Chapter 3

Phylogenetic relationships within *Micaelamys namaquensis* (Rodentia: Muridae) from southern African as inferred from mitochondrial and nuclear genes

Abstract

Evolutionary relationships among members of the murid rodent genus *Micaelamys* (formerly allocated to the subgenus *Micaelamys* within the genus *Aethomys*) are poorly understood. Here I extend my research on the Namaqua rock mouse, *M. namaquensis* from southern Africa by combing existing partial mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) sequences with the nuclear Recombination Activating Gene 1 (RAG1) gene to examine the relationships among 11 of 14 recently identified phylogroups. Cytochrome *b* sequence divergence values ranged between 0.32% and 7.68% while divergence values ranged between 0.08% and 1.44% for the RAG1 gene. Incongruence was shown between the cyt *b* and the combined analyses versus the independent RAG1 analysis possibly as a result of incomplete lineage sorting in the nuclear gene. The combined molecular data supports the polytypic nature of *M. namaquensis* and the fact that most lineages are associated with specific vegetation types of southern Africa.

1. Introduction

African rock rats of the genera *Aethomys* Thomas, 1915, and *Micaelamys* Ellerman, 1941 (see Chimimba and Bennett, 2005), are a diverse group of murid rodents endemic to East, Central and southern Africa and extending marginally into West Africa (Musser and Carleton, 2005). Phylogenetic relationships between these genera and other African murids are uncertain (Musser and Carleton, 2005). Characteristics of the genus (*Aethomys*, *senso lato*) overlap to some extent with *Rattus* and *Arvicanthis* (Ellerman, 1941) added to which the genus has also variously been considered to be closely related to *Mus*, *Mastomys*, *Thallomys*, *Zelotomys* (De Graaff, 1981), *Stochomys*, *Dephomys*, *Dasymys* and *Pelomys* (Bonhomme et al., 1985; Denys, 1990a; b).

Prior to the generic recognition of the genus *Micaelamys* (Chimimba and Bennet, 2005), a phylogenetic study of the genus *Aethomys* using cladistic analysis of all recognised species at that time, based on cranial data suggested the presence of three clades: 1) *A. bocagei* Thomas, 1904, *A. thomasi* De Winton, 1897, *A. silindensis* Roberts, 1938, *A. kaiseri* Noack, 1887, and *A. nyikae* Thomas, 1897; 2) *A. chrysophilus* De Winton, 1897, *A. ineptus* Thomas and Wroughton, 1908, and *A. hindei* Thomas, 1902; 3) *A. granti* Wroughton, 1908, *A. namaquensis* Smith, 1834, and *A. stannarius* Thomas, 1913 (Chimimba, 2005). This study also revealed a sister-taxon relationship between the two cryptic species, *A. chrysophilus* and *A. ineptus* (Chimimba, 2005). In contrast, molecular studies by Ducroz et al. (2001) and Lecompte et al. (2008) revealed no sister-taxon relationship between individuals from *Aethomys* and *Micaelamys*.

Attention was drawn to the close evolutionary relationship between *A. namaquensis* and *A. granti* (currently referred to as *M. namaquensis* and *M. granti*; see Chimimba and Bennett 2005). Similarly, this close relationship has been documented in other studies using dental morphology (Ellerman et al., 1953), karyology (Visser and Robinson, 1986), gross sperm and bacular morphology and their staining properties (Visser and Robinson, 1987) and phenetic analysis (Chimimba et al., 1999). *Micaelamys namaquensis* and *M. granti* were well separated from all other species of *Aethomys*, providing good grounds for the subgeneric separation of *Aethomys* and *Micaelamys* (Chimimba, 2005). Similarly, Davis

(1975) supported the view of a close affinity between *A. namaquensis* and *A. granti* and their inclusion within the subgenus *Micaelamys* (Ellerman et al., 1953).

Palaeontological data of some members of the genera *Aethomys* and *Micaelamys* have been recorded from South Africa (Avery, 1981, 1982, 1985; De Graaff, 1960, 1961; Hendey, 1981; Pocock, 1987) where two Pliocene fossil species, *A. adamanticola* and *A. modernis*, the oldest known representatives of the genus in Africa, have been described (Denys, 1990a, b). *Aethomys adamanticola* shows characteristics reminiscent of *A. namaquensis* (*Micaelamys namaquensis* as currently understood) while *A. modernis* is very similar to *A. chrysophilus* (*sensu lato*) (Denys, 1990a, b). *Aethomys adamanticola* may also represent an advanced stage of an Early Miocene lineage closely related, but not ancestral to *Dasymys* (Denys, 1990a, b). Other palaeontological records include two East African Plio-Pleistocene fossil species, *A. lavocati* (Jaeger, 1976, 1979) and *A. deheinzelini* (Black and Krishtalka, 1986; Denys, 1987; Wesselman, 1984).

Earlier reports (e.g., Roberts, 1951; Meester et al., 1964) recognised 16 subspecies within *M. namaquensis.* Prior to an intraspecific morphometric study within *M. namaquensis* (Chimimba, 2001), the extent of geographic variation within the species remained unknown. Chimimba (2001) suggested the recognition of only four subspecies within *M. namaquensis*: *M. n. namaquensis*, Smith, 1834, *M. n. lehocla* Smith, 1836, *M. n. monticularis* Jameson, 1909 and *M. n. alborarius* Peters, 1852, each associated with one of four of the major biomes of southern Africa (Chimimba, 2001). My subsequent analysis of mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) diversity across the distributional range of the species (Russo, 2003; Chapter 2) confirmed that *M. namaquensis* is polytypic. Fourteen distinct evolutionary lineages were described and most appeared to be associated with major biomes/bioregions of southern Africa. The phylogenetic analysis could only resolve a few of the deeper nodes based on this single mtDNA marker. Coalescent-based dating suggested a major radiation over a relatively short evolutionary period at the end of the Miocene and the beginning of the Pliocene.

In the present study, I utilised a fragment of the Recombination Activating Gene 1 (RAG1) in combination with the protein coding cyt *b* gene in an attempt to obtain a more resolved phylogeny. These genes were used in the present study as phylogenetic analyses of several short stretches from different genes, on average, show a better performance than analyses

based on nearby sites from a single gene fragment (Cummings et al., 1995). Previous studies have shown the RAG1 gene to be useful for phylogenetic analysis of rodents (e.g., Steppan et al., 2004a, b; Steppan et al., 2005; Suzuki et al., 2004). The protein encoded by this gene is involved in activation of immunoglobulin V-D-J recombination (Wenhui et al., 2001). Similarly cyt *b* sequences, when used carefully, have contributed to the investigation of the phylogenetic relationships among murids (e.g., Ducroz et al., 2001; Galewski et al., 2006; Verheyen et al., 1995; Verheyen et al., 1996). In combining RAG1 and cyt *b* characters, it was considered that the latter gene with its faster mutation rate would resolve relationships near the terminal nodes relative to the slower nuclear RAG1 characters which may resolve deeper nodes due to lesser effects of saturation and consequent homoplasy (Suzuki et al., 2004).

A lack of resolution in a phylogenetic tree is represented as a polytomy (i.e., uncertainty about relationships between lineages/groups) in which three or more lineages/groups diverge from a single node. In general, systematics requires resolved trees, which may yield stronger inferences about character evolution and the relationships and biogeography of the species under investigation (McCracken and Sorenson, 2005). The resolution of a polytomy at the species level can be complicated by the incongruence of individual gene trees and the species tree (i.e., the true history of diversification). This lack of resolution in phylogenetic analyses could be as a result of incomplete lineage sorting where the lineage sorting will take longer (3-4 times longer) in any nuclear gene compared to lineage sorting in a mitochondrial gene due to differences in the effective population sizes of the respective genomes (McCracken and Sorenson, 2005; Moore, 1997; Palumbi et al., 2001).

The present study, therefore, represents a phylogenetic analysis using independent cyt *b* and RAG1 sequence data and a combination of both sets of data, based on material from southern Africa and addresses the following questions: 1) What are the phylogenetic relationships among lineages of the southern African *M. namaquensis* (as identified in Chapter 2) based on molecular and nuclear data?; and 2) Does the nuclear RAG1 gene contribute to the overall resolution of the phylogenetic relationships among lineages within *M. namaquensis*?

2. Materials and Methods

2.1 Study area and sampling

Thirty-five individuals representative of 11 of the newly identified lineages of *M. namaquensis* from southern Africa were selected for the phylogenetic analysis (Fig. 3.1; Appendix 3.1). Some of the RAG1 amplifications were unsuccessful and I therefore only included 35 samples. Individuals for the remaining lineages (lineage A, F and L) were not included since, as mentioned before, the RAG1 amplifications were unsuccessful. For sampling protocols, permit numbers and detailed locality information refer to Chapter 2.

2.2 DNA extraction, Polymerase Chain Reaction (PCR) amplification and sequencing

The cyt *b* sequences described in Chapter 2 were re - analysed in the present study. To amplify and sequence the targeted fragment in the RAG1 gene, the primers S70 (5' TCC GAG TGG AAA TTT AAG MTG TT 3'; modified from R13 of Groth and Barrowclough, 1999) and S73 (5' GAG GAA GGT RTT GAC ACG GAT G 3'; Steppan et al., 2004b) were used to amplify the region in five individuals. These *M. namaquensis* sequences were aligned in Clustal X (Thompson et al., 1997) and used to design internal species-specific *M. namaquensis* primers, RAGNam IL (5' GCG TAG GCT CAG CAG CAA GGA 3') and RAGNam IH (5' GAT TTC ACA AAG TGT GCA GGG 3'). The targeted fragment of the remaining individuals was amplified using primers S70/RAGNam IH and S73/RAGNam IL.

Polymerase Chain Reactions (PCR; Saiki et al., 1988) of the RAG1 gene consisted of denaturing at 94° C for 5 min, 35 cycles of the following: 94° C for 30 seconds, primer annealing at 53.5° C (S70/RAGNam IH) and 57° C (S73/RAGNam IL) for 30 seconds and elongation at 72° C for 45 seconds. This was followed by an extended elongation step for 7 minutes at 72° C in a Geneamp® PCR System 9700 (Applied Biosystems). PCR products were purified using the High Pure™ PCR Product Purification Kit (Roche Diagnostics) as prescribed by the manufacturers.

Dye-terminator cycle sequencing was performed with S70/RAGNam IH and S73/RAGNam IL to obtain a 1309 bp fragment of the RAG1 gene. Nucleotide sequences

were determined using an ABI 3130 automated sequencer (Applied Biosystems). Cycle sequencing products were subsequently precipitated using a NaAc salt method.

