Molecular systematics of southern African *Aethomys* (Rodentia: Muridae)

by

Issie Magrieta Russo

Submitted in partial fulfilment of the requirements for the degree of Magister Scientia
in the Faculty of Natural and Agricultural Science

Department of Genetics
University of Pretoria
Pretoria

Supervisor: Dr. P. Bloomer
Co-supervisor: Prof. C.T. Chimimba

July 2003

Declaration

I, the undersigned, hereby declare that this thesis is my own unaided work. It has not been submitted before for any degree or examination at any other university.

Signature:  Date: 06/08/2003

ABSTRACT

Phylogeographic mitochondrial DNA (mtDNA) population structure was determined for *Aethomys namaquensis* and *A. ineptus* from southern Africa and Swaziland. It was evident from the study that *A. namaquensis* reflected a pattern of phylogenetic discontinuity with and without spatial separation between populations. Previously documented mtDNA phylogeographic patterns recorded in the rock hyrax, *Procavia capensis* and the red rock rabbit species, *Pronolagus rupestris* and *P. randensis*, coincided with the phylogeographic break that was detected in one of the mtDNA lineages (C) within *A. namaquensis*. Similar vicariant events may have been responsible for shaping evolutionary processes in the independent *Procavia*, *Pronolagus* and *Aethomys* lineages. In contrast, *A. ineptus* showed a pattern of shallow phylogeographic structuring. The marked genetic differences detected in *A. namaquensis* and *A. ineptus* may reflect the influences of habitat specificity, its fragmentation and the effects of life history on mtDNA gene flow. The study also revealed three genetically well-supported lineages within *A. namaquensis*: a lineage (A) found in the Limpopo valley, a lineage (B) widely distributed across the Karoo and a lineage (C) found across the grasslands of the North-West, Limpopo and Mpumalanga Provinces. These spatial distributions broadly coincided with the ranges of four previously proposed morphologically defined subspecies. From the present sample sizes, there is only good support, from a molecular point of view, for the subspecies *A. n. lehocla* (lineage B). In addition to the three well-supported lineages, six geographically restricted lineages were identified that could not be assigned to any of the four previously proposed subspecies, *A. n. namaquensis*, *A. n. monticularis*, *A. n. alborarius* and *A. n. lehocla*. Molecular techniques, specifically the analysis of the mtDNA cytochrome b gene, have been useful in the identification of sibling species. This technique has also proved to be useful in the identification of two cryptic species, *A. chrysophilus* and *A. ineptus* in this study. Phylogenetic analysis revealed two maternal groups corresponding to *A. chrysophilus* and *A. ineptus*. Distributional data of these two species, suggest that *A. chrysophilus* occupies the low elevations of the Limpopo River drainage, while *A. ineptus* occupies the remainder of South Africa at higher elevations, but expands into lower elevations in the southern portion of its range. Phylogenetic relationships among four southern African species of *Aethomys* suggest the presence of two clades that included: 1) *A. chrysophilus* and *A. ineptus* and 2) *A. namaquensis* and *A. granti*. This study, however,

revealed that *Aethomys* may be paraphyletic, suggesting that the allocation of *A. namaquensis* and *A. granti* to the subgenus *Micaelamys* needs to be investigated further.

For my friend, Frans Lötz Jr.

Thanks for your support and encouragement

ACKNOWLEDGEMENTS

I would firstly like to thank my supervisors Paulette Bloomer and Chris Chimimba for their invaluable advice and expertise. Paulette thanks for always being prepared to help and for always having an open door and greeting me with a smile. Chris thanks for the fast turn around time with the edits and for your very useful comments on the early drafts, thanks also for the emotional support when I needed it.

Secondly I would like to thank Duncan MacFadyen (Transvaal Museum), Peter Taylor (Durban Museum), Nico Avenant and Johan Watson (Bloemfontein Museum) for providing tissue samples. Special thanks also go to Simon and Jackson at the Transvaal Museum for their hard work in preparation of the skeletal samples. To Alicia Linzey, Michael Kesner and Phenda, I appreciate the *A. chrysophilus* and *A. ineptus* samples which greatly helped in filling the gaps in my sampling distribution.

My sincere appreciation goes to all the farmers in the Northern Cape, KwaZulu-Natal, Northern and Mpumalanga Provinces who showed such generous hospitality and friendliness, and for assistance in the field. I would especially like to mention the Stuart, Van Wyk, Theron, Olivier, Liebenberg, Myburgh, and Kruger families. Also my sincere thanks go to the various owners and managers of game farms where I sampled.

To the Lötz and Grant families thanks for providing my sister and me with a roof over our heads when we were homeless. Oom Franz, Tannie Yvonne, Danie, Zelna, Tyler, Ilza and Jabbie thanks for all of your interest and support throughout my studies, I really value the contribution you have made to my life. To Uncle Mike and Aunty Suzanne thanks for all the motivational meals when I had no food in my home (sorry for eating the whole Springbok roast and leaving hardly any for you).

I am very grateful to all my field assistants: Ernst, Ute, Lisel, Chris-Jan, Tyron, Rethana, Jabbie and Maretha. Some of my assistants were more willing and enthusiastic than others but every little bit of help was greatly appreciated. I will always remember the fun times we had in the field.

To all my friends and colleagues from the MEEP lab, thank you for all the laughs and tears. I would particularly like to thank Ute Kryger, Carel Oosthuizen, Marina Alais, Tyron Grant, Wayne Delport, Michael Cunningham, Ernst Swartz, Shaun Forgie, Cathi Lehn, Keith Barnes, Catherine Hughes, James Sakwa, Lisel Solms and Heidi Roos, for their helpful advice and companionship over these past years. For all of their help with my analyses in terms of understanding and processing of my data I would like to thank Ute Kryger and Sarita Maree.

Without my grandparents and parents I would have been stuck in Pretoria and would have had to walk everywhere, thanks Mom, Dad, Ouma Rita and Oupa Corrie for assisting me in buying my green monster. Without my little monster I wouldn't have had half of the samples that I needed to complete this study. To my dearly departed Ouma Rita thanks for the money you willed to me without which I would not have been able to afford a computer to write up my thesis on and which would have meant that I would have had to spend endless hours at the University writing up.

To my parents, I would like to express my gratitude for the financial assistance and for your love. I know at times you wished this thesis would just finish itself and you found it difficult to understand that a thesis could take this long but thanks for your unwavering support. Dad I greatly appreciate your monetary assistance in getting me to this point in my academic career. Mom and Oom Sakkie, thanks for all the breakaways which helped me to refocus my attention on my thesis.

To my sister, Rethana, thank you so much for taking over all the household responsibilities during my busiest times. Thanks also to you and Shaun for all the little things that you helped me with, they may have seemed small and insignificant at the time, but without your assistance I don't think this thesis would have ever been completed. Thank you also for all of your love and support and for not only being my sister but also one of my best friends. My special thanks also go to Ernst Swartz for his valuable comments on various parts of my thesis and for his genuine and steady support. Thanks for being there and for believing in me.

Finally, I would like to thank God for making me aware of nature and for giving me the opportunity to study and appreciate His creation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

General Introduction

General Background

African rock rats of the genus *Aethomys* Thomas, 1915, are long-tailed murid rodents of medium to large size (Meester *et al.* 1986; Skinner & Smithers 1990). The name is derived from the Greek *aithos*, meaning sunburnt and *mys*, meaning mouse (Rosevear 1969; De Graaff 1981). The genus represents a diverse group of rodents endemic to East, Central and southern Africa, with a marginal extension into West Africa (Musser & Carleton 1993). The genus is currently considered to include 11 species, namely, *A. namaquensis*, *A. granti*, *A. silindensis*, *A. chrysophilus*, *A. ineptus*, *A. nyikae*, *A. bocagei*, *A. hindei*, *A. stannarius*, *A. thomasi* and *A. kaiseri* (Musser & Carleton 1993; Chimimba *et al.* 1999) traditionally allocated to two subgenera, *Aethomys* and *Micaelamys*.

Although there is little biological data available, members of the genus are nocturnal, terrestrial and to some extent arboreal, living either in pairs or in small family units (Skinner & Smithers 1990). They utilize a wide variety of covered habitats, ranging from rock crevices, rocky terrain and outcrops, excavated burrows, tree trunks, grassland with some form of cover, to open savanna woodland (Skinner & Smithers 1990). Some species are well adapted to hot, arid environments while some aestivate during drier months (Withers *et al.* 1980). While they do not generally associate with humans to the same extent as, for example, the house mouse (*Mus musculus*), they may do so in agriculturally developed areas (Skinner & Smithers 1990). Little is known about their feeding biology, but all species currently included in the genus are considered to be omnivorous with a preference for seed, fruit, and grain (Watson 1987; Woodall & Mackie 1987; Skinner & Smithers 1990). *Aethomys namaquensis* is also known to forage on flowering heads of *Protea amplexicaulis* and *P. humiflora*, with a preference for their nectar and in the process act as pollinators for these plants (Johnson *et al.* 1999). *Aethomys chrysophilus* on the other hand, has a preference for grain crops and legumes such as sorghum, peas and beans (Skinner & Smithers 1990).

Some members of the genus are generally regarded to breed throughout the year with an average of three offspring per litter, while others show no evidence of breeding during colder months (Rautenbach 1978; Skinner & Smithers 1990). They tend to have unstable population cycles associated with high mortality and high reproductive potential (Withers *et al.* 1980), resulting in population eruptions that have been implicated in causing

epidemiological problems (Gear *et al.* 1966; Hallet *et al.* 1970; Swanepoel *et al.* 1978) and extensive damage to agricultural crops and stored grain (Wilson 1970, 1975; Smithers 1971; De Graaff 1981).

Palaeontology

Palaeontological data of some members of the genus have been recorded from South Africa (De Graaff 1960, 1961; Avery 1981, 1982, 1985; 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* 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* (Wesselman 1984; Black & Krishtalka 1986; Denys 1987).

Systematics

Taxonomically, the genus has undergone a number of nomenclatural changes (Rosevear 1969). Thomas (1915a) proposed *Aethomys* as a subgenus of *Rattus* and later elevated it to a full generic rank (Thomas 1915b). Ever since the genus was proposed, the erection of subgenera, the taxonomic allocation of species and the description of numerous subspecies, have been the source of systematic uncertainty leading to a need for a revision of the genus. This taxonomic uncertainty has been exacerbated by morphological conservatism but a high degree of chromosomal diversity within the genus (Visser & Robinson 1986, 1987). For example, *A. chrysophilus*, traditionally considered a single species was shown to contain two electrophoretically distinct cytotypes (2n = 50 / "slow" haemoglobin; 2n = 44 / "fast" haemoglobin) (Gordon & Rautenbach 1980; Gordon & Watson 1986; Visser & Robinson 1986, 1987; Baker *et al.* 1988). The two cytotypes have also been shown to differ in gross sperm and bacular morphology (Breed *et al.* 1988). Since there was no evidence of hybrids, these findings strongly suggested the presence of two sibling species.

Subsequently, Chimimba (1997) undertook an analysis of both morphometric and morphological patterns of inter- and intraspecific variation in the genus across a more comprehensive geographical coverage in southern Africa than has previously been considered for the genus. The study also included a phylogenetic appraisal of the 11 currently recognised species within the genus. The study recognised *A. namaquensis*, *A. granti* and *A. silindensis* as valid species. Furthermore, the two cytotypes within *A. chrysophilus* were formally allocated to the nominate species, *A. chrysophilus* for the 2n = 50 cytotype and a newly recognised *A. ineptus* for the 2n = 44 cytotype.

The phylogenetic analysis suggested *Aethomys* to be monophyletic with the retention of *Aethomys* and *Micaelamys* as subgenera, the latter consisting of *A. namaquensis* and *A. granti* and all remaining species allocated to the nominate subgenus *Aethomys* (Chimimba 1997). Ducroz *et al.* (2001), however, recently suggested that *Aethomys* may be paraphyletic, with *A. namaquensis* as the most basal member of a clade that would combine arvicanthine rodents and the other African murines. Interestingly, *A. namaquensis* has previously been allocated to the genus *Thallomys* (Ellerman 1941).

Morphometric analysis of intraspecific variation within *A. namaquensis* suggested the recognition of four subspecies, *A. n. namaquensis*, *A. n. lehocla*, *A. n. monticularis* and *A. n. alborarius* (Chimimba 2001a). The morphological discontinuities of these suggested subspecies broadly coincided with the major biomes of southern Africa (Chimimba 2001a). Morphometric variation within *A. chrysophilus* on the other hand, suggested the recognition of two subspecies, *A. c. chrysophilus* and *A. c. imago* whose distributions coincided with an altitudinal limit in the eastern parts of southern Africa (Chimimba 2000). Geographic variation within *A. ineptus* and *A. granti* was shown to be clinal where cranial size within *A. ineptus* was positively and significantly correlated with longitude, while *A. granti* showed a southwestern-northeasterly clinal pattern of variation (Chimimba *et al.* 1998; Chimimba 2001b).

Although the findings in the recent revision may be valid, these need to be independently tested using alternative systematic techniques before robust systematic conclusions can be made. It is clear in modern systematics that the resolution of taxonomic uncertainties is best achieved by using a multidisciplinary approach (Ducroz *et al.* 2001). To this end, the present study attempted to independently test the findings of the morphologically based

systematic revision of Chimimba (1997) by using molecular data and by following both phylogenetic and phylogeographic approaches.

The objectives of the present study therefore were to:

1. Assess molecular patterns of intraspecific variation in members of the genus *Aethomys* from southern Africa (Chapter 2; Chapter 3).

2. Determine of the status of two cryptic species, *A. chrysophilus* and *A. ineptus* from South Africa and Swaziland using molecular data (Chapter 3).

3. Elucidate of phylogenetic relationships within the genus based on molecular data in order to assess supraspecific taxonomic implications, particularly with regard to the subgenera *Aethomys* and *Micaelomys* (Chapter 4).

Approach

Molecular approach

The present investigation was based on the analysis of mitochondrial DNA (mtDNA) data. Mitochondrial DNA data is valuable for understanding evolutionary relationships among species, populations and individuals (Irwin *et al.* 1991). Animal mtDNA is a duplex, covalently closed circular molecule that replicates and transcribes protein-coding genes within the organelle itself (Avise & Lansman 1983; Moritz *et al.* 1987). Its gene content appears to be conserved, with two ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes and 13 protein-coding genes (Moritz *et al.* 1987). A "control" region that lacks structural genes but contains sequences that initiate replication and transcription is present (Moritz *et al.* 1987). Since mtDNA is maternally inherited, however, the history recorded in this molecule is not a complete characterization of the intraspecific phylogeny of a species because relationships may be obscured by gender bias such as in levels of dispersal. Nevertheless two mtDNA genes were used in the present study, namely, cytochrome b and the more conserved ribosomal 16S rRNA genes.

Despite potential limitations of the cytochrome b gene, it has proved useful in addressing questions about relationships among and within species for a range of taxa (Jansen van Vuuren & Robinson 1997; Ducroz *et al.* 1998; Smith 1998; Fumagalli *et al.* 1999; Ohdachi *et al.* 2001). More specifically, the cytochrome b gene has successfully been used to investigate systematic relationships in a number of murid rodents (Patton & Smith 1992; Smith & Patton 1993; Smith & Patton 1999; Ducroz *et al.* 1998). More importantly, its time scale calibrations are available for rodents (Smith & Patton 1993) and its rate of evolution is also well documented (Irwin *et al.* 1991). Sequences of the 16S rRNA gene were also used in the elucidation of the phylogenetic relationships of the southern African *Aethomys* because of its usefulness in resolving relationships at deeper nodes (Moritz *et al.* 1987; Halanych & Robinson 1999).

Phylogeography

The analysis of the cytochrome b gene in this study was also augmented by a phylogeographic approach. Phylogeography is the mtDNA bridge between population genetics and systematics (Avise *et al.* 1987a). It is concerned with principles and processes governing geographic distributions of genealogical lineages among and within closely related species (Avise *et al.* 1987a). Phylogeographic differentiation is usually classified as being the result of either vicariance or dispersal processes. Under vicariance scenarios, populations or taxa become separated when continuous ranges of ancestral forms are split by environmental changes, such as the rise of a mountain range. Under a dispersal scenario, active or passive dispersal from one or more ancestral centres, leads to the establishment of new populations. The population structure is, therefore, determined by the species' potential for gene flow and the environmental influences on that potential (Avise *et al.* 1987a).

Phylogeographic patterns that can be expected range from total phylogenetic discontinuity between populations to genetic continuity due to zoogeographic barriers (Avise *et al.* 1987a). Discontinuous patterns are shaped by extrinsic barriers to gene flow, or by the extinction of intermediate haplotypes in species with limited gene flow and dispersal capabilities while continuous patterns are expected for high gene flow species whose populations have not been separated by long-term biogeographic barriers (Avise *et. al* 1987a; 2000).

It was predicted that *A. namaquensis* from southern Africa may show a pattern of phylogenetic discontinuity. Although the species is widely distributed in southern Africa, it is essentially restricted to rocky habitats. This pattern of phylogenetic discontinuity is illustrated by the deer mouse (*Peromyscus maniculatus*), reflecting possible long-term zoogeographic barriers, which may have prevented deer mice from dispersing (Lansman *et al.* 1983a) and by southern African mammals such as the rock hyrax (*Procavia capensis*; Prinsloo 1993) and the red rock rabbits (*Pronolagus sp.*; Matthee 1993).

In contrast, habitat uniformity in the Springhare (*Pedetes capensis*) is reflected in its phylogeographic pattern, which shows a lack of genetic divergence among samples over a broad geographic range in South Africa (Matthee *et al.* 1997). It was predicted that *A. chrysophilus* and *A. ineptus* from southern Africa may show a pattern similar to that of the Springhare, reflecting a lack of genetic divergence among samples over a broad geographic range. This phylogeographic pattern is also evident in the yellow mongoose from southern Africa (*Cynictis penicillata*; Jansen van Vuuren 1995), the old field mouse (*Peromyscus polionotus*; Avise *et al.* 1983) and the woodrats of the eastern United States (*Neotoma*; Hayes & Harrison 1992).

Phylogenetic analysis

Molecular and phylogeographic data were analysed using phylogenetic techniques. Two distinct approaches were used: maximum parsimony that gives preference to simpler rather than complicated hypotheses and maximum likelihood as they are most widely used methods.

Species and Subspecies Definitions

While the analysis of geographic variation in widely distributed species, may lead to the recognition of distinct aggregates of local populations, a problem arises in deciding whether such aggregates represent species or subspecies (Mayr & Ashlock 1991; Mayr 1997). The problem is exacerbated further by the numerous definitions that have been proposed to define species. For example, the Biological Species Concept defines species as interbreeding natural populations that are reproductively isolated from other such populations (Dobzhansky 1940; Mayr 1942), while the Phylogenetic Species Concept

(Cracraft 1989) considers a species as a cluster of organisms, diagnosably distinct from other such clusters. The Cohesion Species Concept on the other hand, defines species as the most inclusive population of individuals having the potential for phenotypic cohesion through intrinsic cohesion mechanisms (Templeton 1989). The Biological Species Concept and its variations is the most widely used and was followed in the present investigation.

Similarly, several problems exit in defining subspecies because of different criteria that have been proposed (Ryder 1986; Moritz *et al.* 1987; O'Brien & Mayr 1991). These range from the typological subspecies definition to those identifying distinct populations based on the conservation biology concepts of Evolutionarily Significant Units (ESUs) and Management Units (MUs) (Ryder 1986; Ryder *et al.* 1988; Moritz 1994a). The role of ESUs and MUs in biological classification has generally been limited to recognizing intraspecific diversity, which is useful in the conservation and management of endangered species.

For the present study, Lidicker's (1962) subspecies definition was followed. It defines a subspecies as "a relatively homogenous and genetically distinct portion of a species which represents a separately evolving, recently evolved lineage with its own evolutionary tendencies, inhabits a definite geographical area, is usually at least partially isolated, and may integrate with adjacent subspecies". Linked to the Biological Species Concept, this subspecies definition may be useful in cases where it is difficult to test for reproductive isolation.

Relevance of Study

Apart from contributing to small mammal systematics in Africa, the present study may have implications in epidemiological and agricultural research associated with problem rodents. Some members of the genus *Aethomys* have been implicated in causing epidemiological problems (Gear *et al.* 1966; Hallet *et al.* 1970; Swanepoel *et al.* 1978) while others may cause extensive damage to crops and stored grain (Wilson 1970, 1975; Smithers 1971; De Graaff 1981). Consequently, the present study may assist health and agricultural authorities in gaining a better insight into these potentially problem rodents.

Chapter 2

Geographic mtDNA variation within *Aethomys namaquensis* from southern Africa

Abstract

The genetic differentiation among populations of the Namaqua rock mouse, *Aethomys namaquensis*, from southern Africa was examined using cytochrome b sequences. Phylogenetic analysis (neighbour-joining) revealed three genetically well-supported lineages, a lineage (B) widely distributed across the Karoo, a lineage (C) found across the grasslands of the North-West, Limpopo and Mpumalanga Provinces and a lineage (A) found in the Limpopo valley. AMOVA and minimum-spanning networks confirmed these phylogeographic partitions. These spatial distributions broadly coincide with the ranges of four previously proposed morphologically defined subspecies. From the present sample size, there is only good support from a molecular point of view, for the subspecies *A. n. lehocla* (lineage B). In addition, six unique lineages (D-I) were found which could not be assigned with confidence to any of the previously proposed subspecies. The diversity and differentiation detected within *A. namaquensis* in the present study are more complex than the four previously proposed subspecies. Low levels of sequence variation within lineage A and lineage B were detected. Lineage C exhibited substantial mtDNA variability and population structure. The geographic location of the phylogeographic discontinuity detected in this lineage coincided with a break described for other southern African small mammals. The average time of divergence separating lineages A, B and C were estimated at 2.48 Myr ago, suggesting a late Pliocene time of separation.

Introduction

The Namaqua rock mouse, *Aethomys namaquensis* Smith, 1834, is widely distributed in southern Africa. The species shows considerable geographic variation in pelage colouration, tail length, and body size throughout its distributional range (Smithers 1971; De Graaff 1981; Musser & Carleton 1993). This variation led to suggestions that it may reflect either a complex of species (Musser & Carleton 1993) or subspecies (Smithers 1971; De Graaff 1981). More recent studies, however, strongly suggested *A. namaquensis* to be a valid species in southern Africa (Chimimba 1997, 1998; Chimimba *et al.* 1999). This confirmed earlier studies (e.g., Roberts 1951; Meester *et al.* 1964) that, in addition, recognised 16 subspecies. These infraspecific distinctions were, however, made with little or no assessment of patterns of geographic variation over the entire distributional range of the species in the subregion. In addition, these distinctions were primarily based on non-statistical comparisons of type material and/or small geographically restricted samples (Smithers 1971; Smithers & Wilson 1979). Major reviews of the genus by Davis (1975), Meester *et al.* (1986), Skinner & Smithers (1990) and Musser & Carleton (1993) refrained from recognizing subspecies and regarded all described forms as synonyms of *A. namaquensis*.

To understand the infraspecific status of rodents in southern Africa, it is necessary to assess the distribution of lineages through an evaluation of intraspecific variation. Only a few southern African rodents such as species within *Acomys* (Dippenaar & Rautenbach 1986), *Otomys* (Taylor & Meester 1993), *Saccostomus* (Ellison *et al.* 1993) and *Aethomys* (Chimimba 2001a), have been subjected to analysis of intraspecific variation.

Prior to a recent intraspecific morphometric study within *A. namaquensis* from southern Africa (Chimimba 2001a), the nature and extent of geographic variation within the species in the subregion remained virtually unknown. The study, based on a more comprehensive geographical coverage than has hitherto been considered for the species and on a classical infraspecific taxonomic interpretation of the detected variation, suggested the recognition of four subspecies: *A. n. namaquensis* Smith, 1834; *A. n. lehocla* Smith, 1836; *A. n. alborarius* Peters, 1852 and *A. n. monticularis* Jameson, 1909 which differed in both cranial size and shape (Chimimba 2001a). The geographical limits of the proposed

subspecies were suggested to broadly coincide with the major biomes of southern Africa (Rutherford & Westfall 1986; Fig. 2.1). The subspecies *namaquensis* was shown to be largely associated with a combination of the Succulent Karoo, Fynbos and the southern coastal Savanna/Grassland region of the Eastern Cape, KwaZulu-Natal and eastern Mpumalanga Provinces of South Africa, *alborarius* with Savanna, *lehocla* with Nama-Karoo and *monticularis* with Grassland (Chimimba 2001a; Fig. 2.1). These morphometrically defined groups have, however, not been tested using other alternative systematic techniques such as cytogenetics and/or DNA analysis.

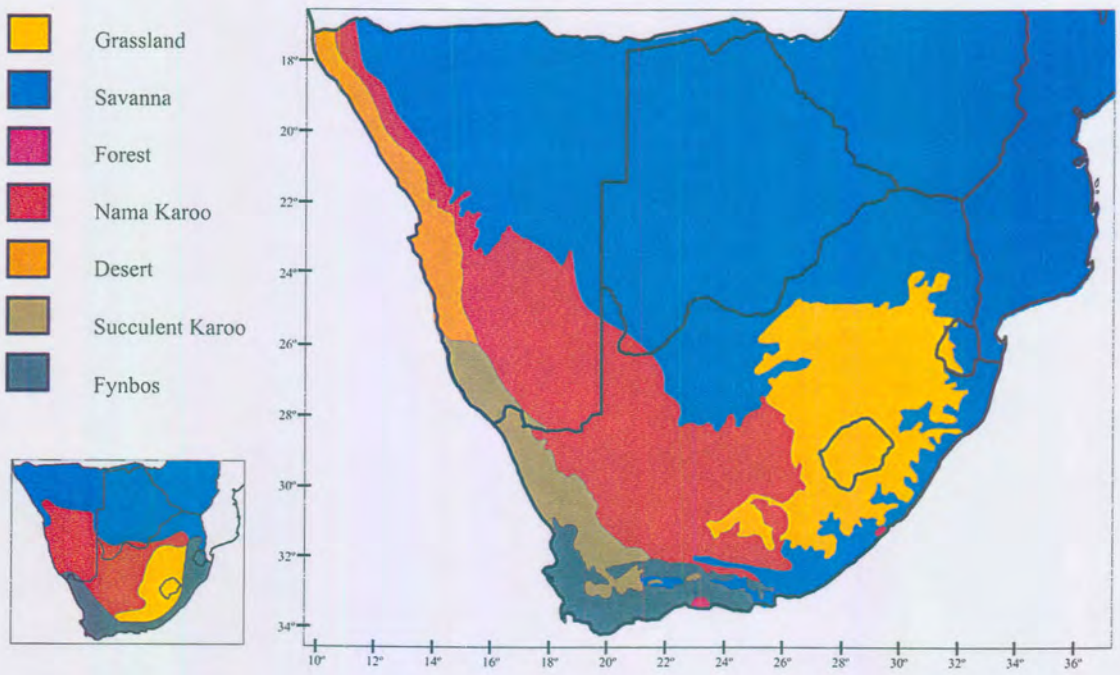


Fig. 2.1 Major biomes of southern Africa (after Rutherford & Westfall 1986). The insert shows the geographic distribution of the phenetic diversity within *A. namaquensis* from southern Africa. The blue, red, yellow and green zones correspond with classical taxonomically defined subspecies: *A. n. alborarius*, *A. n. lehocla*, *A. n. monticularis* and *A. n. namaquensis*, respectively (after Chimimba 2001a).

Of particular relevance in the use of molecular data (such as mtDNA analysis) in assessing intraspecific diversity, more especially in a widely distributed species such as *A. namaquensis*, is the opportunity to determine the phylogeographic structure within the species. Phylogeography affords insight into the processes that lead to geographic distributions of genealogical lineages among and within closely related species (Avise 2000). Limited gene flow and hence evolution in allopatry, have influenced the development of specialised morphological, reproductive and behavioural characteristics found in rock-dwelling mammals (Mares & Lacher 1987). Consequently, traits such as limited dispersal capabilities, strong territoriality, patchiness of the environment and social structuring (communal vs. solitary) may characterize African saxicolous (rock-dwelling) mammalian fauna (Mares & Lacher 1987) such as the rock hyrax, the klipspringer, the red rock rabbits and the Namaqua rock mouse. Molecular studies on saxicolous mammals have shown that population structure is shaped by the island-like nature of their habitat (Patton *et al.* 1996; Kim *et al.* 1998).

In general, rodents have short dispersal distances and rapid generation times (Pusey 1987). Members of the genus *Aethomys* are believed to live between one to two years, during which as many as four litters can be produced by a single female (Chimimba pers. comm.). For *A. namaquensis*, the population size would be expected to be large because of its small body size and short generation period, with breeding during nine months of the year (Smithers 1971). The average number of foetuses ranges between one and seven (Rautenbach 1978). Withers *et al.* (1980) reported that *A. namaquensis* tends to have unstable population cycles associated with high reproductive potential and high mortalities. In mammals with such generation times and large population sizes, the new maternal lineages would disperse quickly from their points of origin. This would create genetic structuring in which the oldest lineages would have the broadest geographical distribution (Neigel & Avise 1993). Over time, such a species would exhibit genetic isolation by distance (Patton *et al.* 1996).

If the landscape matrix between habitat islands severely retards dispersal, maternal lineages would disperse even more slowly. These maternal lineages will have more limited geographical ranges in comparison to species living in the landscape matrix between habitat islands (Kim *et al.* 1998). Habitat islands with sufficient size and stability in resources might reach equilibrium and even accumulate its own set of related lineages.

In contrast, habitat islands with limited resources might become extinct from time to time and then be replaced by dispersers from habitat islands close by (Patton *et al.* 1996). This raises the possibility that local genetic diversity and genetic distances across geographical ranges could reflect dispersal power and demographic stability (Gaggiotti 1996).

Generally, studies of systematic relationships within *Aethomys* have focused on morphological and morphometric variation. Variability in features such as coloration, dentition and properties of the skull, however, are often limited their usefulness (Smithers 1971; Smithers & Wilson 1979). This stressed the need to use new investigatory parameters, such as molecular techniques in combination with the morphological and morphometric information in order to assess the phylogenetic diversity within *A. namaquensis*. The present study, therefore, represented the first analysis of molecular (mtDNA) variation within *A. namaquensis* from southern Africa and addressed the following questions:

1. Is there any concordance between previously morphologically defined and the present study's molecular patterns of variation within *A. namaquensis* from southern Africa?

The results of this part of the study were in turn interpreted with reference to the distribution of the derived groupings within the species.

2. Is there any phylogeographic structuring within *A. namaquensis* from southern Africa?

The results of this part of the study were in turn interpreted with reference to the phylogeographic structure of other rock-dwelling species, the rock hyrax (*Procavia capensis*; Prinsloo 1993) and the rock rabbits of the genus *Pronogalus* (Matthee & Robinson 1996).

