*., 2005). After range expansion or contraction is established, Fu's F_S test was used to test whether the populations are in equilibrium (Fu, 1997).

Neutrality tests were then done to ensure that the pattern of sequence polymorphism observed in the populations conformed to the neutral model. Tajima's D-value was also calculated to evaluate whether the populations are in stasis or expanding using Arlequin ver. 3.11 (Excoffier *et al*., 2005). After an episode of population growth, coalescence theory predicts that external branches are elongated and there is an excess of low-frequency mutations compared with the neutral model (Petit, Balloux & Excoffier, 2000). Tests for goodness-of-fit and Fu's F_S were generated in Arlequin ver. 3.11 using parametric bootstrapping with 10 000 replicates (Felsenstein, 1985).

Niche modelling analysis

Many problems are inherent in using presence-only data (such as museum records) for estimating the predicted distributions of species, but recent models have overcome this limitation. For example, the MaxEnt (Maximum Entropy) method (Phillips *et al*., 2006), which employs a general machine learning algorithm, has been shown to perform well with presence-only data, and with sample sizes as low as five, 10 or 25 occurrences, and to outperform alternative 'climatic envelope' models such as GARP and BIOCLIM (for recent discussions which strongly endorse MaxEnt see Elith *et al*., 2006; Hernandez *et al*., 2006). Based on 41 georeferenced distribution records (precision of 0.001 degrees), which are reasonably evenly spread throughout the latitudinal and longitudinal range of the species (Fig. 1), we used MaxEnt version 3.0 beta to estimate predicted distributions (approximated ecological niches) of the two genetically defined clades of *O. irroratus sensu lato* defined by this study, Clade A ($N = 21$ localities) and Clade B ($N = 20$ localities). We used nine continuous environmental variables (WORLDCLIM version 1.4: http://biogeo.berkeley.edu/worldclim/; Hijmans *et al*., 2005) reflecting means, extremes and seasonal variation of temperature and precipitation: Bio1 (mean annual temperature), Bio4 (temperature seasonality), Bio5 (maximum temperature of warmest month), Bio6 (minimum temperature of coldest month), Bio7 (annual range of temperature), Bio12 (annual precipitation), Bio13 (precipitation of wettest month), Bio14 (precipitation of driest month), and Bio15 (precipitation seasonality). Altitude was not included as preliminary correlation analysis showed it to be significantly correlated ($N = 39$, $P < 0.05$) with all other climatic variables and particularly with Bio6

(minimum temperature of coldest month) ($r = -0.82$, $P << 0.0001$) (data not shown but available from P.J.T.). The environmental data were set to a spatial grid resolution of 2.5 arc minutes. The MaxEnt model was run with all distribution records (100% training), the regularization multiplier was set to 2.0, maximum number of iterations was set to 1000, convergence threshold was set to 1×10^{-5}, 'auto-features' was selected, and output format was set to logistic. Model performance was assessed with proportion of presences correctly classified (sensitivity), proportion of absences correctly classified (specificity), and discrimination ability [area under the curve (AUC) of a receiver operating characteristic plot of sensitivity versus 1 – specificity). As MaxEnt produces a continuous probability (ranging from 0 to 1.0), we transformed the continuous model output to a map representing probabilities. The percentage contribution of each explanatory variable to model performance was evaluated; furthermore, a jackknife procedure implemented in MaxEnt was used, where variables are successively omitted and then used in isolation to measure their relative as well as absolute contribution to the model.

RESULTS

Phylogenetic data analysis

A 953-bp fragment of the cyt *b* locus was obtained from the fresh material ($N = 55$) while a 400-bp fragment was obtained from 43 of the 75 museum specimens (Table 1). Altogether, 98 specimens from 38 localities were successfully amplified, sequenced and deposited in GenBank (accession numbers: HM363654–HM363751). The four sequences that were downloaded from GenBank were included in the analysis, resulting in a total of $N = 102$ specimens, which resolved into 69 haplotypes (see Supporting Information, Table S1).

The partial fragment of the cyt *b* gene (400 bp) obtained from museum material (missing characters were treated as absent base pairs for the museum material) and the larger fragment (953 bp) were first analysed separately to check whether the two datasets were comparable. As identical patterns were obtained the sequences were combined into a single data matrix. MODELTEST selected the HKY + I + G as the model of substitution using the AIC criteria (–lnL = 3405.58; AIC = 6823.16) (Hasegawa, Kishino & Yano, 1985). The proportion of invariable sites (I) was 0.56 and the gamma distribution shape parameter (k) was 0.68 for the among-site variation. The substitution model's rate matrix was R(a) [A–C] = R(c) [A–T] = R(d) [C–G] = R(f) [G–T] = 1.00 while R(b) [A–G] = 10.93 and R(e) [C–T] = 13.32; similar values were retrieved for the two datasets (the 400- and 953-bp fragments) and hence only the combined dataset is discussed. The base frequencies for the combined dataset were A = 31.43%, C = 30.00%, G = 11.50% and T = 27.04%. No significant variation in base composition was evident between sequences as shown by a chi-squared test ($\chi^2 = 15.95$, d.f. = 1, $P = 0.01$). Similar results were found in cyt *b* studies for other rodents (Michaux *et al.*, 2004; Mitsainas *et al.*, 2009). For MP, 157 characters were parsimony informative, and retrieved 100 equally parsimonious tress with a length of 363 steps (CI = 0.60, RI = 0.92) from which a strict consensus was constructed.

The tree topologies retrieved for MP, NJ and BI were nearly identical with regard to the major clades retrieved and hence only the MP tree topology is discussed (Fig. 2). The monophyly of *O. irroratus* was well supported (100%/100%/0.99 p*P*), and two parapatrically distributed clades were retrieved (100%/99%/1.00 p*P*): Clade A occurs in the south-western parts of South Africa (encompassing the Western and Eastern Cape provinces) (96%/100%/0.96 p*P*) while Clade B occurs in the north-eastern parts of South Africa (Eastern Cape, Free State, KwaZulu-Natal, Mpumalanga, Gauteng and Northern Cape provinces) (79%/71%/0.96 p*P*) (Fig. 2). These two clades are separated by 7.0% average sequence divergence.