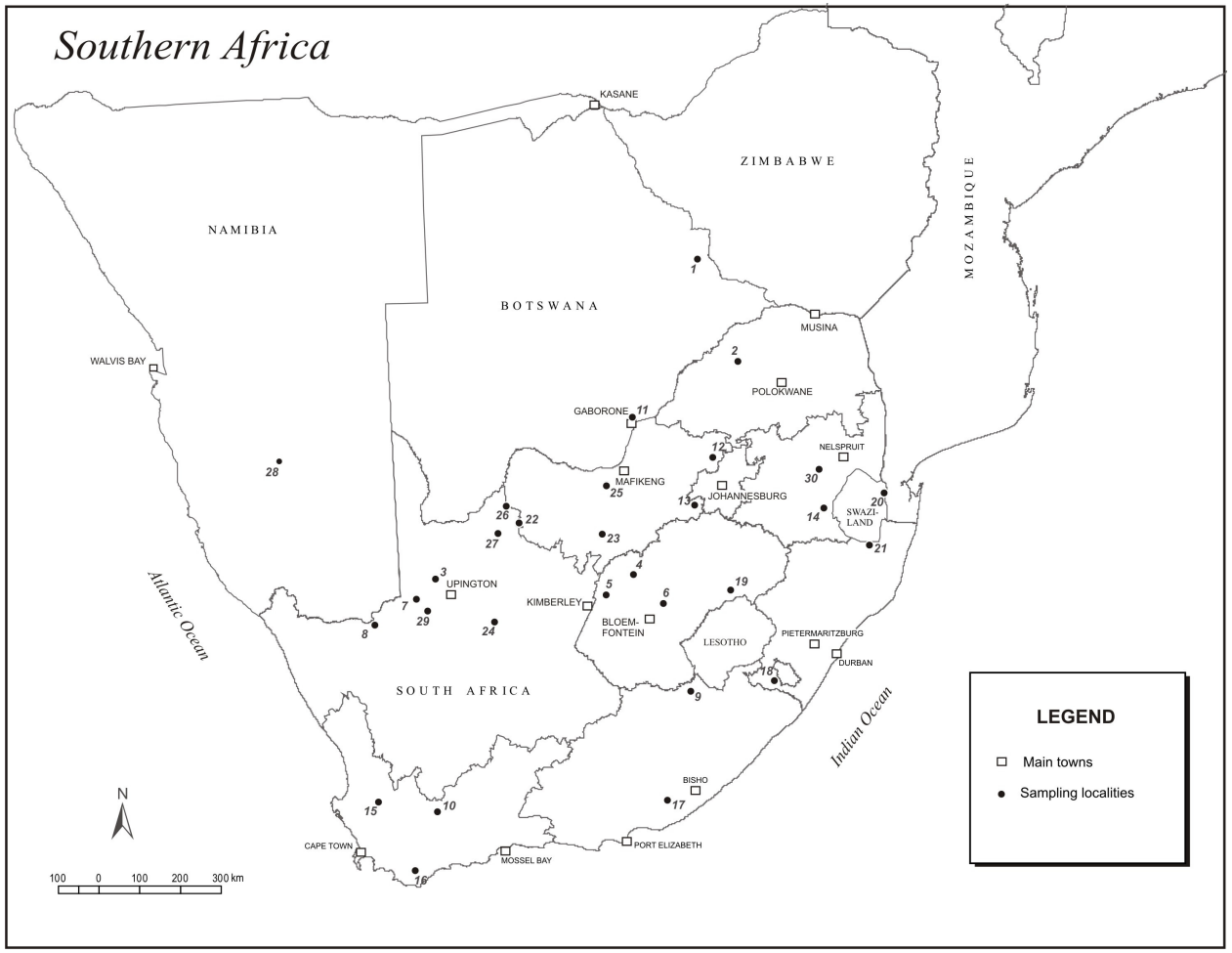

Figure 3.1 Collecting localities of *Micaelamys namaquensis* in South Africa, Swaziland, Botswana and Namibia, representing individuals with sequences for both the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) gene and the nuclear Recombination Activating Gene 1 (RAG1) gene. Numbers correspond to the locality numbers in Appendix 3.1. Also see Appendix 3.1 for locality names.

2.3 Sequencing analysis

The individual RAG1 sequences were imported into either Sequence Navigator, version 1.01 (PE Applied Biosystems) or Vector NTI Advance 10 (Invitrogen). A consensus sequence of each individual was computed by aligning the forward and reverse sequences. The GenBank accession numbers of the relevant samples from Chapter 2 and the newly acquired RAG1 sequences are given in Appendix 3.2. Consensus sequences of all individuals were aligned in Clustal X (Thompson et al., 1997) and sequences were subsequently imported into PAUP, version 4.0b10 (Swofford, 2003) and MRBAYES, version 3.1.2 (Ronquist and Huelsenbeck, 2003) for phylogenetic analyses. Sequences were also imported into MacClade, version 3.0 (Maddison and Maddison, 1992) to translate the cyt *b* and RAG1 nucleotide sequences into amino acids. Amino acid sequences were examined for irregularities (e.g., stop codons in the reading frame) while the transition:transversion (Ti:Tv) ratio was also estimated using MacClade, version 3.0 (Maddison and Maddison, 1992).

2.4 Outgroup choice

The selection of possible outgroups for *M. namaquensis* was difficult because of the previous evolutionary relationships that have been proposed between *Micaelamys* and numerous other murids (see Chimimba, 2005 and references therein). In a preliminary analysis *A. chrysophilus*, *A. ineptus*, *Parotomys brantsi*, *Dasymys incomtus*, *Rattus rattus*, *M. musculus*, *Rhabdomys pumilio* and *Arvicanthis somalicus* were used as outgroups. *Rhabdomys pumilio* was used as outgroup in the PAUP, version 4.0b10 and MRBAYES, version 3.1.2 (Ronquist and Huelsenbeck, 2003) analyses. Sequences for both the cyt *b* and RAG1 genes were available for *R. pumilio* and the species has been shown to be closely related to *M. namaquensis*.

2.5 Phylogenetic analysis

Since the data included sequences from one mtDNA gene and one nuclear gene, two analytical strategies were possible for the treatment of these data due to the controversy surrounding the merits of these two approaches (De Queiroz et al., 1995; Huelsenbeck et al., 1996). The first approach was to analyse the different datasets separately and to construct a consensus tree from these separate analyses, but this approach (Adams, 1972) is considered to be more appropriate if the datasets are heterogeneous. The second approach was to perform the analysis directly using the combined data. This “combined

approach" is considered to often provide phylogenies that are more resolved than consensus trees from separately analysed data (De Queiroz, 1993). The data in the present study were analysed separately and combined.

Modeltest, version 3.06 (Posada and Crandall, 1998) was used to determine optimal substitution models identified by the Akaike Information Criterion (AIC) for the combined and separate datasets. Parameters such as the shape parameter of the gamma distribution of rates among sites (Yang 1996; Yang et al., 1994) and the proportion of invariable sites (I) were also estimated. The chosen model based on only 36 sequences (a subset of individuals representing the diversity within *M. namaquensis,* including the outgroup) was subsequently used in maximum likelihood (ML; Felsenstein, 1973; 1981) analyses as implemented in PAUP, version 4.0b10 (Swofford, 2003) and Bayesian Inference (BI; Ronquist and Huelsenbeck, 2003) phylogenetic analyses. Base frequencies were also estimated in PAUP, version 4.0b10 (Swofford, 2003).

Three independent ML analyses (the two genes as independent datasets and a combined analysis) were conducted using PAUP, version 4.0b10 (Swofford, 2003). The ML analyses were conducted using 100 random addition replicates and were based on a heuristic search using the tree bisection-reconnection (TBR) option with nucleotides as unordered characters. Tree nodal support was assessed by 1 000 bootstrap replicates (Felsenstein, 1985) performed on a computer cluster located at the University of Pretoria. Three independent BI analyses were conducted. Analyses with four chains were run for 5×10^6 generations using random starting trees. Trees and parameters were recorded every 100 generations. Two runs were performed simultaneously and split frequencies were compared every 100^{th} generation to ensure convergence of the runs. All runs used the default heating and swap parameters. The first 5 000 generations (10% burn-in) were excluded as the "burn-in". A 10% burn-in was sufficient to ensure that trees were only sampled from the region of stationarity.

A maximum parsimony tree (Kluge and Farris, 1969; Farris et al., 1970) was generated in PAUP for the combined data, using nucleotides as unordered characters and the tree bisection-reconnection (TBR) method were used to construct the shortest tree. A strict consensus tree was constructed if more than one minimum length tree were obtained. The following were reported: tree-length, consistency index (CI; Kluge and Farris, 1969),

retention index (RI; Farris, 1989), and rescaled consistency index (RC; Archie, 1989; Meier et al., 1991). Support values for internal nodes were determined using bootstrap analysis with 1 000 iterations (Felsenstein 1985).

3. Results

3.1 Sequence statistics

A total of 631 bp of the 5' end of the cyt *b* gene and 1309 bp of the 5' end of the RAG1 gene were analysed. Sequences obtained were either from the mitochondrial cyt *b* or the RAG1 gene since no stop codons were found when sequences were translated into the expected 210 and 436 amino acids, respectively. Of the 112 variable sites in the cyt *b* gene, 78 were phylogenetically informative (Table 3.1). A total of 12 individuals were heterozygous for sites in the RAG1 gene, with a maximum of 8 polymorphic sites in individuals NNH03 and NNH09. Polymorphic sites were included in all analyses and were treated as heterozygous sites. Of the 29 variable positions in the RAG1 gene, only 15 were phylogenetically informative (Table 3.1). The number of invariant site detected in the present study was noticeably higher. Most of the substitutions were silent with four variable amino acid sites between *M. namaquensis* individuals and 13 variable amino acids between the ingroup and the outgroup (*Rhabdomys pumilio*) individual for the cyt *b* gene. Six variable amino acids were detected between *M. namaquensis* individuals and 12 variable amino acids were detected between the ingroup taxa and the outgroup individual for the RAG1 gene.

Mean base compositions at the three different codon positions of the cyt *b* gene and for the RAG1 gene fragment for the whole sample examined are provided in Table 3.1 (excluding the outgroup individual). The overall base composition of cyt *b* showed that the four nucleotides do not occur in equal frequencies, and is similar to that of other mammalian cyt *b* sequences reported in the literature (Ducroz et al., 1998; Irwin et al., 1991). This strong bias in base composition showed an under-representation of guanine at both second (15.70%) and third (2.72%) codon positions. Similar to the sequence statistics in the larger dataset (Chapters 2), a higher representation of adenine at the third codon positions (44.96%) and thymine at the second codon positions (41.12%) was also observed in the present study. RAG1 has a nearly equal average base composition (see overall base

composition in Table 3.1). This nearly equal base composition is partially due to differing composition bias across codon position and gene regions that balance each other (i.e., the divergent versus the conserved regions).

Nucleotide divergence estimates were reported as Tamura-Nei distances with a gamma correction (Gu and Zang, 1997), and a proportion of invariable sites (TrN + Γ + I) for both the cyt *b* and RAG1 genes. Pairwise estimates of the nucleotide sequence divergence for the cyt *b* and the RAG1 gene are indicated in Table 3.2. Corrected cyt *b* Tamura-Nei sequence divergence values ranged between 0.00% and 8.11% within *Micaelamys* and sequence divergence values ranged between 0.08% and 1.44% for the RAG1 gene. Therefore, the RAG1 gene showed less variability as would be expected. Interestingly, some pairwise comparisons between individuals showed no sequence divergence based on the cyt *b* gene but when using the RAG1 gene, differences were detected.

Table 3.1 Variable sites and the average percentage base composition in a 631 bp fragment of the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) gene (A) and in the 1309 bp fragment of the nuclear Recombination Activating Gene 1 (RAG1) gene (B) within *Micaelamys namaquensis* from southern Africa.

A

CYTOCHROME *B*:	VARIABLE SITES	PHYLOGENTIC INFORMATVE	A	G	C	T
First codon position	13	9	28.87	23.31	17.55	30.27
Second codon position	4	2	18.34	15.70	24.84	41.12
Third codon position	95	67	44.96	2.72	32.98	19.33
Overall	112	78	30.72	13.92	25.12	30.24

B

RAG1:	VARIABLE SITES	PHYLOGENTIC INFORMATIVE	A	G	C	T
First codon position	9	2	31.26	25.37	25.69	17.68
Second codon position	9	1	33.99	22.46	18.31	25.24
Third codon position	44	12	21.05	26.15	31.27	21.53
Overall	62	15	28.77	24.66	25.08	21.49