Materials and Methods

Study area and sampling

An attempt was made to collect samples representative of the four proposed subspecies in South Africa. A total of 70 samples were collected from 16 localities in South Africa and

one locality in Botswana, spanning the four major phytogeographic zones in southern Africa (Fig. 2.2). Geographic coordinates for these localities are listed in Table 2.1. Individuals were collected between May 2000 and September 2001. Sherman live traps or snap traps were set out in grids of 50 to 100 traps. Traps were placed 10 to 12 m apart, baited with a mixture of peanut butter, oats and pilchards. Individuals were euthanased with an inhalation anaesthetic, Halothane (Safeline Pharmaceuticals). Individuals were then dissected and heart, liver, kidney and muscle were removed for molecular analyses. In one instance (Kgaswane Mountain Reserve), due to the ecological nature of a related study, toe clipping was used as a means of obtaining tissue samples. Animals were also weighed (Pesola scale in grams) and measured using digital callipers (mm): head and body (from tip of nose to anus), tail length (anus to tip) and foot length. The sex and breeding status of individuals were also recorded with males scored for testicular development, while females were recorded as perforate and scored for nipple development. All tissues and the toe-clippings were stored at room temperature in a saturated NaCl solution supplemented with 20% DMSO (Dimethyl-sulfoxide), a technique developed for the long-term preservation of whale skin (Amos & Hoelzel 1991). Voucher specimens were deposited in the Transvaal Museum (Pretoria). Dental morphology (an M_1 with three cusps in the anterior row, see Chimimba *et al.* 1999) was used for species identification, specifically to distinguish *A. namaquensis* from other murid rodents.

DNA extraction

Total genomic DNA was extracted as follows: 0.2 g of tissue (heart or kidney) or a toe-clipping was mixed with 500 µl extraction buffer (0.05 M Tris-HCl, 0.5 M EDTA-Na_2, 1.0 M NaCl and 10% SDS) and 0.5 mg Proteinase K (Roche Diagnostics). Samples were digested overnight at 55°C, followed by a digestion with 0.1 mg RNase A (Roche Diagnostics) for 30 minutes at 37°C. Impurities were removed by phenol (repeated three times) followed by an extraction step with a chloroform:isoamyl alcohol (24:1) solution (Sambrook *et al.* 1989). Thereafter, DNA was precipitated overnight at -20°C in a solution containing 2.5 volumes of ice-cold 96% EtOH and 0.1 volumes of 3 M NaAc (Sambrook *et al.* 1989). DNA was pelleted in a microcentrifuge at 13000 rpm for 20 minutes after which it was washed with 70% EtOH. Pellets were air-dried and the DNA was resuspended in 50-100µl Sabax® water (Adcock Ingram).

Table 2.1 Geographic coordinates of all collecting localities of *Aethomys namaquensis* in southern Africa analysed in the present study. Numbers 1-17 correspond to those in Figure 2.2.

Locality	Country	Province	Geographic Coordinates
1. Farm: Terrafou, Botswana	Botswana		22 27'29"S 28 45'32"E
2. Messina Nature Reserve	South Africa	Limpopo	22 24'45"S 30 03'01"E
3. Selati Nature Reserve, Hoedspruit	South Africa	Limpopo	24 09'30"S 30 40'50"E
4. Amanita Safaris, Thabazimbi	South Africa	North West	24 09'16"S 26 55'05"E
5. Kgaswane Mountain Reserve	South Africa	North West	24 44'20"S 27 12'56"E
6. Kruisrivier Nature Reserve, Loskop Dam	South Africa	Mpumalanga	25 21'08"S 29 32'26"E
7. Wathaba-Uitkomst, Machadodorp	South Africa	Mpumalanga	25 47'31"S 30 22'28"E
8. Farm: Vlakfontein, Vryburg	South Africa	North West	27 04'22"S 24 46'07"E
9. Farm: Koedoesberg, Pongola	South Africa	Kwa-Zulu Natal	27 26'31"S 31 41'41"E
10. Farm: Steenkampsput, Upington	South Africa	Northern Cape	28 06'13"S 20 54'09"E
11. Farm: Zwartbooisberg, Augrabies National Park	South Africa	Northern Cape	28 02'30"S 20 42'55"E
12. Farm: Rietfontein, Springbok	South Africa	Northern Cape	29 51'40"S 18 11'10"E
13. Gariep Nature Reserve	South Africa	Free State	30 35'56"S 25 32'03"E
14. Farm: Rietpoort, Loxton	South Africa	Northern Cape	31 38'30"S 22 22'34"E
15. Farm: Grootfontein, Porterville	South Africa	Western Cape	32 54'28"S 19 06'31"E
16. Kirkwood	South Africa	Eastern Cape	33 24'20"S 25 25'30"E
17. Farm: Brakrivier, Oudtshoorn	South Africa	Western Cape	33 46'19"S 22 31'45"E

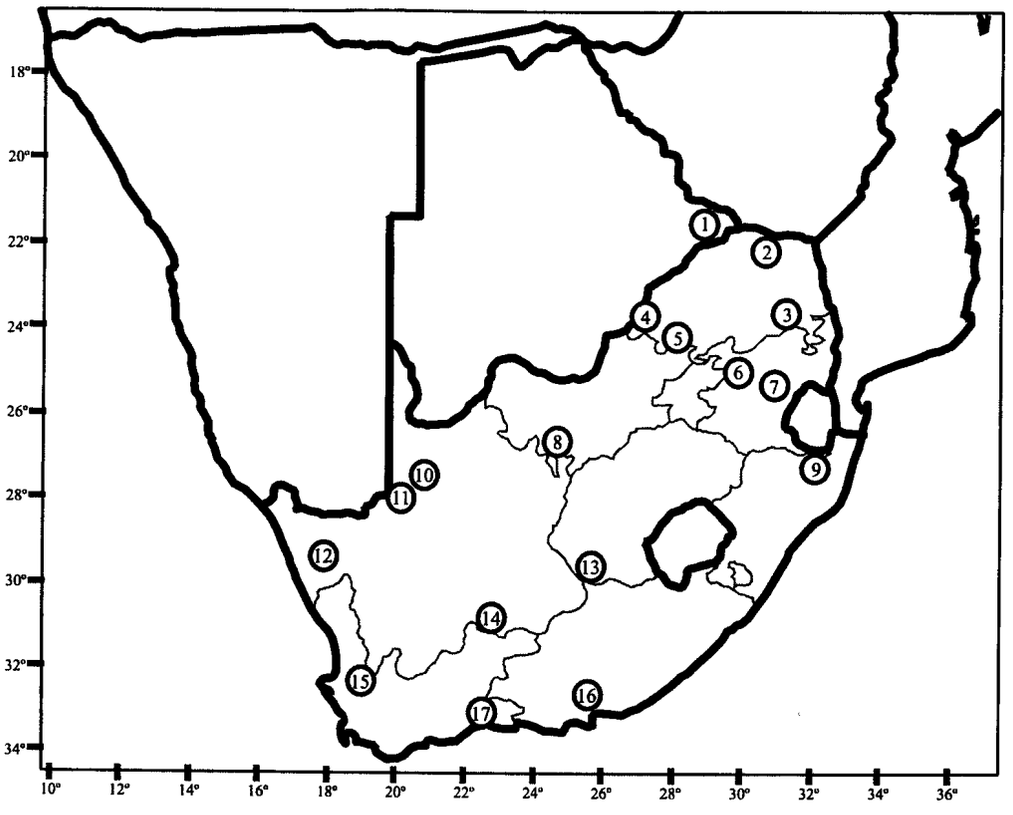


Fig. 2.2 Collecting localities of samples of *Aethomys namaquensis* from southern Africa. Numbers in circles correspond to the locality numbers in Table 2.1. 1 – Farm: Terrafou, Botswana; 2 – Messina Nature Reserve; 3 – Selati Nature Reserve, Hoedspruit; 4 – Amanita Safaris, Thabazimbi; 5 – Kgaswane Mountain Reserve; 6 – Kruisrivier Nature Reserve, Loskop Dam; 7 – Wathaba-Uitkomst, Machadodorp; 8 – Farm: Vlakfontein, Vryburg; 9 – Farm: Koedoesberg, Pongola; 10 – Farm: Steenkampsput, Upington; 11 – Farm: Zwartbooisberg, Augrabies National Park; 12 – Farm: Rietfontein, Springbok; 13 – Gariep Nature Reserve; 14 – Farm: Rietpoort, Loxton; 15 – Farm: Grootfontein, Porterville; 16 –Kirkwood; 17 – Farm: Brakrivier, Oudtshoorn.

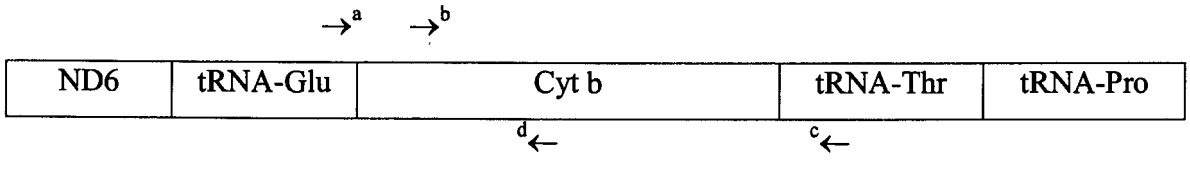
Primer design, PCR amplification and cycle sequencing

A mouse (*Mus musculus*) specific primer was first designed in the tRNA-Thr flanking the cytochrome b gene from a GenBank sequence J01420 (Bibb *et al.* 1981). This primer, H15915 MUS, in combination with the shortened universal vertebrate primer L14724 (Paäbo *et al.* 1988) that anneals in the tRNA-Glu, was used to amplify the cytochrome b gene of two *A. namaquensis* individuals. An *A. chrysophilus* sequence (Ducroz *et al.*

1998, GenBank accession # AF004587), together with these *A. namaquensis* sequences were aligned in Clustal X (Thompson *et al.* 1997) and used to design an internal species-specific *A. namaquensis* primer, Nam IH. This species-specific primer was used in combination with either the shortened L14724 (Paäbo *et al.* 1988) or L14841 (Kocher *et al.* 1989) to amplify the 5' end of cytochrome b (a 537 bp or a 655 bp fragment, respectively). Initially, the complete cytochrome b gene was sequenced using L14724 and H15915 MUS in combination with species-specific internal primers to obtain the complete gene sequence (a 1194 bp fragment). A preliminary analysis showed that the 5' end of the gene yielded considerable levels of variation between different *A. namaquensis* sequences. Consequently, the remaining individuals were only sequenced with the primer combinations that yielded the 5' end of the gene. These primer names, sequences and annealing positions on the mouse mtDNA genome are illustrated in Table 2.2.

Table 2.2 Primer names, sequences (5' to 3') and positions of the primers used in Polymerase Chain Reactions (PCRs; Saiki *et al.* 1988) and cycle sequencing. Rodent specific primer names refer to the position of the 3' nucleotide in the mitochondrial DNA sequence of *Mus musculus* (Bibb *et al.* 1981). Superscripts a-d correspond to the letters in the diagram.

Primer name	Rodent specific name	Primer sequence
L14724[a]	L14115	5' TGAYATGAAAAAYCATCGTTG 3'
L14841[b]	L14233	5' CCAACATCTCAGCATGATGAAA 3'
H15915 MUS[c]	H15309	5' CATTTCAGGTTTACAAGAC 3'
Nam IH[d]	H14769	5' GTCTGCGTCTGAATTTAG 3'



Polymerase Chain Reaction (PCR; Saiki *et al.* 1988) amplifications, including a negative control, were performed in a total volume of 50 µl containing approximately 100 ng of genomic DNA as template or Sabax® water in the negative control reactions, 5 µl 10 x buffer, 2 mM $MgCl_2$, 0.2 mM of each of the four nucleotides (Promega), 25 picamol of each primer and 1.5 U of Super-Therm® DNA polymerase (Southern Cross Biotechnology). The PCR conditions were: denaturing at 94°C for 4 minutes; 35 cycles of denaturing at 94°C for 45 seconds, primer annealing at 55°C for 1 minute, elongation at 72°C for 1 minute; followed by an extended elongation at 72°C for 7 minutes in a Geneamp® PCR System 9700 (PE Applied Biosystems). Subsequently, the PCR products (a 537, 655 or 1194 bp fragment) were purified using the High Pure™ PCR Product Purification Kit (Boehringer Mannheim). The protocol as prescribed by the manufacturers was followed and the purified DNA was eluted in 50 µl of Sabax® water.

Cycle sequencing was performed in a Geneamp® PCR system 9700, making use of the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems). The reaction mix contained the following: 60-80 ng of purified DNA as template, 1.6 or 3.2 picamol primer (Nam IH, L14724 or L14841), and 2 or 4 µl of the Big Dye reaction mix for quarter or half reactions, respectively. Some individuals were sequenced with H15915 MUS followed by sequencing with Nam IH. Nucleotide sequences were determined through electrophoresis on an ABI 377 automated sequencer or through a capillary system on an ABI 3100 automated sequencer (PE Applied Biosystems).

Sequencing analysis

The raw sequence data were checked for quality and background peaks in Sequence Analysis version 3 (PE Applied Biosystems). Thereafter, successful sequences were imported into Sequence Navigator version 1.01 (PE Applied Biosystems). The forward and reverse sequences of each individual were aligned in order to compute a 550 bp consensus sequence for each individual. All *A. namaquensis* sequences were deposited in GenBank. The modified sequences of all individuals were aligned in Clustal X (Thompson *et al.* 1997), after which, the different sequences were imported into PAUP version 4.0b10 (Swofford 2002). MacClade version 3.0 (Maddison & Maddison 1992)

was used in the translation of the nucleotide sequences and to estimate the transition:transversion (Ti:Tv) ratio based on 1000 random trees (Halanych 1996).

Outgroup choice

The selection of possible outgroups for *A. namaquensis* was difficult because of previous proposals of systematic links between *Aethomys* and numerous other rodent species (Chimimba 1997). Preliminary screening of character-state distributions among all possible outgroups, relative to ingroup taxa used by Chimimba (1997), showed evidence of character polarities except for *Dasymys incomtus* and *Rattus rattus*. These two species were considered as outgroups in the elucidation of phylogenetic relationships within the genus. Their suitability as potential outgroups was supported by previous morphological (Ellerman 1941), palaeontological (Denys 1990a, b) and molecular studies (Verheyen *et al.* 1995). Ideally, *R. rattus* should have been used as an outgroup for direct comparison to Chimimba (1997). The available sequences for *R. rattus*, however, are very variable. Instead, other members of the subfamily murinae, *Mus musculus* and *D. incomtus* were chosen as outgroups in all analyses in this study.

Phylogenetic analyses

A maximum likelihood ratio test as implemented in Modeltest version 3.06 (Posada & Crandall 1998) was used to determine the best-fit model of DNA substitution for the sequence data. The best-fit model obtained from this test was then used in all distance analyses. Sequence statistics such as the Ti:Tv ratio, the α parameter of the gamma distribution of rate variation among sites (Yang *et al.* 1994; Yang 1996; Yang & Kumar 1996) and the proportion of invariable sites (I) were also estimated. Genetic variation was assessed by calculating the proportion of pairwise nucleotide transitions and transversions between all pairs of specimens, as well as by estimating the pairwise protein differences. Pairwise estimates of the nucleotide sequence divergence were calculated using the best-fit model of substitution as estimated in Modeltest version 3.06 (Posada & Crandall 1998) and the HKY85 model of substitution (Hasegawa *et al.* 1985).

All distance-based analyses were performed on the different haplotype sequences identified using the pairwise estimates of the nucleotide sequence divergence. Distance trees were constructed using the neighbour-joining algorithm of Saitou and Nei (1987) in

PAUP version 4.0b10 (Swofford 2002). Robustness of the neighbour-joining tree was assessed using bootstrap analysis (Felsenstein 1985) with 1000 replicates.

To avoid the confounding effects (noise caused by multiple changes) of saturated characters (Hackett 1996; Halanych *et al.* 1999) and to optimise the number of informative characters, the observed number of pairwise sequence differences (time) were plotted against the number of pairwise transitions and transversions at different codon positions. Depending on the data, characters were either treated as unweighted if no saturation was evident or weighted according to the level of saturation.

Both maximum parsimony (Kluge & Farris 1969; Farris *et al.* 1970) and maximum likelihood (Felsenstein 1973, 1981) analyses were performed on 19 individuals, representing two or three individuals from each lineage identified in the neighbour-joining tree. Maximum parsimony analysis was based on the heuristic search option using the tree bisection-reconnection (TBR) option and with nucleotides treated as unordered characters. If more than one minimum length tree were obtained, a strict consensus tree was constructed. Tree-length, consistency index (CI; Kluge & Farris 1969), retention index (RI; Farris 1989), and rescaled consistency index (RC; Archie 1989; Meier *et al.* 1991) were calculated for tree descriptions. Phylogenetic signal was assessed by evaluating the tree-length distribution of 1000 randomly generated trees, the g_1 statistic (Hillis & Huelsenbeck 1992). Maximum likelihood trees were generated by first estimating the base frequencies, Ti:Tv ratio, gamma shape parameter (α) and the proportion of invariable sites (I). These estimated values were then used to construct likelihood trees. Support values for internal nodes were determined using bootstrap analysis with 1000 iterations (Felsenstein 1985).

A relative rate test was performed using the Relative-Rate test with a tree (RRTree; Robinson *et al.* 1998) in order to assess differences in third position transversion rates among separate lineages relative to an outgroup (Robinson *et al.* 1998; Robinson-Rechavi & Huchon 2000). The closest relative outgroup was used as a reference taxon, since the amount of noise being introduced on the branch leading to the reference taxon may influence the results. A condition for the selection of an outgroup in this case is that the branch separating outgroups from ingroup taxa must not be considered unresolved (support under threshold) and the outgroup must be monophyletic relative to the ingroup

(Robinson-Rechavi & Huchon 2000). Consequently, *Dasymys incomtus* was used as reference taxon. Repetition of this test, because of all pairwise comparisons, increases the occurrence of false significant results. To address this problem, Robinson-Rechavi & Huchon (2000) proposed adjusting the significance threshold (traditionally 5%) by dividing the latter with the number of pairwise comparisons (see http:/pbil.univ-lyon1.fr/software/rrtree.html). A specific rate of change calibrated on murid rodent data was determined since murid mtDNA has been shown to evolve at a faster rate than that of other rodents (Catzeflis *et al.* 1992). As a calibration point, sequence data from the murid rodents *R. rattus* and *M. musculus*, with a divergence date estimated at 12 Myr based on fossil records, was used (Jacobs & Downs 1994). The number of third position transversions was used as a measure of genetic divergence because it accumulates nearly linearly with time (Irwin *et al.* 1991). The conventional rate of 2% sequence divergence per million years was also applied to the data (Brown *et al.* 1979).

Population level analyses

Population genetic parameters such as haplotype diversity (the probability that two randomly chosen mtDNA sequences are different in the sample) (Nei & Tajima 1981) and nucleotide diversity, π (the average number of nucleotide differences per site between two sequences) (Nei 1987) were calculated using DnaSP version 3.51 (Rozas & Rozas 1999). Minimum-spanning networks were constructed showing the minimum mutational steps between different haplotypes using MINSPNET (Excoffier & Smouse 1994) and Arlequin version 2.0 (Schneider *et al.* 2000). Minimum-spanning networks, in conjunction with frequencies and geographic distributions of different haplotypes, can depict geographical and potential ancestor-descendant relationships among identified haplotype sequences.

A Mantel test as implemented in Mantel version 2.0 (Mantel 1967) was used to test for isolation by distance. The test uses a permutation procedure (1000 permutations) to determine the significance of the correlation between matrices of genetic versus geographic distance.

An analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) was used to assess the extent of differentiation among populations by calculating ϕ_{st}, ϕ_{ct} and ϕ_{sc} values. Statistical significance of the different parameters was tested using a non-parametric

permutation scheme based on 10 100 permutations as implemented in Arlequin version 2.0 (Schneider *et al.* 2000). A 5% level of missing data per site was allowed in all analyses. This method builds on classical analysis of variance to compute molecular variance components at different hierarchical levels. This variance analysis partitions the total variance into covariance components due to intra- and inter-individual differences and/or inter-population differences (Excoffier *et al.* 1992; Weir 1996).

Inference of population history was assessed using mismatch distribution analysis (Rogers & Harpending 1992) as implemented in Arlequin version 2.0 (Schneider *et al.* 2000), allowing for a 5% level of missing data per site. The expansion null-hypothesis was tested using the sum of squared deviation test of significance (P_{Ssd}; Schneider & Excoffier 1999) and Harpending's raggedness index of significance (P_{Rag}; Harpending 1994). Statistical significance was tested using 10 000 permutations. The process of habitat contraction and/or expansion i.e., the presence or absence of geographic expansion, often implies demographic variations (Petit *et al.* 1999). Therefore, episodes of population growth and decline have a strong effect on the pattern of genetic polymorphism, leaving characteristic signatures in the distribution of nucleotide site differences between individuals (Slatkin & Hudson 1991). Unimodal distributions are assumed to be a signature of a previous population expansion event(s) (Rogers & Harpending 1992) because the implications of continued exponential growth cannot be distinguished from those of a sudden burst of growth. On the contrary, populations in equilibrium show multimodal distributions with curves that are free of waves (Rogers & Harpending 1992). These ragged distributions can be interpreted to imply fragmentation and isolation (Slatkin & Hudson 1991).

The neutrality test of Tajima (Tajima 1989a, b, 1993) was performed to determine whether the pattern of sequence polymorphism in populations conformed to the neutral Fisher-Wright model. The test statistic, D, is based on the difference between two estimators of the neutral polymorphism parameter ($4N_e\mu$), one based on the total number of polymorphic nucleotide sites observed and the other on the average number of differences between all pairs of sequences sampled. This method tests the fit of the data to the expected neutral site frequency distribution. A negative D-value indicates an excess of rare frequency variants while positive D-values indicate an excess of intermediate variants (Ford 2002). Although this test is classically used to test the neutrality hypothesis, it can also be used to infer population expansions as negative D-values are predicted for population growth

(Tajima 1989a; Simonsen *et al.* 1995; Knowles *et al.* 1999; Pereira *et al.* 2001; Bernatchez 2001). The test is, however, sensitive to violations of the model's demographic assumptions, which may lead to uncertainty in the interpretation of test results (Tajima 1989b; Simonsen *et al.* 1995). Fu's F_s test (Fu 1997) was also performed to test the neutrality of the data because neutrality tests are often based on different models of evolution. It is therefore more appropriate to use different tests rather than one test (Wall 1999). These two neutrality tests were performed using Arlequin version 2.0 based on resampling with 16 000 replicates (Schneider *et al.* 2000).

Results

Sequence statistics

A 550 bp fragment of the 5' end of the cytochrome b gene was sequenced for 70 specimens of *A. namaquensis* and no insertions or deletions were observed. All sequences obtained were of the mitochondrial cytochrome b gene, because no stop codons were found and the nucleotide sequences corresponded to the expected 183 amino acids of the first 550 bp of the cytochrome b gene, and this was confirmed by a blast search. The 90 variable positions defined 37 haplotypes (Table 2.3; Appendix 2.1). Haplotype NH30 was the most widespread being recorded from four localities (eight individuals) while most haplotypes were locality-specific.

Mean base compositions at the three codon positions for the whole sample studied are provided in Table 2.4. The four nucleotides do not occur in equal frequencies and base composition was similar to that of other mammalian cytochrome b sequences reported in the literature (Irwin *et al.* 1991; Ducroz *et al.* 1998; Santucci *et al.* 1998; Verheyen *et al.* unpubl.). This strong bias in base composition, a feature characteristic of the cytochrome b gene and other mitochondrial protein coding genes, showed a marked under-representation of guanine at both second (16.6%), and especially third codon positions (2.1%). In addition, there was a higher representation of adenine at third (46.7%) and thymine at second codon positions (41.4%). Of the 90 variable sites, 70 were phylogenetically informative (without outgroups). Of the informative sites, nine were at first codon positions, none at second positions and 61 were at third codon positions.

Consequently, most of the substitutions were silent with only eight variable amino acid sites.

The best-fit model was determined to be the HKY85 model of substitution (Hasegawa *et al.* 1985) with a gamma correction (Gu & Zhang 1997) of 0.7287, and a value of 0.6024 for the proportion of invariable sites. Maximum likelihood (the α shape parameter of the gamma distribution of rate variation among sites, the proportion of invariable sites (I), base compositions and the Ti:Tv ratio) and MacClade version 3.0 (Ti:Tv ratio) yielded the same sequence statistics as Modeltest version 3.06. Pairwise estimates of the nucleotide sequence divergence (HKY85 and HKY85 + Γ + I) are given in Table 2.5 where the HKY85 sequence divergence values between the 37 maternal haplotypes ranged between 0.18% and 6.22%. Pairwise estimates of Ti:Tv and the pairwise protein differences are reported in Table 2.6 with the overall Ti:Tv ratio estimated at 15:1. Synonymous changes out-numbered non-synonymous changes (Table 2.6) where the latter ranged between 0.44 and 1.23, while the former changes ranged between 1.95 and 6.94 (Table 2.6). The structural model for the cytochrome b gene (Howell 1989) suggests that most of the variable amino acid positions are located either in the trans-membrane segments, or at the amino or carboxy ends of the protein (Irwin *et al.* 1991). The outer surface of the protein appears to evolve more slowly than either the trans-membrane region or the inner surface of the protein (Esposti *et al.* 1993). Although the amino acid sequences in this study showed more variable amino acid residues in the outer surface of the protein, no variable residues were detected in the previously proposed highly conserved amino acid positions (Howell 1989; Irwin *et al.* 1991; Esposti *et al.* 1993). The reason for detecting more variable amino acid residues in the outer membrane of the protein might be because only 183 amino acids were included in this study.

Table 2.3 Variable sites of 37 mitochondrial DNA cytochrome b haplotypes (550 base pairs) of *Aethomys namaquensis* from southern Africa. Variable positions nine and 549 correspond to positions 14181 and 14721 of *Mus musculus* (Bibb *et al.* 1981). Dots (.) indicate identity to the base in the reference sequence NH01 and question marks (?) indicate undetermined bases. Haplotypes do not follow each other numerically because this specific haplotype sequence corresponds to the different groups in the phylogenetic trees. Haplotype order corresponds to that of Figure 2.4 and Tables 2.5 & 2.6. Population numbers correspond to those in Figure 2.1.

Haplotype No.	Variable positions	No. of individuals	Populations (Population No.)
	111111111112222222222222333333333333333344444444444444444444555555555555555 112233356667778901235678889001124677789901112234455678990122233344455678889000111222344 44 925170491069258248262687369912035340365459258140281732403550362580253951369014703605810379		
NH01	TTCTCCGAAACTATTAGTTACCTCCACCAGTAACATACATCCCAACTTCATTAAAAACCGCCCTACATATGCATCCCTACCCTACCTCGT	1	Farm: Terrafou, Botswana (1)
NH02	......A....C.....A.....	2	Farm: Terrafou, Botswana (1)
NH03	???????.....G.....G.....A.....	2	Farm: Terrafou, Botswana (1)
NH04	....T.....A.....G.....A.....	1	Farm: Terrafou, Botswana (1)
NH05	.....G.....G.....A.....	2	Messina Nature Reserve (2)
NH06	.....C.....A.....	2	Messina Nature Reserve (2)
NH08	......A.....G.....A.....A.	1	Messina Nature Reserve (2)
NH31	........G.T..C..A.....CT.TTTC......CG...T....T........T..T.T....G....C.......C....C...C.AC	4	Farm: Rietpoort, Loxton (14)
NH36	??......G.T..C..A.....CT.TTTG......CG...T....T........T..T.T....G....CT.....TC....C...C.AC	5	Farm: Brakrivier, Oudtshoorn (17)
		1	Gariep Nature Reserve (13)
		1	Kirkwood, middle of town (16)
NH37	........G.T.GC..A.....CT.TTTG......CG...T....T........T..T.T....G....CT.....TC....C...C.AC	1	Farm: Brakrivier, Oudtshoorn (17)
NH34	????????G.T..C..A.C...CT.TTTG......CG...T....T........T..T.T....G....CT.....TC....C...C.AC	1	Fram: Grootfontein, Porterville (15)
		2	Kirkwood, middle of town (16)
NH26	........G.T..C..A.....CT.TTTG.C....CG...T....T........T..T.T.........C......TC..T.C...C.AC	3	Farm: Rietfontein, Springbok (12)
NH27	........G.T..C..A.....CT.TTTG.C....CG...T....T........T..T.T....G....C......TC..T.C...C.AC	1	Farm: Rietfontein, Springbok (12)
NH28	........G.T..C..A.....CT.TTTG......CG...T....T........T..T.T....G....C......TC....C...C.AC	1	Farm: Rietfontein, Springbok (12)
NH29	......A.G.T..C..A.....CT.TTTG......CG...T....T........T..T.T....G....C......TC....C...C.AC	1	Farm: Rietfontein, Springbok (12)
NH30	........G.T..C..A.....CT.TTTG......CG...T....T........T..T.T....G....C.......C....C...C.??	4	Gariep Nature Reserve (13)
		1	Kirkwood, middle of town (16)
		1	Farm: Steenkampsput, Upington (10)
		2	Farm: Zwartbooisberg, Augrabies National Park (11)
NH23	........G.T..C..A.....CT.TTTG......CG...T....T........T..T.T....G....C.....T.C....C...C.AC	1	Farm: Steenkampsput, Upington (10)
NH14	???????????...C....G...T.TTT....G.......T..G.T..TG.......T.T...C.....CA..CT..C.T...G..C.A.	1	Kruisrivier Nature Reserve (6)
NH15	.C.C......T........G...T.TTT....GG.......T..G.T..TG.......T.T...C..G..CA..CT..C.T...G..C.A.	1	Kruisrivier Nature Reserve (6)
NH16	.C.C......TC.......G...T.TTT....G.......T..G.T..TG.......T.T...C....?????????????????????	1	Matbeste Nature Reserve, Machadodorp (7)
NH07	.C.C......T.....T.TTT.AC.G...G...T..G.T.CT........T.C...C.....CA..C...C.T......C.A.	1	Messina Nature Reserve (2)
NH10	.C.C......T.....T.TTT..C.G...G...T..G.T..T........T.C...C.....CA..C...C.T.....TC.??	1	Selati Nature Reserve (3)
NH09	...C......T.....T.TTT..C.G...G...T..G.T..T........T.C...C.....CA..C...C.T......????	1	Selati Nature Reserve (3)
NH12	CC.C......TC......C....T.TTT....G...G...TT.G.T..T.......GT.C...C.....C...C...C.T.T....C.A.	1	Rustenburg Nature Reserve (5)
NH13	CC.C......TC.....T.TTT....G...G...TT.G.T..T.......GT.C...C.....C...C...C.T.T....C.A.	1	Rustenburg Nature Reserve (5)
NH11	?????????.T.....T.TTT....G...G...T....T..T........T.C...C.....CA..C...C.T.T....????	1	Amanita Safaris, Thabazimbi (4)
NH32	.....TT.T.TTT......GA....T.T..T..T...G...GTTT........G.AT.C...C.T......????	2	Fram: Grootfontein, Porterville (15)
NH33	.....TT.T.TTT......G.....T.T..T..T...C...GTTT........G.AT.C...C.T......C.A.	2	Fram: Grootfontein, Porterville (15)
NH35	.....TT.T.TTT......G.....T.T..T..T...C....TTT........G.AT.C...C.T......C.A.	1	Fram: Grootfontein, Porterville (15)
NH18	..T......G.......C.....TTT.TG.......G...T....T.........G.T.T..T......C..G....CG.......C.A.	1	Farm: Vlakfontein, Vryburg (8)
NH20	..T.....C.....TTT.TG.......G...T....T...........T.T..T......C..G....CG.......C...	1	Farm: Vlakfontein, Vryburg (8)
NH22	..T.....C.....TTT.TG..G....G...T...GT.........G.T.T..T......C..G....C.........C...	4	Farm: Steenkampsput, Upington (10)
NH19	........G..C.....C.....T.TTT.......CG.G.TTT..T..T....G..GT.T.....G...C.......C.T..C...C.A.	2	Farm: Vlakfontein, Vryburg (8)
NH17	.C.....T.TTTG...........T..G.TC.T........T.TA........CA..C...C.........CTA.	1	Matbeste Nature Reserve, Machadodorp (7)
NH21	...C.T.G.....CT.TTTG...........T..G.T..T........T.T.T.......CA..C...C.T......C???	1	Farm: Koedoesberg, Pongola (9)
NH24	.C......G.......A......T.TT......T...T.C.....T..T........T.T.......C.C.......C.T....TTC.AC	3	Farm: Zwartbooisberg, Augrabies National Park (11)
NH25	.C....A.G.......A......T.TT.G....T..GT.C...G.T..T.C......T.T.......C.C.......C.T....TTC.AC	2	Farm: Zwartbooisberg, Augrabies National Park (11)

Table 2.4 Variable sites and the average percentage base composition (A, G, C & T) in a 550 base pair fragment of the cytochrome b gene in *Aethomys namaquensis* from southern Africa.