Within Clade A two groups are evident, the first comprising 13 specimens which are exclusive to the Eastern Cape province (100%/67/0.99 p*P*), and the second group having 39 specimens which occur in the Western and Eastern Cape provinces (100%/ < 50%/0.96 p*P* (Fig. 2). Clade A is characterized by low sequence divergences between specimens (0.2–1.91%). Similarly, Clade B is subdivided into two groups (subclades B1 and B2) that are separated by 4.8% (Fig. 2). Subclade B1 comprises ten individuals derived from a wide geographical range in the northern parts of South Africa (Mpumalanga, Northern Cape, KwaZulu-Natal and Free State provinces) (77%/75%/1.0 p*P*). Subclade B2 has 33 specimens predominantly distributed in the southeastern half of South Africa (Eastern Cape, KwaZulu-Natal) (100%/86%/0.96 p*P*). Specimens from Hogsback in the Eastern Cape clustered within the two subgroups comprising Cade B; similarly specimens from Fort Nottingham (KwaZulu-Natal) were present in the two subgroups of Clade B. In turn Clades A and B appear to have an area of sympatry at Alice in the Eastern Cape province as specimens ($N = 6$) from this locality grouped in both Clades A and B. Divergence time estimates dated the split between the two major clades at 1.1 Ma (95% confidence intervals: 0.4338–2.7652 Ma), suggesting a late Pliocene/Pleistocene

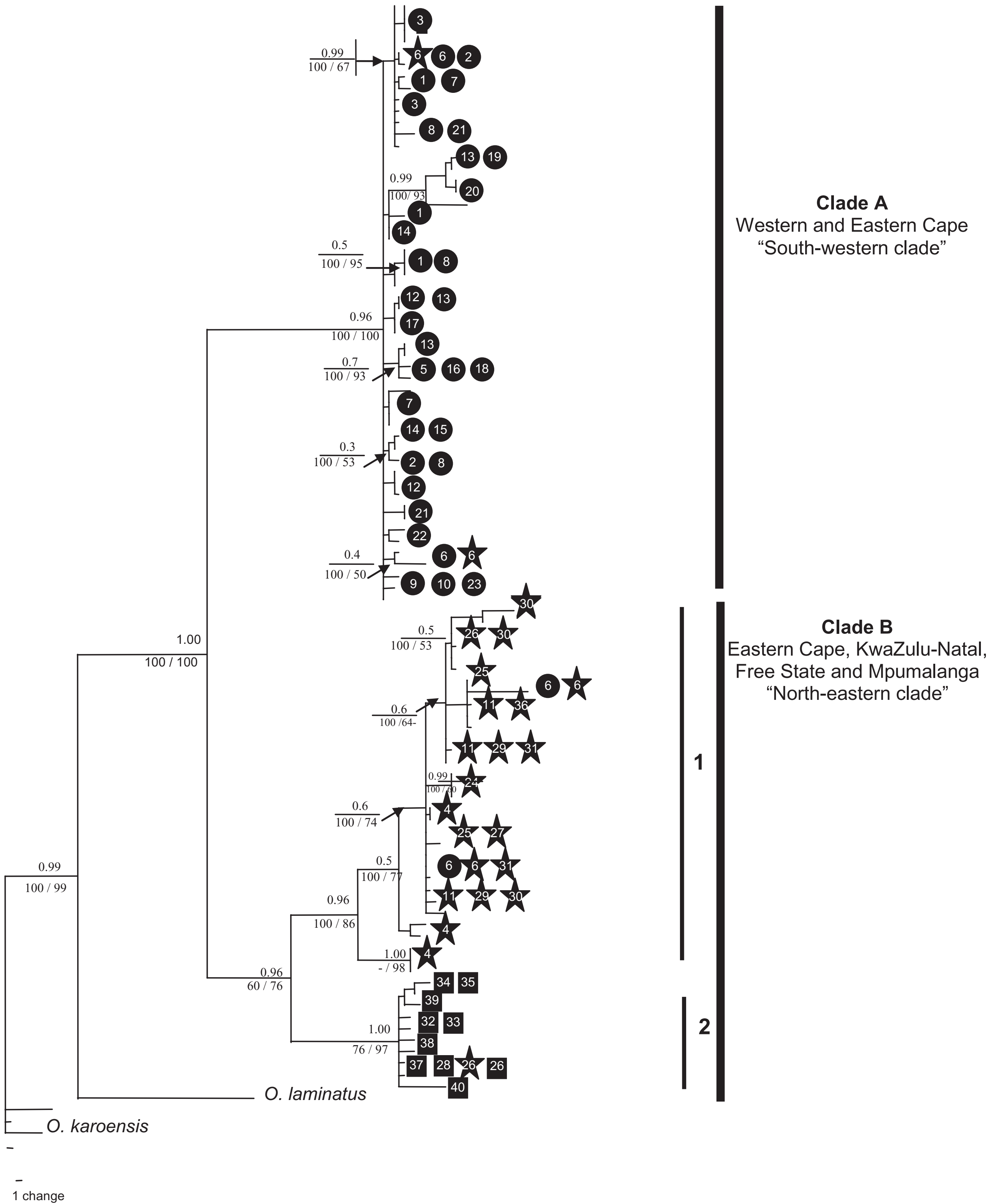

Figure 2. The maximum-parsimony phylogram for *Otomys irroratus* for the cyt *b* gene ($N = 102$) based on 69 haplotypes. Statistically well-supported clades have posterior probabilities which are > 0.95 p*P* (above nodes), while bootstrap values above 95% are indicated below supported nodes (maximum-parsimony values on the left and neighbour-joining values on the right). The two major clades A and B that were retrieved are indicated by the vertical lines.

cladogenesis. Divergence within Clade A is dated at 0.62 Ma (0.11330–0.90300 Ma) and for clade B it is dated at 0.66 Ma (0.1349–1.1906 Ma).