Table 3.2 Pairwise estimates of percent Tamura-Nei (TrN) + Γ (1.97) + I (0.66) cytochrome *b* (cyt *b*) sequence divergence (below the diagonal) and percent Tamura-Nei (TrN) + Γ (0.87) + I (0.83) Recombination Activating Gene 1 (RAG1) sequence divergence (above the diagonal) between 35 *Micaelamys namaquensis* individuals from southern Africa. Outgroup individual: *Rhab* = *Rhabdomys pumilio*. Names follow that of the RAG1 alleles (Appendix 3.2). Corrected cyt *b* sequence divergence values for the outgroup are not given since undefined distances were obtained.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1	NNH01	-	0.55	0.08	0.24	0.08	0.32	0.33	0.31	0.32	0.24	0.08	0.33	0.32	0.42	0.42	0.47	0.00	0.24	0.24	0.57	0.40	0.34	0.41	0.43	0.24	0.24	0.16	0.16	0.16	0.33	0.23	0.08	0.16	0.18	0.65	4.50
2	NNH02	2.12	-	0.94	0.95	0.82	1.16	1.04	1.04	1.16	0.94	0.94	0.88	1.03	1.16	1.16	0.68	0.74	0.96	0.95	1.11	1.06	1.16	1.26	0.95	0.93	0.93	0.85	0.85	0.75	1.16	0.93	0.93	0.75	0.75	1.38	5.86
3	NNH03	5.96	4.97	-	0.24	0.24	0.49	0.24	0.48	0.50	0.41	0.16	0.54	0.41	0.25	0.25	0.48	0.08	0.50	0.52	0.89	0.57	0.56	0.59	0.44	0.24	0.24	0.24	0.32	0.24	0.52	0.32	0.16	0.32	0.27	0.45	4.63
4	NNH04	5.95	4.97	0.00	-	0.24	0.49	0.08	0.48	0.49	0.41	0.16	0.52	0.32	0.33	0.33	0.18	0.16	0.49	0.51	0.88	0.57	0.54	0.58	0.34	0.00	0.00	0.24	0.32	0.24	0.51	0.24	0.08	0.32	0.27	0.56	4.45
5	NNH05	5.99	5.00	0.00	0.00	-	0.32	0.24	0.16	0.33	0.16	0.08	0.35	0.32	0.59	0.58	0.47	0.08	0.08	0.08	0.36	0.24	0.36	0.41	0.51	0.24	0.24	0.16	0.08	0.08	0.34	0.23	0.08	0.08	0.19	0.85	4.73
6	NNH06	5.99	4.97	0.32	0.33	0.33	-	0.58	0.16	0.00	0.08	0.32	0.00	0.00	0.88	0.87	0.59	0.41	0.33	0.34	0.68	0.40	0.00	0.08	0.71	0.48	0.48	0.16	0.16	0.16	0.00	0.32	0.32	0.32	0.44	1.18	5.48
7	NNH07	6.54	5.50	0.34	0.34	0.34	0.67	-	0.58	0.59	0.50	0.24	0.62	0.40	0.33	0.33	0.27	0.24	0.60	0.61	0.88	0.68	0.64	0.67	0.43	0.08	0.08	0.32	0.41	0.32	0.60	0.32	0.16	0.41	0.36	0.55	4.77
8	NNH08	6.22	5.20	0.16	0.16	0.16	0.49	0.50	-	0.16	0.00	0.32	0.17	0.16	0.87	0.86	0.56	0.40	0.08	0.08	0.36	0.16	0.17	0.24	0.52	0.48	0.48	0.16	0.08	0.08	0.16	0.32	0.32	0.24	0.44	1.16	5.28
9	NNH09	6.01	4.98	0.32	0.33	0.32	0.33	0.67	0.49	-	0.08	0.33	0.00	0.00	0.88	0.87	0.59	0.41	0.33	0.34	0.68	0.41	0.00	0.08	0.71	0.49	0.49	0.16	0.16	0.16	0.00	0.32	0.32	0.33	0.44	1.18	5.60
10	NNH10	6.82	5.70	0.67	0.67	0.49	0.49	1.03	0.84	0.66	-	0.24	0.08	0.08	0.78	0.77	0.47	0.33	0.08	0.08	0.37	0.24	0.08	0.16	0.52	0.41	0.41	0.08	0.08	0.08	0.08	0.24	0.24	0.16	0.35	1.05	5.33
11	NNH11	7.09	6.01	0.67	0.67	0.67	0.66	1.03	0.84	0.67	1.00	-	0.34	0.24	0.42	0.41	0.36	0.08	0.34	0.34	0.68	0.42	0.35	0.41	0.52	0.16	0.16	0.08	0.16	0.08	0.33	0.16	0.00	0.16	0.09	0.64	4.65
12	NNH12	6.23	5.21	0.49	0.50	0.50	0.16	0.85	0.65	0.51	0.65	0.84	-	0.00	0.93	0.93	0.59	0.44	0.36	0.35	0.70	0.45	0.00	0.08	0.75	0.52	0.52	0.17	0.17	0.17	0.00	0.25	0.33	0.35	0.36	1.22	5.98
13	NNH13	7.32	7.07	5.59	5.47	5.43	6.12	6.14	5.88	5.67	6.22	6.62	6.37	-	0.69	0.69	0.37	0.33	0.32	0.33	0.67	0.40	0.00	0.08	0.52	0.32	0.32	0.16	0.16	0.16	0.00	0.24	0.24	0.24	0.33	0.94	5.31
14	NNH14	7.04	6.21	6.05	6.03	6.04	6.04	6.29	6.32	6.04	6.56	7.25	6.34	2.56	-	0.00	0.58	0.33	0.90	0.89	1.22	0.97	0.96	0.97	0.35	0.16	0.16	0.59	0.69	0.59	0.89	0.60	0.33	0.68	0.55	0.18	5.02
15	NNH15	6.71	5.92	5.80	5.79	5.85	5.78	6.02	6.05	5.80	6.57	6.93	6.06	2.54	0.00	-	0.58	0.33	0.89	0.89	1.22	0.96	0.96	0.96	0.35	0.16	0.16	0.59	0.68	0.59	0.89	0.59	0.33	0.68	0.55	0.18	4.94
16	NNH16	6.54	6.00	5.08	5.07	5.08	5.09	5.55	5.34	5.09	5.59	6.19	5.36	2.06	1.50	1.49	-	0.37	0.58	0.58	0.92	0.58	0.59	0.69	0.38	0.17	0.17	0.28	0.38	0.28	0.59	0.17	0.26	0.38	0.38	0.78	5.72
17	NNH17	6.31	6.05	5.03	5.04	5.05	5.08	5.63	5.28	5.10	5.86	6.09	5.28	1.05	1.92	2.04	1.64	-	0.33	0.34	0.69	0.49	0.46	0.50	0.34	0.16	0.16	0.16	0.24	0.16	0.43	0.24	0.08	0.16	0.18	0.54	4.45
18	NNH18	6.55	5.65	6.42	6.38	6.39	6.47	6.88	6.73	6.48	7.01	7.68	6.74	5.11	4.81	4.77	4.65	4.51	-	0.00	0.28	0.16	0.37	0.42	0.52	0.49	0.49	0.16	0.00	0.00	0.35	0.32	0.33	0.16	0.46	1.18	5.44
19	NNH19	6.09	5.15	5.96	5.93	5.94	5.95	6.18	6.25	5.96	6.17	7.21	6.27	5.77	5.29	5.30	5.03	4.86	0.38	-	0.28	0.17	0.37	0.42	0.52	0.51	0.51	0.17	0.00	0.00	0.34	0.34	0.34	0.17	0.46	1.18	5.59
20	NNH20	6.40	5.21	5.82	5.87	5.82	5.88	6.49	6.10	5.88	5.77	7.02	6.12	5.00	4.92	4.81	4.54	4.40	0.54	0.37	-	0.38	0.68	0.78	0.88	0.88	0.88	0.48	0.28	0.28	0.68	0.60	0.68	0.49	0.79	1.44	6.52
21	NNH21	7.05	5.26	4.50	4.50	4.52	4.96	5.03	4.74	4.56	5.48	5.50	5.18	4.92	5.19	4.99	4.34	4.88	3.41	3.92	3.17	-	0.36	0.50	0.61	0.57	0.57	0.24	0.16	0.16	0.43	0.40	0.40	0.32	0.46	1.18	4.92
22	NNH22	7.86	5.86	5.41	5.40	5.42	5.42	5.97	5.66	5.43	6.18	6.46	5.67	5.16	5.18	5.00	4.78	4.76	3.38	3.87	3.15	1.01	-	0.09	0.75	0.54	0.54	0.18	0.18	0.18	0.00	0.34	0.34	0.37	0.46	1.18	6.36
23	NNH23	6.19	5.67	6.20	6.19	6.23	6.27	6.82	6.50	6.29	7.15	7.40	6.50	4.71	4.70	4.51	4.36	3.83	4.48	5.08	4.15	3.80	3.78	-	0.80	0.58	0.58	0.25	0.25	0.25	0.08	0.41	0.41	0.41	0.54	1.28	5.79
24	NNH24	5.52	4.05	3.72	3.71	3.74	3.71	3.93	3.94	4.21	4.16	4.72	3.95	3.88	3.09	3.00	2.89	3.28	3.71	3.94	3.26	3.21	3.21	3.73	-	0.32	0.32	0.43	0.43	0.34	0.71	0.42	0.42	0.52	0.65	0.56	5.48
25	NNH25	5.91	5.36	5.00	4.98	4.73	5.06	5.44	5.26	5.50	5.24	6.09	5.27	3.36	3.21	3.30	2.93	3.26	3.05	3.49	2.88	4.13	4.14	3.24	1.30	-	0.00	0.24	0.32	0.24	0.51	0.24	0.08	0.32	0.27	0.34	4.50
26	NNH26	5.78	5.23	4.87	4.85	4.85	4.93	5.32	5.14	5.38	5.41	5.97	5.15	3.40	2.75	2.72	2.49	2.89	3.46	3.96	3.34	3.58	3.58	3.12	0.74	0.70	-	0.24	0.32	0.24	0.51	0.24	0.08	0.32	0.27	0.34	4.50
27	NNH27	7.34	6.11	3.94	3.80	3.79	4.38	4.33	4.18	3.99	4.21	4.76	4.60	3.95	5.49	5.30	4.60	4.17	4.68	4.76	3.80	3.57	3.97	5.03	4.21	4.39	4.27	-	0.00	0.00	0.17	0.08	0.08	0.08	0.17	0.85	4.79
28	NNH28	8.11	6.75	4.72	4.71	4.71	5.22	5.16	4.98	4.76	4.91	5.79	5.48	4.83	5.78	5.70	5.37	4.81	4.88	4.98	4.18	4.41	4.36	5.52	4.57	4.49	4.52	0.53	-	0.00	0.17	0.16	0.16	0.08	0.26	0.96	4.93
29	NNH29	7.52	6.31	4.13	4.00	4.00	4.57	4.52	4.36	4.17	4.39	4.95	4.79	4.14	6.29	6.03	5.31	4.74	4.42	4.47	3.57	3.75	4.20	5.24	4.43	4.15	4.50	0.51	0.71	-	0.17	0.08	0.08	0.00	0.17	0.86	4.79
30	NNH30	7.76	6.81	4.78	4.78	4.80	5.29	5.06	5.05	4.83	5.26	5.88	5.05	5.53	6.61	6.43	5.24	5.44	5.92	5.82	5.05	4.75	5.44	6.14	4.68	5.05	4.91	1.25	1.49	1.06	-	0.33	0.33	0.34	0.45	1.18	5.89
31	NNH31	5.26	4.72	4.60	4.47	4.24	5.08	4.99	4.85	5.08	4.88	5.47	5.32	5.91	6.34	6.64	5.58	5.82	6.31	6.21	5.79	5.83	6.81	6.94	4.49	4.62	4.78	4.90	5.18	5.11	5.51	-	0.16	0.16	0.17	0.74	5.02
32	NNH32	5.90	4.89	4.69	4.67	4.47	5.17	4.88	4.93	5.19	5.12	5.72	5.42	6.28	6.36	6.66	5.66	6.07	6.33	6.18	6.07	6.05	7.08	7.18	4.49	4.69	4.83	5.20	5.27	5.40	5.57	0.72	-	0.16	0.09	0.54	4.52
33	NNH33	5.24	4.70	4.61	4.47	4.24	5.06	4.98	4.84	5.08	4.87	5.45	5.30	5.91	6.34	6.62	5.58	5.80	6.31	6.21	5.79	5.81	6.80	6.93	4.47	4.61	4.78	4.90	5.18	5.09	5.51	0.00	0.72	-	0.26	0.96	4.91
34	NNH34	6.74	6.85	5.19	5.17	4.89	5.29	5.77	5.47	5.69	4.86	5.96	5.48	6.50	5.80	5.83	5.45	5.85	6.88	6.32	5.76	6.94	7.44	5.76	4.87	4.56	4.70	5.48	6.18	5.66	7.01	6.20	6.40	6.18	-	0.75	5.33
35	NNH35	4.92	4.49	4.11	4.10	4.10	4.14	4.35	4.09	4.60	4.61	5.15	4.37	3.98	3.20	3.17	3.01	3.21	3.55	3.64	3.38	3.62	3.62	3.65	1.75	1.88	1.53	4.32	4.63	4.56	5.36	4.07	4.10	4.07	4.77	-	5.89
36	*Rhab*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

3.2 Phylogenetic analyses

The results of saturation analyses for both cyt *b* and RAG1 are summarised in Figs. 3.2 and 3.3, respectively. First and second position changes accumulated slowly for both genes. The rate of transversional substitutions was lower than that of transitions (Ti:Tv ratio = 13:1 and 2:1 for the cyt *b* and RAG1 genes, respectively). Transitions normally outnumber transversions but this high Ti:Tv ratio in the cyt *b* gene is rather unusual and has also been reported in vlei rats of the genus *Otomys* (Maree, 2002). Transversions and transitions at all positions were not saturated (Figs. 3.2 and 3.3). This result has also been reported for other rodents (Lecompte et al., 2002).