	Variable sites	A	G	C	T
First position	13	28.2	24.9	15.0	31.9
Second position	1	16.8	16.6	25.1	41.4
Third position	76	46.7	2.1	31.4	19.8
Overall	90	30.6	14.5	23.8	31.0

Phylogenetic analyses

The results of saturation analysis are summarised in Figure 2.3. First and second position changes accumulated slowly and transitions out-numbered transversions at first and third codon positions. The rate of transversional substitution was lower than that of transitions (Ti:Tv ratio of 15:1) with no transversions at first, and only one at second codon positions. Transitions normally outnumber transversions but this high Ti:Tv ratio is rather unusual and has also been reported in vlei rats of the genus *Otomys* (Maree 2002). Transversions at all positions and transitions at the first codon position were not saturated. Transitions at the third codon position showed minimal saturation, therefore, it was not necessary to remove or reduce the weight of this type of substitution.

Table 2.5 Pairwise estimates of percent HKY85 + Γ (0.7287) + I (0.6024) sequence divergence between 37 maternal haplotypes (above the diagonal) and percent HKY85 nucleotide divergence (nucleotide substitutions per site) between the detected haplotypes within *Aethomys namaquensis* (below the diagonal) from southern Africa. Outgroup individuals: *Mus* = *Mus musculus* and *Das* = *Dasymys incomtus*. The order of haplotypes is as indicated in Figure 2.4 and Tables 2.3 & 2.6. Grey blocks indicate lineages identified in Figure 2.4 (A-I). Lineages F-H are grouped in one block.

		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	37	38	39
1	NH01		0.57	0.60	0.76	0.57	0.57	0.76	5.72	6.59	6.73	7.58	6.12	6.47	6.47	6.62	5.21	6.12	7.06	7.27	5.17	6.57	6.02	5.06	7.28	6.91	5.29	5.71	5.62	5.53	4.90	4.03	4.59	6.48	4.66	5.32	5.24	6.99	-	-
2	NH02	0.56		0.82	0.97	0.78	0.57	0.57	6.75	7.36	7.48	8.04	7.20	6.85	7.60	7.98	6.20	7.20	7.13	7.67	5.18	6.95	6.38	5.38	7.68	7.30	5.33	6.06	6.17	5.68	5.88	4.91	5.53	6.62	5.18	5.85	6.25	7.39	-	-
3	NH03	0.58	0.60		0.40	0.39	0.51	0.40	7.45	7.85	8.25	8.32	7.94	7.95	8.42	8.83	8.82	7.94	7.38	7.52	5.73	6.86	6.20	5.57	8.03	7.59	5.56	5.90	6.38	6.02	6.01	5.01	4.98	8.32	5.33	5.50	6.51	8.07	-	-
4	NH04	0.75	0.92	0.39		0.57	0.76	0.76	6.40	6.98	7.10	7.59	6.83	7.21	7.21	7.58	5.67	6.83	7.49	8.06	6.43	7.32	6.75	5.72	6.91	6.49	5.85	6.43	6.51	6.23	6.20	5.22	5.21	7.95	5.49	6.00	5.90	7.78	-	-
5	NH05	0.56	0.73	0.38	0.55		0.57	0.57	6.73	7.32	7.45	8.37	7.18	7.58	7.56	7.94	6.20	7.18	7.45	7.64	6.00	6.92	6.37	5.36	6.45	6.05	5.64	6.06	6.15	5.86	5.85	4.90	4.69	7.54	5.18	5.85	6.23	6.15	-	-
6	NH06	0.57	0.55	0.59	0.73	0.55		0.76	6.39	6.97	7.10	8.01	6.83	7.20	7.20	7.57	5.86	6.83	7.11	7.27	5.57	6.58	6.01	5.06	8.07	7.67	5.32	5.71	5.62	5.54	5.53	4.80	5.21	7.18	4.86	5.32	5.90	7.77	-	-
7	NH08	0.73	0.55	0.40	0.73	0.55	0.73		6.40	6.99	7.11	7.61	6.84	6.49	7.22	7.59	6.22	6.84	6.71	7.29	6.01	6.59	6.40	5.39	8.08	7.69	5.31	6.08	5.83	5.55	5.54	5.22	5.21	7.19	4.87	5.87	5.91	7.02	-	-
8	NH31	4.55	5.15	5.56	4.95	5.15	4.95	4.95		0.56	0.75	0.92	0.37	0.56	0.56	0.75	0.18	0.37	6.20	6.53	5.24	5.54	5.04	4.45	6.19	5.65	4.23	6.47	6.49	6.20	4.59	4.31	4.89	3.97	4.56	4.41	4.82	5.24	-	-
9	NH36	5.07	5.49	5.80	5.28	5.48	5.28	5.29	0.55		0.18	0.20	0.18	0.38	0.77	0.57	0.38	0.57	6.59	6.72	5.12	5.72	5.14	4.89	6.37	6.02	4.57	7.09	7.11	6.72	5.02	4.72	5.33	4.65	4.41	4.55	5.08	5.38	-	-
10	NH37	5.15	5.55	5.99	5.35	5.55	5.35	5.35	0.73	0.18		0.39	0.37	0.56	0.95	0.75	0.57	0.75	6.98	7.26	5.73	6.21	5.89	5.06	7.28	6.90	4.87	7.22	7.22	6.91	5.20	4.90	5.52	4.84	4.85	4.71	5.55	5.89	-	-
11	NH34	5.61	5.85	6.04	5.62	6.04	5.65	5.62	0.79	0.20	0.39		0.59	0.39	1.09	0.82	0.61	0.82	6.93	7.14	5.52	6.17	5.52	5.29	8.10	8.46	4.84	6.05	6.05	7.62	5.37	5.08	5.70	5.30	5.04	4.23	5.88	5.86	-	-
12	NH28	4.78	5.36	5.78	5.18	5.38	5.18	5.18	0.37	0.18	0.37	0.59		0.18	0.57	0.37	0.18	0.37	6.88	6.96	5.33	5.95	5.44	4.81	8.63	8.27	4.60	6.93	6.95	6.84	4.65	4.36	4.95	4.30	4.81	4.47	4.99	5.31	-	-
13	NH29	4.98	5.16	5.77	5.36	5.58	5.36	4.96	0.55	0.37	0.55	0.38	0.18		0.79	0.57	0.38	0.57	6.89	7.37	5.73	6.30	5.79	5.13	7.00	6.84	4.60	7.33	7.33	7.02	4.96	4.66	5.27	4.60	4.92	4.78	5.31	5.00	-	-
14	NH26	4.96	5.56	6.01	5.38	5.56	5.36	5.36	0.92	0.75	0.92	0.99	0.55	0.73		0.18	0.77	0.97	7.14	7.37	5.33	5.62	5.11	4.51	7.00	6.64	4.99	7.34	7.33	7.02	4.96	4.66	5.27	4.60	4.92	4.78	5.31	5.65	-	-
15	NH27	5.18	5.77	6.22	5.58	5.77	5.56	5.56	0.73	0.56	0.73	0.79	0.37	0.55	0.16		0.57	0.76	7.54	7.74	5.73	5.95	5.44	4.81	7.37	6.99	5.33	7.73	7.71	7.39	5.27	4.98	5.59	4.89	5.23	5.09	5.63	5.98	-	-
16	NH30	4.20	4.60	5.18	4.60	4.81	4.60	4.81	0.18	0.37	0.55	0.60	0.18	0.37	0.74	0.55		0.19	5.98	6.37	5.33	5.38	5.11	4.51	6.01	5.88	4.27	6.56	6.34	6.04	4.13	3.56	4.14	3.60	4.09	4.17	4.74	5.06	-	-
17	NH23	4.76	5.38	5.78	5.16	5.36	5.16	5.16	0.37	0.56	0.73	0.79	0.37	0.55	0.92	0.73	0.18		6.68	6.98	5.33	5.95	5.44	4.81	6.63	6.27	4.60	6.93	6.95	6.64	4.65	4.36	4.95	4.30	4.81	4.47	4.99	5.31	-	-
18	NH14	5.28	5.28	5.45	5.48	5.47	5.28	5.05	4.78	5.02	5.23	5.21	5.01	5.01	5.26	5.48	4.81	5.01		0.81	0.56	2.42	2.17	1.92	3.23	2.92	2.20	5.35	4.76	4.42	5.90	5.89	6.62	5.81	2.93	2.44	5.88	6.42	-	-
19	NH15	5.37	5.57	5.52	5.78	5.57	5.37	5.37	4.96	5.08	5.37	5.32	5.18	5.38	5.38	5.56	4.83	5.18	0.59		0.71	2.29	2.12	2.11	3.29	3.03	2.32	5.90	5.30	5.02	7.07	6.71	6.70	6.61	3.28	3.38	6.04	7.16	-	-
20	NH16	4.18	4.14	4.51	4.91	4.67	4.40	4.65	4.21	4.08	4.46	4.34	4.22	4.46	4.21	4.46	4.22	4.22	0.35	0.59		2.02	1.49	1.74	2.03	1.75	1.85	4.96	4.57	4.24	5.84	5.00	6.27	5.33	2.56	3.24	4.83	5.89	-	-
21	NH07	4.97	5.17	5.16	5.37	5.17	4.97	4.97	4.37	4.48	4.76	4.75	4.57	4.77	4.37	4.57	4.22	4.57	2.14	2.04	1.42		0.57	0.56	2.09	1.83	1.16	5.57	5.00	4.72	8.02	5.89	6.36	5.60	3.01	3.10	5.73	6.10	-	-
22	NH10	4.86	4.86	4.79	5.07	4.87	4.85	4.87	4.06	4.13	4.46	4.38	4.27	4.48	4.07	4.27	4.07	4.27	1.95	1.90	1.38	0.58		0.39	1.56	1.64	0.81	5.35	4.85	4.53	5.85	5.19	5.86	5.43	2.83	2.88	4.58	4.88	-	-
23	NH09	4.07	4.26	4.40	4.47	4.27	4.06	4.27	3.87	3.96	4.07	4.22	3.86	4.08	3.68	3.88	3.88	3.88	1.74	1.90	1.58	0.56	0.39		1.56	1.65	0.86	4.56	4.21	3.95	5.20	4.58	5.21	4.61	2.56	2.35	5.22	5.59	-	-
24	NH12	5.37	5.57	5.79	6.19	5.96	5.78	5.78	4.76	4.87	5.37	4.72	4.98	5.18	5.18	5.38	4.62	4.98	2.77	2.81	1.83	1.80	1.60	1.71		0.18	1.35	6.26	5.63	6.01	6.71	6.36	7.07	4.35	4.08	4.22	6.40	6.78	-	-
25	NH13	5.17	5.37	5.56	5.96	5.76	5.57	5.57	4.56	4.86	5.17	4.93	4.77	4.96	4.96	5.16	4.42	4.77	2.54	2.62	1.60	1.67	1.50	1.52	0.16		1.09	5.91	5.30	5.67	6.35	6.01	6.70	4.07	3.80	3.93	6.05	6.42	-	-
26	NH11	4.21	4.23	4.40	4.43	4.43	4.23	4.21	3.53	3.76	3.95	3.94	3.74	3.73	3.99	4.20	3.52	3.74	1.97	2.07	1.52	1.09	0.87	0.94	1.26	1.03		4.52	4.18	3.87	4.92	4.26	4.94	4.54	2.70	2.43	4.70	5.43	-	-
27	NH32	4.50	4.70	4.83	4.92	4.71	4.50	4.71	4.95	5.30	5.37	5.82	5.17	5.38	5.38	5.59	4.97	5.17	4.21	4.54	3.97	4.33	4.20	3.71	4.75	4.54	3.86		0.38	0.56	7.02	6.28	7.02	4.83	3.91	4.30	6.32	6.38	-	-
28	NH33	4.56	4.76	4.90	4.96	4.76	4.56	4.56	4.96	5.30	5.36	5.81	5.17	5.37	5.37	5.56	4.83	5.17	3.64	4.18	3.72	3.98	3.88	3.47	4.37	4.18	3.45	0.38		0.18	8.31	5.97	6.86	4.62	3.77	3.91	6.01	7.93	-	-
29	NH35	4.38	4.59	4.70	4.79	4.58	4.38	4.39	4.78	5.10	5.19	5.80	5.00	5.20	5.20	5.40	4.68	5.00	3.62	4.00	3.49	3.80	3.88	3.29	4.60	4.40	3.24	0.57	0.19		6.01	5.68	6.35	4.95	3.53	3.85	5.72	7.55	-	-
30	NH18	3.97	4.58	4.66	4.78	4.56	4.36	4.37	3.77	4.05	4.17	4.28	3.78	3.98	3.98	4.17	3.43	3.78	4.52	5.19	4.49	4.58	4.48	4.09	4.99	4.78	3.90	5.17	4.78	4.60		0.57	0.96	4.95	4.08	4.64	6.40	6.77	-	-
31	NH20	3.38	3.97	4.04	4.17	3.97	3.77	4.17	3.58	3.85	3.97	4.07	3.58	3.78	3.78	3.98	3.04	3.58	4.51	4.99	3.97	4.38	4.08	3.69	4.78	4.58	3.49	4.78	4.57	4.40	0.55		0.76	4.85	3.80	4.24	6.05	6.42	-	-
32	NH22	3.77	4.36	4.03	4.17	3.97	4.17	4.17	3.97	4.25	4.38	4.48	3.98	4.17	4.17	4.37	3.44	3.98	4.93	4.98	4.75	4.78	4.50	4.10	5.19	4.99	3.92	5.18	4.98	4.80	0.92	0.73		5.26	4.37	4.85	8.75	7.14	-	-
33	NH19	4.98	5.18	5.97	5.77	5.56	5.36	5.36	3.37	3.84	3.96	4.26	3.56	3.77	3.77	3.97	3.23	3.58	4.51	4.97	4.22	4.37	4.27	3.86	3.56	3.39	3.70	3.92	3.77	3.99	3.98	3.76	4.17		4.91	5.09	5.29	6.32	-	-
34	NH17	3.96	4.16	4.26	4.36	4.16	3.96	3.96	3.77	3.87	3.96	4.09	3.77	3.97	3.97	4.17	3.42	3.77	2.56	2.61	2.26	2.61	2.46	2.26	3.39	3.19	2.38	3.29	3.19	3.01	3.39	3.19	3.56	3.97		1.83	4.39	4.72	-	-
35	NH21	4.23	4.43	4.35	4.64	4.44	4.22	4.43	3.64	3.74	3.84	3.52	3.65	3.85	3.85	4.05	3.45	3.85	2.16	2.86	2.76	2.86	2.50	2.09	3.45	3.25	2.15	3.52	3.26	3.08	3.88	3.48	3.87	4.05	1.68		4.86	5.21	-	-
36	NH24	4.17	4.77	4.93	4.57	4.77	4.57	4.57	3.77	4.07	4.37	4.56	3.98	4.18	4.18	4.37	3.82	3.98	4.37	4.58	3.71	4.36	3.68	4.09	4.79	4.58	3.78	4.78	4.58	4.41	4.79	4.58	4.99	4.16	3.58	3.85		0.98	-	-
37	NH25	5.20	5.41	5.80	5.62	5.82	5.61	5.21	4.16	4.26	4.56	4.57	4.19	3.99	4.39	4.59	4.04	4.19	4.62	5.22	4.50	4.81	3.89	4.32	5.01	4.81	4.22	5.66	5.63	5.44	5.01	4.81	5.21	4.79	3.81	4.08	0.93		-	-
38	*Mus.*	20.32	20.33	20.66	20.85	20.06	19.81	19.25	20.33	19.82	19.50	20.87	20.26	20.52	22.57	20.78	19.93	19.49	21.04	21.01	20.78	19.83	21.56	20.24	18.51	19.92	21.02	21.28	20.78	21.02	20.50	20.45	19.04	19.53	21.26	19.72	21.06	20.73		-
39	*Das.*	15.75	15.78	16.35	16.00	16.23	15.52	16.00	15.92	16.14	16.19	16.50	16.17	16.42	16.88	16.65	15.83	16.17	16.81	17.16	17.59	16.20	16.75	16.60	15.95	16.19	17.64	16.64	16.18	16.07	16.40	16.15	16.65	17.38	15.70	16.88	16.20	17.03	16.18	

Table 2.6 Pairwise estimates of transition:transversion ratio (Ti:Tv) between 37 maternal haplotypes detected within *Aethomys namaquensis* from southern Africa (above the diagonal) and pairwise protein differences between the different haplotypes (below the diagonal). The order of haplotypes is as indicated in Figure 2.4 and Table 2.3 & 2.5. Grey blocks indicate lineages identified in Figure 2.4 (A-I). Lineages F-H are grouped in one block.

		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	37
1	NH01		3/0	3/0	4/0	3/0	2/0	4/0	20/4	22/4	23/4	23/4	22/3	23/3	23/3	24/3	19/3	22/3	22/2	26/2	16/2	24/2	22/2	19/2	26/2	25/2	17/2	20/3	21/3	20/3	19/2	16/2	18/2	23/3	18/3	20/2	20/2	25/2
2	NH02	1		3/0	5/0	4/0	3/0	3/0	23/4	24/4	25/4	24/4	25/3	24/3	26/3	27/3	22/3	25/3	22/2	27/2	16/2	25/2	23/2	20/2	27/2	26/2	17/2	21/3	22/3	21/3	22/2	19/2	21/2	24/3	19/3	21/2	23/2	26/2
3	NH03	0	0		2/0	2/0	3/0	2/0	23/4	24/4	25/4	25/4	25/3	25/3	26/3	27/3	22/3	25/3	23/2	25/2	18/2	23/2	21/2	19/2	26/2	25/2	18/2	19/3	21/3	20/3	21/2	18/2	18/2	26/3	18/3	19/2	22/2	26/2
4	NH04	1	2	1		3/0	4/0	4/0	22/4	23/4	24/4	23/4	24/3	25/3	25/3	26/3	21/3	24/3	23/2	28/2	19/2	26/2	24/2	21/2	30/2	29/2	18/2	22/3	23/3	22/3	23/2	20/2	20/2	27/3	20/3	22/2	22/2	27/2
5	NH05	1	2	1	2		3/0	3/0	23/4	24/4	25/4	25/4	25/3	26/3	26/3	27/3	22/3	25/3	23/2	27/2	18/2	25/2	23/2	20/2	29/2	28/2	18/2	21/3	22/3	21/3	22/2	19/2	19/2	26/3	19/3	21/2	23/2	28/2
6	NH06	0	1	0	1	1		4/0	22/4	23/4	24/4	24/4	24/3	25/3	25/3	26/3	21/3	24/3	22/2	26/2	17/2	24/2	22/2	19/2	28/2	27/2	17/2	20/3	21/3	20/3	21/2	18/2	20/2	25/3	18/3	20/2	22/2	27/2
7	NH08	2	1	1	3	3	2		22/4	23/4	24/4	23/4	24/3	23/3	25/3	26/3	22/3	24/3	21/2	26/2	18/2	24/2	23/2	20/2	28/2	27/2	17/2	21/3	21/3	20/3	21/2	20/2	20/2	25/3	18/3	21/2	22/2	25/2
8	NH31	3	4	3	2	4	3	3		1/2	2/2	2/2	1/1	2/1	4/1	3/1	0/1	1/1	20/2	24/2	16/2	21/2	19/2	17/2	23/2	22/2	14/2	22/3	23/3	22/3	18/2	17/2	19/2	15/3	17/3	17/2	18/2	20/2
9	NH36	2	3	2	1	3	2	2	1		1/0	1/0	0/1	1/1	3/1	2/1	1/1	2/1	21/2	24/2	16/1	21/2	19/2	18/2	23/2	22/2	15/2	23/3	24/3	23/3	19/2	18/2	20/2	17/3	16/3	17/2	19/2	20/2
10	NH37	2	3	2	1	3	2	2	1	0		2/0	1/1	2/1	4/1	3/1	2/1	3/1	22/2	26/2	16/1	23/2	21/2	19/2	26/2	25/2	16/2	24/3	25/3	24/3	20/2	19/2	21/2	18/3	18/3	18/2	21/2	22/2
11	NH34	2	2	2	1	3	2	1	1	0	0		1/1	1/1	4/1	3/1	2/1	3/1	22/2	24/2	16/1	21/2	19/2	18/2	21/2	22/2	16/2	24/3	25/3	24/3	19/2	18/2	20/2	18/3	17/3	15/2	20/2	20/2
12	NH28	2	3	2	1	3	2	2	1	0	0	0		1/0	3/0	2/0	1/0	2/0	22/1	26/1	17/1	23/1	21/1	19/1	25/1	24/1	16/1	24/2	25/2	24/2	19/1	18/1	20/1	17/2	18/2	18/1	20/1	21/1
13	NH29	3	2	2	2	4	3	1	2	1	1	0	1		4/0	3/0	2/0	3/0	22/1	27/1	18/1	24/1	22/1	20/1	26/1	25/1	16/1	25/2	26/2	25/2	20/1	19/1	21/1	18/2	19/2	19/1	21/1	20/1
14	NH26	2	3	2	1	3	2	2	1	0	0	0	0	1		1/0	4/0	5/0	23/1	27/1	17/1	22/1	20/1	18/1	26/1	25/1	17/1	25/2	26/2	25/2	20/1	19/1	21/1	18/2	19/2	19/1	21/1	22/1
15	NH27	2	3	2	1	3	2	2	1	0	0	0	0	1	0		3/0	4/0	24/1	28/1	18/1	23/1	21/1	19/1	27/1	26/1	18/1	26/2	27/2	26/2	21/1	20/1	22/1	19/2	20/2	20/1	22/1	23/1
16	NH30	1	2	1	0	2	1	2	1	0	0	0	0	1	0	0		1/0	20/1	24/1	17/1	21/1	20/1	16/1	23/1	22/1	15/1	23/2	23/2	22/2	17/1	15/1	17/1	15/2	16/2	17/1	19/1	20/1
17	NH23	2	3	2	1	3	2	2	1	0	0	0	0	1	0	0	0		22/1	26/1	17/1	23/1	21/1	19/1	25/1	24/1	16/1	24/2	25/2	24/2	19/1	18/1	20/1	17/2	18/2	18/1	20/1	21/1
18	NH14	1	1	1	2	2	1	0	2	1	1	1	1	1	1	1	1	1		3/0	2/0	10/0	9/0	8/0	13/0	12/0	9/0	18/1	17/1	16/1	21/0	21/0	23/0	20/1	11/1	10/0	20/0	22/0
19	NH15	2	3	2	3	3	2	2	3	2	2	2	2	3	2	2	2	2	1		3/0	11/0	10/0	10/0	15/0	14/0	10/0	22/1	21/1	20/1	27/0	26/0	26/0	25/1	14/1	15/0	24/0	27/0
20	NH16	0	1	0	1	1	0	1	2	1	1	1	1	2	1	1	1	1	0	1		8/0	6/0	7/0	8/0	7/0	6/0	16/1	15/1	14/1	19/0	17/0	20/0	17/1	9/1	12/0	16/0	19/0
21	NH07	2	3	2	3	3	2	2	3	2	2	2	2	3	2	2	2	2	1	2	1		3/0	3/0	10/0	9/0	5/0	21/1	20/1	19/1	24/0	23/0	25/0	22/1	13/1	14/0	23/0	24/0
22	NH10	0	1	0	1	1	0	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1		2/0	9/0	8/0	4/0	20/1	19/1	18/1	23/0	21/0	23/0	21/1	12/1	13/0	19/0	20/0
23	NH09	0	1	0	1	1	0	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0		9/0	8/0	3/0	18/1	17/1	16/1	21/0	19/0	21/0	19/1	11/1	11/0	21/0	22/0
24	NH12	1	2	1	2	2	1	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0		1/0	6/0	23/1	22/1	23/1	26/0	25/0	27/0	18/1	17/1	18/0	25/0	26/0
25	NH13	1	2	1	2	2	1	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0		5/0	22/1	21/1	22/1	25/0	24/0	26/0	17/1	16/1	17/0	24/0	25/0
26	NH11	0	0	0	1	1	0	0	2	1	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0		16/1	15/1	14/1	18/0	16/0	18/0	16/1	10/1	10/0	17/0	19/0
27	NH32	0	1	0	1	1	0	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0		0/2	1/2	25/1	23/1	25/1	18/2	15/2	17/1	23/1	28/1
28	NH33	1	2	1	2	2	1	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0		1/0	24/1	23/1	25/1	18/2	15/2	16/1	23/1	28/1
29	NH35	1	2	1	2	2	1	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0		23/1	22/1	24/1	19/2	14/2	15/1	22/1	27/1
30	NH18	1	2	1	2	2	1	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0		3/0	5/0	20/1	17/1	20/0	25/0	26/0
31	NH20	0	1	0	1	1	0	2	3	2	2	2	2	3	2	2	1	2	1	2	0	2	0	0	1	1	0	0	1	1	1		4/0	19/1	16/1	18/0	24/0	25/0
32	NH22	0	1	0	1	1	0	2	3	2	2	2	2	3	2	2	1	2	1	2	0	2	0	0	1	1	0	0	1	1	1	0		21/1	18/1	20/0	26/0	27/0
33	NH19	2	3	2	3	3	2	2	3	2	2	2	2	3	2	2	2	2	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2		19/2	20/1	21/1	24/1
34	NH17	1	2	1	2	2	1	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1	1		8/1	18/1	19/1
35	NH21	0	1	0	1	1	0	1	2	1	1	1	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0		20/0	21/0
36	NH24	2	3	2	1	3	2	2	1	0	0	0	0	1	0	0	0	0	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2	2	1	1		5/0
37	NH25	3	2	2	2	4	3	1	2	1	1	0	1	0	1	1	1	1	1	3	2	3	2	2	2	2	1	2	2	2	2	3	3	3	2	2	1	

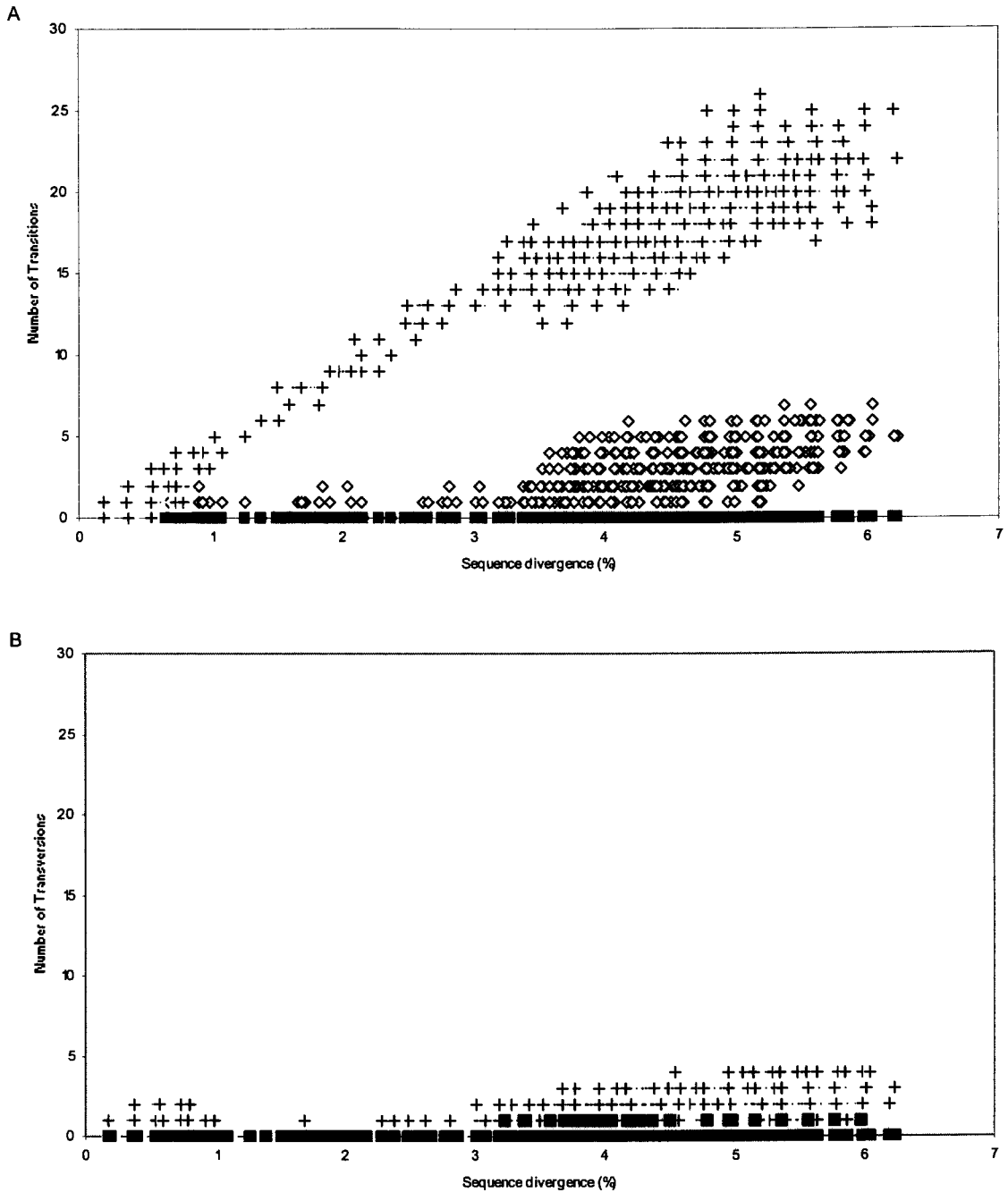


Fig. 2.3 Number of transitions (A) and transversions (B) plotted against HKY85 sequence divergence (%) estimates between 37 *Aethomys namaquensis* mitochondrial DNA haplotypes from southern Africa. (◊) First codon positions, (■) second codon positions and (+) third codon positions.