Population genetics

The AMOVA retrieved a Φ_{ST} (among group variation) of 77.10% d.f. = 36, sum of squares 408.71, variance component = 1.38) and the two major clades are separated by significant pairwise $F_{ST} = 0.87$ ($P < 0.01$). Within the two clades some genetic substructure was found with F_{ST} values ranging from 0.40 ($P < 0.01$) in Clade A to 0.44 ($P < 0.01$) in Clade B. The Tajima's D-value over all of the sampled localities was 0.52, suggesting balancing selection or a decrease in the population size, while a P-value of 0.83 was reported by Fu's F_S test for all the sampled localities, indicating that the entire population is in stasis. A bell-shaped (unimodal) distribution of the frequency of the pairwise differences was found in Clade A, which is indicative of a recently expanded population (supporting Fig. S1A). In contrast, Clade B has a multimodal pairwise difference distribution of a stable population (Fig. S1B).

Genetic diversity statistics

The haplotype diversity for the combined dataset was high ($h = 0.99$) with a low nucleotide diversity of $\pi = 0.045$ (pairwise nucleotide differences $k = 17.165$). The reduced data had comparable indices: $h = 0.95$, $\pi = 0.037$, $k = 13.957$. As two main clades were retrieved, we calculated genetic indices for each. Haplotype diversity for Clade A was high ($h = 0.898$) but was identical to the two subclades A1 ($h = 0.882$) and A2 ($h = 0.878$) and the nucleotide diversity very low, $\pi = 0.0006$ ($\pi = 0.004$ for subClade A1 and 0.0046 for A2; supporting Table S3). Clade B similarly showed a high haplotype diversity ($h = 0.92$; Clade B1 $h = 0.86$, Clade B2 $h = 0.8$) yet a higher nucleotide diversity ($\pi = 0.046$) (pairwise nucleotide differences $k = 18.87$ and variable haplotype diversity = 0.00055). The results emanating from this analysis indicate that Clade A experienced a recent population expansion while Clade B seems to be in stasis and older than Clade A, as indicated by the higher nucleotide diversity.

Nested clade analysis

Sixty-nine haplotypes were retrieved from 102 specimens and these were grouped into a single network at the 95% confidence level (supporting Fig. S2, Table S1). The two major clades which were revealed in the phylogeny were also evident from the haplotype network. The number of haplotypes (Nh) per locality ranged from one to four and most of the haplotypes were separated from each other by one to three mutational steps. Two genetically distinct groups are apparent from this network, as indicated by the high number of mutational steps (18 in total) separating them. The two subgroups in Clade B are separated by eight mutational steps (supporting Fig. S2).

Forty-five one-step clades, 16 two-step clades, six three-step clades and two four-step clades were evident from nested clade analysis. GeoDis (ver. 2.0) also retrieved two divergent groups which correspond to the two lineages obtained using phylogenetic analysis (clades 4-1 and 4-2). Clade A occurs in the south western parts of South Africa (4-2), while Clade B occurs in the north-eastern parts of South Africa. The analysis indicated that eight of the clades had a statistically significant relationship between nested clades and geography (supporting Table S2). Clade A-15 was the only clade on the first nesting level to have a statistically significant P-value. Inadequate geographical sampling, however, prevented any conclusive interpretation. Three clades on the second nesting level showed statistically significant P-values (2-6, 2-8 and 2-15). For Clade B-6, however, inadequate sampling precluded any robust conclusion, and isolation by distance as a result of short-distance dispersal vs. long-distance dispersal could not be confirmed. For Clade B-8 restricted gene flow with isolation by distance led to the current genetic structure. This haplotype occurs in the Oudtshoorn region, in the little Karoo of the Western Cape province. The analysis retrieved an inconclusive outcome for Clade B-15. On the third nesting level two clades had statistical support (3-2 and 3-5). Clade 3-2 showed restricted gene flow with some long-distance dispersal; clade 3-2 includes most of the Western Cape province samples (Clade A). Clade 3–5 has haplotypes which originate from Hogsback in the Eastern Cape province, and restricted gene flow with isolation by distance is implicated as the driving force leading to the contemporary genetic structure in this clade. Only one clade on the fourth level had sufficient statistical support (4-1); however, an inconclusive outcome was retrieved from the inference key for this clade. The total cladogram showed that allopatric fragmentation led to the current genetic structure within *O. irroratus*.

Niche modelling analysis

The MaxEnt algorithm converged after 240 (Clade A) and 300 (Clade B) iterations with a regularized training gain of 1.521 (Clade A) and 1.420 (Clade B). Modelled potential distributions (MaxEnt) of Clades 1 and 2 are shown in Figure 3A, B. Localities based on