Tamura-Nei sequence divergence (α-value for the gamma shape parameter and a value for the proportion of invariable sites) estimates were used to infer relationships among the 35 *M. namaquensis* individuals for the cyt *b* and RAG1 genes. Data for the outgroup were also included in these analyses. The average separation between *M. namaquensis* and the outgroup was estimated to be between 16.48% and 18.70% based on uncorrected p-distances (Tamura-Nei + Γ + I resulted in undefined distances) for the cyt *b* gene, while the average separation between *M. namaquensis* and the outgroup was estimated at 6.05% (range 4.45% - 6.52%) for the RAG1 gene.

The BI phylograms (Figs. 3.4, 3.5 and 3.6) depict the relationships between 11 of the 14 lineages (as identified in Chapter 2) and is characterised by short internal branches possibly due to a rapid radiation of lineages that have also been reported for other rodents. Individuals from lineage A (Koppies Dam), lineage F (Volksrust) and lineage L (Sub-Escarpment Grassland bioregion) were not included in the BI analyses (see Chapter 2). Six well-supported (posterior probability values of ≥ 0.95) lineages with clear geographical patterns were identified in the combined (cyt *b* and RAG1) analysis (Fig. 3.6): 1) B, Grassland; 2) G, Albany Thicket; 3) H, western Fynbos; 4) I, Bushmanland/Upper Karoo bioregion; 5) J, Nama-Karoo; and 6) M, Eastern Kalahari Bushveld. While lineages D and N appeared to be associated with the Lowveld bioregion and the Savanna biome, respectively, these nodes had no support. Lineages C, E, and K were only represented by one individual each. It was also evident from the combined analysis that there may be a clade comprising lineages B (Grassland biome), C (Machadodorp and Malelane) and D (Lowveld bioregion) although lineages C (with only one individual) and D were not supported (Figs. 3.6 and 3.7). Likewise, there was support (posterior probability value of ≥

0.90) for an association between lineages E (Fouriesburg), G (Albany Thicket bioregion) and H (Fynbos biome) (Figs. 3.6 and 3.7). Some of the associations between lineages and biome/bioregions were not apparent (lineage B - Grassland biome and lineage G - Albany Thicket bioregion). The independent BI analysis using cyt *b* sequences (Fig. 3.4) gave the same topology as the combined (cyt *b* and RAG1) BI phylogram (Fig. 3.6). The independent RAG1 BI tree had a different topology and some lineages were not monophyletic (Fig. 3.5). More nodes were supported using the combined approach (14 nodes compared to nine and four nodes for cyt *b* and RAG1, respectively). Most lineages showed good support (posterior probability values of $\geq$ 0.95) in the combined analysis (Fig. 3.6) and this was also evident in the ML analysis (not shown). The ML analyses gave the same topologies as their respective BI phylograms (not shown). The combined parsimony analysis gave the same topology as the combined BI tree (Figs. 3.6). The following parsimony statistics were obtained: 102 parsimony informative characters; tree length = 252; CI = 0.46; RI = 0.74; RC = 0.34. These nodes showed high bootstrap support (indicated in squares in Fig. 3.6). The parsimony analysis did not support the two groupings that were identified in the combined BI analysis: 1) Lineages B, C, and D and 2) Lineages E, G and H. Lineage N (Savanna biome) was not supported in the combined BI analysis, while a support of 71% was detected in the maximum parsimony analysis.

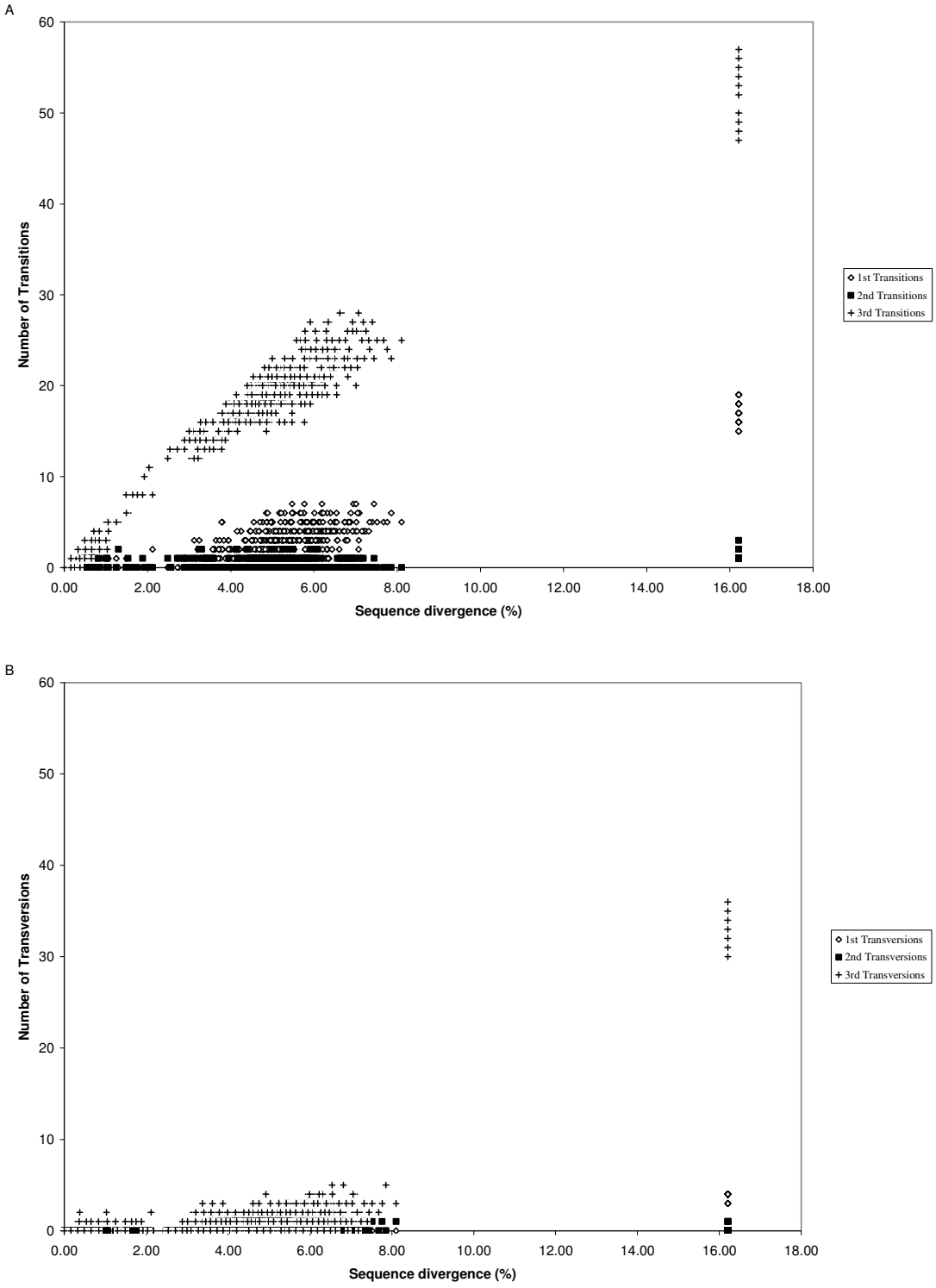

Figure 3.2 The number of transitions (A) and transversions (B) plotted against cytochrome *b* Tamura-Nei + Γ (1.97) + I (0.66) sequence divergence (%) estimates between 35 individuals of *Micaelamys namaquensis* and one outgroup (*Rhabdomys pumilio*) from southern Africa. (◊) = First codon positions, (■) = second codon positions and, (+) = third codon positions.

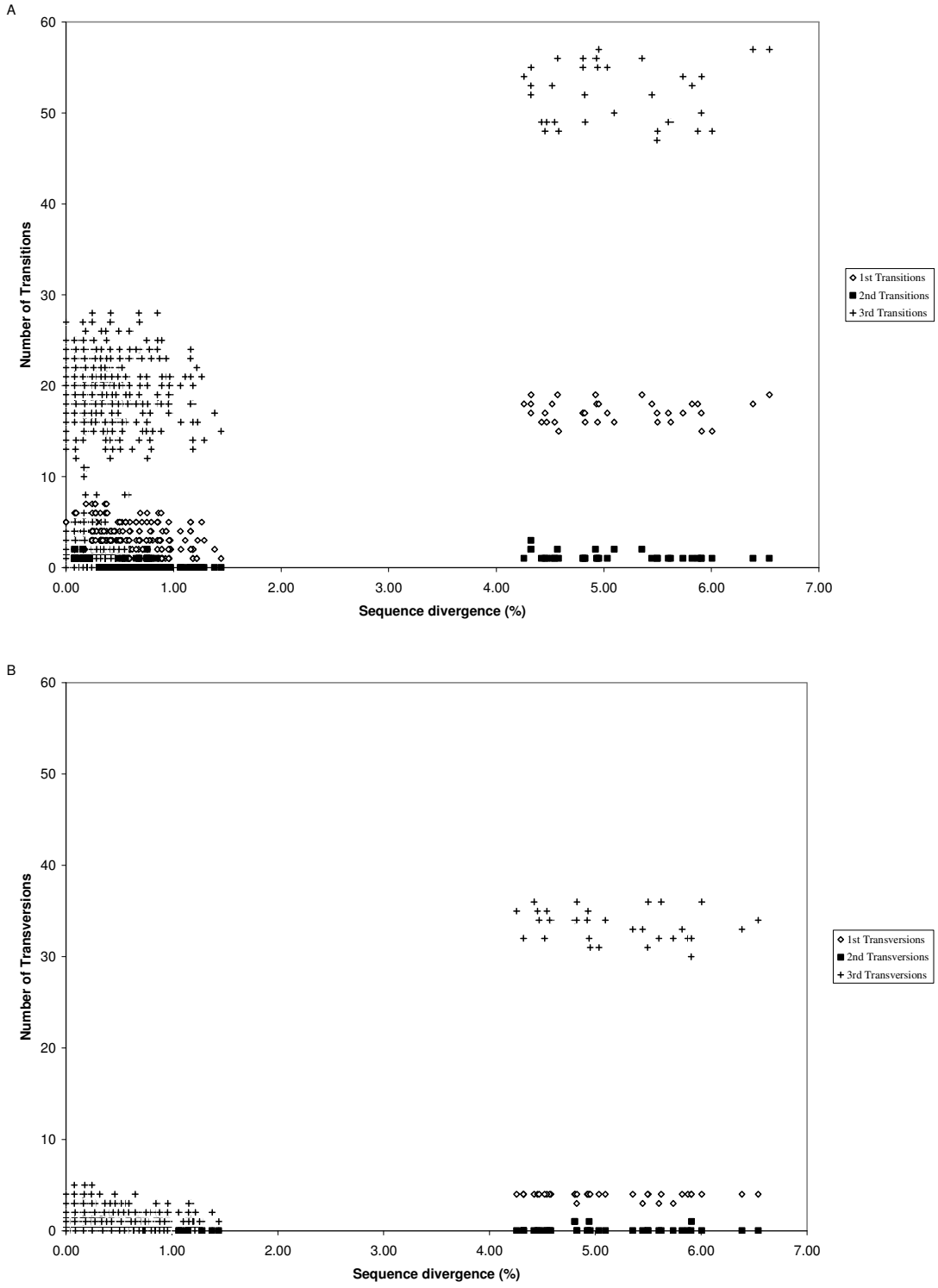

Figure 3.3 The number of transitions (A) and transversions (B) plotted against RAG1 Tamura-Nei + Γ (0.87) + I (0.83) sequence divergence (%) estimates between 35 individuals of *Micaelamys namaquensis* and one outgroup (*Rhabdomys pumilio*) from southern Africa. (◊) = First codon positions, (■) = second codon positions, and (+) = third codon positions.