Corrected HKY85 sequence di[illegible]ed to summarise relationships among the 37 haplotypes using neighbour-joining trees (Fig. 2.4). Three phylogenetically well-defined (high bootstrap support) lineages (A-C) are evident from the present analysis (Fig. 2.4). The phylogram (Fig. 2.4A) was characterised by short internal nodes possibly due to the rapid radiation of lineages. The average separation within lineage A was 0.59% (range 0.37-0.92%), 0.53% (range 0.18-0.99%) for lineage B and 1.52% (range 0.18-2.81%) for lineage C (Fig. 2.4; Table 2.5). Average sequence divergence between lineages A and B was 5.33% (range 4.20-6.22%), 5.02% (range 4.07-6.19%) between lineages A and C and 4.55% (range 3.52-5.58) between lineages B and C (Fig. 2.4; Table 2.5). Lineages D to I (Fig. 2.4) differed, on average, by 4.26% (range 2.09-5.97%) from lineages A, B and C and the average separation within these lineages was 1.40% (range 0.18-4.05%). The average sequence divergence among lineages D to I was 4.26% (range 3.01-5.86%). *Mus musculus* and *D. incomtus* differed, on average, by 18.58% from ingroup taxa.

Parsimony and maximum likelihood analyses (α = 0.85; I = 0.56; Ti:Tv ratio = 15:1; proportion of A = 0.30, C = 0.24, G = 0.14 and T = 0.31) (not illustrated), generated the same topology as neighbour-joining trees, with the following statistics for the parsimony analysis: 88 parsimony informative characters (with outgroups), number of trees = 1; tree length = 170 steps; CI = 0.62; RI = 0.71 and RC = 0.44. The data were also significantly more structured than random (g_1 = -1.07, $P < 0.01$; see Hillis & Huelsenbeck, 1992). These results suggest that some of the genetic units may, depending on interpretation, represent valid taxonomic units at either the species or subspecies levels.

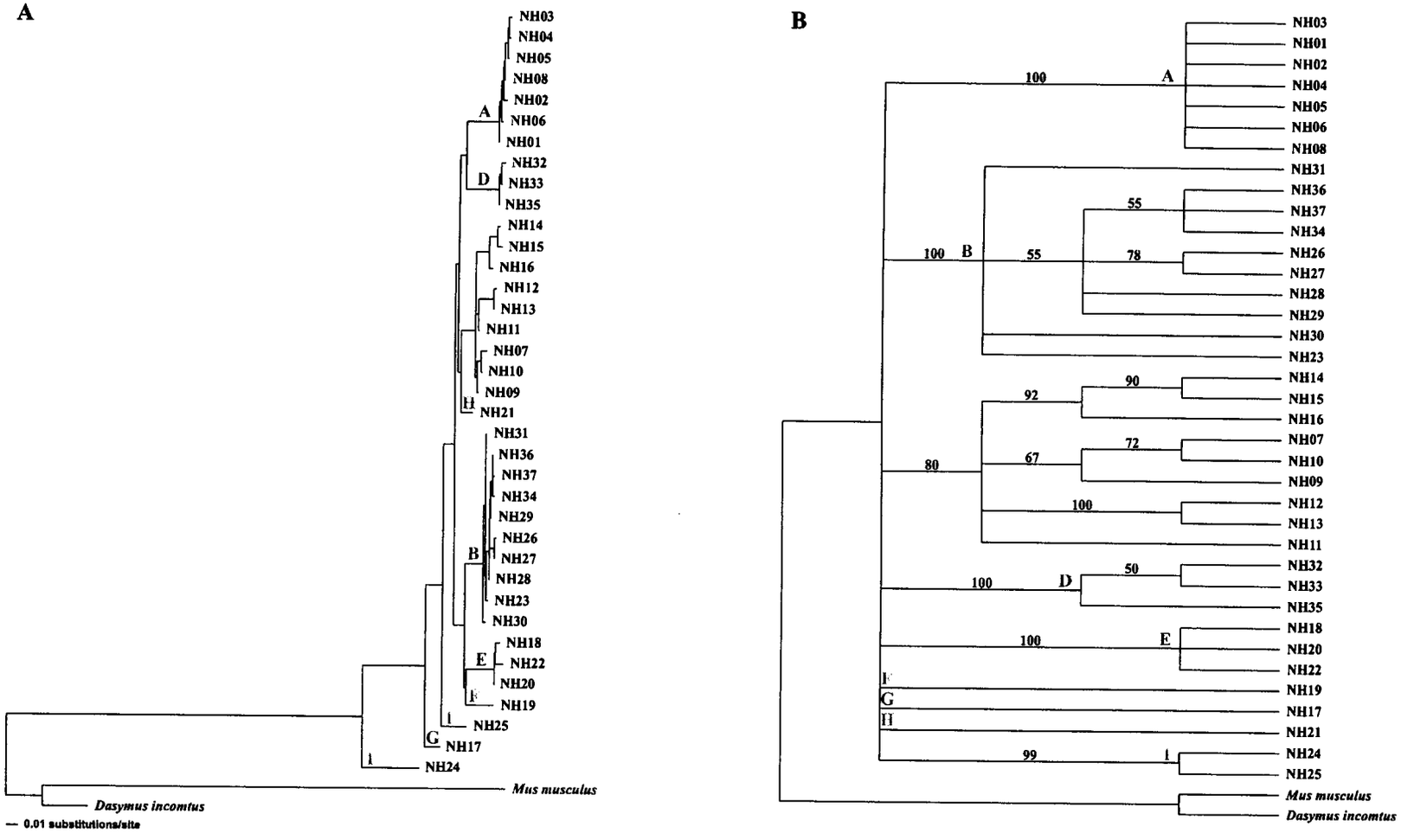


Fig. 2.4 A neighbour-joining phylogram (A) and a neighbour-joining tree (B) based on percentage nucleotide sequence divergence between mitochondrial DNA cytochrome b sequences within *Aethomys namaquensis* from southern Africa. Letters A-I represent different lineages identified in the tree. Colours of lineages A to I correspond to the colours used in Figures 2.5 & 2.8. The bootstrap confidence limits (% occurrence in 1000 replicates) for internal branches are given at each node. *Mus musculus* and *Dasymys incomtus* were used as outgroups.

The geographic distribution of the mtDNA cytochrome b diversity within *A. namaquensis* from southern Africa is illustrated in Figure 2.5. Of the three phylogenetically well-defined lineages (A-C; Fig. 2.4), two (B in red & C in yellow; Fig. 2.5) were also geographically well defined. Lineage A was not geographically well represented, while lineages A and C were found in sympatry at Messina (A in blue & C in yellow; Fig. 2.5). The other six lineages (D-I; Fig. 2.4) were identified from only a few individuals, from one or two localities and did not show a geographically discernible pattern.

Several of these lineages were found in sympatry with lineage B (e.g., lineage D at Porterville, lineage I at Augrabies National Park and lineage E (haplotype N22) at Upington; Fig 2.5). Lineage E was also recorded at Vryburg where two divergent lineages were recorded in sympatry (lineage E & F; Fig. 2.5). Apart from the overlap with lineage A (blue) at Messina, lineage C (yellow) was also found in sympatry with a unique haplotype (lineage G) at Machadodorp, the only locality where lineage G was recorded (Fig. 2.5). Pongola was only represented by one individual (NH21) where the unique lineage H was found.

The geographic distribution of the morphometrically defined and mtDNA cytochrome b diversity within *A. namaquensis* is illustrated in Figure 2.6. From the phylogenetic relationships and geographic distributions, three geographically and genetically distinct lineages were evident: lineage A was associated with Savanna, lineage B with Nama-Karoo and lineage C mostly with Grassland, roughly corresponding to the biome-related distributional patterns proposed by Chimimba (2001a) for the subspecies *alborarius*, *lehocla* and *monticularis*, respectively (see Fig. 2.4, Fig. 2.5 & Fig. 2.6). In general, the distribution of lineage B seems to extend further south than the proposed distribution for *lehocla* by Chimimba (2001a), while lineage C was only identified from areas north of the proposed distribution of *monticularis*. Lineage A was only identified from two localities, both within the proposed distributional range of *A. n. alborarius*. This lineage was found in sympatry with lineage C at Messina. Six of the sampling sites were potentially within the proposed distributional range of the subspecies *A. n. namaquensis*. Of the five lineages found (Fig. 2.6), only lineages D, G or H can potentially be equated to *A. n. namaquensis*, since the main distributional range of lineages B and C are outside of the proposed distribution of *A. n. namaquensis*.

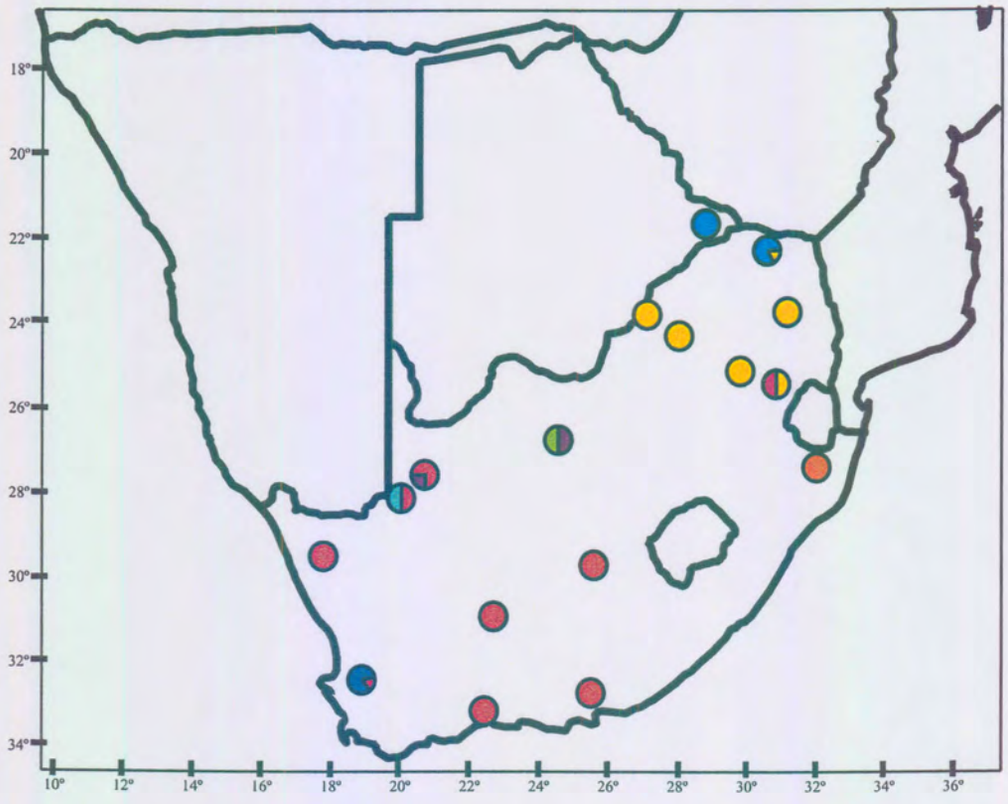


Fig. 2.5 Geographic distribution of mitochondrial DNA cytochrome b diversity within *Aethomys namaquensis* from southern Africa. Circles represent cytochrome b variability: blue, red and yellow represent well-supported lineages A to C, respectively (see Fig. 2.4), while the other lineages are represented in purple blue (D), purple (E), green (F), pink (G), orange (H) and mint green (I) (see Fig. 2.4). The size of the pies within circles represents haplotype frequencies.

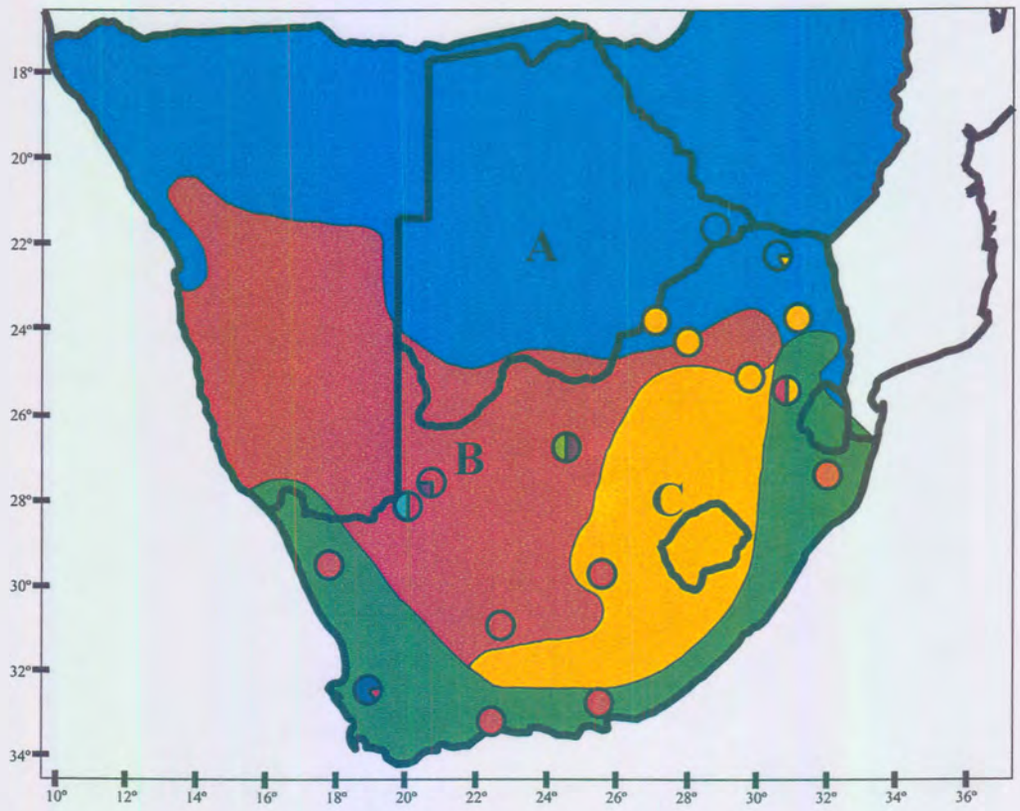


Fig. 2.6 Geographic distribution of morphometrically defined (Chimimba 2001a) and mitochondrial DNA cytochrome b diversity in *Aethomys namaquensis* from southern Africa. The blue, red, yellow and green zones correspond with morphometrically defined subspecies, *A. n. alborarius*, *A. n. lehocla*, *A. n. monticularis* and *A. n. namaquensis*, respectively (Chimimba 2001a). Circles represent cytochrome b variability: red, blue and yellow represent well-supported lineages, while the other lineages are represented in purple blue (D), purple (E), green (F), pink (G), orange (H) and mint green (I) (see Fig. 2.4 & Fig. 2.5).

For multiple pairwise comparisons of lineages relative to an outgroup, there were no significant differences in substitution rates at the adjusted significance level of 0.029% ($P \geq 0.0003$), indicating that the different lineages are evolving at the same rate in comparison to the outgroup. The same approach as Ducroz *et al.* (1998) was followed to date the different lineages. There are 38 third position transversions between *Mus* and *Rattus* (Bibb *et al.* 1981; Gadaleta *et al.* 1989) in the 550 bp fragment of the 5' end of the cytochrome b gene sequenced in the present study. This resulted in a sequence divergence estimate of

20.77% that corresponds to a rate of 1.73% third position transversions per Myr. This rate of third position transversions is similar to an estimate of 1.7% per Myr reported by Smith and Patton (1993) based on an 801 bp fragment of the same gene sequenced in South American rodents. Based on this calibration, the divergence time separating lineages A & B was calculated at 1.07 (range 948 000-1.26) Myr ago, 632 000 years ago for lineages A & C and 432 000 (range 316 000-632 000) years ago for lineages B & C (Fig. 2.4). These dates may be underestimations because of the low number of third position transversions detected between lineages identified within *A. namaquensis* (Table 2.6). When using the conventional substitution rate of 2% per Myr, the divergence time separating lineages A & B was estimated at 2.67 (range 2.10-3.02) Myr ago, 2.51 (2.04-3.10) Myr ago for lineages A & C and 2.27 (1.76-2.79) Myr ago for lineages B & C, suggesting a late Pliocene time of separation. The divergence time, based on the conventional substitution rate of 2% per Myr, within lineage A was estimated to have occurred 297 000 (185 000-460 000) years ago, the separation within lineage B at 263 000 (90 000-495 000) years ago, whereas the separation within lineage C was estimated at 758 000 (90 000 years-1.41 Myr) years ago. These results suggest that the differentiation within lineages A and B is of similar age and is relatively more recent than the separation within Group C. The divergence time separating lineages D to I (Fig. 2.4) from lineage A, B and C was estimated at 2.13 (range 1.05-2.99) Myr ago and within these lineages at 700 000 (range 90 000-2.03 Myr) years ago. The divergence between lineages D to I was estimated at 2.13 (range 1.51-2.93) Myr ago. The divergence time between *Mus* and the ingroup taxa was estimated at 10.20 (range 9.26-11.29) Myr ago and that between *Dasymys* and the ingroup at 8.38 (range 7.76-9.42) Myr ago.

Population level analyses

A minimum-spanning network supported the major well-supported lineages (A, B & C) derived by the neighbour-joining tree (Fig. 2.4) and these lineages, on average, differed from each other by 17 mutational steps (results not illustrated). Lineages D to I also, on average, differed from lineages A, B and C by 17 mutational steps (range 10-21 mutational steps). Minimum-spanning networks of the three lineages were illustrated separately because there were no shared haplotypes between these genetically distinct groups (Fig. 2.7). Lineage B (Fig. 2.7B) was characterised by ten closely related haplotypes over a wide geographic area, with haplotypes differing by only one or two mutational steps from each other. Two to three mutational steps were found between the seven haplotypes in

lineage A but, the differentiation was within samples from South Africa (Messina) and Botswana, rather than between them (Fig. 2.7A). In contrast, more differentiation was detected within lineage C, with six mutational steps between haplotypes NH14-NH16, NH07 and NH09-NH13; Fig. 2.7C).

No isolation by distance was detected in lineages B and C (Fig. 2.4). Mantel test results for these lineages were r = -0.22; $P = 0.19$ and r = 0.27; $P = 0.15$, respectively. In both analyses, the standard normal variate (g) was also smaller than the critical value at $P \leq 0.01$ ($g = -1.11$ (lineage B) and $g = 0.91$ (lineage C) at a critical value of 2.33), indicating that the null-hypothesis (no association between elements in the two matrices) could not be rejected.

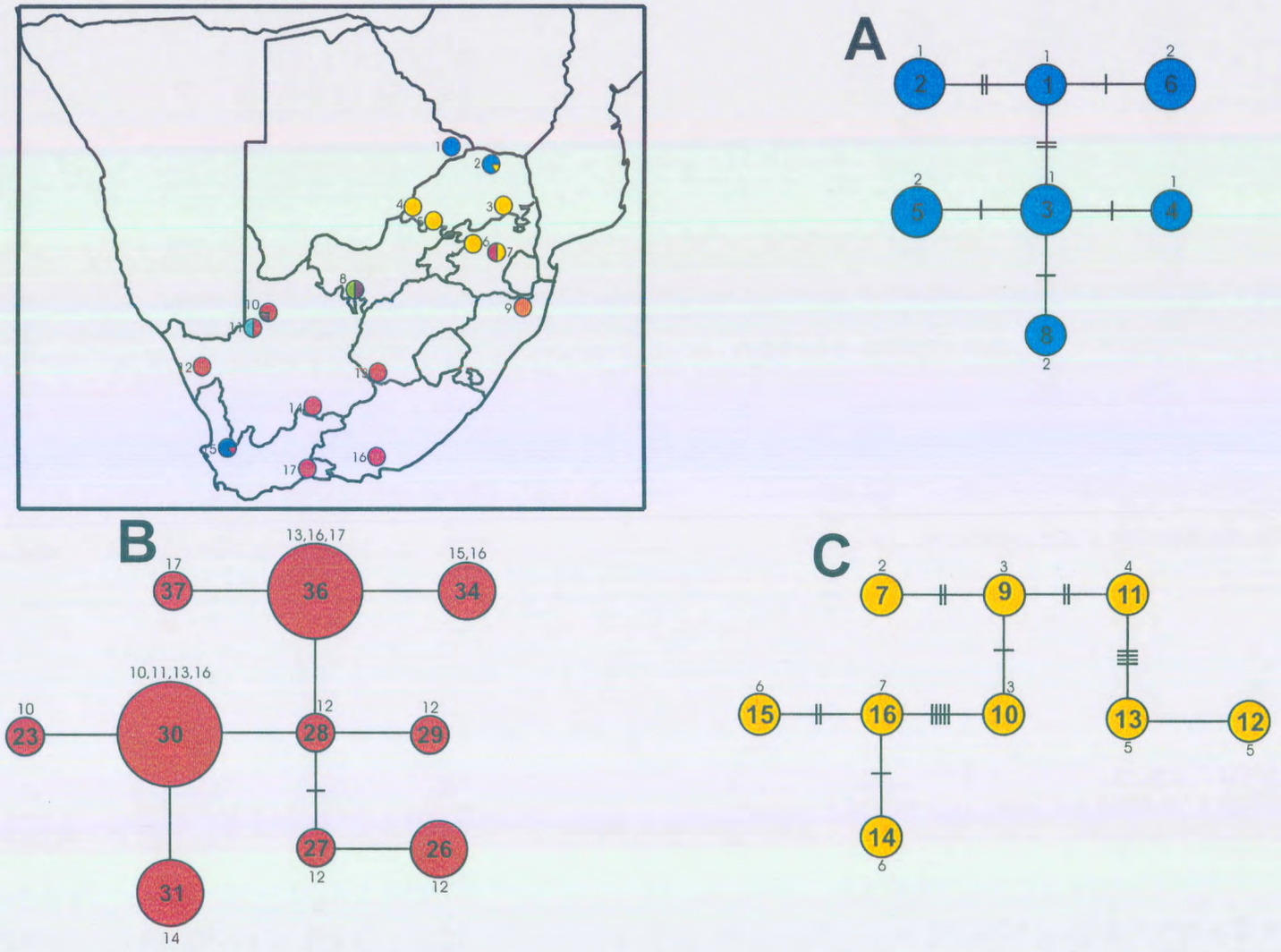


Fig. 2.7 Minimum-spanning networks indicating the least number of mutational steps between composite haplotypes within *Aethomys namaquensis* from southern Africa. The size of the circles represents the haplotype frequencies. Horizontal lines connecting haplotypes represent one mutational step and cross-hatching along branches designates additional number of detected changes. Numbers inside circles represent the haplotype designations and correspond to those in Table 2.3. Numbers outside circles correspond to locality numbers (see insert for localities). Letters A-C correspond to lineages identified in the neighbour-joining tree (Fig. 2.4).

In the DnaSP analysis, all 70 sequences were considered as one population and the overall nucleotide diversity was estimated at 3.10%, while the derived haplotypic diversity value of 0.92 corresponded to those reported for other rodents (Table 2.7) (Avise *et al.* 1989; Matthee & Robinson 1997). This may reflect the high incidence of locality-specific haplotypes detected at the different sampling sites. The high haplotype diversity relative to nucleotide diversity may also be indicative of a population bottleneck followed by rapid population growth (also see mismatch distribution results) (Grant & Bowen 1998; Avise 2000).

Based on geographic information, 50 samples were separated into three groups corresponding to the three lineages (Fig. 2.4): A (11 individuals), B (30) and C (nine) for AMOVA and DnaSP analyses. These groups were analysed in different combinations in order to test for significant genetic structuring among the samples (Table 2.7). Estimates of nucleotide diversity (Table 2.7) confirmed the average HKY85 sequence divergence within the three different lineages, given that DnaSP excludes sites with alignment gaps or missing data. The nucleotide diversity when the three lineages were combined (16 populations, 50 individuals) was 2.55%, with values of 0.44%, 0.44% and 1.40% (Table 2.7) for populations within lineage A, B and C, respectively (Fig. 2.4). These rather low nucleotide diversity values are indicative of shallow divergences (Grant & Bowen 1998; Avise 2000). Lineage A and C were characterised by high haplotypic diversity values, indicating the high incidence of locality-specific haplotypes detected within these lineages. In contrast, lineage B showed a haplotypic diversity value of 0.86, a value lower than that of lineages A and B, showing a reduction in the number of locality-specific haplotypes detected.

An AMOVA incorporated in Arlequin showed strong phylogeographic structuring of haplotypes. Eighty-four percent of the total variance was at the regional level when all geographic regions were considered. The variance contribution among populations within groups was 9.89%, while only 5.81% of the variance was contained among individuals within populations. When lineages B and C (Fig. 2.4) were analysed separately, most of the variance was contained among populations (Table 2.7). In contrast, when lineage A (Fig. 2.4) was analysed separately, most of the variance was within populations (Table 2.7).

Mismatch distribution analyses were performed on the three lineages as identified in the neighbour-joining tree: lineage A (Fig. 2.8A), lineage B (Fig. 2.8B) and lineage C (Fig. 2.8C). Lineage B ($P_{Ssd} \geq 0.65$ & $P_{Rag} \geq 0.90$) and lineage C ($P_{Ssd} \geq 0.63$ & $P_{Rag} \geq 0.61$) showed approximately unimodal distributions (Figs. 2.8B & 2.8C), a result compatible with a population expansion in the recent past (Rogers & Harpending 1992). Lineage A also showed a unimodal distribution (Fig. 2.8A), but it differed significantly from the expected Poisson distribution ($P_{Ssd} \geq 0.002$ & $P_{Rag} \geq 0.02$). However, the results for lineages A and C need to be treated with caution, as only few individuals were included in these analyses.

Table 2.7 Estimates of genetic differentiation (nucleotide and haplotypic diversity) and a hierarchical analysis of molecular variance (AMOVA) of mitochondrial DNA cytochrome b sequences between lineages identified within *Aethomys namaquensis* from southern Africa.

	Overall	Lineage A, B & C	Lineage A	Lineage B	Lineage C	Lineage A & B	Lineage A & C	Lineage B & C
Number of: Populations	17	16	2	8	6			
Sample size (*n*)	70	50	11	30	9			
DnaSP:								
Nucleotide diversity	3.10%	2.55%	0.44%	0.44%	1.40%	2.21%	2.64%	1.59%
Haplotypic diversity	0.92	0.86	0.93	0.86	0.97	0.92	0.97	0.78
AMOVA:								
Among groups		84.29%	-	-	-	90.08%	75.70%	80.62%
Among populations within groups		9.89%	9.22%	63.66%	76.60%	4.99%	15.13%	14.07%
Within populations		5.81%	90.78%	36.34%	23.40%	4.92%	9.17%	5.31%
Φ_{ST}		0.94***	0.09	0.64***	0.77**	0.95***	0.91***	0.95***
Φ_{CT}		0.84***	-	-	-	0.90*	0.76*	0.81***
Φ_{SC}		0.63***	-	-	-	0.50***	0.62***	0.73***
V_a		9.19	0.12	2.18	0.76	11.23	8.21	7.34
V_b		1.09	-	-	-	0.62	1.64	1.28
V_c		0.63	-	-	-	0.61	0.99	0.48

Statistical significance: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$

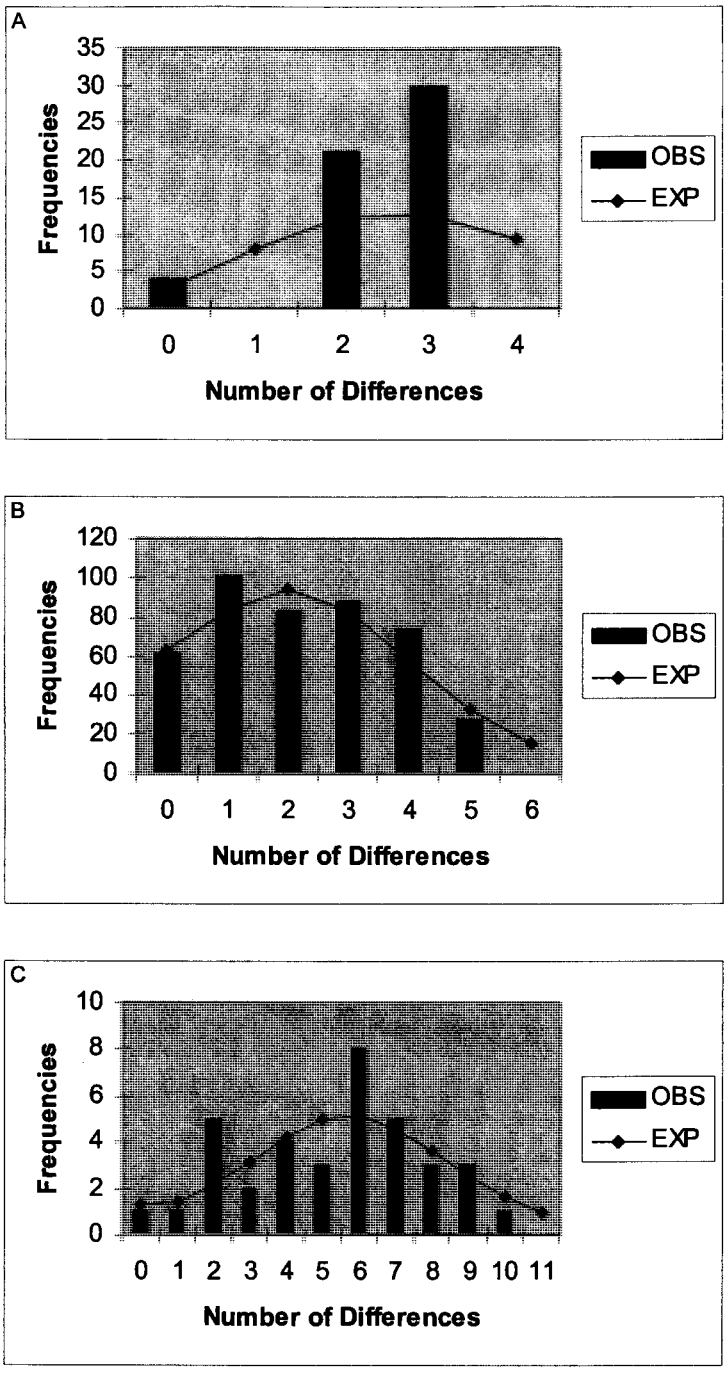


Fig. 2.8 Frequency distributions of pairwise genetic differences between individuals of *Aethomys namaquensis* from southern Africa examined within (A) lineage A (Sum of squared deviation = 0.15; $P = 0.002$; Harpending's raggedness index = 0.48; $P = 0.02$), (B) lineage B (Sum of squared deviation = 0.004; $P = 0.65$; Harpending's raggedness index = 0.03; $P = 0.90$) and (C) lineage C (Sum of squared deviation = 0.02; $P = 0.63$; Harpending's raggedness index = 0.06; $P = 0.61$), as identified in the neighbour-joining tree (Fig. 2.4).