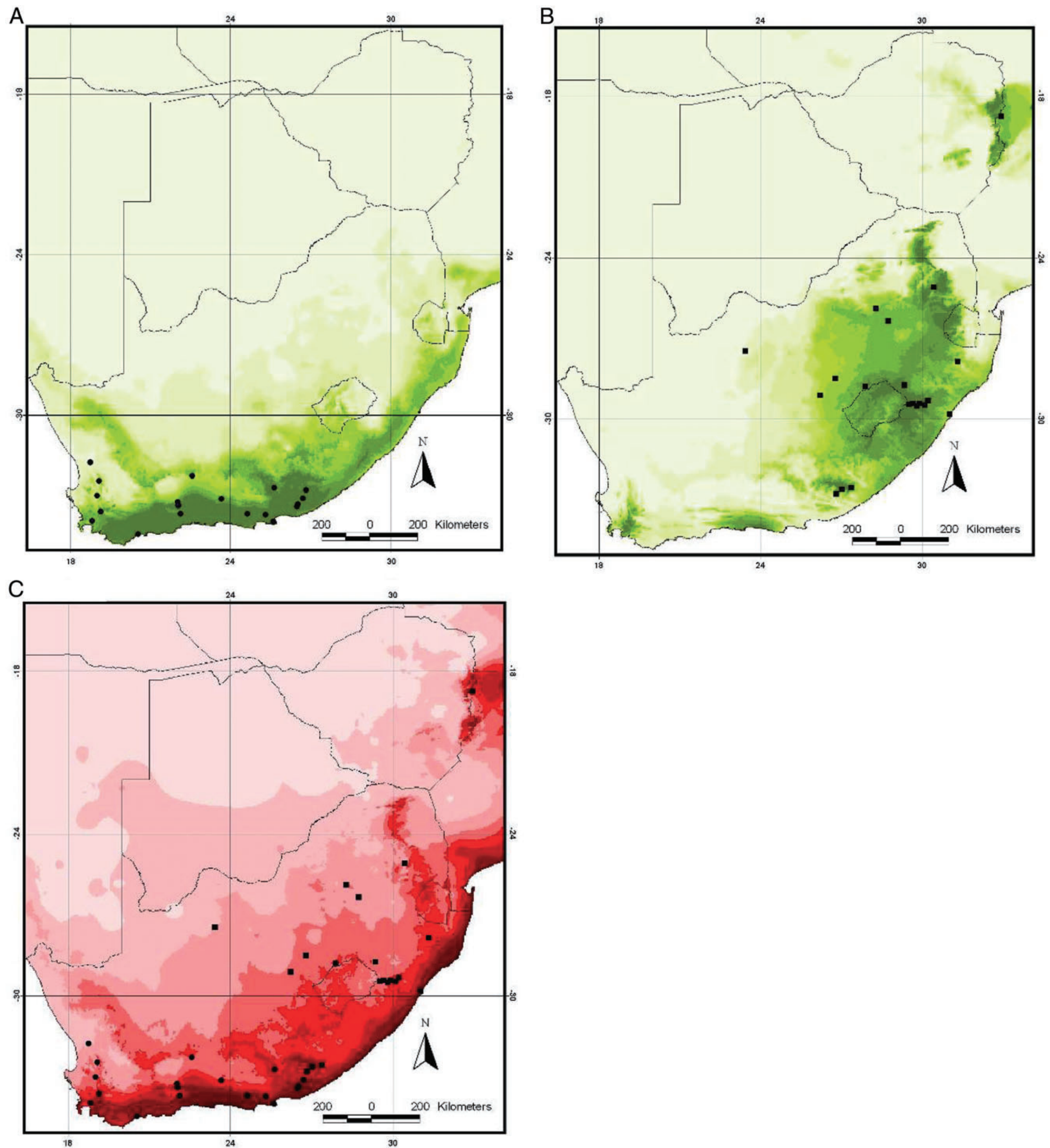

Figure 3. Modelled potential distributions (MaxEnt) of genetically defined Clades A (A) and B (B) of *Otomys irroratus* based on 41 records of occurrence from southern Africa. Shading represents ranges of probability of occurrence (i.e.habitat suitability). (C) Distribution of one bioclimatic variable (maximum precipitation of driest month: Bioclim 14) in southern Africa in relation to sampled localities of Clade A (closed circles) and Clade B (closed squares).

museum collecting records are shown as filled circles. Shading represents ranges of probability of occurrence (i.e. habitat suitability) with darker shading indicating higher suitability. Model performance as assessed by the AUC was very high (0.940 and 0.944 for Clades 1 and 2, respectively), indicating efficient classification of suitable versus unsuitable habitats. For both clades, the environmental variable which

© 2011 The Linnean Society of London,

provided the greatest relative contribution to the MaxEnt model was precipitation of the driest month (Bio14) (variable contribution was 73 and 49% for Clades A and B, respectively). Figure 3C shows the distribution of this variable in southern Africa, with the sampled localities of the two clades superimposed; localities assigned to Clade A occupy regions having relatively high precipitation in the driest month whilst Clade B localities from the high plateau of South Africa and eastern highlands of Zimbabwe are characterized by a more seasonal rainfall regime in which the driest month receives relative low precipitation. The environmental variables with the highest explanatory power (gain) when used in isolation were precipitation seasonality (Bio15; Clade A) and maximum temperature of warmest month (Bio5; Clade B). The environmental variables that decreased the gain the most in the overall model when omitted (and therefore had the most information that was not present in the other variables) were Bio14 (precipitation of the driest month) (Clade A) and Bio12 (annual precipitation) (Clade B). The MaxEnt models (Fig. 3A, B) predicted distributions for the two clades in southern Africa that were largely non-overlapping (apart from areas including the coastal zone of KwaZulu-Natal and mountains of the Cape Fold Mountains and the Great Escarpment), with Clade A mostly associated with the Fynbos and Thicket Biomes associated with coastal locations in the southern and eastern Cape and Clade B mostly associated with the higher-altitude Grassland Biome of South Africa and eastern highlands of Zimbabwe.

DISCUSSION

The topology inferred from the cyt *b* mtDNA sequence data clearly demonstrates the presence of two monophyletic lineages each containing two sub-clades within *Otomys irroratus*. These two major clades are characterized by elevated sequence divergence values and separated by chromosomal differences (data not shown). Clade A includes the distribution of the type locality of *O. irroratus* (Uitenhage, Eastern Cape; Roberts, 1929) while Clade B represents a novel lineage, referred to as *O. auratus* Wroughton (1906) by Taylor *et al*. (2009b). Clades A and B are sympatric at Alice in the Eastern Cape, and are genetically distinct, as evident from the general absence of shared haplotypes. Cladogenesis within *O. irroratus* coincides with a Pliocene/Pleistocene divergence event. Furthermore, mismatch distributions indicate that Clade A underwent a recent expansion while Clade B has a mismatch distribution indicative of a stable population, suggesting that the two groups have different evolutionary histories. Allopatric fragmentation was inferred from NCA, suggesting that historical events led to the current genetic structure within *O. irroratus*. Ecological differentiation between the two clades was demonstrated by the MaxEnt models, associated mainly with environmental variables relating to the more seasonally equitable precipitation regime of the Fynbos and Thicket Biomes compared with the more pronounced precipitation seasonality experienced by the Grassland Biome.