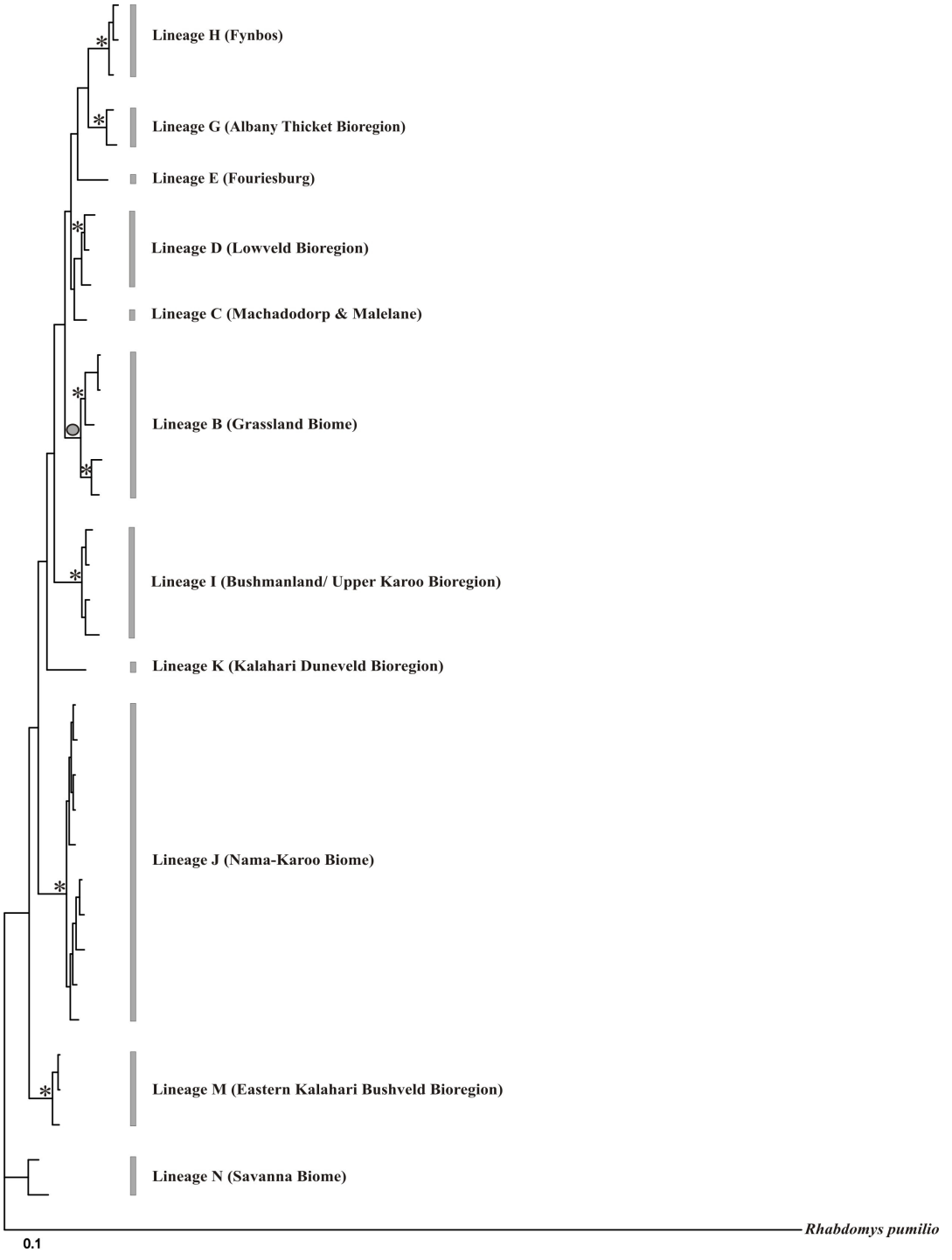

Figure 3.4 A Bayesian Inference (BI) tree derived from mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) sequences of *Micaelamys namaquensis* from southern Africa. Different lineages correspond to the lineages identified in Chapter 2 (see Fig. 2.4). The BI posterior probability values for internal branches are given at each node with either an asterisk (*) or a circle (•) where asterisks indicate BI posterior probability values ≥ 0.95, while circles indicate BI posterior probability values ≥ 0.90. *Rhabdomys pumillio* was used as an outgroup.

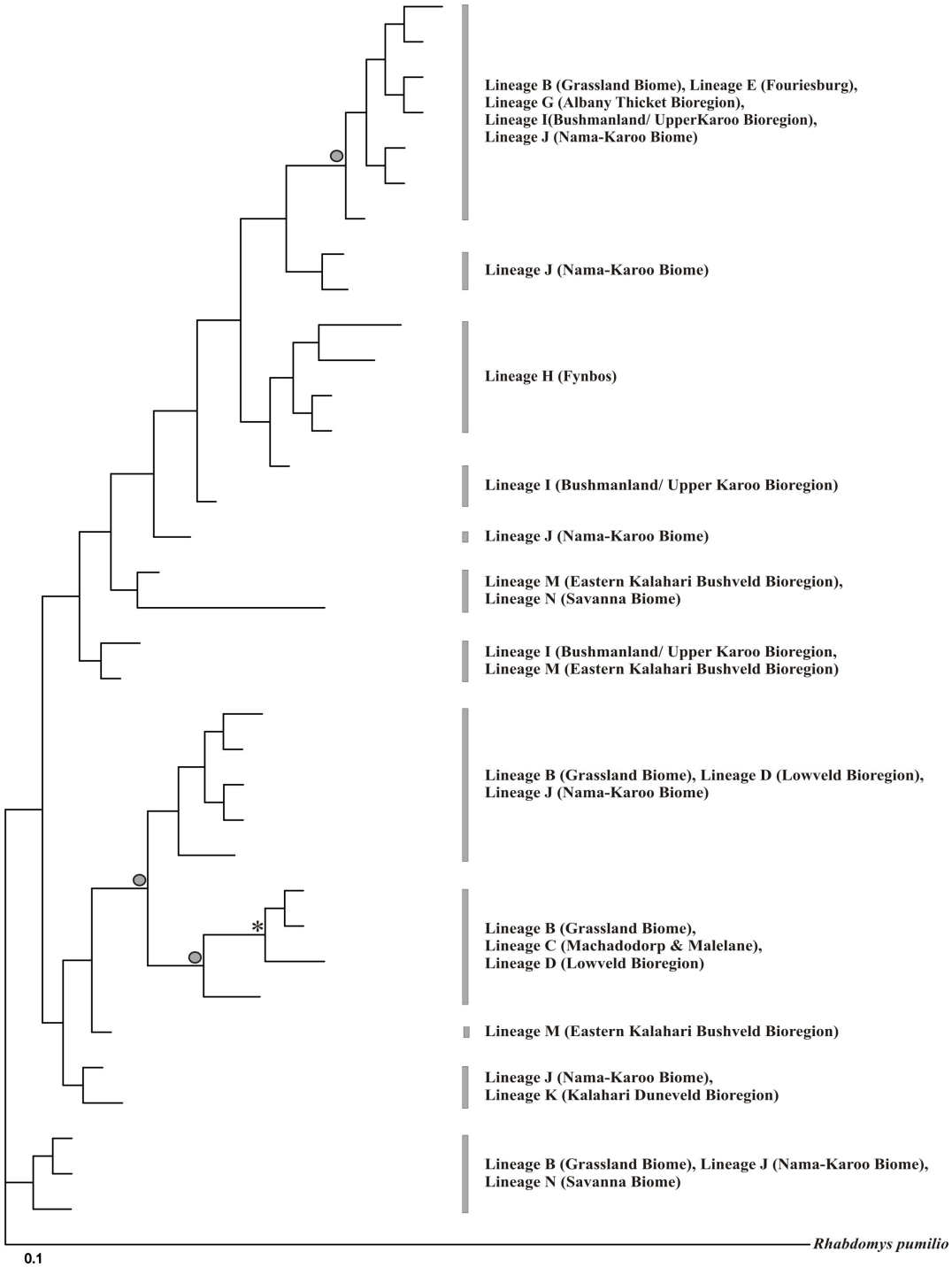

Figure 3.5 A Bayesian Inference (BI) tree derived from Recombination Activation Gene 1 (RAG1) sequences of *Micaelamys namaquensis* from southern Africa. Different lineages correspond to the lineages identified in Chapter 2 (Fig. 2.4). The BI posterior probability values for internal branches are given at each node with either an asterisk (*) or a circle (•) where asterisks indicate BI posterior probability values ≥ 0.95, while circles indicate BI posterior probability values ≥ 0.90. *Rhabdomys pumillio* was used as an outgroup.

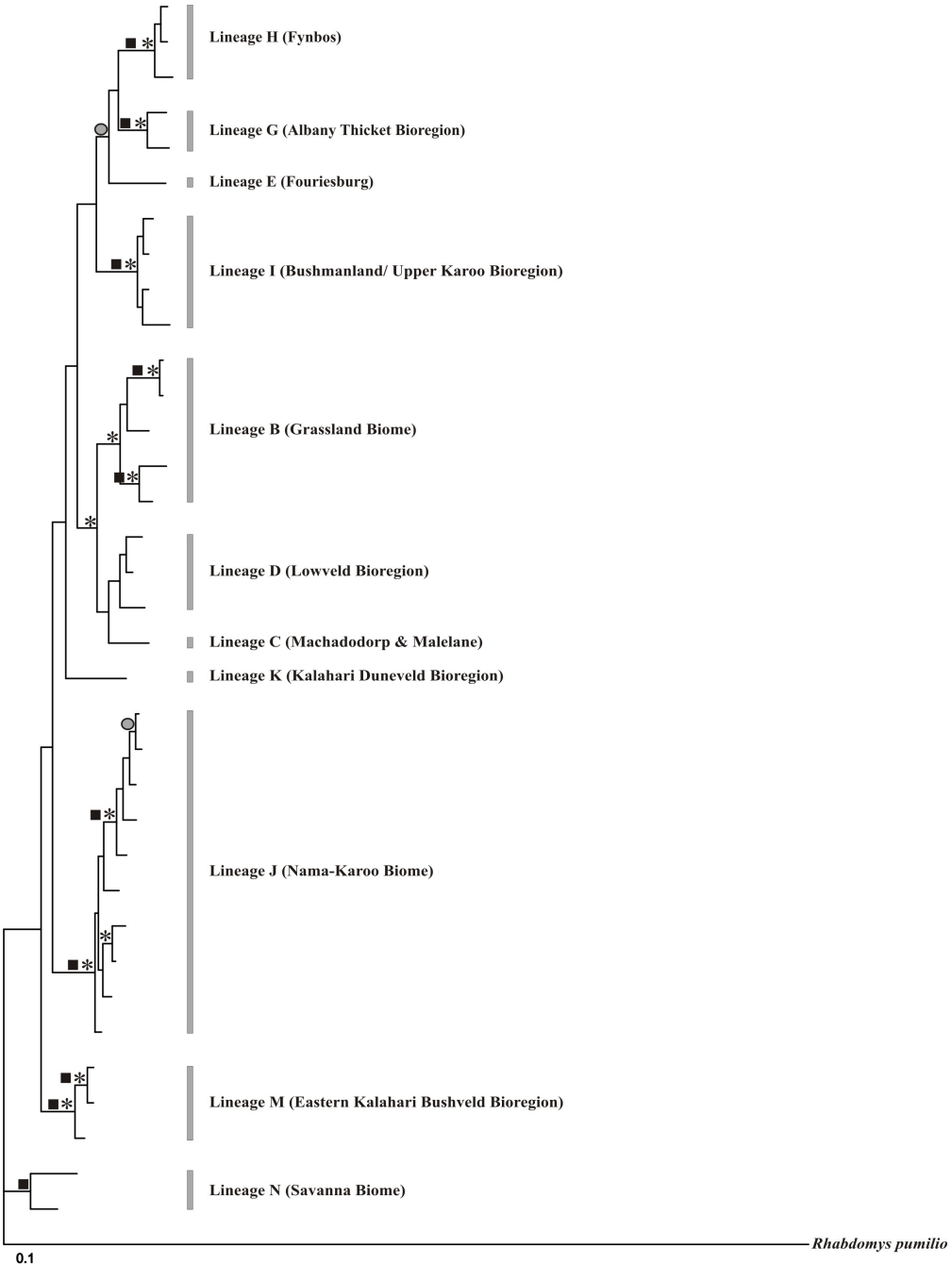

Figure 3.6 A Bayesian Inference (BI) tree derived from mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) and Recombination Activation Gene 1 (RAG1) sequences of *Micaelamys namaquensis* from southern Africa. Different lineages correspond to the lineages identified in Chapter 2 (Fig. 2.4). The BI posterior probability values for internal branches are given at each node with either an asterisk (*) or a circle (•) where asterisks indicate BI posterior probability values ≥ 0.95, while circles indicate BI posterior probability values ≥ 0.90. Maximum parsimony bootstrap confidence limits (above 70% occurrence in 1 000 replicates) for internal branches are given at each node and indicated by squares (■). *Rhabdomys pumillio* was used as an outgroup.