The neutrality test was conducted using the same groupings of lineages as in the mismatch distribution analyses (Table 2.8). The Tajima *D* test for neutrality (Tajima 1989a, b, 1993) was not statistically significant ($P > 0.05$), suggesting that the gene examined is neutral. The Tajima *D* test results for lineages A, B and C were negative, suggesting population expansions (Tajima 1989b). Fu's F_s test (Fu 1997) rejected the null hypothesis of neutrality when lineages A, B and C were analysed separately ($P < 0.05$) and this test also suggested population expansions. Mismatch distribution analysis is probably a more appropriate test of historic population demography and indicated that population expansions may be one of the dominant processes shaping genetic diversity within *A. namaquensis* from southern Africa. The cytochrome b gene in *A. namaquensis* seems to be neutral, since all combinations tested negative with Tajima' s *D* test. The only exceptions were when lineages A, B and C were subjected to Fu's F_s test individually, possibly because this method may not be sensitive enough.

Table 2.8 A summary of Tajima's (Tajima 1989a, b, 1993) and Fu's (Fu 1997) neutrality tests for *Aethomys namaquensis* from southern Africa. Lineages A to C correspond to the different lineages identified in the neighbour-joining tree (Fig. 2.4). $P(D) = P$ value for D; $P(F_s) = P$ value for F_s.

	Lineage A	Lineage B	Lineage C
Sample size (*n*):	11	30	9
Tajima's *D* value	-0.50	-0.08	-0.16
$P(D)$	0.34	0.50	0.45
Fu's F_s value	-2.31	-2.71	-4.66
$P(F_s)$	0.05	0.00	0.01

Discussion

The analysis of a 550 bp fragment of the mtDNA cytochrome b gene supported earlier suggestions that *A. namaquensis* in southern Africa is polytypic (Roberts 1951; Meester *et al.* 1964; Chimimba 2001a). The present phylogenetic analysis, however, suggested that greater diversity exists than the four subspecies (*A. n. namaquensis*; *A. n. lehocla*; *A. n. alborarius* and *A. n. monticularis*) proposed by Chimimba (2001a). The deeper nodes in the phylogenetic trees were unresolved possibly due to the rapid radiation of lineages. Similar rapid radiations have been reported in other rodent species (Martin *et al.* 2000).

In so far as geographic distributions of the lineages identified in this study are concerned, it is not clear whether morphometric variation (Chimimba 2001a) is as a result of historical differentiation or as a result of ecological selection. None of the genetic lineages identified in this study showed concordance with the distributions proposed for the four subspecies. Chimimba (2001a) noted that the geographic boundaries of these proposed subspecies need to be refined using alternative systematic techniques such as mtDNA analysis. While the geographic limits delineated by the molecular approach may be valid, these need to be investigated further using more samples.

When taking molecular evidence into consideration, species or subspecies classification should be based on the geographic distribution of genetic diversity and not the extent of sequence divergence (Buth & Mayden 1981). Compared to other mammals, however, the differentiation between the *A. namaquensis* lineages is comparable to species differences often reported in the literature. Sequence divergence values of 0.58% have been reported between yellow mongoose populations from southern Africa (Jansen van Vuuren 1995). It has been suggested that values in this range may represent subspecies (Girman *et al.* 1993; Jansen van Vuuren 1995). The average Kimura two-parameter sequence divergence between *Arvicanthis* species (Rodentia: Muridae) was 10.2% (Ducroz *et al.* 1998), but lower sequence divergence values (3.3%) were also detected between some of the species (*A. dembeensis* and *A. niloticus*). This study suggested that values in this range usually correspond to conspecific populations rather than to separate species (Ducroz *et al.* 1998). In contrast, sequence divergence values of 4% have been reported to reflect interspecific comparisons in rodents (Hayes & Harrison 1992). Low sequence divergence values of

between 1.6 and 3.8% were also recorded between two cryptic species *A. chrysophilus* and *A. ineptus* in southern Africa (see Chapter 3). Average HKY85 sequence divergence estimates between the three *Aethomys* lineages identified in Figure 2.4 varied from 4.55 to 5.33%. These patterns of variation are comparable to those reported for other intraspecific comparisons in muroid rodents (Patton & Smith 1992; Matthee & Robinson 1997; Ducroz *et al.* 1998). If the gamma shape parameter and the proportion of invariable sites were taken into account, sequence divergences as high as 8.91% were detected between the *A. namaquensis* lineages. These lineages are currently classified as subspecies, but the high sequence divergence detected between these lineages should be investigated further as they may reflect the presence of more than one species within *A. namaquensis* from southern Africa.

The first step in testing the biological species concept in *A. namaquensis* would be to narrow the taxon distributional boundaries and mapping areas of overlap through further sampling and genetic analyses. In areas of sympatry or parapatry, the presence or absence of hybrids (identified with diagnostic nuclear markers) could be used to test the breeding integrity of these groups. Lack of hybrids could indicate the existence of separate biological species. For example, several authors were able to show breeding isolation between the cryptic species *A. chrysophilus* and *A. ineptus* using various techniques such as karyology and sperm and bacular morphology (Gordon & Rautenbach 1980; Gordon & Watson 1986; Visser & Robinson 1986; Breed *et al.* 1988).

The remarkable diversity of karyotypes found in vertebrate species suggests that speciation may be associated with karyotypic changes (Robinson *et al.* 1986). This is certainly true for some rodents that include variation in diploid numbers within the North American genera *Spermophilus* and *Marmota* (Nadler & Hoffmann 1970) and the two cryptic sister species *A. crysophilus* and *A. ineptus* (Gordon & Rautenbach 1980; Visser & Robinson 1986), and heterochromatic differences between South African ground squirrels (Robinson *et al.* 1986). Other examples exist within southern African rodent genera, *Mastomys* and *Saccostomus* where species complexes of morphologically similar, though genetically distinct, species have been identified using karyotypes and protein electromorph mobility (Gordon & Watson 1986).

Individuals of *A. chrysophilus* were characterised by 2n = 50 and 2n = 44 (Gordon & Rautenbach 1980; Visser & Robinson 1986). No intermediate chromosomal forms (no evidence of hybridisation) were detected in sympatric populations, suggesting a lack of gene flow between the two forms and therefore, the presence of two good biological species, *A. chrysophilus* and *A. chrysophilus sp. B* (Gordon & Rautenbach 1980; Gordon & Watson 1986; Visser & Robinson 1986). Subsequently, morphometric analysis of cytogenetically identified individuals as well as more broadly sampled *A. chrysophilus* revealed the existence of two sympatric species referred to as the nominate, *A. chrysophilus*, 2n = 50 (De Winton, 1897) and a newly recognised *A. ineptus*, 2n = 44 (Thomas & Wroughton, 1908) (Chimimba 1997, 1998; Chimimba *et al.* 1999). These species also show differences in protein (haemoglobin) mobility, as well as gross sperm and bacular morphology (Baker *et al.* 1988; Breed 1997; Gordon & Watson 1986).

It is possible that similar scenarios as outlined above may also occur in other members of the genus *Aethomys*. Visser & Robinson (1986) found one diploid number in *A. namaquensis* in samples from eight localities. These localities mostly represented areas where lineages A, B and C (Fig. 2.4) were recorded in the present study. Since a further six mtDNA lineages were found, it may be premature to assume that there is no variation in diploid number within *A. namaquensis*. It is possible that a geographically more representative sample may reveal differences in chromosome number and/or structure. In addition, differences in protein electromorph mobility and gross sperm and bacular morphology warrant further investigation as these differences may prevent gene flow between lineages leading to reproductive isolation and subsequently to speciation.

The average time of separation between the lineages within *A. namaquensis* varied between 2.67 and 2.13 Myr ago suggesting a late Pliocene time of divergence. It is possible that the lineages within *A. namaquensis* were already separated at the beginning of the Pleistocene.

Since there may be an association between the lineages identified and vegetation types, ecological speciation may have played an important role in diversification within *A. namaquensis*. Ecological speciation occurs when divergent natural selection on traits between populations or subpopulations in different environments leads to the evolution of reproductive isolation (Schluter 2001). Ecological speciation might occur in either

allopatry or sympatry (Schluter 2001), but may have occurred within *A. namaquensis* in both sympatry and allopatry. For *A. namaquensis*, several behavioural or physiological adaptations may be a result of adaptation to vegetation type (food availability), climate and rainfall. For example, the smaller body size of *A. namaquensis* in the Karoo (Chimimba 2001a) may represent a form of adaptation to this biome. Bergmann's and Allen's rules suggest that homeotherms from cold environments should be larger (Bergmann 1847) and have smaller appendages (Allen 1877) than those from warmer climates. Studies on *Saccostomus* (Rodentia: Cricetidae) have shown that temperature has little effect on morphology, as there is no clear correlation between the size of appendages (tail or ear length) and temperature (Ellison *et al.* 1993) but geographic variation in body size was significantly correlated with temperature and seasonality, appearing to be largely the result of a positive correlation between rainfall and body size (Ellison *et al.* 1993).

In order to test the hypothesis of ecological speciation in *A. namaquensis*, one would have to sample extensively on a micro-geographic scale to test for changeover of *Aethomys* lineages associated with vegetation changes. If ecological speciation occurred it may be expected to find sympatric species of *Aethomys* in areas where habitat types overlap. More evidence is needed on the distribution of the lineages identified here to test their association with certain vegetation types, but allopatric speciation alone seems unlikely. From a topographic perspective, the Drakensberg mountain range, which forms part of the Great Escarpment, may act as a barrier to gene flow between populations east of the escarpment and those situated westwards of this barrier (Chimimba 2001a). This barrier has also been proposed to affect the biogeography of other small mammals, such as the yellow mongoose (Jansen van Vuuren 1995).

Aethomys namaquensis is known to utter soft piercing calls when disturbed (De Graaff 1997), but nothing is known about the geographic variation in these calls. Alarm vocalizations are often species-specific and can, therefore, be used to distinguish between closely related species (Bradbury & Vehrencamp 1998). Habitat dissimilarities, as is the case with *A. namaquensis*, could play a role in the evolution of interspecific call variation (Brandbury & Vehrencamp 1998). Different habitats should uniquely shape the vocalizations of species that inhabit them (Morton 1975, Ryan & Brenowitz 1985). Morton (1975) proposed that vocalizations of animals in a more densely vegetated habitat would be lower in frequency and fewer rapid frequency modulations than those of animals

in a more open habitat. These predictions are based on environmental factors that differ between habitats such as scattering by obstructions, echoes, background noise and irregular amplitude fluctuations (Morton 1975; Ryan & Brenowitz 1985). It has been reported by Gannon & Lawlor (1989) that related taxa diverge in vocalizations as much as they do in other characteristics. Reproductive character displacement in alarm calls has been reported, with vocal differences being enhanced at parapatric sites of some populations of the Townsend's chipmunks (Gannon & Lawlor 1989).

In contrast, neither species of whistling rats show reproductive character displacement in their alarm calls but there are consistent differences in the alarm whistles of the two species (Le Roux *et al.* 2002). *Parotomys littledalei's* calls are lower in frequency and wider in frequency bandwidth than *P. brantsii*, which supports Morton's hypothesis (Le Roux *et al.* 2002). This implies that the closed habitat species emits calls that are less pure in tone than those of the open habitat species (Le Roux *et al.* 2002), which is contrary to the predictions of the acoustic adaptation hypothesis (Morton 1975). In a similar study on marmots of the genus *Marmota*, Daniel & Blumstein (1998) also found results that did not support Morton's hypothesis. Morton's hypothesis is at least applicable only partially to rodent calls and it may explain the differences in call structure between the two whistling rat species (Le Roux *et al.* 2002). It is not known if vocalization plays a role in mate-recognition or mate-choice in *Aethomys*, but because *A. namaquensis* occurs in different habitat types which could have very different acoustic properties, it would be interesting to investigate whether the variation in vegetation played a role in their vocalization and subsequently their speciation.

Global temperature changes of considerable magnitude have occurred during the last 50 million years, more especially the last three million years and may have influenced the evolution of *A. namaquensis*. These changes had the effect of breaking up the African landmass into a series of discontinuous habitats (Brain 1985). A progressive decline in temperature from the Paleocene until the end of the Miocene, was followed by a series of subsequent oscillations. Between 6.5 and 5 million years ago, a severe temperature plunge resulted in a rapid and dramatic sea level drop of over 100 m worldwide. The temperature during the succeeding period appeared to have fluctuated, until 2.6-2.5 million years ago. Since then, temperatures have oscillated at least 17 times during the last 1.7 million years

(Brain 1985). These temperature changes certainly affected vegetation and may have contributed to speciation in southern African mammals (Axelrod & Raven 1978).

The Pliocene and early Pleistocene were characterised by uplifts which raised the interior plateaus of South Africa more than 1 800 m above the Miocene level (Baker & Wohlenberg 1971). There were also major land elevations during the later Tertiary in South Africa that resulted in the elevation of the central Highveld and the outward tilting of the marginal regions (King 1963). Maximum uplift took place along a line outside the Great Escarpment, while the central Karoo plateau developed as a basin. Aridity soon spread as a result of the Karoo basin. Superimposed on this were minor upwarps (in the form of long ridge-like axes) and basins, which led to the elevation of the central plateau far above its former level (King 1963). These deformations modified both the topography and the climate of the region (Axelrod & Raven 1978).

The broad warping and uplift of Africa during the Miocene and Plio-Pleistocene resulted in two major episodes of rapid speciation in South Africa. The first episode of speciation evidently commenced in the Miocene and the second burst of speciation was evident during the Plio-Pleistocene as a result of deformation and accompanying fluctuation of climate. Mountains were elevated and broad basins developed over the interior, the low areas became drier and the mountains moister. Erosion also sculptured many new small basins. Climate changed between wet and dry so that populations were shifting continuously (Axelrod & Raven 1978). Moist climate vegetation invaded the area of the present semi-desert, only to return to their original areas as conditions became drier. These changes had an influence not only on the vegetation but also on the large mammal fauna as new habitat zones spread (Axelrod & Raven 1978). These changes in climate, vegetation and the deformation of South Africa might have contributed to the differentiation within *A. namaquensis*. Many of these geological and climatic changes could have contributed to speciation, either through physical isolation (allopatric speciation) or changes in habitat (sympatric or ecological speciation).

The history of climate and vegetation change in South Africa is also relevant from a phylogeographic perspective. Phylogeography deals with genealogical relationships within and between closely related species. If *A. namaquensis* indeed represents a complex of closely related species, categories I and II (discontinuous genetic divergence

patterns) of Avise *et al.* (1987a) would apply. For example, some may prove to be spatially separated i.e., category I (e.g., lineages A and B) while others are not spatially separated (e.g., lineages A and C; B and D; B and E; B and I).

Lineage B appeared to be associated with semi-arid and arid vegetation types with low annual rainfall (Low and Rebelo 1996). On the periphery of the distribution of lineage B, several geographically unique lineages were identified in sympatry within lineage B. It is interesting to note that these are areas (Porterville, Augabies National Park, Upington and Vryburg) where different semi-arid or arid vegetation types merge or are in contact with each other. These areas are not classified as Nama Karoo but Karoo elements are evident (Low & Rebelo 1996), which may explain the occurrence of lineage B at these localities, excluding Vryburg. Both the Augrabies National Park and Upington localities are situated in areas where two vegetation types are in close contact with each other, which may explain the occurrence of divergent lineages. The Orange River Nama Karoo veld type and the Karriod Kalahari Bushveld veld type are in close contact with each other in the Upington and the Augrabies area (Low & Rebelo 1996).

Porterville (lineages B and D) is also located in an area where two vegetation types are in contact with each other namely, the Lowland Succulent Karoo and Mountain Fynbos (Low & Rebelo 1996). Fynbos burns at between 6 and 45 years of age and in the process sustains its plant species leading to many species to re-sprout, but the majority rely on the predictability of fires and only regenerate after the fire form seeds (Low & Rebelo 1996). If the vegetation does not burn frequently enough at low rainfall, the Fynbos is replaced by Succulent Karoo below 200mm annual rainfall (Low & Rebelo 1996).

The Vryburg locality is situated in the Vryburg Shrub Bushveld veld type and through the process of overgrazing this veld type has been broken down, allowing the Karoo to invade from the south (Acocks 1975). The overlap/invasion of different biomes may explain why two divergent lineages were found in Vryburg (Fig. 2.4). From Chimimba's (2001a) proposed distributions and because of the invasion of Karoo elements (Acocks 1975), lineage B may have been expected to occur at Vryburg. This lineage was, however, not recorded in the Vryburg area, possibly because of the small sample size or because of displacement by the lineages that were identified from this area. Instead, lineages E and I were recorded from this locality.

The available data showed that *A. namaquensis* falls within the discontinuous end (category I) of the spectrum of categories proposed by Avise *et al.* (1987a) to classify phylogeographic structuring. Category I entails phylogenetic discontinuities, spatial separation, long-term extrinsic barriers to gene flow and/or extinctions of intermediate genotypes in the species with limited gene flow. This type of genetic structure has been reported in other taxa showing discontinuous intraspecific mtDNA phylogenetic networks, with a strong geographic orientation such as the pocket gopher (*Geomys pinetis*) (Avise *et al.* 1979b) and deer mouse (*Peromyscus maniculatus*) (Lansman *et al.* 1983a). This pattern of phylogenetic discontinuity (category I) has also been detected in southern African small mammals, such as the rock hyrax (Prinsloo 1993), the red rock rabbit species (Matthee & Robinson 1996) and the scrub hare (Kryger 2002).

The phylogeographic profiles within lineages A, B and C differed markedly from each other. No evidence of isolation by distance was detected within lineages B and C. More specifically, there was more genetic differentiation within lineage C compared to lineage B, despite analysing less samples over a geographically more restricted area for the former lineage (Figs. 2.7B & 2.7C; Fig. 2.4). Lineage C had no shared haplotypes between localities and differentiation within this lineage occurred approximately 758 000 years ago. Within lineage B, differentiation occurred more recently (about 263 000 years ago) with three haplotypes being shared among localities over large geographical distances, for example, haplotype 30 was shared between Springbok and Gariep Nature Reserve (700 km). Lineage B may, therefore, have undergone a more recent expansion after a bottleneck (Fig. 2.8B), whereas the population history within lineage C may have been more complex with waves of colonization or dispersal among regions (Fig. 2.8C). It is possible that large population sizes are responsible for the maintenance of the high haplotypic diversity (0.97) and maintaining the existence of divergent haplotypes. Only a small part of the probable distributional range of lineage A was covered in the present study. As many as seven haplotypes were identified within lineage A from only two localities and only 11 individuals, suggesting that this lineage shows more diversity than lineage B. It seems as if the haplotypes sampled within lineage A diverged at least 297 000 years ago.

Lineage A was only recorded from two localities in the Savanna Biome and similar to lineage C (Fig. 2.4), it was characterised by a low intraspecific sequence divergence

estimate (0.59%) and a low number of mutational steps (Fig. 2.7A). This may reflect fairly close common female ancestry, possibly due to an expansion in the recent past (Fig. 2.8A). The Tajima's *D* test and Fu's F_s test supported this finding. The mismatch distribution test, however, did not support expansion possibly due to a small sample size. The variance within the populations of lineage A (90.78%) was far greater than that detected between populations (9.22%), thus emphasizing the lack of geographic genetic partitioning in this lineage. Four maternal haplotypes were detected in the six samples collected from Botswana and four haplotypes were detected in the six samples collected from Messina, which may also indicate relatively old, historically large populations, or it may simply point to recent waves of colonization events.

All the individuals collected within the proposed distributional range of the subspecies *lehocla* (Upington, Augrabies National Park and Loxton) grouped within lineage B, lending support to the recognition of this subspecies. It is evident from the previous study by Chimimba (2001a) that the subspecies *lehocla* is restricted to the Nama Karoo but it seems from the present study that this subspecies may also occur in neighbouring semi-arid areas (Kirkwood, Oudtshoorn, Springbok, Gariep Nature Reserve and Porterville). Although not all of these localities fall within the recognised boundaries of the Nama Karoo, Karoo elements and/or the invasion of the Nama Karoo are evident. The annual rainfall is relatively low, between 100 – 520 mm per year (Low & Rebelo 1996); with the possible exception of Porterville, localities can all be described as semi-arid or arid. The Oudtshoorn (Spekboom Succulent Thicket), Kirkwood (Xeric Succulent Thicket) and Springbok (Upland Succulent Karoo) localities are not situated in the Nama Karoo biome but these localities are in close contact with Nama Karoo vegetation types (Little Succulent Karoo, Great Nama Karoo, Central Lower Karoo and Bushman land). The Gariep Nature Reserve and Loxton localities are, however, both situated in the Nama Karoo biome.

The limited genetic structure within lineage B (Fig. 2.7B) was also reflected by the relatively low intraspecific nucleotide diversity (0.53%). Haplotypes within lineage B were generally separated by few mutational steps, even between haplotypes that were geographically distant (Fig. 2.7B). Rare haplotypes are more likely to be mutational derivations of common haplotypes (Excoffier & Langaney 1989). Therefore, the presence of the closely related haplotypes NH30 (4 localities), NH34 (two localities) and NH36 (three localities) within almost all populations of lineage B (Table 2.3), might reflect close

common female ancestry, probably due to recent range expansion from a source population(s) (Fig. 2.8B). However, four maternal haplotypes were detected in the six samples collected from Springbok. This might indicate a relatively old, historically large population, or it may simply reflect a recent colonization event.

Sequences could, unfortunately, not be generated from museum specimens from the central parts of South Africa, leaving a sampling gap across most of the Grassland Biome. A geographically widespread lineage (C) associated with Grassland/Savanna, was however identified. Kruisrivier Nature Reserve (Loskop Dam) and Wathaba-Uitkomst (Machadodorp) were the only localities that clearly form part of the Grassland Biome. It has been proposed that Grassland expanded during the glacial periods because it was colder, allowing frost that is an important factor in the distribution of Grassland (Brain 1985). If this correspondence between Grassland and winter frost with the added effects of fire is significant, then it may be reasonable to suggest that when the area of greater winter cold expanded periodically in the past, it may have caused an expansion of Grassland at the expense of other vegetation types. There is evidence from Aliwal North and Florisbad that during the last glacial maximum, pure Grasslands replaced other forms of vegetation (Coetzee 1978; Van Zinderen Bakker 1957). If Grasslands expanded in response to suitable climate, it may be assumed that animals with a preference for this kind of habitat such as the white-tailed rat (*Mystomys albicaudatus*) would have similarly expanded their range. The expansion of Grassland during the last glacial maximum might explain why the distribution of lineage C may have extended further north to Thabazimbi, Kgaswane Mountain Reserve and Messina than expected (Chimimba 2001a). Hoedspruit, Kgaswane Mountain Reserve, Thabazimbi and Messina represent areas recognised as Savanna with Grassland elements that have changed both historically and naturally.

In contrast to lineage B, lineage C was characterised by high sequences divergences (1.52%) between the sampled mtDNA haplotypes (Fig. 2.7C & Table 2.5). This suggests that the populations have been separated from each other for a long period, thus allowing for more site changes to have accumulated. All the populations were characterised by two haplotypes except for one haplotype identified at Messina. In one instance, a closer affinity was shown to haplotypes from other localities rather than to those from the same locality (NH14 and NH15). No haplotypes were shared between populations and the presence of several discrete haplotypes within populations may reflect evolutionarily old

populations, which may have served as core regions from where colonization into the adjacent areas may have occurred.

Lineage C (Fig. 2.7C) showed two distinct mtDNA groups, separated by six mutational steps. Interestingly, this mtDNA discontinuity coincides broadly with that of the rock hyrax, *Procavia capensis* (Prinsloo 1993) and the red rock rabbits, *P. rupestris* and *P. randensis* (Matthee & Robinson 1996) in the Mpumalanga region. The discontinuity in the rock hyrax is considered to reflect dispersal along two separate routes corresponding to two mtDNA clades: along the mountain ranges comprising the Great Escarpment and along the Soutpansberg-Magaliesberg axis (Prinsloo & Robinson 1992). The south-eastern clade of *P. rupestris* may also have dispersed along the Great Escarpment of South Africa, while the north-western assemblage was not so tightly constrained. It seems likely that the current observed patterns of genetic variation within this assemblage, have probably resulted from a combination of vicariance and dispersal events (Matthee & Robinson 1996). Whether the correspondence in the zone of contact between the two *Pronolagus* species, the two mtDNA clades within the rock hyrax and the two mtDNA lineages (C) within *A. namaquensis* is real or merely by chance, is subject to speculation. Avise (1992) has proposed that the presence of congruence in phylogeographic patterns in evolutionarily independent lineages might reflect similar vicariant events. This hypothesis may also hold for many other mammals with similar habitat and dispersal capabilities.

Sequence divergence estimates based on mtDNA data indicate that the events responsible for the divergence within lineage C occurred 1.16 (0.80-1.60) Myr ago. In contrast, the divergence within the red rock rabbit and the rock hyrax assemblages occurred earlier, at 4 Myr and 2 Myr ago, respectively. This suggests that three different events caused differentiation in these taxa or the same barrier to gene flow caused isolation at different times. Conversely, the differences in the divergence times referred to above may be due to the differences in molecular clock calibrations for the three evolutionary lineages representing the Rodentia, the Lagomorpha and the Hyracoidea. The rate of nucleotide substitution in rodents is shown to be at least 1.5 times higher than those for lagomorphs (Li *et al.* 1990). It has also been shown that mtDNA of mice evolves at a rate at least three times faster than the 2% per million years currently found in primates and other mammals (She *et al.* 1990).

Chapter 3

Molecular systematics of *Aethomys chrysophilus* and *A. ineptus* from southern Africa

Abstract

The genetic differentiation among populations of the Red veld rat, *Aethomys chrysophilus*, and the Tete veld rat, *A. ineptus* from South Africa and Swaziland was examined using cytochrome b sequences. Neighbour-joining analysis showed reciprocal monophyly between the species, which suggested that mitochondrial DNA (cytochrome b) analysis can be used to distinguish the two cryptic species from one another. There seems to be much less structuring within *A. ineptus* and *A. chrysophilus* than other southern African small mammals. A database that included all positively identified specimens of both species, data collected by previous authors and distributional data from the present study, suggested that the distribution of *A. chrysophilus* reaches its southern limit in northern South Africa, primarily in low elevations of the Limpopo River drainage. *Aethomys ineptus* occupies the remainder of South Africa, which occurs at higher elevations but extends into lower elevation habitats in the southern portion of its range. The first confirmed record of *A. ineptus* in Swaziland was also evident from the distributional data in the present study. The divergence time separating the two sibling species was estimated at 1.36 Myr ago, suggesting an early Pleistocene separation.

Introduction

The Red veld rat, *Aethomys chrysophilus* De Winton, 1897, conventionally regarded as a single species, is widely distributed in eastern and southern Africa ranging from Kenya southward to southern KwaZulu-Natal in South Africa and into Namibia (Meester *et al.* 1986; Skinner & Smithers 1990). Studies in the 1980s provided the first indications that *A. chrysophilus* may not be a homogeneous assemblage in southern Africa (Gordon & Rauthenbach 1980; Gordon & Watson 1986; Visser & Robinson 1986). These studies revealed the occurrence of two cytotypes within *A. chrysophilus*, 2n = 50 referred to as *A. chrysophilus* and 2n = 44 referred to as *A. chrysophilus sp. B*, suggesting that these cytotypes represent two distinct species. Although the two cytotypes were indistinguishable in external appearance, they differed in haemoglobin protein electromorphs, and in gross sperm and bacular morphology (Gordon & Rautenbach 1980; Gordon & Watson 1986; Visser & Robinson 1986, 1987; Baker *et al.* 1988; Breed *et al.* 1988). Haemoglobin electromorph patterns showed that the 2n = 50 cytotype had a "slow" double band and the 2n = 44 cytotype had a "fast" double band. It was also found that sperm heads of the 2n = 50 cytotype were hook-shaped while those of the 2n = 44 cytotype were spatulate-shaped. The absence of hybrids in areas of sympatry suggested reproductive isolation between the two cytotypes.

In a recent morphological and morphometric revision of the genus in southern Africa, the 2n = 50 cytotype was referred to the nominate species while the 2n = 44 cytotype was formally designated as *A. ineptus* Thomas & Wroughton, 1908 (Chimimba 1997, 1998; Chimimba *et al.* 1999). Additional intraspecific morphometric analysis during the systematic revision of the genus led to seven of the nine previously recognised subspecies within *A. chrysophilus* (*sensu lato*), namely *tzaneenensis* Jameson, 1909; *pretoriae* Roberts, 1913; *capricornis* Roberts, 1926; *magalakuini* Roberts, 1926; *tongensis* Roberts, 1931; *harei* Roberts, 1946 and *fouriei* Roberts, 1946 to be re-allocated to *A. ineptus* but provisionally treated as junior synonyms (Chimimba 1997; Chimimba *et al.* 1999). The two remaining previously recognised subspecies within *A. chrysophilus* (*sensu lato*), namely *acticola* Thomas & Wroughton, 1908 and *imago* Thomas, 1927 were provisionally retained within *A. chrysophilus* as junior synonyms (Chimimba 1997; Chimimba *et al.* 1999).

The phenetic discontinuity between the two proposed subspecies within *A. chrysophilus* coincided with an altitudinal limit (Clark 1967) with *A. c. chrysophilus* occurring below 500m above sea level while *A. c. imago* occurred higher than 500m above sea level in the eastern parts of southern Africa (Chimimba 2000). It was hypothesized that this altitudinal limit in geographic distribution might act as a potential barrier to gene flow between these two subspecies (Chimimba 2000). The recognition of these subspecies was also supported by both size- and shape-related morphometric characters, with no evidence of clinal patterns of variation (Chimimba 2000). In contrast, variation within *A. ineptus* suggested a clinal pattern of variation, where overall cranial size was positively and significantly correlated with longitude (Chimimba 2001b). These intraspecific patterns of variation within both *A. chrysophilus* and *A. ineptus* from southern Africa were only provisionally considered pending further investigation using other alternative systematic techniques such as DNA analysis (Chimimba 2001b).