Population demography, biogeography

The Eastern Cape and the Western Cape provinces consist of heterogeneous biomes containing forests (Afrotemperate forest and Scarp forests), fynbos and grasslands (which comprise the preferred habitat of *O. irroratus*). Lawes *et al*. (2007) found that the two major forest types have undergone expansion and contraction cycles since the Last Glacial Maximum. In particular, these forests contracted to isolated refugia in KwaZulu-Natal and the Eastern Cape and expanded when moist conditions prevailed again (Eeley, Lawes & Piper, 1999; Lawes *et al*., 2007). The expansion and the retreat of the forests would have also impacted the grassland biome. This in turn would also affect demographics of species occurring in these areas such as *O. irroratus*, which would then be evident in population genetic structures. The mismatch analysis supports this by showing a recent expansion for Clade A, which is distributed in the south-western parts of South Africa. The recent expansion could have been due to favourable climatic conditions which prevailed during the Pleistocene, which enabled the expansion of fynbos, northwards up the west coast of South Africa, thus enabling *O. irroratus* to expand in this direction (Moreau, 1962). Our molecular dating indicates that the expansion occurred almost at the same time as the forest expansions, which are dated at 18 000 years BP (Lawes *et al*., 2007). In contrast, Clade B shows a mismatch distribution of a stable population and pronounced substructure at the sequence level. The multimodal pattern in Clade B reflects the high haplotype diversity and elevated sequence divergences separating these haplotypes. According to Lawes *et al*. (2007), the forest contraction and expansion regime in the KwaZulu-Natal area is much older than in the southern–western seaboard. The estimated divergence time between the two major clades found in *O. irroratus* is 1.1 Ma, which is in line with the climate changes that took place since the late Pliocene between 2.9 and 0.8 Ma (Ellery, Scholes & Mentis, 1991; de Menocal, 2004; Lawes *et al*., 2007).

The phylogeographical pattern observed in this investigation is similarly retrieved in other co-distributed taxa. These include the four-striped fieldmouse *Rhabdomys pumilio* (Rambau *et al*., 2003),

© 2011 The Linnean Society of London,

which is a generalist species also inhabiting grasslands, and the forest shrew *Myosorex varius* (Willows-Munro & Matthee, 2009), which is a montane grassland specialist. In each of these species multiple lineages were retrieved with all of them sharing the phylogeographical break in the Eastern Cape. The area of sympatry (Alice) between these two putative species is in the vicinity of a known vegetation cross-over zone between the fynbos and grassland biomes in southern Africa (Mucina & Rutherford, 2006). Therefore, the distribution of the two main clades closely follows the biomes in the southern parts of South Africa. The divergent lineages within *O. irroratus* are therefore not unique with respect to other taxa.

Cape Floristic Region

Part of our sampling included the Cape Floristic Region especially across the Cape Fold Mountains (CFM) which acts as a genetic barrier to many species (Tolley, Chase & Forest, 2008, 2009; Swart, Tolley & Matthee, 2009). Interestingly, this barrier is not reflected in the specimens that we sampled on either side of the mountains comprising the CFM. For instance, specimens from Porterville, Van Rhynsdorp, and Algeria, which occur north of the Hottentots Holland range, and specimens from Stellenbosch and De Hoop, which are on the south of the Langeberg mountains, all occur in the same clade and are separated by sequence divergences averaging 0.2%. Our NCA indicates that long-distance dispersal could have resulted in gene flow between distant populations. Clearly, high topography may not be a dispersal barrier to *O. irroratus* as it is known that they occur at altitudes up to 2000 m above sea level in the Drakensburg mountain range (Lynch & Watson, 1992). Furthermore, Taylor *et al*. (2009b) argue that during unfavourable periods (such as dry conditions) *O. irroratus* would seek refuge in highland areas and expand when conditions are permissible again. Apart from the dry conditions, the grassland habitat of *O. irroratus* may have been able to expand during higher temperature cycles as woodlands would contract during this time (Ellery *et al*., 1991; Lawes *et al*., 2007). The lack of genetic structure within Clade A could therefore be an artefact of this process.

Species boundaries

Since the inception of the biological species concept (Mayr, 1963), a plethora of species concepts have been developed (Cracraft, 1989, 1992; Mayden, 1997; Sites & Marshall, 2003). More recently, there is a tendency to infer species distinction when genetically divergent lineages are detected within putatitive species (Rambau *et al*., 2003; Daniels *et al*., 2007; Daniels, Heideman & Hendricks, 2009), particularly when cyt *b* sequence divergence exceeds 5% in rodent taxa (Bradley & Baker, 2001; Baker & Bradley, 2006). In doing so, the general approach appears to apply the genetic species concept, which has several limitations when applied in a phylogenetic framework (Bond & Sierwald, 2002; Monaghan *et al*., 2009). The alternative approach (most acceptable in our view) is to use an integrated approach wherein using multiple diagnostic characters (including genetic characters) is advocated as proposed by Crandall *et al*. (2000) and Bond & Stockman (2008).

In the case of *O. irroratus*, the elevated mitochondrial sequence divergence separating the two major clades is supported by several datasets. In the first instance, these lineages inhabit geographical areas characterized by different biomes. Clade A occurs in the CFM and Albany thicket biome and Clade B occurs in the northern grassland biome. This is underslined by divergent ecological parameters which were described using niche modelling predictions. Altogether, our mitochondrial data derived from a wide geographical range and distribution predictions derived from niche modelling provide strong grounds for two species within *O. irroratus*, which were provisionally recognized by Taylor *et al*. (2009b) as *O. irroratus* (Clade A) and *O. auratus* (Clade B).

This study showed the presence of a potential contact zone at Alice in the Eastern Cape, although more sampling is required to establish the size of the contact zone. Furthermore, additional nuclear genotyping would help determine whether gene flow occurs across the contact zone. The phenotypic and pelage conservatism displayed in *O. irroratus* adds to the growing number of cryptic species occurring in the southern African subregion (Denys & Jeager, 1986; Britton-Davidian *et al*., 1995; Taylor *et al*., 1995; Rambau *et al*., 2003; Mullin, Pillay & Taylor, 2004; Willows-Munro & Matthee, 2009). Considering the results presented here, biodiversity indices for southern Africa may currently grossly underestimate the number of taxa present as many widespread taxa characterized by morphological conservatism may have diverged at the molecular level. This clearly demonstrates the importance of molecular approaches when investigating population differences (and demographics) within rodents, particularly in taxa that do not display morphological diagnostic characters. Although museum material is renowned for its difficulty to amplify, this investigation clearly underlines the utility of museum material (and museum collections) in addressing species demarcations between taxa (for instance see Goodman *et al*., 2006; Smit *et al*., 2007), and certainly helped us to comprehensively sample (without further invasive sampling)

throughout the major parts of the distribution of *O. irroratus*.