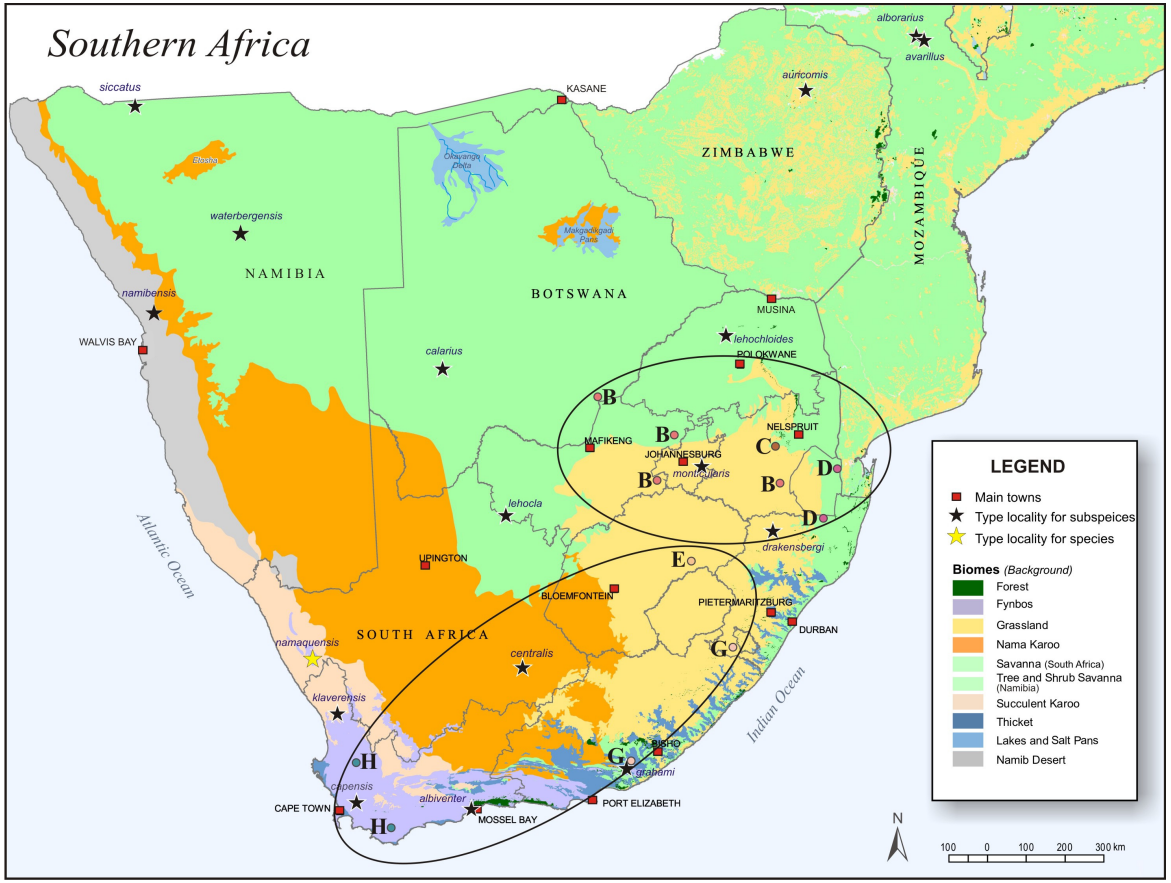

Figure 3.7 Major biomes of southern Africa (Low and Rebelo, 1996). Background colours for the major biomes are indicated in the legend; major lakes and saltpans are also included. The yellow star indicates the type locality for *Micaelamys namaquensis* and the blue stars indicate the type localities for its 16 previously described subspecies. Lineages B, C, D, E, G and H (see Chapter 2 for identified lineages) are indicated on the map (only well-supported associations between lineages are indicated). The encircled areas indicate the two groupings: (1) Lineage B, C and D and (2) Lineage E, G and H that were identified in the combined Bayesian Inference (BI) phylogram (Fig. 3.6).

4. Discussion

The phylogenetic analyses presented in the present study offer several insights into the evolutionary relationships within *M. namaquensis*. In this study, 35 individuals representing 11 of the 14 lineages from southern Africa as identified in Chapter 2 were analysed. The samples were subjected to molecular systematic analyses using the mitochondrial cyt *b* and nuclear RAG1 genes. Based on the independent cyt *b* and combined (cyt *b* and RAG1) ML and BI analyses, six lineages were recognised with strong geographical patterns showing an association with different vegetation types of southern Africa. Although the remaining lineages were not supported, they also were associated with biomes/bioregions (vegetation types) of southern Africa. These unsupported lineages were, in contrast, supported by the ML, BI and BEAST analyses in Chapter 2. It was not surprising that many of the relationships within the genus *Micaelamys* were not well-resolved, even using the combined approach. This lack of posterior support may well be a reflection of incomplete lineage sorting in the nuclear RAG1 gene (Belfiore et al., 2008).

A relationship between lineages B (Grassland biome), C (Machadodorp and Malelane) and D (Lowveld bioregion) was evident from the combined analysis although lineage C (represented by only one individual) and D were not supported (Figs. 3.6 and 3.7). Likewise, there was support (posterior probability value of ≥ 0.90) for some association between lineages E (Fouriesburg), G (Albany Thicket bioregion) and H (Fynbos biome) (Figs. 3.6 and 3.7). It was also evident that there is a geographic association between these lineages. Lineages B, C and D occur within the Grassland biome while lineage E, G and H are associated with the Great Escarpment. The geographic association of lineages may explain the groupings of these respective lineages. The sister-relationship between lineages B, C and D was also evident in the BEAST analysis but the relationship between lineages E, G and H was not supported by the BEAST analysis (see Chapter 2, Fig. 2.6). Some of the associations between lineages and biome/bioregions (lineage B - Grassland biome and lineage G - Albany Thicket bioregion) were not apparent; these discrepancies might be an indication that some lineages have been expanded in recent times. Associations between lineages might not hold considering more individuals representing each lineage. Results should therefore be interpreted with caution.

The incongruence between the independent RAG1 analysis and the independent cyt *b* and combined analysis may be due to incomplete lineage sorting within the nuclear RAG1 gene. Different historical and demographic scenarios have been described as being responsible for incomplete lineage sorting in rapidly radiating rodent species within the genus *Thomomys* (Thaeler, 1968). The different scenarios may be as follow: 1) some species have diverged recently through vicariant events, as a result of partial niche overlap, such that lineage sorting is incomplete but will become complete with time (Thaeler, 1968); 2) other species have arisen rapidly by peripheral isolation from the "parent" species, a portion of the genome of the new species is then identical to that of the "parent" species, without additional selective forces these lineages will not become completely sorted unless an extinction of one of the species occurs (Rogers, 1991a, b) and 3) species have diverged and become effectively reproductively isolated, but neither drift nor selection has been strong enough to eliminate the shared gene histories (Patton, 1990; Patton et al., 1972). *Micaelamys namaquensis* is currently classified as a single species and although the above scenarios have been described at a higher taxonomic level, the divergences that have been estimated within *M. namaquensis* are comparable with these higher level (between species) divergences (up to 7.68% divergence in cyt *b*, see Table 3.2).

Time elapsed since the dichotomy event should also be considered as a key factor in the evolution of lineage sorting. If enough time has passed then lineage sorting should be complete. In contrast, a signal of incomplete lineage sorting would be evident in species that have evolved in recent times. Divergences between *M. namaquensis* lineages are fairly old (if the estimated divergence dates presented in Chapter 2 is correct) and the result of incomplete lineage sorting is somewhat surprising. The incomplete lineage sorting detected within the species may be as a result of the selection specific to the gene (RAG1) and the mutation rate for this gene.

Mitochondrial DNA trees have a greater chance of being congruent with speciation history because its effective population size (N_e) is only one-quarter of any nuclear locus (McCracken and Sorenson, 2005; Moore, 1997). Therefore, complete lineage sorting will take longer in any nuclear gene. In general, a lack of complete lineage sorting would not have been evident without the use of multiple genes and more than one individual per taxon (Belfiore et al., 2008).

So far, only a few studies have compared the utility of mitochondrial and nuclear sequence data for phylogenetic analysis (e.g., Adkins et al., 2001; Matthee et al., 2001; McCracken and Sorenson, 2005). Most of these studies have been conducted on groups with older divergences where the slower rate of nuclear genes was more informative (Steppan et al., 2005). Steppan et al. (2005) showed that within recently evolved subfamilies, the mitochondrial DNA appears to be less informative at the deeper nodes. This was surprising because the basal node was only 12 MYA, an age younger than many of the mitochondrial phylogenetic studies that have been conducted on mammals (Catzeflis et al., 1995; Honeycutt et al., 1995; Irwin et al.; 1991; Yang and Yoder, 2003). Steppan et al. (2005) also demonstrated that the decline in bootstrap values for deeper nodes was as a result of the mtDNA whereas the nuclear genes showed no significant loss of robustness with increasing depth. The lower bootstrap support for the deeper mtDNA nodes may be due to the shorter deep branches in the mtDNA tree. Likewise, the mtDNA cyt *b* BI phylogram (Fig. 3.4) in the present study was characterised by short deep branches. In contrast, the RAG1 BI phylogram (Fig. 3.5) did not show such short deep branches. The poor resolution that was obtained for higher-level relationships (deeper nodes) made it impossible to ascertain the phylogenetic affinities between the *M. namaquensis* lineages. This may be due to the rapid radiation of rodents entering sub-Saharan Africa from Eurasia, possibly through the Middle East (Ducroz et al., 2001), and more samples and other DNA markers (e.g., more nuclear genes) may assist in resolving these relationships. The Murinae has therefore originated in Asia and colonised both Europe and Africa during dispersal events at about 11.8 MYA (Lecompte et al., 2008). Animals followed the establishment of a vegetation corridor across the Arabian peninsula connecting Africa and Asia. It has also been suggested that lineages moving into Africa were differentiated prior to their dispersal into Africa (Lecompte et al., 2008).

It is also evident from Steppan et al. (2005) that mtDNA data deteriorates measurably for murine nodes older than about 6 MYA. Figure 2.6 in Chapter 2 represents estimated times of divergence within and between *M. namaquensis* lineages. Within-lineage divergences differed from 940 000 years (lineage G, Albany Thicket bioregion) to 3.42 MYA (lineage N, Savanna biome). The well-supported nodes in the ML phylogram (Chapter 2, Fig. 2.4) may be a result of these younger divergences. The divergence time separating different lineages differed from 2.70 MYA (between lineages F and G) to 7.26 MYA (between

lineages B, C and D). Time to the most recent common ancestor for *M. namaquensis* was estimated at 9.44 MYA. Subsequent divergence of two groups followed: 1) a group more confined to the mesic habitats of southern Africa (lineages A-H), and 2) a group found in the more arid habitats of southern Africa (lineages I-N). The major diversification within this species-group appears to have occurred during the Late Miocene, between 7.87 MYA and 5.30 MYA (see Chapter 2, Fig. 2.4). These deeper nodes were not well-supported in the ML phylogram (Chapter 2, Fig. 2.4) possibly due to ages older than 6 MYA. Therefore, the utility of mtDNA should be extended to more recent divergences (younger than 6 MYA). In contrast, mitochondrial DNA was still informative when divergence dates older than 6 MYA were used in the present study (see Chapter 2, Fig. 2.6).