While *A. chrysophilus* (*sensu lato*) has traditionally been considered to be widely distributed in southern Africa, initial studies suggested an extensive overlap between the two sibling species (Gordon & Rautenbach 1980). The distributions derived from morphological and morphometric data suggested a broad overlap between the two cryptic species including local sympatry (Chimimba *et al.* 1999). Locality data from positively identified specimens (cytogenetics, protein electrophoresis, sperm morphology and mtDNA analysis) in South Africa, however, strongly suggested that *A. chrysophilus* primarily occupies the low-lying areas of the Limpopo River drainage system, reaching its southern limit in northern South Africa (Linzey *et al.* 2003). In contrast, *A. ineptus* occupies the remainder of South Africa at higher elevations (800–1375 m above sea level) in the north-central parts but expands into lower elevations in the southern parts of its range, suggesting that within South Africa, *A. ineptus* tolerates a broader array of ecological conditions than *A. chrysophilus* (Linzey *et al.* 2003). Only one sympatric locality (Langjan Nature Reserve, Limpopo Province) was identified, despite their ranges being reported to overlap in the north-western part of Pretoria (Gauteng Province) and west of the Soutpansberg (Limpopo Province) (Linzey *et al.* 2003).

A number of studies over the past 25 years have drawn attention to the presence of sibling species of rodents in southern Africa (Gordon & Watson 1986; Visser & Robinson 1986, 1987; Chimimba 2000; 2001b). These studies have, however, largely been based, either

independently or in combination, on cytogenetics, protein electrophoresis, sperm/bacular morphology, cranial morphology and/or morphometrics, but have remained untested using alternative systematic techniques (Gordon & Rautenbach 1980; Gordon & Watson 1986; Robinson *et al.* 1986). Consequently, the present study represents a case study in which results obtained by the above techniques were independently tested using molecular and cladistic techniques.

Molecular techniques, specifically the analysis of the mtDNA cytochrome b gene, have been useful in the identification of sibling species (Peppers & Bradley 2000). Bradley & Baker (2001) proposed that sequence divergence values in the mtDNA cytochrome b gene lower than 2% were indicative of intraspecific variation; values between 2% and 11% had a high probability of being indicative of either conspecific populations or valid species but merited additional studies to evaluate species status. Values above 11% were indicative of valid species suggesting that the higher the value, the greater the probability for the existence of biological species (Bradley & Baker 2001). Other examples where cytochrome b sequence divergence data have been used to identify species are in cotton rats of the genus *Sigmodon* (Peppers & Bradley 2000) and rodents within *Peromyscus* (Bradley *et al.* 2000) and *Reithrodontomys* (Bell *et al.* 2001). However, caution needs to be exercised when using mtDNA sequence divergence data in cases where hybrids occur. Instead, data from nuclear genes may assist in resolving this problem, but since no hybridisation has to date been detected between *A. chrysophilus* and *A. ineptus*, cytochrome b data were used to address the following questions:

1. Is there any molecular difference between *A. chrysophilus and A. ineptus* from southern Africa?

2. Are intraspecific patterns of molecular variation within *A. chrysophilus* and *A. ineptus* concordant with those obtained using morphometric data (Chimimba 2000; 2001b)?

The results obtained by molecular techniques were in turn interpreted with reference to:

1. The phylogeographic structure of other southern African small mammal species for which data are available, namely, the rock hyrax (*Procavia capensis*; Prinsloo & Robinson 1992), red rock rabbits (*Pronolagus* sp.; Matthee & Robinson 1996), the springhare

(*Pedetes capensis*; Matthee 1993; Matthee & Robinson 1997) and the yellow mongoose (*Cynictis penicillata*; Jansen van Vuuren & Robinson 1997).

2. The distributional data obtained from both morphometric data (Chimimba 1997, 1998; Chimimba *et al.* 1999) and positively identified specimens using protein electrophoresis, karyology and/or sperm morphology (Linzey *et al.* 2003).

Materials and Methods

Study area and sampling

Samples were obtained from 17 localities in South Africa and Swaziland (Fig. 3.1). Geographic coordinates for these localities are listed in Table 3.1. The hypothesized distribution of *A. chrysophilus* (grey circles) and *A. ineptus* (grey vertical lines) based on positively identified individuals (Linzey *et al.* 2003) are indicated in Figure 3.1. Samples were collected between July 1999 and June 2002 according to the procedure described in Chapter 2. Some of these individuals were positively identified using karyology (chromosome counts). Heart, liver, kidney and muscle tissue were removed and stored according to the methods described in Chapter 2. Some tissues were also stored in 100% EtOH. Voucher specimens were deposited in the Transvaal Museum, Pretoria, South Africa.

Table 3.1 Geographic coordinates of all collecting localities of *Aethomys chrysophilus* and *A. ineptus* from South Africa and Swaziland analysed in the present study. Numbers 1 – 17 correspond to the numbers indicated in Figure 3.1.

Locality	Country	Province	Geographic coordinates
1. Messina Nature Reserve	South Africa	Limpopo	22 24'45"S 30 03'01"E
2. Pafuri, Kruger National Park	South Africa	Western Cape	22 25'49"S 31 10'25"E
3. Langjan Nature Reserve	South Africa	Limpopo	22 51'30"S 29 14'30"E
4. Happy Rest Nature Reserve	South Africa	Limpopo	23 01'30"S 29 44'30"E
5. Percy Fyfe Nature Reserve	South Africa	Limpopo	24 02'30"S 29 10'30"E
6. Selati Nature Reserve, Hoedspruit	South Africa	Limpopo	24 09'30"S 30 40'50"E
7. Orpen Gate, Kruger National Park	South Africa	Mpumalanga	24 34'30"S 31 06'30"E
8. Vaalkop Dam Nature Reserve	South Africa	North West	25 19'25"S 27 25'52"E
9. Kruisrivier Nature Reserve, Loskop Dam	South Africa	Mpumalanga	25 21'08"S 29 32'26"E
10. Botsalano Game Reserve	South Africa	North West	25 34'30"S 25 41'30"E
11. Roodeplaat Dam Nature Reserve	South Africa	Gauteng	25 38'30"S 28 22'30"E
12. Wathaba-Uitkomst, Machadodorp	South Africa	Mpumalanga	25 47'27"S 30 22'31"E
13. Malolotja Nature Reserve	Swaziland		26 10'30"S 31 11'32"E
14. Farm: Vlakfontein, Vryburg	South Africa	North West	27 04'22"S 24 46'07"E
15. Farm: Koedoesberg, Pongola	South Africa	Kwa-Zulu Natal	27 26'31"S 31 41'41"E
16. Albert Falls Nature Reserve	South Africa	Western Cape	29 28'18"S 30 23'53"E
17. Ashburton	South Africa	Northern Cape	29 38'56"S 30 27'16"E

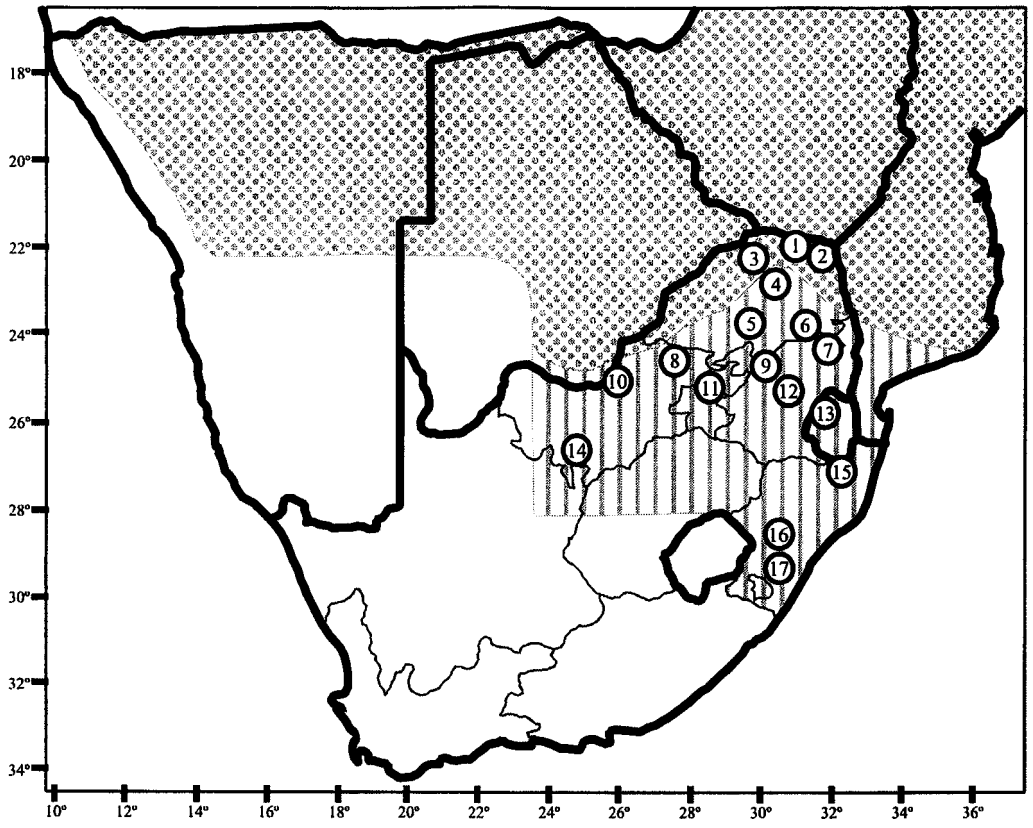


Fig. 3.1 Collecting localities of *Aethomys chrysophilus* and *A. ineptus* in South Africa and Swaziland. Numbers in circles correspond to the locality numbers in Table 3.1. The area shaded with grey dots and the one shaded with vertical lines indicate the hypothesized distribution of *A. chrysophilus* and *A. ineptus*, respectively (from Linzey *et al.* 2003). 1 – Messina Nature Reserve; 2 – Pafuri, Kruger National Park; 3 – Langjan Nature Reserve; 4 – Happy Rest Nature Reserve; 5 – Percy Fyfe Nature Reserve; 6 – Selati Nature Reserve, Hoedspruit; 7 – Orpen gate, Kruger National Park; 8 – Vaalkop Dam Nature Reserve; 9 – Kruisrivier Nature Reserve, Loskop Dam; 10 – Botsalano Game Reserve; 11 – Roodeplaat Dam Nature Reserve; 12 – Wathaba-Uitkomst, Machadodorp; 13 – Malolotja Nature Reserve, Swaziland; 14 – Farm: Vlakfontein, Vryburg; 15 – Farm: Koedoesberg, Pongola; 16 – Albert Falls Nature Reserve; 17 – Ashburton. Localities 1 and 2 represent localities where positively identified individuals of *A. chrysophilus* were collected while localities 8, 12, 16 and 17 represent localities where positively identified specimens of *A. ineptus* were obtained.

DNA extraction, PCR amplification, cycle sequencing and sequencing

Total genomic DNA was extracted from 0.2 g of tissue (a toe or ear clip) using a standard phenol/chloroform protocol (Sambrook *et al.* 1989; Chapter 2). DNA pellets were air-dried and the DNA was resuspended in 50-100µl Sabax® water (Adcock Ingram).

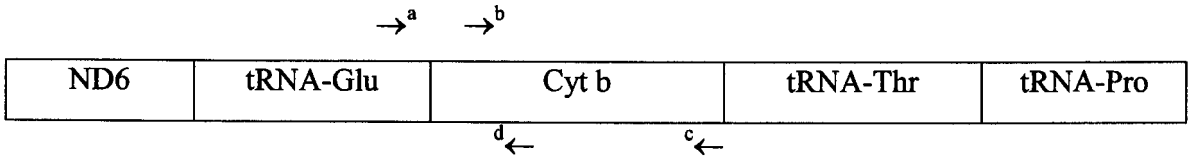
An *A. chrysophilus* sequence (Ducroz *et al.* 1998; GenBank accession # AF004587) together with a *Mus musculus* sequence (Bibb *et al.* 1981; GenBank accession # J01420) were aligned in Clustal X (Thompson *et al.* 1997) and used to design an internal species-specific primer (MUSINTH) for *A. chrysophilus* and *A. ineptus*. This internal primer was used in combination with either the shortened L14724 (Paäbo *et al.* 1988) or L14841 (Kocher *et al.* 1989) to amplify either a 594 or a 476 bp fragment of the 5' end of the cytochrome b gene, respectively. In some individuals, a longer fragment was amplified with the shortened L14724 and H15915 MUS (a 1194 bp fragment). These primer names, sequences and annealing positions on the mouse mtDNA genome are indicated in Table 3.2.

Polymerase Chain Reaction (PCR; Saiki *et al.* 1988) amplifications were performed as described in Chapter 2. The PCR conditions were: denaturing at 94°C for 4 minutes; 35 cycles of denaturing at 94°C for 45 seconds, primer annealing at 58°C for 1 minute, and elongation at 72°C for 1 minute; and an extended elongation at 72°C for 5 minutes. PCR products were purified using the High Pure™ PCR Product Purification Kit (Boehringer Mannheim) as prescribed by the manufacturers.

Dye-terminator cycle sequencing was performed for both the light (L14724 or L14841) and heavy strands (MUSINTH). Some individuals were also sequenced with the shortened L14724 and H15915 MUS, after which individuals were sequenced with MUSINTH in a Geneamp® PCR system 9700 (PE Applied Biosystems). Nucleotide sequences were determined using an ABI 377 automated sequencer or by a capillary system on an ABI 3100 sequencer (PE Applied Biosystems).

Table 3.2 Primer names, sequences (5' to 3'), and positions of primers used in Polymerase Chain Reactions (PCRs; Saiki *et al.* 1988) and cycle sequencing. Rodent specific primer names refer to the position of the 3' nucleotide position in the mtDNA sequence of *Mus musculus* (Bibb *et al.* 1981). Superscripts a-d correspond to the lettering in the diagram below.

Primer name	Rodent specific name	Primer sequence
L14724[a]	L14115	5' TGAYATGAAAAAYCATCGTTG 3'
L14841[b]	L14233	5' CCAACATCTCAGCATGATGAAA 3'
H15915 MUS[c]	H15309	5' CATTTCAGGTTTACAAGAC 3'
MUSINTH[d]	H14709	5' AGGTGAACGATTGCTAGGGC 3'



Sequencing analysis

The raw sequence data were checked for quality in Sequence Analysis version 3 and sequences were imported into Sequence Navigator version 1.01 (PE Applied Biosystems). A consensus sequence for each individual was computed by aligning the forward and reverse sequences of each individual (370 bp consensus sequence). Although the initial sequences were on average 530 bp long, consensus sequences of only 370 bp were computed because most of the missing data were excluded. All sequences were deposited in GenBank. Consensus sequences of all individuals were aligned in Clustal X (Thompson *et al.* 1997) after which sequences were imported into PAUP version 4.0b10 (Swofford 2002) for phylogenetic analyses and MacClade version 3.0 (Maddison & Maddison 1992) in order to translate the nucleotide sequences into amino acids. Amino acid sequences were examined for irregularities (e.g., stop codons in the reading frame).

Transition:transversion ratio (Ti:Tv) was estimated using MacClade based on a 1000 random trees (Halanych 1996).

Outgroup choice

The selection of possible outgroups for *A. chrysophilus* and *A. ineptus* was difficult because of previous systematic links of the genus to numerous other murids (Chimimba 1997). Additionally, the phylogenetic analysis of *A. chrysophilus* and *A. ineptus* is herein treated as an intraspecific phylogeny since these two cryptic species are closely related. This posed a problem since, when dealing with intraspecific phylogenies, the specified outgroup such as *Dasymys incomtus* may often be too distant to be meaningful, (Maddison *et al.* 1992; Crandall & Templeton 1993). For example, *A. namaquensis* was even more distantly related to *A. chrysophilus* and *A. ineptus* than *D. incomtus* (17-19% sequence divergence in the cytochrome b gene, see Chapter 4). Consequently, midpoint rooting in PAUP was used.

Phylogenetic analyses

A maximum likelihood ratio test performed in Modeltest version 3.06 (Posada & Crandall 1998) was used to determine the best-fit model of DNA substitution for the sequence data. The best-fit model obtained from this test was then used in distance analyses. Sequence statistics such as the Ti:Tv ratio, the α parameter of the gamma distribution of rate variation among sites (Yang *et al.* 1994; Yang 1996; Yang & Kumar 1996) and the proportion of invariable sites (I) were also estimated. Pairwise estimates of the nucleotide sequence divergence were calculated according to the best-fit model of substitution and the HKY85 model of substitution (Hasegawa *et al.* 1985) since this model corrects for unequal base frequencies and different rates of transition and transversion substitutions. Genetic variation was also assessed by calculating the number of pairwise nucleotide Ti and Tv substitutions between all pairs of specimens as well as pairwise protein differences using PAUP.

All distance analyses were performed on the different haplotype sequences identified based on the pairwise estimates of the nucleotide sequence divergence derived from the best-fit model of substitution. Distance trees were constructed using the neighbour-joining

algorithm (Saitou & Nei 1987) by applying the midpoint rooting option in PAUP. Tree nodal support was assessed by 1000 bootstrap replicates (Felsenstein 1985).

Parsimony trees (Kluge & Farris 1969; Farris *et al.* 1970) were generated using nucleotides as unordered characters and the heuristic search (500 heuristic searches) option with tree bisection-reconnection (TBR). If more than one minimum length tree was obtained, a strict consensus tree was constructed. Tree length, consistency index (CI; Kluge & Farris 1969), retention index (RI; Farris 1989) and a rescaled consistency index (RC; Archie 1989; Meier *et al.* 1991) were calculated as tree description data. Phylogenetic signal was assessed by evaluating the tree-length distribution of 1000 randomly generated trees using the g_1 statistic (Hillis & Huelsenbeck 1992). Different weighting schemes were applied in an attempt to reduce the number of minimum length trees by reweighting with reference to the CI and RI (Farris 1969). For maximum likelihood analyses (Felsenstein 1973; 1981), the base frequencies and the Ti:Tv ratio were estimated. To adjust for the possible effects of among site rate variation, the α parameter of the gamma distribution (Yang & Kumar 1996) and the proportion of invariable sites (I) were also calculated in maximum likelihood. These estimated values were then used to construct likelihood trees. Parsimony analyses were performed on all unique haplotypes identified while maximum likelihood analysis was based on representatives of the lineages identified in the neighbour-joining tree. The data were also interpreted in terms of the molecular clock hypothesis.

A relative rate test was performed using RRTree in order to assess differences in third position transversion rates among separate lineages relative to *D. incomtus* as a reference taxon (Robinson *et al.* 1998; Robinson-Rechavi & Huchon 2000). A specific rate of change calibrated on murid rodent data was determined since murid mtDNA evolves at a faster rate than that of other rodents (Catzeflis *et al.* 1992). Sequence data from *Rattus rattus* and *Mus musculus*, with a divergence time estimated at 12 Myr, was used as a calibration point (Jacobs & Downs 1994; Chapter 2). The number of third position transversions was used as a measure of genetic divergence because it accumulates nearly linearly with time (Irwin *et al.* 1991). The conventional rate of 2% sequence divergence per million years was also applied to the data (Brown *et al.* 1979).

Population level analyses

Measures of population genetic parameters such as haplotype diversity (Nei & Tajima 1981) and nucleotide diversity (Pi (π); Nei 1987) were calculated using DnaSP version 3.51 (Rozas & Rozas 1999; Chapter 2). The minimum number of base substitutions between maternal lineages was determined using MINSPNET (Excoffier & Smouse 1994) in combination with Arlequin version 2.0 (Schneider *et al.* 2000), with the latter allowing for a 5% level of missing data per site. Minimum spanning networks in combination with frequency data and the geographic distribution of haplotypes were used to depict geographical and potential ancestor-descendant relationships among identified haplotypes.

The extent to which sequence variation is partitioned between specified groups, between populations within groups as well as within populations were evaluated using ϕ_{st}, ϕ_{ct} and ϕ_{sc} estimates derived from an analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) algorithm in Arlequin version 2.0 (Schneider *et al.* 2000). Statistical significance was tested using a non-parametric permutation scheme using 10000 permutations with an allowance for a 5% level of missing data per site (Chapter 2).

Mismatch distribution analysis as implemented in Arlequin version 2.0 (Schneider *et al.* 2000) was used to examine the frequency distribution of pairwise nucleotide differences (Slatkin & Hudson 1991; Rogers & Harpending 1992). The fit of the observed distribution of mismatches to a sudden model of population expansion was tested by bootstrap analysis with 10000 permutations. This distribution is usually unimodal for lineages that have undergone a recent bottleneck or population expansion and multimodal for populations at equilibrium (Rogers & Harpending 1992). A 5% level of missing data per site was allowed for mismatch distribution analyses (see Chapter 2).

Geographic distribution

A database of collecting localities for *A. chrysophilus* and *A. ineptus* that included positively identified specimens from the literature (see Linzey *et al.* 2003) and data from the present study was assembled. The majority of records were based on specimens identified by chromosome number and sperm morphology while samples in this study were identified based on mtDNA analysis. Geographic coordinates were used to plot the geographic distribution of *A. chrysophilus* and *A. ineptus*, represented by half degree grid

squares, on a map of southern Africa and compared with published data based on positively identified specimens (Linzey *et al.* 2003). Although only three *A. chrysophilus* localities were included in the present study, the geographic distribution of this species was discussed with reference to previous positively identified specimens.

Results

Sequence statistics

A total of 370 bp of the 5' end of the cytochrome b gene were successfully sequenced in all individuals of *A. chrysophilus* and *A. ineptus* examined. All sequences obtained were of the mitochondrial cytochrome b gene, since no stop codons were found when sequences were translated into amino acids and the nucleotide sequences corresponded to the expected 123 amino acids. A blast search also found the highest homology to cytochrome b sequences of *Aethomys*. There were 25 variable positions defining 21 haplotypes (Table 3.3; Appendix 3.1). The most widespread haplotype (H01) was recorded from nine localities (30 individuals). Of the 25 variable sites (Table 3.4), 13 were phylogenetically informative (excluding outgroups). Three informative sites were recorded from the first, one from second and ten from third codon positions. Most of the substitutions were silent with only three variable amino acid sites between the most divergent lineages.

Mean base compositions at the three different codon positions for the whole sample studied are provided in Table 3.4. The four nucleotides did not occur in equal frequencies with base composition similar to that of the *A. namaquensis* cytochrome b sequences (Chapter 2). This strong bias in base composition showed a marked under-representation of guanine at both second (18.0%) but particularly third (1.9%) codon positions. Similar to the sequence statistics in Chapter 2, a higher representation of adenine at the third position (53.0%) and thymine at the second position (40.7%) was also observed in these sequences.

Table 3.3 Variable sites of the 21 mtDNA cytochrome b haplotypes (370 bp) of *Aethomys chrysophilus* and *A. ineptus* from South Africa and Swaziland. Variable positions five and 346 correspond to position 14247 and 14588 of *Mus musculus* (Bibb *et al.* 1981). Dots (.) indicate identity to the base in the reference sequence H01 and question marks (?) indicate undetermined bases. Population numbers correspond to those in Figure 3.1.

Haplotype No.	Variable positions	No. of individuals	Populations (Sample size, Population No.)
	1111112222223333 11144679022247145569024 4 51390113531485244041245 06		
H01	?GTATCTGACTATCCTAGCCTACTT	30	Langjan Nature Reserve (1,3), Happy Rest Nature Reserve, Lower (2,4), Percy Fyfe Nature Reserve (1,5), Albert Falls Nature Reserve (1,17), Kruisrivier Nature Reserve (1,9), Vaalkop Dam Nature Reserve (2,8), Roodeplaat Dam Nature Reserve (10,11), Botsalano Game Reserve (3,10) Happy Rest Nature reserve, Upper (9,4)
H02	C.........G.....	3	Langjan Nature Reserve (1,3), Selati Nature Reserve (1,6), Happy Rest Nature Reserve (1,4)
H03	C.........CG.....	1	Langjan Nature Reserve (1,3)
H04	T.........G.....	2	Happy Rest Nature Reserve, Upper (2,4)
H05	CA........G.....	3	Happy Rest Nature Reserve, Upper (3,4)
H06	C.........G.....T.....	1	Happy Rest Nature Reserve, Lower (1,4)
H07	C.........C.....	1	Happy Rest Nature Reserve, Lower (1,4)
H08	C.......G.....	1	Percy Fyfe Nature Reserve (1,5)
H09	?......A.....C	1	Percy Fyfe Nature Reserve (1,5)
H10	C.........G..T..A...G...	1	Kruger National Park, Orpen Gate (1,7)
H11	C.........G..T.....T..	1	Wathaba-Uitkomst, Machadodorp (1,12)
H12	C...C......G..T.....	1	Malolotja Nature Reserve, Swaziland (1,13)
H13	C.....C....	1	Farm: Vlakfontein, Vryburg (1,14)
H14	C....T.....C....	1	Farm: Vlakfontein, Vryburg (1,14)
H15	C......A......T.....	3	Farm: Koedoesberg, Pongola (3,15)
H16	C.........G..T.....	1	Ashburton (1,17)
H17	C........T.GCT.C.CT....C.	1	Messina Nature Reserve (1,1)
H18	C.C......T..CT...AT.C..C.	1	Kruger National Park, Pafuri (1,2)
H19	C.C...C..T..CT..GAT.C..??	1	Kruger National Park, Pafuri (1,2)
H20	CACG.....T..CT....T.C..C.	1	Langjan Nature Reserve (1,3)
H21	C........T.GCT...CT....C.	2	Langjan Nature Reserve (2,3)

The best-fit model was determined to be the HKY85 model of substitution (Hasegawa *et al.* 1985) with a gamma correction (Gu & Zhang 1997) of 0.7273 and the proportion of invariable sites of 0.7041. Maximum likelihood (the α shape parameter of the gamma distribution of rate variation among sites, the proportion of invariable sites (I), base compositions and the Ti:Tv ratio) and MacClade version 3.0 (Ti:Tv ratio) yielded the same sequence statistics as Modeltest version 3.06. Pairwise estimates of the nucleotide sequence divergence (HKY85 and HKY85 + Γ + I) are indicated in Table 3.5. HKY85 sequence divergence values between the 21 haplotypes ranged from 0.27% to 3.75%. Pairwise estimates of Ti:Tv and the pairwise protein differences are reported in Table 3.6 where the overall Ti:Tv ratio was estimated at 32:1.

Table 3.4 Variable sites and the average percentage base composition (A, G, C & T) in a 370 bp fragment of the cytochrome b gene in *Aethomys chrysophilus* and *A. ineptus* from South Africa and Swaziland.