ACKNOWLEDGEMENTS

Sampling permits were obtained from the Department of Economic Development and Environmental Affairs of the Eastern Cape (Permit no. CRO – 100/08CR and CRO – 101/08CR), Cape Nature (Permit no. AAA004-00269-0035) and Mpumalanga Tourism and Parks Agency (Permit no. MPB. 5240). Ethical clearance was approved by Sub Committee B of the University of Stellenbosch (Ethics clearance certificate No. 2008B01004). Funding for this project was provided by the National Research Fund grants to R.V.R. and S.R.D. We would also like to thank Dr Sonja Matthee, Dr Nico Avenant (The National Museum Bloemfontein, South Africa) and Dr Frans Radlof for providing us with some of the samples used in this study. We thank two anonymous reviewers for constructive comments.

REFERENCES

Akaike H. 1973. Information theory as an extension of maximum likelihood principle. In: Petrov BN, Csake F, eds. *Second International Symposium on Information Theory*. Budapest: Akademiai Kaido, 267–281.

Avery DM. 1991. Late Quaternary incidence of some micro mammalian species in Natal. *Durban Museum Novitates* **16:** 1–11.

Avise JC. 1998. The history and purview of phylogeography: a personal reflection. *Molecular Ecology* **7:** 371–379.

Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.

Baker RJ, Bradley RD. 2006. Speciation in mammals and the genetic species concept. *Journal of Mammalogy* **87:** 643–662.

Bond JE, Sierwald P. 2002. *Eurhinocricus rosenbergi*, a new species of rhinocricid from the Caribbean island of Jamaica (Spirobolida: Rhinocricidae). *Proceedings of the Biological Society of Washington* **155:** 670–675.

Bond J, Stockman A. 2008. An integrative method for delimiting cohesion species: finding the population–species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology* **57:** 628–646.

Bradley RD, Baker RJ. 2001. A test of genetic species concept: cytochrome *b* sequences and mammals. *Journal of Mammalogy* **82:** 960–973.

Britton-Davidian J, Catalan J, Granjon L, Duplantier J-M. 1995. Chromosomal phylogeny and evolution in the genus *Mastomys* (Mammalia, Rodentia). *Journal of Mammalogy* **76:** 248–262.

Clement MD, Posada D, Crandall KA. 2001. TCS. A computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657–1660.

Contrafatto G, Meester JA, Bronner G, Taylor PJ, Willan K. 1992b. Genetic variation in the African rodent subfamily Otomyinae (Muridae). IV. Chromosome G banding analysis of *Otomys irroratus* and *O. angoniensis*. *Israel Journal of Zoology* **38:** 277–291.

Contrafatto G, Meester JA, Willan K, Taylor PJ, Roberts MA, Baker CM. 1992c. Genetic variation in the African rodent subfamily Otomyinae (Muridae). II. Chromosomal changes in some populations of *Otomys irroratus*. *Cytogenetics and Cell Genetics* **59:** 293–299.

Contrafatto G, Van Der Berg JR, Grace JH. 1997. Genetic variation in the African rodent subfamily Otomyinae (Muridae): immuno-electron transfer of liver proteins of some of *Otomys irroratus* (Brants, 1827) populations. *Tropical Zoology* **10:** 157–171.

Cracraft J. 1989. Speciation and its ontology: the empirical consequences of alternative species concepts for understanding the pattern and process of differentiation. In: Ottc D, Endler JA, eds. *Speciation and its consequences*. Sunderland, MA: Sinauer, 28–59.

Cracraft J. 1992. The species of the birds-of-paradise (Paradisaeidae): applying the phylogenetic species concept to a complex pattern of diversification. *Cladistics* **8:** 1–43.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK. 2000. Considering evolutionary process in conservation biology. *Tree* **15:** 290–295.

Crochet PA. 2000. Genetic structure of avian populations allozymes revisited. *Molecular Ecology* **9:** 1463–1469.

Daniels SR, Hofmeyer MD, Henen BT, Crandall AK. 2007. Living with the genetic signature of Miocene induced change: evidence from the phylogeographic structure of the endemic angulate tortoise *Chersina angulata*. *Molecular Phylogenetics and Evolution* **45:** 915–926.

Daniels SR, Heideman NJ, Hendricks MGJ. 2009. Examination of evolutionary relationships in a cape fossorial skink species complex (Acontiinae: *Acontias meleagris meleagris*) reveals the presence of five cryptic lineages. *Zoologica Scripta* **38:** 449–463.

Denys C. 2003. Evolution du genre *Otomys* (Rodentia: Muridae) au Plio-Pléistocène d'Afrique orientale et australe. In: Radulescu C, Petculecus PMS, eds. *Advances in paleontology 'Henth to Pantha'* . Memorial Volume, 75–84.

Denys C, Jeager J-J. 1986. A biostratigraphic problem: the case of the East-African Plio-Pleistocene rodent faunas. *Modern Geology* **10:** 215–233.

Drummond AJ, Rambaut A. 2006. BEAST v. 1.4. Bayesian evolutionary analysis sampling trees. Available at http://beast.bio.ed.ac.uk/Main_Page

Eeley HAC, Lawes MJ, Piper SE. 1999. The influence of climate change on the distribution of indigenous forest in KwaZulu–Natal, South Africa. *Journal of Biogeography* **26:** 595–617.

**Elith J, Graham CH, Anderson RP, Dudı'k M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Mc Overton JC, Peterson AT, Phillips SJ, Richardson KS,

Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann N. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29:** 129–151.