Steppan et al. (2005) suggested that slower evolving nuclear exons should be used more often in phylogenetic studies even for relatively recent divergence dates (younger than 5 MYA). In addition, it has been shown that more promising results are coming from analyses of less rapidly evolving nuclear genes such as LCAT (Michaux and Catzeflis, 2000; Robinson et al., 1997), vWF (Jansa and Weksler, 2004; Michaux et al., 2001) and IRBP (deBry and Sagel, 2001; Jansa and Weksler, 2004; Suzuki et al., 2004). Similarily, intron sequences have been shown to be informative at a phylogenetic level (deBry and Seshadri, 2001; Robinson-Rechavi et al., 2000; Steppan et al., 2004a). An added advantange of intron sequences is that it provides an ideal source of nuclear non-coding DNA that are flanked by protein-coding regions that may allow for robust PCR primers (deBry and Seshadri, 2001).

It was also shown in the present study that fewer nodes in the independent cyt *b* and RAG1 analyses were supported. Once the datasets were combined the supported nodes were increased from nine (cyt *b*) and four (RAG1) to 14 in the combined analysis suggesting that combined approaches (multiple genes, mitochondrial and nuclear) may be useful for phylogenetic analysis. Cummings et al. (1995) noted that phylogenetic analyses of several short stretches from different genes show a better performance than analyses based on single gene fragments. Future research should include a multidisciplinary approach (both faster and slower evolving nuclear genes, chromosomes, morphology) on more samples.

5. Conclusion

Micaelamys namaquensis is a polytypic species with more variation than previously thought. The diversity detected within *M. namaquensis* appears to be indicative of a species complex. Only 11 lineages were included in the phylogenetic analyses that were supported in the BI analyses. Of these six lineages were well-supported with strong geographic patterns. Although the remaining lineages were not supported, they nevertheless are associated with biomes/bioregions (vegetation types) of southern Africa. This suggests that the taxonomic status of *M. namaquensis* needs further investigation and the species is in need of a taxonomic revision based on a multidisciplinary approach and extensive sampling.

6. Acknowledgements

I am grateful to all field assistance for helping with the collection of samples (M.J. Russo, T.J. Grant, C.J. Oosthuizen, S.M.R dos Santos, W. Delport, E.R. Swartz, H. Roos, F.J. Lötz, A.V. Linzey, A. Hulse, M. Kesner, I. Stermann, M. Cunningham, N. Maputla, V. Coetzee, M. de Castro, U. Kryger, H. Smit, S.L. Gardner). Thanks also to Albé van der Merwe for computational support. C.L. Sole is thanked for the constructive comments on earlier drafts of this chapter. This study was funded by the National Research Foundation grant to PB and CTC (Grant number: 2073181) and the University of Pretoria.

References

Adams, E.N., 1972. Consensus techniques and the comparison of taxonomic trees. Syst. Zool. 21, 390-397.

Adkins, R.M., Gelke, E.L., Rowe, D., Honeycutt, R.L., 2001. Molecular phylogeny and divergence time estimates for major rodent groups; evidence from multiple genes. Mol. Biol. Evol. 18, 777-791.

Archie, J.W., 1989. Homoplasy excess ratios: new indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. Syst. Zool. 38, 253-269.

Avery, D.M., 1981. Holocene micromammalian faunas from the northern Cape Province, South Africa. S. Afr. J. Sci. 77, 265-273.

Avery, D.M., 1982. Micromammals as palaeoenvironmental indicators and an interpretation of the late Quaternary in the southern Cape Province, South Africa. Ann. S. Afr. Mus. 85, 183-374.

Avery, D.M., 1985. A preliminary assessment of the micromammalian remains from Gladysvale cave, South Africa. Palaeontol. Afr. 32, 1-10.

Belfiore, N.M., Liu, L., Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the genus Thomomys (Rodentia: Geomyidae). Syst. Biol. 57, 294-310.

Black, C.G., Krishtalka, L., 1986. Rodents, bats and insectivores from Plio-Pleistocene sediments to the east of Lake Turkana, Kenya. Contrib. Sci. 372, 1-15.

Bonhomme, F., Iskander, D., Thaler, L., Petter, F., 1985. Electromorphs and phylogeny of muriod rodents. In: Luckett, P.W., Hartenberger, J-L. (Eds.), Evolutionary relationships among rodents, a multidisciplinary analysis. Plenum Press, New York, pp. 671-683.

Catzeflis, F.M., Hanni, C., Sourrouille, P., Douzery, E., 1995. Molecular systematics of hystricognath rodents: the contribution of sciurognath mitochondrial 12S rRNA sequences. Mol. Phylogenet. Evol. 4, 357-360.

Chimimba, C.T., 2001. Infraspecific morphometric variation in *Aethomys namaquensis* (Rodentia: Muridae) from southern Africa. J. Zool. Lond. 253, 191-210.

Chimimba, C.T., 2005. Phylogenetic relationships in the genus *Aethomys* (Rodentia: Muridae). Afr. Zool. 40, 271-284.

Chimimba, C.T., Dippenaar N.J., Robinson, T.J., 1999. Morphometric and morphological delineation of southern African species of *Aethomys* (Rodentia: Muridae). Biol. J. Linn. Soc. 67, 501-527.

Chimimba, C.T., Bennett, N., 2005. Order: Rodentia. In: Skinner, J.D., Chimimba, C.T. (Eds.), The mammals of the southern African. Cambridge Univ. Press, pp 156-163.

Cummings, M.P., Otto, S.P., Wakeley, J., 1995. Sample properties of DNA sequence data in phylogenetic analysis. Mol. Biol. Evol. 12, 814-822.

Davis, D.H.S., 1975. Genus *Aethomys*. In: Meester, J., Setzer, H.W. (Eds.), The mammals of Africa: an identification manual. Smithsonian Institution Press, Washington, D.C. Part 6.6, pp. 1-5.

deBry, R.W., Sagel, R.M., 2001. Phylogeny of Rodentia (Mammalia) inferred from the nuclear-encoded gene IRBP. Mol. Phylogenet. Evol. 19, 290-301.

deBry, R.W., Seshadri, E., 2001. Nuclear intron sequences for phylogenetics of closely related mammals: an example using th phylogeney of *Mus*. J. Mammal. 82, 280-288.

De Graaff, G., 1960. A preliminary investigation of the mammalian microfauna in Pleistocene deposits of caves in the Transvaal system. Palaeontol. Afr. 7, 79-118.

De Graaff, G., 1961. On the fossil mammalian microfauna collected at Kromdraai by Dapper in 1895. S. Afr. J. Science 56, 259-260.

De Graaff, G., 1981. The Rodents of southern Africa. Butterworths, Pretoria.

Denys, C., 1987. Micromammals from the West Natron Pleistocene deposits (Tanzania): biostratigraphy and paleoecology. Sci. Geol. Bull. 40, 185-201.

Denys, C., 1990a. Implications paleoecologiques et paleobiogeographiques de l'etude de rongeurs plio-pleistocenes d'Afrique orientale et australe. Phd thesis, Universite Paris, Paris.

Denys, C., 1990b. Deux nouvelles especes d'*Aethomys* (Rodentia: Muridae) a Langebaanweg (Pliocene, Afrique du Sud): implications phylogenetiques et paleoecologiques. Ann. Paleontol. (Vertebrates and Invertebrates) 76, 41-69.

De Queiroz, A., 1993. For consensus (sometimes). Syst. Biol. 42, 368-372.

De Queiroz, A., Donoghue, M.J., Kim, J., 1995. Separate versus combined analysis of phylogenetic evidence. Annu. Rev. Ecol. Syst. 26, 657-681.

Ducroz, J-F., Volobouev, V., Granjon, L., 1998. A molecular perspective on the systematics and evolution of the genus *Arvicanthis* (Rodentia: Muridae): inferences from complete cytochrome *b* gene sequences. Mol. Phylogenet. Evol. 10, 104-117.

Ducroz, J-F., Volobouev, V., Granjon, L., 2001. An assessment of the systematics of Arcicanthine rodents using mithochondrial DNA sequences: evolutionary and biogeographical implications. J. Mamm. Evol. 8, 173-206.

Ellerman, J.R., 1941. The families and genera of living rodents. British Museum Natural History, London.

Ellerman, J.R., Morrison-Scott, T.C.S., Hayman, R.W., 1953. Southern African mammals 1758 to 1951: a reclassification. British Museum Natural History, London.

Farris, J.S., 1989. The retention index and rescaled consistency index. Cladistics 5, 417-419.

Farris, J.S., Kluge, A.G., Eckardt, M.J., 1970. A numerical approach to phylogenetic systematics. Syst. Zool. 19, 172-179.

Felsenstein, J., 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. Syst. Zool. 22, 240-249.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368-376.

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

Galewski, T., Tilak, M-K., Sanchez, S., Chevret, P., Paradis, E., Douzery, E.J.P., 2006. The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. BMC Evol. Biol. 6, 80.

Groth, J.G., Barrowclough, G.F., 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG1 gene. Mol. Phylogenet. Evol. 12, 115-123.

Gu, X., Zhang, J., 1997. A simple method for estimating the parameters of substitution rate variation among sites. Mol. Biol. Evol. 14, 1106-1113.

Hendey, Q.B., 1981. Paleoecology of the late tertiary fossil occurrences in “E” Quarry, Langebaanweg, South Africa and a re-introduction of their geological context. Ann. S. Afr. Mus. 84, 1-104.

Honeycutt, R.L., Nedbal, M.A., Adkins, R.M., Janecek, L.L., 1995. Mammalian mitochondrial DNA evolution: a comparison of the cytochrome b and cytochrome c oxidase II genes. J. Mol. Evol. 40, 260-272.

Huelsenbeck, J.P., Bull, J.J., Cunningham, C.W., 1996. Combining data in phylogenetic analysis. TREE 11, 152-158.

Irwin, D.M., Kocher, T.D., Wilson, A.C., 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. 32, 128-144.

Jaeger, J.J., 1976. Les rongeurs (Mammalia: Rodentia) du Pleistocene inferieur d'Olduvai Bed I (Tanzanie), 1e parti: les Murides. In: Savage, R.J.G., Coryndon, S.C. (Eds.), Fossil vertebrates of Africa. Academic Press, New York and London, pp. 58-120.

Jaeger, J.J., 1979. Les faunes de rongeurs et de lagomorphes du Pliocene et du Pleistocene d'Afrique orientale. Bull. Soc. Geol., France 7, 301-308.

Jansa, S.A., Weksler, M., 2004. Phylogenetic studies on didelphid marsupials I. Introduction and preliminary results from nuclear IRBP gene sequences. J. Mammal. Evol. 7, 43-77.

Kluge, A.G., Farris, J.S., 1969. Quantitative phylogenetics and the evolution of anurans. Syst. Zool. 18, 1-32.

Lecompte, E., Aplin, K., Denys, C., Catzeflis F., Chades, M., Chevret, P., 2008. Phylogeney and biogeography of African Murinae base don mitocondrial and nuclear gene sequences, wiht a new tribal classification of the subfamily. BMC Evol. Biol. 8, 1-21.

Low, A.B., Rebelo, A.G. (Eds.), 1996. Vegetation of South Africa, Lesotho and Swaziland. Department of Environmental Affairs and Tourism, Pretoria.

Maddison, W.P., Maddison, D.R., 1992. MacClade: Analysis of phylogeny and character evolution, Version 3. Sinauer, Sunderland, Massachusetts.

Meester, J., Davis, D.H.S., Coetzee, C.G., 1964. An interim classification of southern African mammals. Mimeograph of the Zoological Society of southern Africa and the Council of Scientific and Industrial Research, South Africa.

Maree, S., 2002. Phylogenetic relationships and mitochondrial DNA sequence evolution in the African rodent subfamily Otomyinae (Muridae). PhD thesis, University of Pretoria, Pretoria.

Matthee, C.A., Burzlaff, J.D., Taylor, J.F., Davis, S.K., 2001. Mining the mammalian genome for artiodactyl systematics. Syst. Bio. 50, 367-390.

McCracken, K.G., Sorenson, M.D., 2005. Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (Nomonyx - Oyxyura)? Syst. Biol. 54, 35-55.