	Variable sites	A	G	C	T
A. chrysophilus & *A. ineptus*:					
First position	6	27.7	27.0	15.2	30.2
Second position	1	17.1	18.0	24.3	40.7
Third position	18	53.0	1.9	24.3	20.8
Overall	25	32.6	15.5	21.3	30.5

Table 3.5 Pairwise estimates of HKY85 + G (0.7287) + I (0.6024) sequence divergence (%) between 21 maternal haplotypes (above the diagonal) and HKY85 nucleotide divergence (% nucleotide substitutions per site) between haplotypes (below the diagonal) within *Aethomys chrysophilus* and *A. ineptus* from South Africa and Swaziland.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	H01		0.28	0.58	0.28	0.57	0.57	0.29	0.28	0.58	1.20	0.88	0.89	0.29	0.58	0.57	0.57	2.66	2.74	3.64	3.16	2.26
2	H02	0.27		0.28	0.28	0.28	0.28	0.57	0.57	0.88	0.87	0.57	0.58	0.57	0.88	0.88	0.28	2.25	3.13	4.10	3.57	1.88
3	H03	0.55	0.28		0.57	0.57	0.57	0.28	0.87	1.21	1.19	0.87	0.87	0.87	1.19	1.19	0.57	2.59	3.52	4.58	3.98	2.20
4	H04	0.27	0.27	0.54		0.57	0.57	0.87	0.87	0.88	1.19	0.87	0.87	0.87	1.19	1.19	0.57	2.59	3.52	4.59	3.99	2.21
5	H05	0.55	0.27	0.54	0.54		0.57	0.87	0.87	1.20	1.19	0.87	0.87	0.87	1.19	1.19	0.57	2.59	3.53	4.59	3.08	2.21
6	H06	0.55	0.27	0.54	0.54	0.54		0.87	0.87	1.20	1.19	0.87	0.87	0.87	1.19	1.19	0.57	2.59	3.52	4.59	3.99	2.21
7	H07	0.28	0.55	0.27	0.82	0.82	0.82		0.57	0.89	1.53	1.19	1.19	0.57	0.87	0.87	0.87	2.99	3.08	4.06	3.52	2.59
8	H08	0.27	0.54	0.82	0.82	0.82	0.82	0.54		0.88	1.53	1.19	1.19	0.57	0.87	0.87	0.87	2.99	3.08	4.02	3.51	2.59
9	H09	0.55	0.82	1.11	0.82	1.10	1.10	0.83	0.82		1.90	1.54	1.55	0.89	1.21	0.57	1.20	3.51	3.63	4.18	4.10	3.06
10	H10	1.10	0.82	1.09	1.09	1.09	1.09	1.37	1.37	1.66		0.87	0.87	1.53	1.89	1.19	0.57	2.99	3.52	4.59	5.02	2.59
11	H11	0.82	0.55	0.82	0.82	0.82	0.82	1.09	1.09	1.38	0.82		0.57	1.19	1.53	0.87	0.28	3.00	4.00	5.20	4.49	2.59
12	H12	0.83	0.55	0.82	0.82	0.82	0.82	1.09	1.09	1.39	0.82	0.54		1.19	1.53	0.87	0.28	2.99	3.99	5.19	4.48	2.59
13	H13	0.28	0.55	0.82	0.82	0.82	0.82	0.54	0.54	0.83	1.37	1.09	1.09		0.28	0.87	0.87	2.99	2.26	3.03	2.66	2.59
14	H14	0.56	0.83	1.09	1.09	1.09	1.09	0.82	0.82	1.11	1.65	1.37	1.37	0.27		1.19	1.19	3.43	2.66	3.53	3.08	3.00
15	H15	0.55	0.82	1.09	1.09	1.09	1.09	0.82	0.82	0.55	1.09	0.82	0.82	0.82	1.09		0.57	3.44	3.54	4.64	4.01	3.01
16	H16	0.55	0.27	0.54	0.54	0.54	0.54	0.82	0.82	1.10	0.54	0.27	0.27	0.82	1.09	0.54		2.59	3.52	4.59	3.99	2.21
17	H17	2.26	1.96	2.21	2.21	2.21	2.21	2.49	2.49	2.83	2.49	2.49	2.49	2.49	2.78	2.78	2.21		1.50	2.49	2.20	0.28
18	H18	2.27	2.54	2.79	2.79	2.79	2.79	2.50	2.50	2.84	2.79	3.07	3.07	1.93	2.21	2.79	2.79	1.37		0.63	0.87	1.17
19	H19	2.86	3.14	3.42	3.42	3.42	3.42	3.12	3.10	3.18	3.42	3.75	3.75	2.47	2.79	3.44	3.42	2.14	0.60		1.68	2.06
20	H20	2.55	2.82	3.07	3.07	2.50	3.07	2.79	2.79	3.13	3.65	3.36	3.36	2.21	2.50	3.07	3.07	1.93	0.82	1.49		1.84
21	H21	1.97	1.68	1.93	1.93	1.93	1.93	2.21	2.21	2.54	2.21	2.21	2.21	2.21	2.49	2.49	1.93	0.27	1.09	1.82	1.65	

Table 3.6 Pairwise estimates of Ti/Tv ratios between 21 maternal haplotypes (above the diagonal) and pairwise protein differences between the different haplotypes (below the diagonal) within *Aethomys chrysophilus and A. ineptus* from South Africa and Swaziland.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	H01		1/0	2/0	1/0	2/0	2/0	1/0	1/0	2/0	4/0	3/0	3/0	1/0	2/0	2/0	2/0	7/1	8/0	9/0	9/0	6/1
2	H02	0		1/0	1/0	1/0	1/0	2/0	2/0	3/0	3/0	2/0	2/0	2/0	3/0	3/0	1/0	6/1	9/0	10/0	10/0	5/1
3	H03	0	0		2/0	2/0	2/0	1/0	3/0	4/0	4/0	3/0	3/0	3/0	4/0	4/0	2/0	7/1	10/0	11/0	11/0	6/1
4	H04	0	0	0		2/0	2/0	3/0	3/0	3/0	4/0	3/0	3/0	3/0	4/0	4/0	2/0	7/1	10/0	11/0	11/0	6/1
5	H05	1	1	1	1		2/0	3/0	3/0	4/0	4/0	3/0	3/0	3/0	4/0	4/0	2/0	7/1	10/0	11/0	9/0	6/1
6	H06	1	1	1	1	2		3/0	3/0	4/0	4/0	3/0	3/0	3/0	4/0	4/0	2/0	7/1	10/0	11/0	11/0	6/1
7	H07	0	0	0	0	1	1		2/0	3/0	5/0	4/0	4/0	2/0	3/0	3/0	3/0	8/1	9/0	10/0	10/0	7/1
8	H08	1	1	1	1	2	2	1		3/0	5/0	4/0	4/0	2/0	3/0	3/0	3/0	8/1	9/0	10/0	10/0	7/1
9	H09	0	0	0	0	1	1	0	1		6/0	5/0	5/0	3/0	4/0	2/0	4/0	9/1	10/0	10/0	11/0	8/1
10	H10	0	0	0	0	1	1	0	1	0		3/0	3/0	5/0	6/0	4/0	2/0	8/1	10/0	11/0	13/0	7/1
11	H11	0	0	0	0	1	1	0	1	0	0		2/0	4/0	5/0	3/0	1/0	8/1	11/0	12/0	12/0	7/1
12	H12	0	0	0	0	1	1	0	1	0	0	0		4/0	5/0	3/0	1/0	8/1	11/0	12/0	12/0	7/1
13	H13	0	0	0	0	1	1	0	1	0	0	0	0		1/0	3/0	3/0	8/1	7/0	8/0	8/0	7/1
14	H14	0	0	0	0	1	1	0	1	0	0	0	0	0		4/0	4/0	9/1	8/0	9/0	9/0	8/1
15	H15	0	0	0	0	1	1	0	1	0	0	0	0	0	0		2/0	9/1	10/0	11/0	11/0	8/1
16	H16	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0		7/1	10/0	11/0	11/0	6/1
17	H17	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0		4/1	6/1	6/1	1/0
18	H18	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0		2/0	3/0	13/1
19	H19	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0		5/0	5/1
20	H20	1	1	1	1	0	2	1	2	1	1	1	1	1	1	1	1	1	1	1		5/1
21	H21	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	

HKY85 sequence divergence estimates based on the gamma shape parameter and proportion of invariable sites were used to summarize relationships among the 21 haplotypes in a neighbour-joining tree (Fig. 3.2). Based on the inclusion of positively identified individuals using karyology, haplotypes H01-H16 represented *A. ineptus* and haplotypes H17-H21 represented *A. chrysophilus*. The average separation between *A. chrysophilus* and *A. ineptus* was 2.72% (range 1.93-3.75%). Pairwise estimates of sequence divergence between haplotypes within *A. ineptus* was generally low, ranging from 0.27-1.66%. More divergence was detected within *A. chrysophilus* with the average sequence divergence estimated at 1.32% (range 0.27-2.14%). Parsimony and maximum likelihood analyses (α-shape parameter = 0.7349; I = 0.6958; Ti:Tv ratio = 29:1; proportion of A = 0.33; C = 0.22; G = 0.15 and T = 0.30) gave the same topology as the neighbour-joining tree with the following parsimony statistics: 13 parsimony informative characters (excluding outgroups); number of trees = 26; tree length = 20; CI = 0.70; RI = 0.86; RC = 0.61. The data were also significantly more structured than random data (g_1 = -1.22, $P < 0.01$; see Hillis & Huelsenbeck 1992). In an attempt to reduce the number of trees generated, reweighting according to the consistency index (CI) yielded 17 trees with the same topology as the neighbour-joining and parsimony trees. In all analyses, there was high bootstrap support for the two genetically distinct lineages associated with *A. ineptus* (lineage A; Fig. 3.2) and *A. chrysophilus* (lineage B; Fig. 3.2).

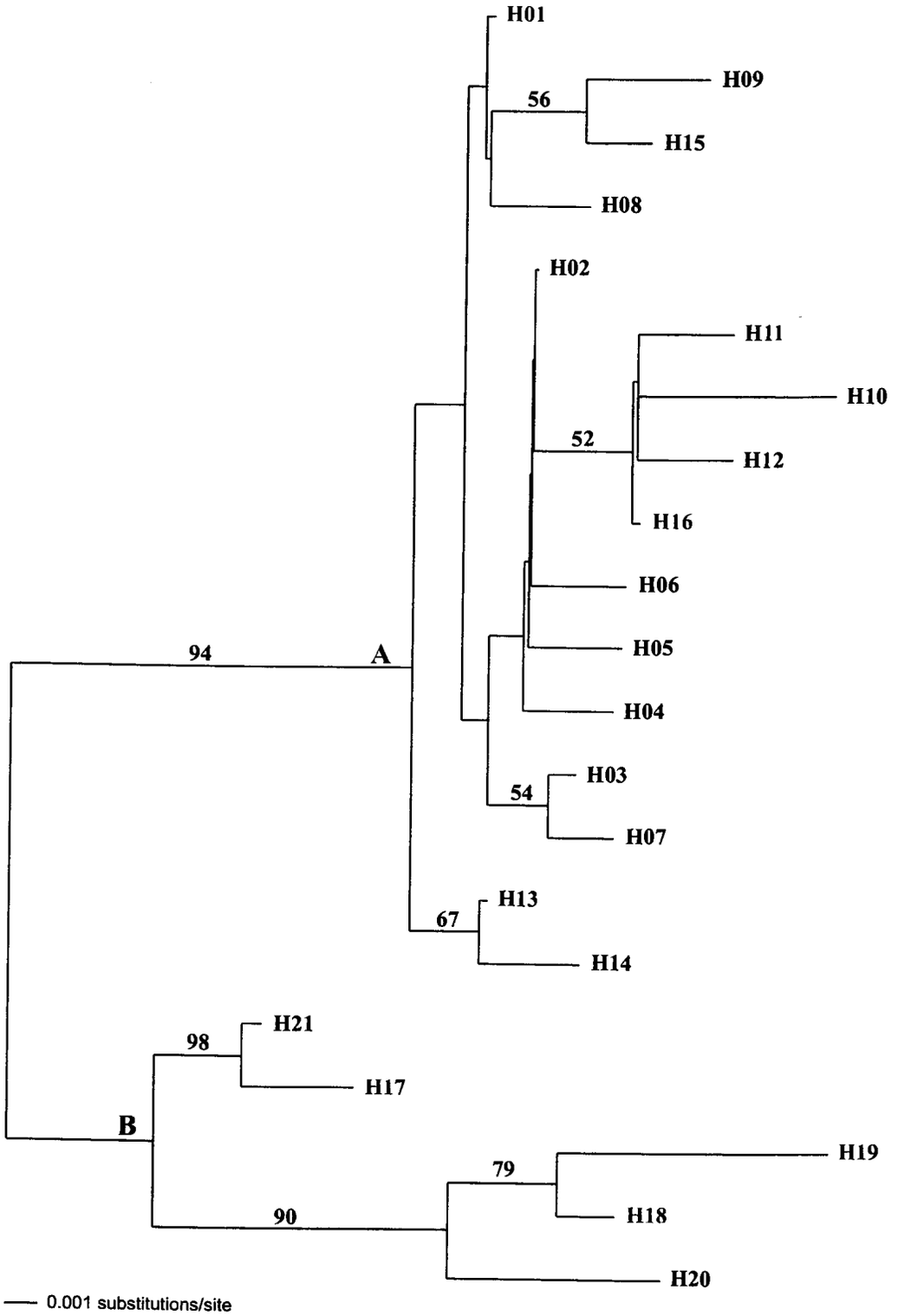


Fig. 3.2 Midpoint rooted neighbour-joining phylogram depicting the clustering of 21 mtDNA haplotypes within *Aethomys chrysophilus* and *A. ineptus* from South Africa and Swaziland into two lineages. Based on reference individuals identified using chromosome number (Linzey *et al.* 2003), lineage A represents *A. ineptus* while lineage B represents *A. chrysophilus*. The values at nodes indicate bootstrap support via rooting against H01. Colours of the different lineages (A & B) correspond to designations used in Figure 3.3. Haplotypes positively identified for *A. chrysophilus* and *A. ineptus* are indicated in green and blue, respectively.

The relative rate test showed no significant differences in the substitution rates between lineages relative to a reference taxon at the adjusted significance level of 0.024% ($P \geq 0.0002$). The same approach as for *A. namaquensis* was followed to date the different lineages (Chapter 2). There are 28 third position transversions between *Mus* and *Rattus* (Bibb *et al.* 1981; Gadaleta *et al.* 1989) in the 370 bp fragment sequenced at the 5' end of the cytochrome b gene in the present analysis. This resulted in a divergence of 22.71% between *Mus* and *Rattus*, corresponding to a rate of change in third position transversions of 1.89% per Myr. Based on this calibration, the divergence time separating *A. ineptus* and *A. chrysophilus* was estimated at 116 000 years ago. This date however may be an underestimation because of the low number of third position transversions detected between *A. chrysophilus* and *A. ineptus* (Table 3.6). When using the conventional substitution rate of 2% per Myr, separation between the two sibling species was estimated at 1.36 Myr ago (range 965 000-1.88), suggesting an early Pleistocene separation. Three third position transversions were detected within *A. chrysophilus* and none within *A. ineptus*. Because of this low number in third position transversions it was also decided to use the 2% per Myr substitution rate to estimate the time of divergence within *A. chrysophilus* and *A. ineptus*. The divergence time within *A. ineptus* was estimated at 400 000 years ago (range 135 000-830 000), whereas the separation within *A. chrysophilus* was estimated at 660 000 years ago (range 135 000-1.07 Myr). Differentiation within *A. ineptus* seems to be relatively more recent than the divergence within *A. chrysophilus* and the differentiation within *A. namaquensis* appears to have occurred earlier than within the two sibling species (Chapter 2). The divergence time separating *Dasymys* and the ingroup taxa was estimated at 8.68 Myr ago (range 8.40-9.17).

Population level analyses

A minimum spanning network supported the two major lineages revealed in the neighbour-joining phylogram (Fig 3.2). These lineages differed by six mutational steps, with no shared haplotypes between them (Fig. 3.3). Lineage A (Fig. 3.3) was characterised by sixteen closely related haplotypes over a wide geographic area with haplotype H01 as the most widespread. This haplotype was shared by 30 individuals from nine different sampling localities covering a geographic range of over 900 km and was identified as a likely ancestral haplotype. Adjacent haplotypes in the minimum spanning network differed by only one or two mutational steps. In contrast, haplotypes in lineage B, on average, differed by one to four mutational steps, with haplotypes from different

localities being more closely related than haplotypes from the same locality (H20 and H21) (Fig. 3.3). Although lineage B (*A. chrysophilus*) was only represented by six individuals covering a small geographic area, more divergence appears to exists within it than within lineage A (*A. ineptus*).

AMOVA and DnaSP analyses were conducted with species combined as well as treated independently. Strong phylogeographic structuring was evident in the combined analysis. Seventy-six percent of the total variance was contained at the regional level indicating the distinctness of the two species. The contribution of the variance among populations within groups was 11.28% while 11.90% of the variance was contained among individuals within populations (Table 3.7). When *A. ineptus* was analysed separately, most of the variance was contained among populations but there was almost as much variance among populations than within populations (Table 3.7). In contrast, when *A. chrysophilus* was analysed separately, most of the variance was contained within rather than among populations, suggesting limited genetic structure, such that populations could not be distinguished from each other.

Both *A. chrysophilus* (1.30%) and *A. ineptus* (0.39%) showed rather low nucleotide diversity values when compared to that of *A. namquensis* (Chapter 2). This is indicative of shallow divergences and in combination with high haplotype diversities (such as in the case of *A. chrysophilus*), it may also be indicative of a population bottleneck followed by rapid population growth (Grant & Bowen 1998; Avise 2000). *Aethomys ineptus* showed a relatively low haplotypic diversity value (0.66), suggesting a low frequency of locality-specific haplotypes (haplotype H01 was shared among nine localities and H02 was shared among three localities). In contrast, *A. chrysophilus* was characterised by a high haplotypic diversity value (0.93), reflecting a high incidence of locality-specific haplotypes at sampling sites. This value approximated those reported for other rodents (Avise *et al.* 1989; Matthee & Robinson 1997; Table 3.7).

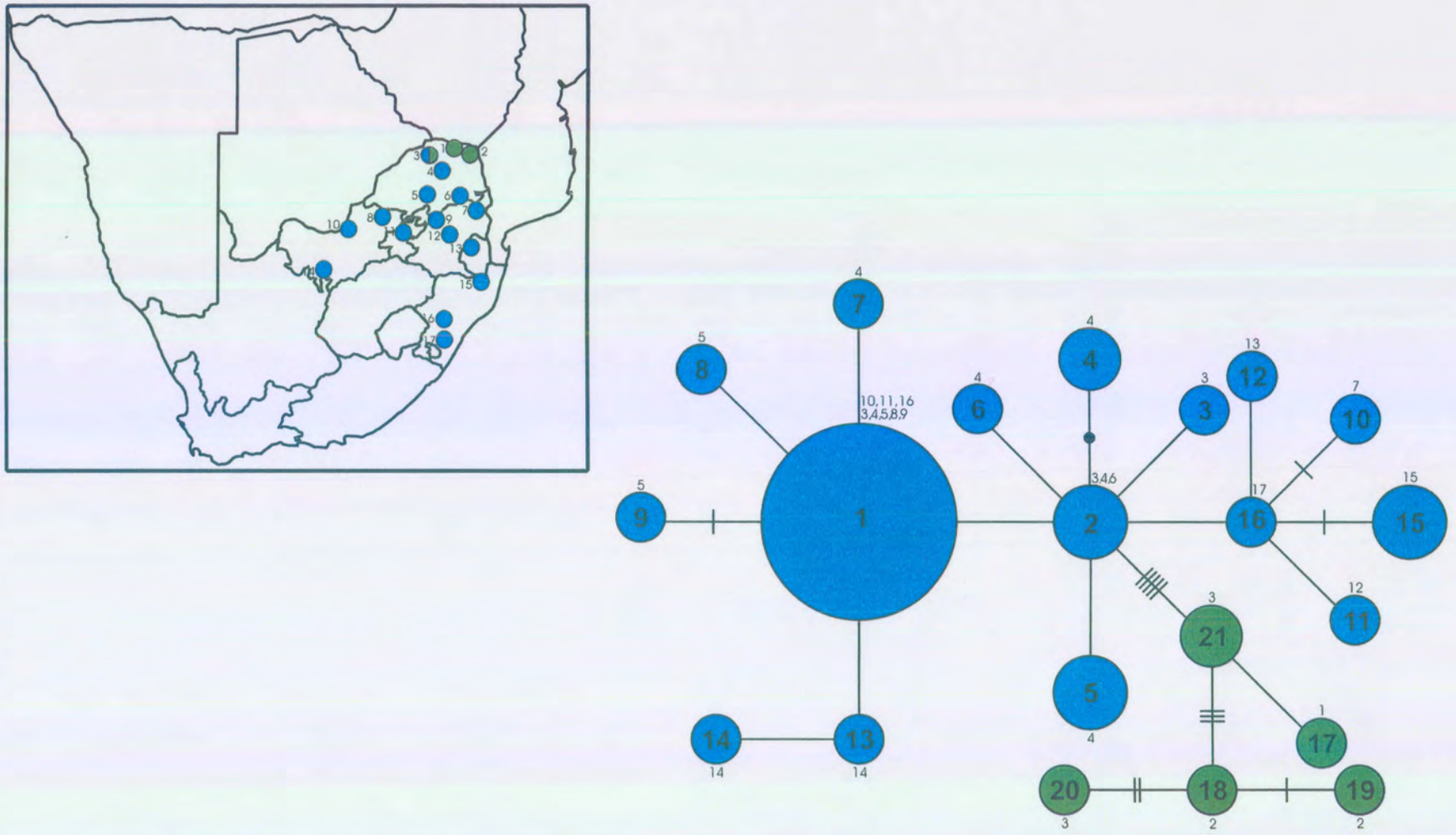


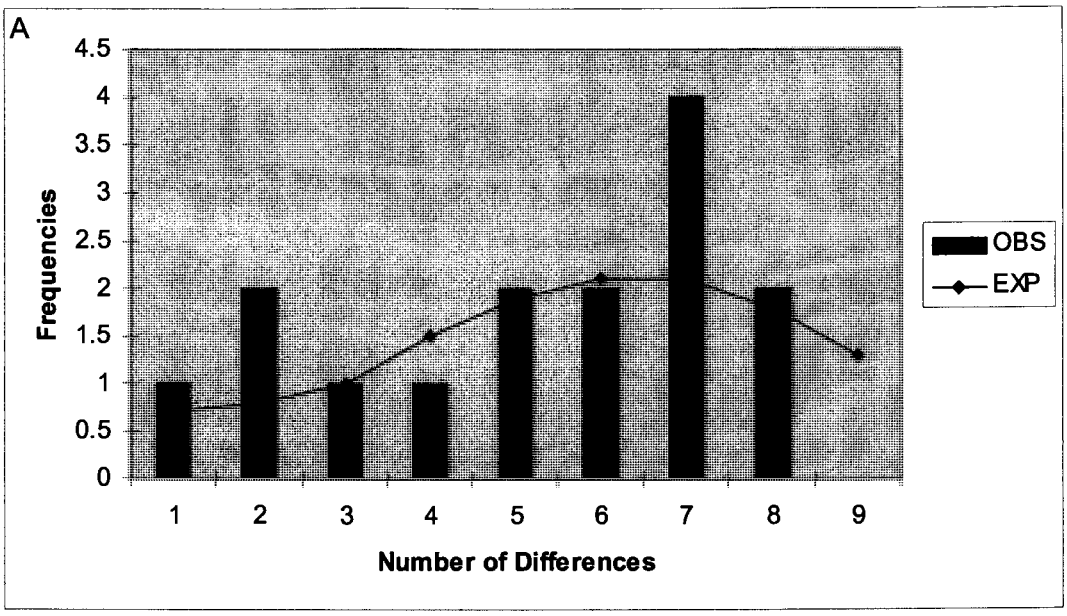
Fig. 3.3 Minimum-spanning network indicating the least number of mutational steps between composite haplotypes. Blue circles indicate the different haplotypes in *Aethomys ineptus* while green circles indicate the haplotypes in *A. chrysophilus* from South Africa and Swaziland. The size of circles represents haplotype frequencies. Horizontal lines connecting haplotypes represent one mutational step and cross-hatching along branches designates additional number of changes detected. The dot (•) indicates identity between H02 and H04 because most of the missing data was excluded in this analysis. Numbers inside circles represent haplotype designations and correspond to those in Table 3.3. Numbers outside circles correspond to locality numbers (see insert for localities) in Table 3.2.

Table 3.7 Estimates of genetic differentiation (nucleotide and haplotypic diversity) and hierarchical analysis of molecular variance (AMOVA) of mtDNA cytochrome b sequences among samples of *Aethomys chrysophilus* and *A. ineptus* from South Africa and Swaziland.

	Overall	*A. chrysophilus*	*A. ineptus*
Number of populations	18	3	15
Sample size (*n*)	58	6	52
DnaSP:			
Nucleotide diversity	-	1.30%	0.39%
Haplotypic diversity	-	0.93	0.66
AMOVA:			
Among groups	76.82%	-	-
Among populations within groups	11.28%	26.17%	53.09%
Within populations	11.90%	73.83%	46.91%
Φ_{ST}	0.88***	0.26	0.53***
Φ_{CT}	0.77***	-	-
Φ_{SC}	0.49***	-	-
V_a	2.85	0.59	0.38
V_b	0.42		
V_c	0.44		

Statistical significance: *** = $P < 0.001$

Aethomys chrysophilus and *A. ineptus* samples were analysed separately in the mismatch distribution analysis resulting in the detection of approximately unimodal distributions (Fig. 3.4A & 3.3B). The distributions were, however, not significantly different from an expected Poisson distribution (*A. chrysophilus*: $P_{Ssd} \geq 0.67$, $P_{Rag} \geq 0.92$; *A. ineptus*: $P_{Ssd} \geq 0.49$, $P_{Rag} \geq 0.68$), therefore, supporting a model of population expansion. These results should, however, be treated with caution as they are based on limited sample sizes, particularly for *A. chrysophilus* where the analysis was based on six individuals only.



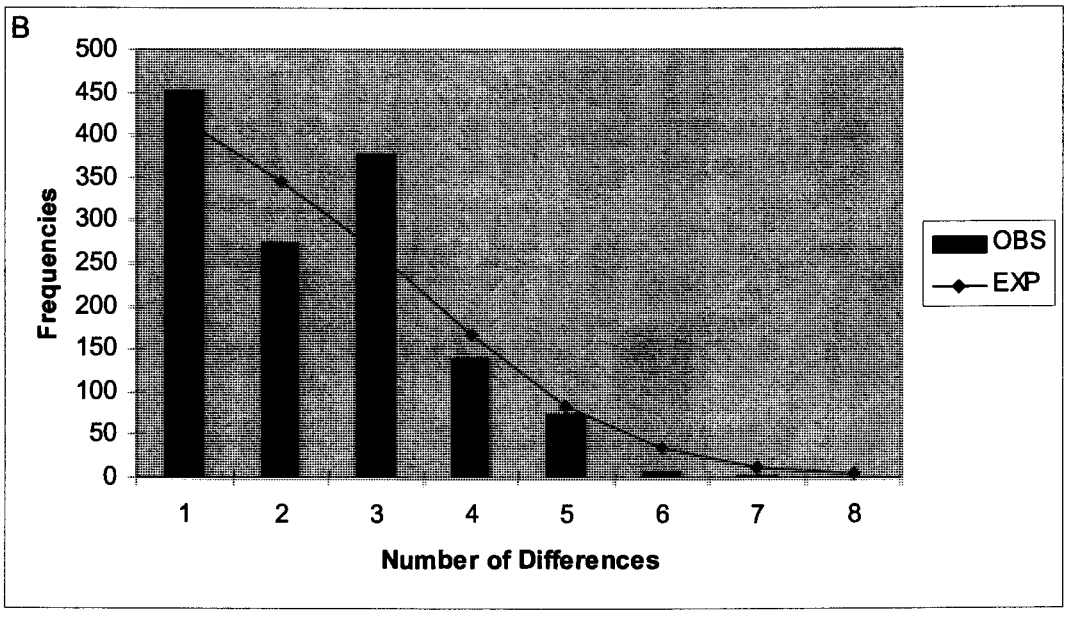


Fig. 3.4 Mismatch frequency distribution of pairwise nucleotide differences in *Aethomys chrysophilus* (A) (Sum of squared deviation = 0.025; $P = 0.67$; Harpending's raggedness index = 0.066; $P = 0.92$) and *A. ineptus* (B) (Sum of squared deviation = 0.012; $P = 0.49$; Harpending's raggedness index = 0.061; $P = 0.68$) from South Africa and Swaziland.

Geographic distribution

The distribution of *A. chrysophilus* and *A. ineptus* in South Africa and Swaziland is indicated in Figure 3.5. Half-degree grid squares represent individuals that were positively identified using either mtDNA analysis and sperm morphology/chromosome number (*A. chrysophilus*: pink, *A. ineptus*: blue), mtDNA analysis (*A. chrysophilus*: purple, *A. ineptus*: yellow) or sperm morphology/chromosome number (*A. chrysophilus*: red, *A. ineptus*: green). Based on the mtDNA analysis, an additional five half-degree grid squares were added to the present verifiable distribution of *A. ineptus*. The other three half-degree grid squares (Langjan Nature Reserve, Orpen gate, Kruger National Park and Percy Fyfe Nature Reserve) identified based on mtDNA were based on previously published data (Linzey *et al.* 2003). The occurrence of *A. chrysophilus* at Langjan Nature Reserve has already been documented by Linzey *et al.* (2003), therefore, this study did not significantly contribute to the current distributional records of this species. The present study, however, confirmed the presence of *A. chrysophilus* in the northern parts of South Africa.

The distributional pattern of *A. chrysophilus* strongly suggested that *A. chrysophilus* primarily occupies the low-lying areas of the Limpopo River drainage system, reaching its southern limit in northern South Africa (Fig. 3.5). In contrast, *A. ineptus* occupies the remainder of South Africa at higher elevations in the north-central parts but extends into lower elevations in the southern parts of its range. Only one sympatric locality (Langjan Nature Reserve, Limpopo Province) was identified. The distributional ranges of *A. chrysophilus* and *A. ineptus* have also been reported to overlap in the north-western part of Pretoria (Gauteng Province) and west of the Soutpansberg (Limpopo Province) (Linzey *et al.* 2003).

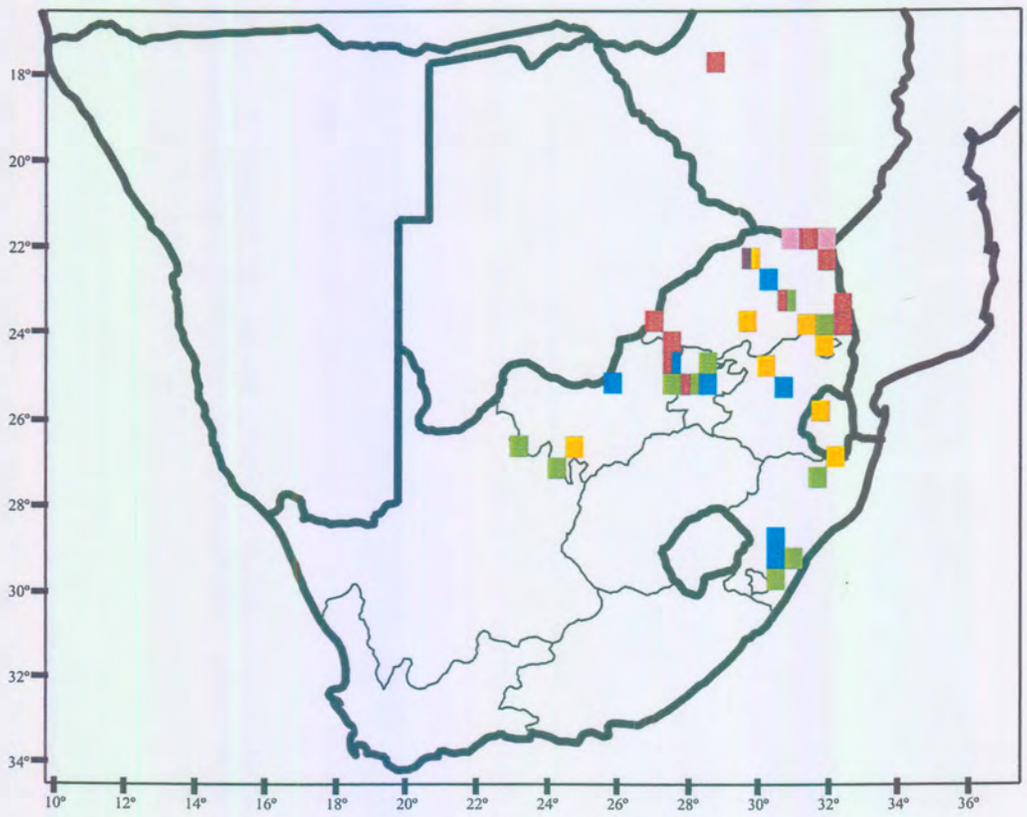


Fig. 3.5 The distribution of *Aethomys ineptus* and *A. chrysophilus* in South Africa and Swaziland indicated by half-degree grid squares. *Aethomys ineptus*: blue squares represent individuals that were identified using both mtDNA analysis and sperm structure/chromosome number, green squares represent individuals that were identified using sperm structure/chromosome number and yellow squares represent individuals that were positively identified using mtDNA analysis. *Aethomys chrysophilus*: pink squares represent individuals that were identified using both mtDNA analysis and sperm structure/chromosome number, red squares represent individuals that were identified using sperm structure/chromosome number and purple squares represent individuals that were positively identified using mtDNA analysis.

Discussion

The mtDNA divergence between *A. chrysophilus* and *A. ineptus* was relatively low (1.93-3.75%) compared to differentiation between, for example, subspecies within *A. namaquensis* (see Chapter 2). Although differentiation within *A. chrysophilus* (0.27-2.14%) and *A. ineptus* (0.27-1.66%) was not much lower than the divergence between them, it was however, possible to distinguish two clear monophyletic lineages. All positively identified specimens of *A. chrysophilus* and *A. ineptus* grouped within lineage B and A, respectively. Therefore, the two lineages were delineated to represent *A. chrysophilus* and *A. ineptus* and were strongly supported by the inclusion of independently positively identified individuals in the mtDNA analysis. The nodes within *A. ineptus* were largely unresolved suggesting that phylogenetic analysis cannot accurately assess shallow intraspecific variation. Although no hybridisation has been detected in areas of sympatry using karyology, this needs to be tested with nuclear genes. The analyses in this study, however, strongly suggested that *A. chrysophilus* and *A. ineptus* can be distinguished based on divergence in the cytochrome b gene. Such molecular distinctions may prove useful for ecological studies between sibling rodent species where animals need not be sacrificed.

Based on morphometric analysis, Chimimba (2000) suggested that *A. chrysophilus* in southern Africa may be polytypic and separated into two subspecies: *A. c. chrysophilus* and *A. c. imago*. The phenetic discontinuity separating the two proposed subspecies broadly coincides with an altitudinal limit (Clark 1967) of either below or above 500 m above sea level in the eastern parts of southern Africa (Chimimba 2000). The few samples collected above and below 500 m above sea level in the present study, however, did not show any separation within *A. chrysophilus* associated with altitudinal differences. It is clear that more samples representative of the distributional range of *A. chrysophilus* in southern Africa are needed to address the question of molecular variation within *A. chrysophilus*.