Ellery WN, Scholes RJ, Mentis MT. 1991. An initial approach to predicting the sensitivity of the grassland biome to climate change. *South African Journal of Science* **87:** 499–503.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetica* **131:** 479–491.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evoutionary Bioinformaics Online* **1:** 47–50.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39:** 783–791.

Frankham R, Ballou JD, Briscoe DA. 2003. *Introduction to conservation genetics*. Cambridge: Cambridge University Press.

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915–925.

Goodman SM, Raxworthy CJ, Maminirnua CP, Olson LE. 2006. A new species of shrew tenrec (*Microgale jobihely*) from northern Madagascar. *Journal of Zoology* **270:** 384–398.

Harpending HC. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66:** 591–601.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22:** 160–174.

Hernandez PA, Graham CH, Master LL, Albert DL. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* **29:** 773–785.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Hopley PJ, Latham AJ, Marshall JD. 2006. Palaeoenvironments and palaeodiets of mid Pliocene micromammals from Makapansgat Limeworks, South Africa: a stable isotope and dental microwear approach. *Palaeogeography, Palaeoclimatology, Palaeoecology* **233:** 235–251.

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinfomatics* **17:** 754–755.

Irwin DM, Kocher TD, Wilson AC. 1991. Evolution of cytochrome *b* gene of mammals. *Journal of Molecular Evolution* **32:** 128–144.

Jansa SA, Barker FK, Heaney LR. 2006. The pattern and timing of diversification of Philippine endemic rodents: evidence from mitochondrial and nuclear gene sequences. *Systematic Biology* **55:** 73–88.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablancatt FX, Wilson AC. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Science USA* **86:** 6196–6200.

Lawes MJ, Eeley HAC, Findlay NJ, Forbes D. 2007. Resilient forest faunal communities in South Africa: a legacy of palaeoclimatic change and extinction filtering? *Journal of Biogeography* **34:** 1246–1264.

Lynch CD, Watson J. 1992. The distribution and ecology of *Otomys sloggetti* (Mammalia: Rodentia) with notes on its taxonomy. *Navorsinge van die Nasionale Museum Bloemfontein* **8:** 141–158.

Matthews T, Denys C, Parkington JE. 2005. The palaeoecology of the micromammals from the late middle Pleistocene site of Hoedjiespunt 1 (Cape Province, South Africa). *Journal of Human Evolution* **49:** 432–451.

Mayden RL. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge ME, Dawah HA, Wilson MR, eds. *Species: the units of biodiversity*. New York: Chapman and Hall, 381–424.

Mayr E. 1963. *Animal species and evolution*. Cambridge, MA: Harvard University Press.

Meester JAJ, Taylor PJ, Contrafatto G, Campbell GK, Willan K, Lamb JM, Pillay N. 1992. Chromosomal speciation in Southern African Otomyinae (Rodentia: Muridae): a review. *Durban Museum Novitates* **17:** 58–63.

de Menocal P. 2004. African climate change and faunal evolution during the Pliocence–Pleistocene. *Earth and Planetary Science Letters* **220:** 3–24.

Michaux JR, Libois R, Paradis E, Filippucci MG. 2004. Phylogeographic history of the yellow-necked fieldmouse (*Apodemus flavicollis*) in Europe and in the near and middle east. *Molecular Phylogenetics and Evolution* **32:** 788–798.

Mitsainas GP, Tryfonopoulos GA, Thanoub EG, Bisac R, Fraguedakis-Tsolisb SE, Chondropoulos BP. 2009. New data on the distribution of *Mus spicilegus* Petenyi, 1882 (Rodentia, Muridae) and a distinct mtDNA lineage in the southern Balkans. *Mammalian Biology* **74:** 351–360.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology* **58:** 298–311.

Moreau RE. 1962. Vicissitudes of the African biomes in the late Pleistocene. *Proceedings of the Zoological Society London* **141:** 395–421.

Mucina L, Rutherford MC. 2006. *The vegetation of South Africa, Lesotho and Swaziland*. Strelitzia 19. Pretoria: South African National Biodiversity Institute.

Mullin SK, Pillay N, Taylor PJ. 2004. Cranial variation and geographic patterns within the *Dasymys rufulus* complex (Rodentia: Muridae). *Journal of Mammalogy* **85:** 911–923.

Nylander JAA, Ronquist F, Heulsenbeck JP, Nieves-Alderey JL. 2004. Bayesian phylogenetic analysis of combined data. *Systematic Biology* **53:** 47–67.

Pääbo S, Wilson AC. 1988. Polymerase chain reaction reveals cloning artifacts. *Nature* **334:** 387–388.

Palumbi SR, Martin AP, Romano S, McMillan WO, Stice L, Grabowski G. 1991. *The simple fool's guide to PCR*. Honolulu, HI: Department of Zoology Special Publication, University of Hawaii.

Petit E, Balloux F, Excoffier L. 2000. Mammalian population genetics: why not Y? *Trends in Ecology and Evolution* **17:** 28–33.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231–259.

Piel WH, Nutt KJ. 2000. One species or several? Discordant patterns of geographic variation between allozymes and mtDNA sequences among spiders in the genus *Metepeira* (Araneae: Araneidae). *Molecular Phylogenetics and Evolution* **15:** 414–418.

Pillay N, Willan K, Meester J. 1992. Post-zygotic reproductive isolation in the African vlei rat *Otomys irroratus* (Muridae: Otomyinae). *Israel Journal of Zoology* **38:** 307–313.

Pillay N, Willan K, Cooke J. 1995. Evidence of pre-mating reproductive isolation in two populations of the vlei rat *Otomys irroratus*: experiments of inta- and interpopulation male-female encounters. *International Journal of Mammalian Biology* **60:** 352–360.

Pocock TN. 1976. Pliocene mammalian microfauna from Langebaanweg: a new fossil genus linking the Otomyinae with the Murinae. *South African Journal of Science* **72:** 58–60.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinfomatics* **14:** 817–818.