Michaux, J.R., Catzeflis, F.R., 2000. The bushlike radiation of muroid rodents is exemplified by a phylogenetic analysis of LCAT nuclear sequences. Mol. Phylogenet. Evol. 17, 280-293.

Michaux, J.R., Reyes, A., Catzeflis, F., 2001. Evolutionary history of the most speciose mammals: molecular phylogeny of muroid rodents. Mol. Biol. Evol. 18, 2017-2031.

Meier, R., Kores, P., Darwin, S., 1991. Homoplasy slope ratio: a better measurement of observed homoplasy in cladistic analyses. Syst. Zool. 40, 78-88.

Moore, W.S., 1997. Mitochondrial-gene trees versus nuclear-gene trees, a reply to Hoelzer. Evolution 51, 627-629.

Musser, G.G., Carleton, M.D., 2005. Family Muridae. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal species of the world: a taxonomic and geographic reference. Smithsonian Institution Press in association with the American Society of Mammalogists, London and Washington D.C., pp. 501-755.

Palumbi, S.R., Cipriano, F., Hare, M.P., 2001. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. Evolution 55, 859-868.

Patton, J.L., 1990. Geomyid evolution: the historical, selective and random basis for divergence patterns within and among speices. In: Nevo, E., Reig, O.A. (Eds.), Evolution of subterranean mammals at the organismal and molecular levels. Wiley-Liss, New York, pp. 49-69.

Patton, J.L., Selander, R.K., Smith, M.H., 1972. Genic variation in hybridising populations of gophers (genus *Thomomys*). Syst. Zool. 21, 263-270.

Pocock, T.N., 1987. Plio-Pleistocene mammalian microfauna in southern Africa - a preliminary report including description of two new fossil muriod genera (Mammalia: Rodentia). Paleontol. Afr. 26, 69-71.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Roberts, A., 1951. The mammals of South Africa. Trustees of "The mammals of South Africa" book fund, Johannesburg.

Robinson, M., Catzeflis, F., Briolay, J., Mouchiroud, D., 1997. Molecular phylogeny of rodents, with special emphasis on murids: evidence from nuclear gene LCAT. Mol. Phylogenet. Evol. 8, 423-434.

Robinson-Rechavi, M., Ponger, L., Mouchiroud, D., 2000. Nuclear gene LCAT supports rodent monophyly. Mol. Biol. 17, 1410-1412.

Rogers, M.A., 1991a. Evolutionary differentiation within the Northern Great Basin pocket gopher *Thomomys townsendii* I. Morphological variation. Great Basin Nat. 51, 109-126.

Rogers, M.A., 1991b. Evolutionary differentiation within the Northern Great Basin pocket gopher *Thomomys townsendii* II. Genetic variation and biogeographic consequences. Great Basin Nat. 51, 127-152.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Russo, I.M., 2003. Molecular systematics of southern African *Aethomys* (Rodentia: Muridae). MSc thesis, University of Pretoria, Pretoria.

Steppan, S.J., Adkins, R.M., Anderson, J., 2004a. Phylogeny and divergence-date estimates of rapid radiations in Muroid rodents based on multiple nuclear genes. Syts. Biol. 53, 533-553.

Steppan, S.J., Storz, B.L., Hoffmann, R.S., 2004b. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. Mol. Phylogenet. Evol. 30, 703-719.

Steppan, S.J., Adkins, R.M., Spinks, P.Q., Hale, C., 2005. Multigene phylogeny of the Old World mice Murinae reveals distinct geographic lineages and the declining utility of mitochondrial genes compared to nuclear genes. Mol. Phylogenet. Evol. 37, 370-388.

Swofford, D.L., 2003. PAUP*: phylogenetic analysis using parsimony (* and other methods), Version 4. Sinauer Associates, Sunderland, Massachusetts.

Suzuki, A., Bisordi, I., Levis, S., Garcia, J., Pereira, L.E., Souza, R.P., Sugahara, T.K., Pini, N., Enria, E., Souza, L.T., 2004. Identifying rodent hantavirus reservoirs, Brazil. Emerg. Infect. Dis. 10, 2127-2134.

Thaeler, C.S., 1968. An analysis of the distribution of pocket gopher species in northeastern California (genus Thomomys), vol. 86. University of California Publications in Zoology. University of Carlifonia Press, Berkley, California.

Thompson, J.D., Gibson, T.J., Plewniak, T.F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876-4882.

Verheyen, E., Colyn, M., Verheyen, W., 1995. The phylogeny of some African muriods (Rodentia) based upon partial mitochondrial cytochrome *b* sequences. Belg. J. Zool. 125, 403-407.

Verheyen, E., Colyn, M., Verheyen, W., 1996. A mitochondrial cytochrome *b* phylogeny confirms the paraphyly of the Dendromurinae Alston, 1896 (Rodentia: Muridae). Mammalia 60, 780-785.

Visser, D.S., Robinson, T.J., 1986. Cytosystematics of the South African *Aethomys* (Rodentia: Muridae). S. Afr. J. Zool. 21, 264-268.

Visser, D.S., Robinson, T.J., 1987. Systematic implications of spermatozoan and bacular morphology for the South African *Aethomys*. Mammalia 51, 447-454.

Wenhui, L., Chang, F.C., Desiderio, S., 2001. RAG1 mutations with B-Cell-Negative SSCID dissociate the nicking and transesterification steps of V(D)J recombination. Mol. Cell. Biol. 21, 3935-3946.

Wesselman, H., 1984. Omo micromammals. In: Hecth, M.K., Szalay, F.S. (Eds.), Contributions to vertebrate evolution. S Karger, New York, pp. 219.

Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analyses. TREE 144, 1941-1950.

Yang, Z., Goldman, N., Friday, A., 1994. Comparrison of models from nucleotide substitution used in maximum likelihood phylogenetic estimation. Mol. Biol. Evol. 11, 316-324.

Yang, Z.H., Yoder, A.D., 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst. Biol. 52, 705-716.

Appendix 3.1 Geographic coordinates of all collecting localities of *Micaelamys namaquensis* in South Africa, Swaziland, Botswana and Namibia representing individuals with sequences for both the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) gene and the nuclear Recombination Activating Gene 1 (RAG1) gene that were analysed in the present study. Numbers 1 - 30 correspond to those in Fig. 3.1.

LOCALITY	COUNTRY	PROVINCE	GEOGRAPHIC COORDINATE
Savanna Biome			
1. Francistown, just outside town (municipal grounds)	Botswana		21°11'15"S 27°23'22"E
2. Farm: Goedgelegen, Baltimore	South Africa	Limpopo	23°26'27"S 28°23'02"E
Nama-Karoo Biome			
3. Farm: Steenkampsput, Upington	South Africa	Northern Cape	28°06'13"S 20°54'10"E
4. Farm: Warmhoek, Hoopstad	South Africa	Free State	28°10'08"S 25°49'11"E
5. Farm: Viljoenshof, Boshof	South Africa	Free State	28°34'45"S 25°04'33"E
6. Farm: Palmietfontein, Brandfort	South Africa	Free State	28°48'07"S 26°33'32"E
7. Farm: Tierkoppen, Augrabies	South Africa	Northern Cape	28°34'06"S 20°26'05"E
8. Farm: Boomrivier, Pofadder	South Africa	Northern Cape	29°04'33"S 19°18'24"E
9. Lady Grey, just outside town (municipal grounds)	South Africa	Eastern Cape	30°45'00"S 27°15'00"E
10. Matjiesfontein, just outside town (municipal grounds)	South Africa	Western Cape	33°15'00"S 20°34'48"E
Grassland Biome			
11. Gaborone, just outside town (municipal grounds)	Botswana		24°40'12"S 25°49'48"E
12. Brits Agricultural School, Brits	South Africa	North West	25°34'29"S 27°46'02"E
13. Farm: Ratzegaai, Ventersdorp	South Africa	North West	26°20'30"S 26°44'01"E
14. Farm: Uitspanning, Amsterdam	South Africa	Mpumalanga	26°39'56"S 30°31'26"E

Fynbos Biome			
15. Farm: Grootfontein, Porterville	South Africa	Western Cape	32°54'28"S 19°06'31"E
16. Farm: Fairfield, Napier	South Africa	Western Cape	34°27'27"S 19°45'10"E
Albany Thicket Bioregion			
17. Andries Vosloo Kudu Reserve, Grahamstown	South Africa	Eastern Cape	33°10'55"S 26°38'10"E
18. Mount Currie Nature Reserve, Kokstad	South Africa	KwaZulu-Natal	30°29'36"S 29°23'18"E
Fouriesburg/Kasane			
19. Wynford Guest Farm, Fouriesburg	South Africa	Free State	28°30'30"S 28°15'42"E
Lowveld Bioregion			
20. Mantenga Nature Reserve	Swaziland		26°26'37"S 31°10'22"E
21. Farm: Koedoesberg, Pongola	South Africa	KwaZulu-Natal	27°26'31"S 31°41'41"E
Bushmanland/Upper Karoo Bioregion			
22. Farm: Karlsrühe, Hotazel	South Africa	Northern Cape	26°58'34"S 22°59'57"E
23. Farm: Donkerpoort, Schweizer-Reneke	South Africa	North West	27°14'46"S 25°06'01"E
24. Farm: Rooidam, Groblershoop	South Africa	Northern Cape	29°08'33"S 22°19'34"E
Eastern Kalahari Bushveld Bioregion			
25. Farm: Rus en Vrede, Stella	South Africa	North West	26°10'23"S 25°13'27"E
26. Farm: Jones, Severn	South Africa	Northern Cape	26°35'22"S 22°41'46"E
27. Tswalu Kalahari Reserve, Sonstraal	South Africa	Northern Cape	27°12'51"S 22°'27'22"E
Kalahari Duneveld Bioregion			

28. Gibeon	Namibia		25°20'42"S 17°15'13"E
29. Farm: Zwartbooisberg, Kakamas	South Africa	Northern Cape	28°02'30"S 20°42'55"E
Machadodorp/Malelane			
30. Wathaba-Uitkomst, Machadodorp	South Africa	Mpumalanga	25°47'31"S 30°22'28"E

Appendix 3.2 *Micaelamys namaquensis* individuals (29 alleles for mitochondrial DNA and 35 for nuclear DNA) examined in this chapter and their GenBank accession numbers for both the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) and nuclear Recombination Activating Gene 1 (RAG1) genes. Note that the cytochrome *b* accession numbers correspond with those indicated in Chapter 2. Although some individuals shared the same cytochrome *b* allele, they had different RAG1 alleles.

CYT *B* ALLELE	CYT *B* ACCESSION NUM.	RAG1 ALLELE	RAG1 ACCESSION NUM
NH001	GQ471959	NNH01	GU139424
NH006	GQ471964	NNH02	GU139425
NH014	GQ471972	NNH03	GU139426
NH014	GQ471972	NNH04	GU139427
NH014	GQ471972	NNH05	GU139428
NH019	GQ471977	NNH06	GU139429
NH024	GQ471982	NNH07	GU139430
NH026	GQ471984	NNH08	GU139431
NH032	GQ471990	NNH09	GU139432
NH035	GQ471993	NNH10	GU139433
NH041	GQ471999	NNH11	GU139434
NH044	GQ472002	NNH12	GU139435
NH051	GQ472009	NNH13	GU139436
NH058	GQ472016	NNH14	GU139437
NH058	GQ472016	NNH15	GU139438
NH063	GQ472021	NNH16	GU139439
NH072	GQ472030	NNH17	GU139440
NH073	GQ472031	NNH18	GU139441
NH078	GQ472036	NNH19	GU139442
NH080	GQ472038	NNH20	GU139443
NH084	GQ472042	NNH21	GU139444
NH086	GQ472044	NNH22	GU139445
NH090	GQ472048	NNH23	GU139446
NH093	GQ472051	NNH24	GU139447

NH095	GQ472053	NNH25	GU139448
NH096	GQ472054	NNH26	GU139449
NH097	GQ472055	NNH27	GU139450
NH099	GQ472057	NNH28	GU139451
NH100	GQ472058	NNH29	GU139452
NH106	GQ472064	NNH30	GU139453
NH113	GQ472071	NNH31	GU139454
NH113	GQ472071	NNH32	GU139455
NH115	GQ472073	NNH33	GU139456
NH120	GQ472078	NNH34	GU139457
NH131	GQ472089	NNH35	GU139458