Posada D, Crandall KA, Templeton AL. 2000. GeoDIS. A program for nested cladistic analysis of the geographic distance of genetic haplotypes. *Molecular Ecology* **9:** 487–488.

Rambau RV, Elder FFB, Robinson TJ. 2001. Chromosomal evolution in the vlei rat, *Otomys irroratus* (Muridae: Otomyinae): a compound chromosomal rearrangement separates two major cytogenetic groups. *Cytogenetic and Cell Genetics* **93:** 253–257.

Rambau RV, Robinson TJ, Stanyon R. 2003. Molecular genetics of *Rhabdomys pumilio* subspecies boundaries: mtDNA phylogeography and karyotypic analysis by fluorescence *in situ* hybridization. *Molecular Phylogenetics and Evolution* **28:** 564–575.

Roberts A. 1929. New forms of African mammals. *Annals of the Transvaal Museum* **13:** 82–121.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9:** 552–569.

Sénégas F. 2001. Interpretation of the dental pattern of the South African fossil *Euryotomys* (Rodentia, Murinae) and origin of otomyine dental morphology. In: Denys C, Granjon L, Poulet A, eds. *African small mammals*. Paris: IRD, 151–160.

Sénégas F, Avery M. 1998. New evidence for the murine origins of the Otomyinae (Mammalia, Rodentia) and the age of Bolt's Farm (South Africa). *South African Journal of Science* **94:** 503–507.

Sites JW Jr, Marshall JC. 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution* **18:** 462–470.

Skinner JD, Chimimba CT. 2005. *The mammals the Southern African Subregion*, 3rd edn. Cambridge: Cambridge University Press, 174–175.

Smit HA, Robinson TJ, Jansen van Vuuren B. 2007. Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Molecular Ecology* **16:** 2680–2692.

Stuart BL, Dugan KA, Allard MW, Kearney M. 2006. Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. *Systematic Biology* **4:** 133–136.

Swart BL, Tolley KA, Matthee CA. 2009. Climate change drives speciation in the Southern rock Agama (*Agama atra*) in the Cape Floristic Region, South Africa. *Journal of Biogeography* **36:** 78–87.

Swofford DL. 2002. *PAUP* phylogenetic analysis using parsimony (and other methods), version 4.10*. Champaign, IL: Illinois Natural History Survey.

Taylor PJ, Campbell GK, Van Dyk D, Meester J, Willan K. 1992. Genetic variation in the African vlei rat *Otomys irroratus*. *Israel Journal of Zoology* **38:** 293–305.

Taylor PJ, Contrafatto G, Willan K. 1994. Climatic correlates of chromosomal variation in the African vlei rat, *Otomys irroratus*. *Mammalia* **58:** 623–634.

Taylor PJ, Rautenbach IL, Gordon DH, Sink K, Lotter P. 1995. Diagnostic morphometrics and southern African distribution of two sibling species of tree rat, *Thallomys paedulcus* and *Thallomys nigricauda* (Rodentia: Muridae). *Durban Museum Novitates* **20:** 49–62.

Taylor PJ, Denys C, Mukerjee M. 2004. Phylogeny of the African murid tribe Otomyini (Rodentia), based on morphological and allozyme evidence. *Zoologica Scripta* **33:** 389–402.

Taylor PJ, Maree S, van Sandwyk J, Kerbis Peterhans JC, Stanley WT, Verheyen E, Kaliba P, Verheyen W, Kaleme P, Bennett NC. 2009a. Speciation mirrors geomorphology and palaeoclimatic history in African laminate-toothed rats (Muridae: Otomyini) of the *Otomys denti* and *O. lacustris* species-complexes in the 'Montane Circle' of East Africa. *Biological Journal of the Linnean Society* **96:** 913–941.

Taylor PJ, Maree S, van Sandwyk J, Baxter R, Rambau RV. 2009b. When is a species not a species? Uncoupled phenotypic, karyotypic and genotypic divergence in two species of South African laminate-toothed rats (Murinae: Otomyini). *Journal of Zoology* **277:** 317–332.

Templeton AR. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* **7:** 381–397.

Templeton AR. 2004. Statistical phylogeography: methods of evaluating and minimizing inference error. *Molecular Ecology* **13:** 798–809.

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred

from restriction endonuclease mapping and DNA sequence data. III Caladogram estimation. *Genetics* **132:** 619–633.

Templeton AR, Routman E, Phillips CA. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* **140:** 767–782.

Tolley KA, Chase BM, Forest F. 2008. Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern Africa chameleons. *Journal of Biogeography* **35:** 1402–1414.

Tolley KA, Makokha JS, Houniet DT, Swart BL, Matthee CA. 2009. The potential for predicted climate shifts to impact genetic landscapes of lizards in the South African Cape Floristic Region. *Molecular Phylogenetics and Evolution* **51:** 120–130.

Vega R, Vazquez-Domingueza E, Mejia-Puentea A, Cuaron AD. 2007. Unexpected high levels of genetic variability and the population structure of an island endemic rodent (*Oryzomys couesi cozumelae*). *Biological Conservation* **137:** 210–222.

Willows-Munro S, Matthee CA. 2009. The evolution of the southern African members of the shrew genus *Myosorex*: understanding the origin and diversification of a morphologically cryptic group. *Molecular Phylogenetics and Evolution* **51:** 394–398.

Yuasa T, Nagata J, Hamasaki S, Tsuruga H, Furubayashi K. 2007. The impact of habitat fragmentation on genetic structure of the Japanese sika deer (*Cervus nippon*) in southern Kantoh, revealed by mitochondrial d-loop sequences. *Ecological Research* **22:** 97–106.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Mismatch distribution for the cyt *b* gene of the two *O. irroratus* clades.
Figure S2. A haplotype network of 69 haplotypes retrieved from 102 sequences at the 95% confidence interval.
Table S1. The frequency of haplotypes ($N = 102$ specimens obtained from 40 localities)
Table S2. Nested clade analyses results for 41 populations of *Otomys irroratus* collected throughout South Africa
Table S3. Genetic diversity and neutrality test estimates for the four sub-clades of *O. irroratus* retrieved in Figure 2.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.