Despite limited samples, *A. chrysophilus* was characterised by high levels of genetic diversity over a small geographic range. On the one hand, this suggests that there may be more variation within *A. chrysophilus* than detected in the present study and may lead to

the recognition of subspecies. On the other hand, this high level of variation may also be indicative of populations that have been separated for a long time by the Limpopo River as has been detected in the Scrub hare, *Lepus saxatilis* (Kryger 2002). It is possible that the Limpopo River experienced recurrent climatic episodes of drying out alternating with periods of high water levels during the Pleistocene and Holocene oscillations. Scrub hares have a tendency to avoid arid habitats but could have dispersed across the riverbed during drought. The river may have functioned as a permanent/temporary barrier to gene flow between populations south and north of the river resulting in the disruption of localised gene flow (Kryger 2002). It is, therefore, possible that in a period of extended isolation the fragmented populations might have adapted to the drier conditions typical of the north of the Limpopo and to the more humid conditions south of the river. It is possible that this barrier to gene flow may also have contributed to the high levels of variation detected in *A. chrysophilus*.

The mtDNA cytochrome b gene tree (neighbour-joining tree) of *A. ineptus* from South Africa and Swaziland was characterised by a shallow division without any major genetic gaps corresponding to the phylogeographic category IV of Avise *et al.* (1987a) and Avise (2000). Similarly, other more widely distributed species such as the springhare (*Pedetes capensis*) (Matthee & Robinson 1997) and the yellow mongoose (*Cynictis penicillata*) (Jansen van Vuuren & Robinson 1997) have also been characterised by weak geographic structuring. In the present study, shallow geographic structuring was evident from haplotype H01 that occurred throughout the distributional range of *A. ineptus* at nine of the 15 localities. This was the geographically most widely sampled haplotype found within the genus *Aethomys*. Haplotype H02 was also common and was found in three localities. It has been suggested that geographic populations of species exhibiting this category of intraspecific phylogeographic structure have had recent historical interconnections through gene flow (Avise *et al.* 1987a). This would require the absence of long standing zoogeographic barriers to movement and a life history conducive to dispersal (Avise *et al.* 1987a). Not much is known about the life history of *A. ineptus* but it is evident from other rodents that these animals can move distances of up to 400 km over long periods of time (Kim *et al.* 1998). It has also been reported that during 'outbreak' years some rodent species exhibit rapid maturation of young and minimal intervals between litters, which results in rapid population growth (Pearson 1975). This may influence large-scale

movements to new locations in search of suitable feeding grounds (Jaarola & Tegelström 1995).

Aethomys ineptus seems to be a habitat generalist and not only confined to rocky habitats, as is the case with *A. namaquensis* (Skinner & Smithers 1990; Linzey & Kesner 1997). They occur in both grassland, with some scrub cover, and savanna woodland (Skinner & Smithers 1990). The divergence time separating *A. ineptus* and *A. chrysophilus* was, in this study, estimated at 1.42 (range 1.02-1.98) Myr ago. This time of divergence suggests an early Pleistocene separation, which coincides with the glacial cycles of the last 1.7 million years. Glacial cycles (17 oscillations during the past 1.7 million years and individual cycles with a mean duration of about 100 000 years) have altered species distributions significantly (Avise 1989). Global temperature fluctuations profoundly influenced southern African climate, which in turn altered the vegetation (Van Zinderen Bakker 1978). Grassland is associated with a combination of winter frost and fire. As temperatures dropped in the past and the area of winter cold expanded, it caused a simultaneous expansion of grassland at the expense of other vegetation types (Brain 1985). There are indications from Florisbad (Van Zinderen Bakker 1957) and Aliwal North (Coetzee 1978) that grassland replaced other forms of vegetation during the last glacial maximum. The mismatch distribution analysis suggested evidence of an expansion in *A. ineptus*. Since the species is associated with grassland, it may be possible that the species distributional range expanded (Fig. 3.4B) as grassland vegetation became available.

The geographic differentiation observed in several southern African rock-dwelling mammals such as the rock hyrax (Prinsloo & Robinson 1992), the red rock rabbits (Matthee & Robinson 1996) and the Namaqua rock mouse (see Chapter 2), was not evident in *A. ineptus*. *Aethomys ineptus* is much younger than the rock hyrax, the red rock rabbits and the Namaqua rock mouse, which may account for the different pattern. Furthermore, it is possible that differences in habitat preferences may account for the lack of complementarity. The rock hyrax, red rock rabbits and the Namaqua rock mouse are restricted to rocky habitats, whereas *A. ineptus* is not only confined to rocky habitats but occupies habitats associated with grassland and savanna woodland.

Morphometric data suggested that *A. chrysophilus* and *A. ineptus* are broadly sympatric in southern Africa (Chimimba 2001b). Based on molecular and cytogenetic analyses (Linzey

et al. 2003 and the present study), very few localities have, however, been confirmed where the two species occur in sympatry. Subsequently, distributions based on positively identified specimens found *A. ineptus* rather than *A. chrysophilus* to be widespread in South Africa (Linzey *et al.* 2003). The hypothesized distributional patterns suggested by Linzey *et al.* (2003) were also evident from the present study (Fig. 3.5). There has been no previous confirmation of the identity of the *Aethomys* caught in Swaziland but it is evident from the present study that *A. ineptus* is present (Fig. 3.5). In addition, the present study confirmed a further five half degree grid squares where *A. ineptus* occurs within South Africa. Despite limited sampling, Linzey *et al.* (2003) and the present study strongly suggest that the range of *A. chrysophilus* may reach its southern limits in northern South Africa, following the Limpopo River Valley.

The Limpopo River basin represents an area characterised by low elevations and low rainfall, suggesting that *A. chrysophilus* is adapted to habitats characteristic of warmer climates while *A. ineptus* prefers cooler climates (Linzey *et al.* 2003). This proposed distribution of *A. chrysophilus* is associated with the Savanna biome, specifically with the distribution of Mopane (*Colephospermum mopane*) and Baobab (*Adansonia digitata*) trees (Low & Rebelo 1996). In contrast, *A. ineptus* occurs at low elevations in the southern portions of its distribution, but is predominantly associated with the cooler high elevation grassland in the north (Fig. 3.5). This ecological differentiation is also shown by a variety of species that have distributional patterns similar to that of *A. chrysophilus* and *A. ineptus*. It has been suggested that the Limpopo basin acts as a barrier to dispersal in *Myosorex varius* (Forest shrew) and *Heterohyrax brucei granti* (Yellowspotted hyrax), both species with a preference for humid, high-lying areas (Meester 1958; Bothma 1966), a pattern similar to that of *A. ineptus*. Other examples of mammals with southern ranges that reach their northern limit in the central high elevation grasslands include *Otomys irroratus* (Vlei rat), and *Pelea capreolus* (Grey rhebok) (Skinner & Smithers 1990). The Scrub hare also has a tendency to avoid arid habitats and is replaced in such areas by the Cape hare, *Lepus capensis* (Kryger 2002). Examples of mammals with northern ranges that extend southward along the Limpopo River Valley include bats (e.g., Butterfly bat, *Chalinolobis variegates*; Schlieffen's bat, *Nycticeius schlieffenii*) and ungulates (e.g., Waterbuck, *Kobus ellipsiprymnus*; Warthog, *Phacochoerus aethiopicus*) (Skinner & Smithers 1990).

Chapter 4

Phylogenetic relationships within the genus *Aethomys* from southern African

Abstract

Evolutionary relationships among species within *Aethomys* are poorly understood. Phylogenetic analyses of partial mitochondrial cytochrome b and 16S rRNA sequences were used to examine the relationships among four species of *Aethomys* from southern Africa and other African murines. The molecular data suggested that *Aethomys* from southern Africa comprises two monophyletic lineages, one including *A. namaquensis* and *A. granti* and the other lineage comprising two cryptic species, *A. chrysophilus* and *A. ineptus*. This study, however, suggested that *Aethomys* may be paraphyletic, such that the taxonomic status of *A. namaquensis* and *A. granti* within the subgenus *Micaelamys* needs further investigation.

Introduction

African rock rats of the genus *Aethomys* are a diverse group of murid rodents endemic to East, Central and southern Africa. Phylogenetic relationships between the genus and other African murids are uncertain (Musser & Carleton 1993). Characteristics of the genus overlap to some extent with *Rattus* and *Arvicanthis* (Ellerman 1941) and the genus has also variously been considered to be closely related to *Mus* De Winton, 1897, *Mastomys*, *Thallomys*, *Zelotomys* (De Graaff 1981), *Stochomys*, *Dephomys*, *Dasymys* and *Pelomys* (Bonhomme *et al.* 1985; Denys 1990a, b).

Little is generally known about evolutionary relationships among the 11 currently recognised species of *Aethomys*, namely *A. bocagei*, *A. chrysophilus*, *A. hindei*, *A. kaiseri*, *A. ineptus*, *A. granti*, *A. namaquensis*, *A. nyikae*, *A. silindensis*, *A. stannarius* and *A. thomasi* (Musser & Carleton 1993; Chimimba 1997). Previous postulations of evolutionary relationships within the genus have largely been based on previous taxonomic treatments of the genus. For example, studies in the 1980s revealed that *A. chrysophilus* comprised two distinct cytotypes: 2n = 44 and 2n = 50, that also differed in gross sperm and bacular morphology (Gordon & Rauthenbach 1980; Gordon & Watson 1986; Visser & Robinson 1986; Baker *et al.* 1988; Breed *et al.* 1988). The 2n = 50 cytotype was recognised as *A. chrysophilus* while the 2n = 44 cytotype was designated as *A. ineptus* (Chimimba 1997, 1998; Chimimba *et al.* 1999), suggesting that these cryptic species may be phylogenetically closely related (Chimimba 2002; Chapter 3).

Fossils representing *A. namaquensis* and *A. chrysophilus* (*senso lato*) have been described from South Africa (De Graaff 1960, 1961; Avery 1981, 1982, 1985; Hendey 1981; Pocock 1987). Recently, two fossil species, a small-sized *A. modernis* and a large-sized *A. adamanticola*, the oldest known representatives of the genus in Africa, have been reported from Langebaanweg, Western Cape Province, South Africa (Denys 1990a, b).

Aethomys modernis is very similar to extant *A. chrysophilus* (*senso lato*), while *A. adamanticola* is different from any other known *Aethomys*, but shows characteristics reminiscent of *A. namaquensis* and *A. hindei* (Denys 1990a, b). Denys (1990a, b) suggested that this species may represent an advanced stage of an early Miocene lineage

closely related to *Dasymys*. Other fossil records include two East African Plio-Pleistocene species, *A. lavocati* (Jaeger 1976, 1979) from Lake Natron and *A. deheinzelini* (Wesselman 1984) from Lake Turkana (Black & Krishtalka 1986; Denys 1987). There is, however, no close relationship between these East African species and those from South Africa to allow speculation on the origin and time of divergence of the genus (Denys 1990a, b).

A phylogenetic study of the genus *Aethomys* using cladistic analysis of all currently recognised species, based on cranial data suggested the presence of three clades: 1) *A. bocagei*, *A. thomasi*, *A. silindensis*, *A. kaiseri* and *A. nyikae*; 2) *A. chrysophilus*, *A. ineptus* and *A. hindei*; 3) *A. granti*, *A. namaquensis* and *A. stannarius* (Chimimba 1997). This study also revealed a sister-taxon relationship between the two cryptic species, *A. chrysophilus* and *A. ineptus* (Chimimba 1997).

Attention was drawn to the close evolutionary relationship between *A. namaquensis* and *A. granti*. Similarly, this close relationship has been documented in other studies using dental morphology (Ellerman *et al.* 1953), karyology (Visser & Robinson 1986), gross sperm and bacular morphology and their staining properties (Visser & Robinson 1987) and phenetic analysis (Chimimba *et al.* 1999). *Aethomys namaquensis* and *A. granti* were separated from all other species of *Aethomys*, providing good grounds for the subgeneric separation of *Aethomys* and *Micaelamys* (Chimimba 1997). Similarly, Davis (1975) supported the view of a close affinity between *A. namaquensis* and *A. granti* and their inclusion in the subgenus *Micaelamys* (Ellerman *et al.* 1953).

Davis (1975) subdivided the subgenus *Aethomys* into "*kaiseri*" and "*hindei*" groups. The former included *A. kaiseri*, while the "*hindei*" group included *A. hindei*, *A. bocagei*, *A. chrysophilus* (*sensu lato*), *A. thomasi*, *A. nyikae* and *A. silindensis*, excluding *A. stannarius* that was then considered a subspecies of *A. hindei* (Ellerman 1941; Rosevear 1969; Davis 1975; Hutterer & Joger 1982). In contrast, this subdivision of the subgenus *Aethomys* into "*kaiseri*" and "*hindei*" groups was not supported by cladistic analyses (Chimimba 1997).

In the present study, in addition to morphological data, molecular data was used in the elucidation of phylogenetic relationships within southern African *Aethomys*. A fragment of the ribosomal gene, 16S rRNA and the protein coding cytochrome b gene, were used to infer the phylogeny of *Aethomys*. This approach was adopted because 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 that this mitochondrial ribosomal gene is useful for phylogenetic studies of rodents (Allard & Honeycutt 1992; Sullivan *et al.* 1995; Frye & Hedges 1995). Similarly, cytochrome b sequences when used carefully, have contributed to the investigation of the phylogenetic relationships among murids at the subfamily level (Verheyen *et al.* 1995; Verheyen *et al.* 1996; Ducroz *et al.* 2001). In combining these genes, it was considered that the cytochrome b gene with its faster mutation rate would possibly resolve relationships near the terminal branches relative to the slower 16S rRNA characters which, based on other investigations, could be predicted to be distributed more basally (Moritz *et al.* 1987; Halanych & Robinson 1999).

The present study, therefore, represents the first phylogenetic analysis using a total evidence approach (cytochrome b, 16S rRNA and morphological characters). Although the morphological study was pan African in context, the present study is based on material from southern Africa only for which data was available and addressed the following questions:

1. What are the phylogenetic relationships among the southern African *Aethomys* species based on molecular data and how do these compare with previously postulated relationships based on morphological data?

2. Can these postulated relationships be useful in resolving the status of the traditional subgeneric separation of *Aethomys* and *Micaelamys*?

Materials and Methods

Study area and sampling

Samples of *A. namaquensis*, *A. chrysophilus*, *A. ineptus* and *A. granti* were obtained from localities in South Africa and Swaziland (Fig. 4.1, Table 4.1). Tissue samples were preserved according to the methods described in Chapters 2 and 3 and voucher specimens

deposited in the Transvaal Museum, Pretoria. In some of the analyses, GenBank sequences from other murines were also included.

DNA extraction, PCR amplification, cycle sequencing and sequencing

Total genomic DNA was prepared by proteinase K-phenol/chloroform extraction (Sambrook *et al.* 1989) followed by ethanol precipitation (Chapter 2 & 3). Fragments of the protein coding cytochrome b and the ribosomal 16S rRNA genes, from the mitochondrial genome were used to infer the phylogeny of southern African *Aethomys*. The same primer combinations as described in Chapters 2 and 3 were used to amplify the cytochrome b gene of 16 individuals. Individuals of *A. granti* were sequenced with the same primer combinations used for *A. namaquensis*. To amplify and sequence the targeted fragment in the 16S rRNA gene, the primers 16S ar-L and 16S br-H (Palumbi *et al.* 1991) were used. The primer names, sequences and annealing positions on the mouse mtDNA genome for the 16S rRNA are given in Table 4.2 while those for the cytochrome b gene are given in Chapters 2 and 3.

Polymerase Chain Reaction (PCR; Saiki *et al.* 1988) amplifications of the cytochrome b gene were performed as described in Chapters 2 and 3. The PCR reactions of the 16S rRNA gene consisted of 35 cycles at 94°C for 45 seconds, 55°C for 1 minute and 72°C for 1 minute. PCR products were purified using the High Pure™ PCR Product Purification Kit (Boehringer Mannheim) as prescribed by the manufacturers.

Dye-terminator cycle sequencing was performed as described in Chapters 2 and 3 using the shortened L14724 or L14841 in combination with the internal Mus IH or Nam IH primers. Individuals were also sequenced with the 16S ar-L and 16S br-H to obtain a 596 bp fragment of the 16S rRNA gene. Nucleotide sequences were determined using either an ABI 377 automated sequencer (PE Applied Biosystems) or an ABI 3100 sequencer (PE Applied Biosystems).

Table 4.1 Geographic coordinates for all collection localities of samples of *Aethomys namaquensis*, *A. granti*, *A. chrysophilus* and *A. ineptus* in South Africa and Swaziland analysed in the present study. Numbers 1 – 11 correspond to the numbers indicated in Figure 4.1. Numbers in parenthesis indicate sample sizes collected per locality.

Locality and sample size	Country	Province	Geographic coordinates
A. namaquensis			
1. Wathaba-Uitkomst, Machadodorp (2)	South Africa	Mpumalanga	25 47'27"S 30 22'31"E
2. Farm: Rietfontein, Springbok (1)	South Africa	Northern Cape	29 51'40"S 18 11'10"E
3. Gariep Nature Reserve (1)	South Africa	Free State	30 35'56"S 25 32'03"E
A. granti			
4. Farm: Rietpoort, Loxton (2)	South Africa	Northern Cape	31 38'30"S 22 22'34"E
5. Farm: Droëkloof, Calvinia (2)	South Africa	Northern Cape	31 55'30"S 20 05'32"E
A. chrysophilus			
6. Pafuri, Kruger National Park (2)	South Africa	Limpopo	22 25'49"S 31 10'25"E
A. ineptus			
7. Vaalkop Dam Nature Reserve (1)	South Africa	North West	25 19'25"S 27 25'52"E
8. Wathaba-Uitkomst, Machadodorp (1)	South Africa	Mpumalanga	25 47'27"S 30 22'31"E
9. Malolotja Nature Reserve (1)	Swaziland		27 04'22"S 24 46'07"E
10. Farm: Koedoesberg, Pongola (2)	South Africa	Kwa-Zulu Natal	27 26'31"S 31 41'41"E
11. Ashburton (1)	South Africa	Kwa-Zulu Natal	29 38'56"S 30 27'16"E

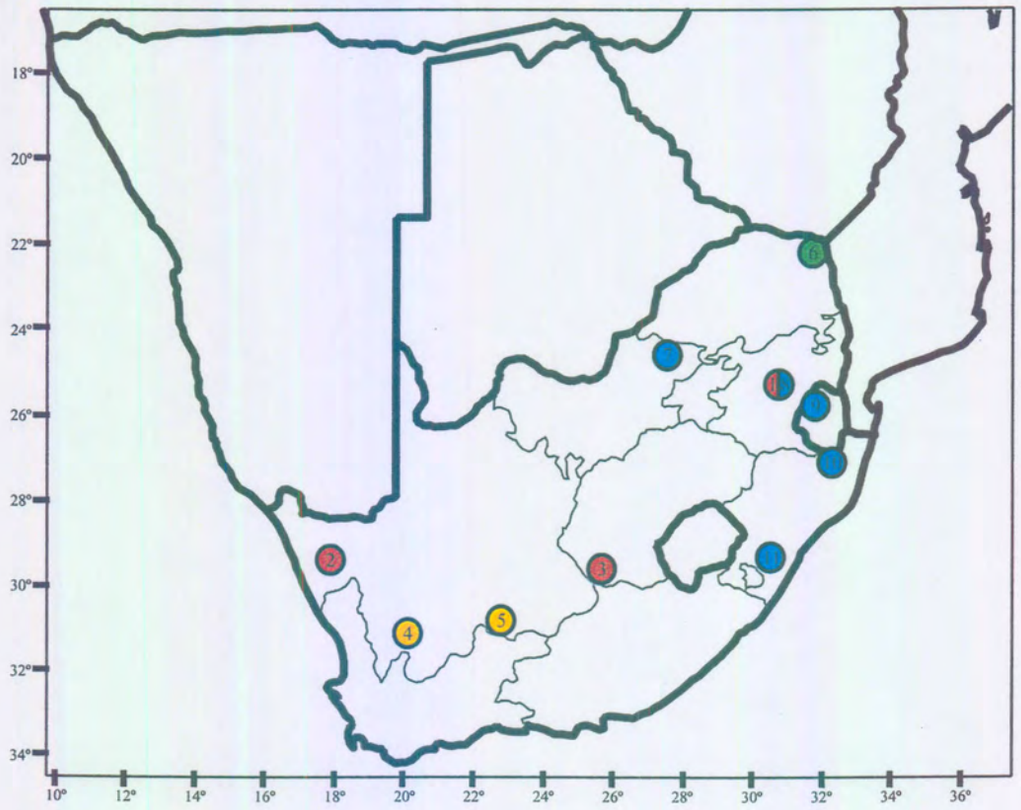
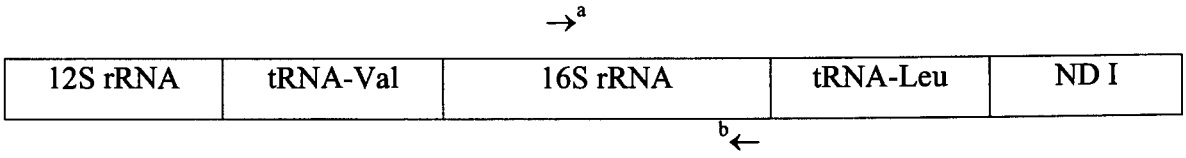


Fig. 4.1 Collection localities of *Aethomys namaquensis* (red), *A. granti* (yellow), *A. chrysophilus* (green) and *A. ineptus* (blue) in South Africa and Swaziland. Numbers in circles refer to the locality numbers in Table 4.1. 1 – Wathaba-Uitkomst, Machadodorp; 2 – Farm: Rietfontein, Springbok; 3 – Gariep Nature Reserve; 4 – Farm: Calvinia; 5 – Farm: Rietpoort, Loxton; 6 – Pafuri, Kruger National Park; 7 – Vaalkop Dam Nature Reserve; 8 – Wathaba-Uitkomst, Machadodorp; 9 – Malolotja Nature Reserve, Swaziland; 10 – Farm: Koedoesberg, Pongola; 11 – Ashburton.

Table 4.2 Primer names, sequences (5' to 3') and positions. Rodent specific primer names refer to the position of the 3' nucleotide position of the mitochondrial DNA sequence of *Mus musculus* (Bibb *et al.* 1981). Superscripts a and b correspond to the letters in the figure.

Primer name	Rodent specific name	Primer sequence
16S ar-L[a]	L1960	5' CGCCTGTTTATCAAAAACAT 3'
16S br-H[b]	H2575	5'CCGGTCTGAACTCAGATCACGT 3'



Sequencing analysis

The individual cytochrome b and 16S rRNA sequences were imported into Sequence Navigator version 1.01 (PE Applied Biosystems). A consensus sequence of each individual was computed by aligning the forward and reverse sequences for the different genes. Consensus sequences of all individuals were aligned in Clustal X (Thompson *et al.* 1997) and sequences subsequently imported into PAUP version 4.0b10 (Swofford 2002) for phylogenetic analyses and MacClade version 3.0 (Maddison & Maddison 1992) to translate the cytochrome b 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 & Maddison 1992).

Outgroup choice

The suitability of *D. incomtus* as an outgroup to *Aethomys* has been supported by previous morphological (Ellerman 1941), palaeontological (Denys 1990a, b) and molecular

(Verheyen *et al.* 1995) studies. *Dasymys incomtus* was therefore used as an outgroup and for comparison with a morphological phylogeny obtained by Chimimba (1997). *Gerbillus nigeriae* was used as an outgroup in the phylogenetic analyses that included other murids.

Phylogenetic analysis

Quantitative estimates of nucleotide divergence among all individuals were evaluated by the Kimura two-parameter distance (Kimura 1980) as well as the number of Ti and Tv using PAUP version 4.0b10. Pairwise protein differences were also estimated for the cytochrome b gene. Some of the phylogenetic analyses in this study also included 22 qualitative morphological data obtained from a previous phylogenetic study of the genus based on qualitative cranial and dental morphology (Chimimba 1997).

Since the data included sequences from two mtDNA genes and morphological characters (Appendix 4.1), 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 data sets separately and to construct a consensus tree from these separate analyses, but this approach (Adams 1972) is considered to be more appropriate if the data sets are heterogeneous. The second approach was to perform the analysis directly using all combined data. This "total evidence" approach is considered to often provide phylogenies that are more resolved than consensus trees from separately analysed data (De Queiroz 1993). The "total evidence" approach was followed in this study and the data were analysed in using the following data combination: cytochrome b; cytochrome b + 16S rRNA and cytochrome b + 16S rRNA + morphological data. The congruence between these data combinations was assessed by a partition homogeneity test as implemented in PAUP version 4.0b10 using 1000 search replicates (Farris *et al.* 1995). In some of the analyses, samples of other murines that have been postulated to be closely related to *Aethomys* were also included. This included the African *Dasymys* and *Arvicanthis*, several Eurasian murines (*Mus* and *Rattus*), as well as representatives of the murine subfamilies Acomyinae, Gerbillinae, Otomyinae and Arvicolinae.

All distance-based analyses were performed on the different haplotype sequences identified using pairwise estimates of the nucleotide sequence divergence derived from the Kimura two-parameter model of substitution. Distance trees were constructed using the

neighbour-joining algorithm of Saitou and Nei (1987) available in PAUP version 4.0b10. Tree nodal support was assessed by 1000 bootstrap replicates (Felsenstein 1985) with outgroup rooting.

To avoid the confounding effects of saturated characters (Hackett 1996; Halanych *et al.* 1999) and to optimise the number of informative characters, the observed numbers of pairwise sequence differences with reference to the time of divergence were plotted against the number of pairwise transitions and transversions at different codon positions. Depending on the data, saturated characters were excluded from the maximum parsimony analyses (Kluge & Farris 1969; Farris *et al.* 1970). An assessment of the saturation in the 16S rRNA gene was not considered in this study because Ducroz *et al.* (2001) detected no saturated characters in this gene.

Maximum parsimony (Kluge & Farris 1969; Farris *et al.* 1970) analyses were performed on all haplotypes identified. Parsimony trees were generated using nucleotides as unordered characters and 500 heuristic searches using the tree bisection-reconnection (TBR) method were used to construct the shortest tree. If more than one minimum length tree were obtained, a strict consensus tree was constructed. Tree-length, consistency index (CI; Kluge & Farris 1969), retention index (RI; Farris 1989), and rescaled consistency index (RC; Archie 1989; Meier *et al.* 1991) were derived for tree descriptions. Phylogenetic signal was assessed within the different data sets by evaluating the skewness of the tree-length distributions (g_1 statistic), generated from 1000 random trees (Hillis & Huelsenbeck 1992). Different weighting schemes were applied to reduce the number of minimum length trees by reweighting with reference to the CI and RI (Farris 1969). Support values for internal nodes were determined using bootstrap analysis with 1000 iterations (Felsenstein 1985).

The Relative-Rate test with a tree (RRTree; Robinson *et al.* 1998)) was used to assess differences in third position transversion rates in 489 bp of the cytochrome b gene and differences in 526 bp of the 16S rRNA gene among separate lineages relative to an outgroup (Robinson *et al.* 1998; Robinson-Rechavi & Huchon 2000). *Gerbillus nigeriae* was used as a reference taxon in both analyses. For dating of divergence times, sequence data from *Rattus rattus* and *Mus musculus*, with a time of divergence estimated at 12 Myr based on fossil records, was used as a calibration point (Jacobs & Downs 1994) for both

the cytochrome b and 16S rRNA data (Chapter 2). The number of third position transversions was used as a measure of genetic divergence because it accumulates nearly linearly with time (Irwin *et al.* 1991). The conventional rate of 2% sequence divergence per million years was also applied to the data (Brown *et al.* 1979).

Results

Sequence statistics

A total of 489 bp of the 5' end of the cytochrome b gene and 526 bp of the 3' end of the 16S rRNA gene were sequenced in all individuals examined. All sequences obtained were of the mitochondrial cytochrome b gene since no stop codons were found when sequences were translated into amino acids and the nucleotide sequences corresponded to the expected 163 amino acids. The 110 variable positions in the cytochrome b gene defined 13 haplotypes, while the 64 variable positions in the 16S rRNA gene defined nine haplotypes (Tables 4.3 & 4.4; Appendices 4.2 & 4.3). Of the 110 variable sites in cytochrome b (Table 4.5A), 101 were phylogenetically informative. Of the 64 variable positions in 16S rRNA (Table 4.5B), 61 were phylogenetically informative. Eighteen informative sites were recorded from first codon positions, two from second codon positions and 81 from third positions of the cytochrome b gene. Most of the substitutions were silent with eight variable amino acid sites between *A. namaquensis* and *A. granti* and between *A. chrysophilus* and *A. ineptus*. Four variable amino acids were detected between *A. namaquensis* and *A. granti* and none between *A. chrysophilus* and *A. ineptus*.

Mean base compositions at the three different codon positions of cytochrome b gene and for the 16S rRNA fragment for the whole sample examined are provided in Table 4.5A and 4.5B. The overall base composition of cytochrome b showed that the four nucleotides do not occur in equal frequencies, and is similar to that of other mammalian cytochrome b sequences reported in the literature (Irwin *et al.* 1991; Ducroz *et al.* 1998). This strong bias in base composition showed an under-representation of guanine at both second (18.0%) and third (1.7%) codon positions. Similar to the sequence statistics in intraspecific analyses (Chapters 2 & 3), a higher representation of adenine at the third position (50.0%) and thymine at the second position (39.0%) was also observed in the

present sequences. The 16S rRNA fragment showed an overrepresentation of adenine, as previously described in other mammalian mitochondrial ribosomal genes (Nedbal *et al.* 1996).

Nucleotide divergence estimates were reported as Kimura two-parameter distances for the cytochrome b and 16S rRNA genes. Pairwise estimates of the nucleotide sequence divergence for the cytochrome b and 16S rRNA genes are indicated in Tables 4.6 and 4.7, respectively. The availability of sequences (cytochrome b or 16S rRNA) on GenBank determined the number of individuals that were included in the pairwise Kimura two-parameter sequence divergence tables. Corrected cytochrome b Kimura two-parameter sequence divergence values ranged between 0.24 and 16.56% within *Aethomys* and sequence divergence values ranged between 0.20 and 12.33% for the 16S rRNA gene. Sequence divergence values in this range suggest a separation at the genus level. This may support the previous allocation of *A. namaquensis* to the genus *Thallomys* (Ellerman 1941). For both genes, more divergence was detected between some of the species of *Aethomys* than between these species and some of the outgroups. Pairwise estimates of Ti:Tv ratios for both genes and pairwise protein differences for the cytochrome b gene are reported in Tables 4.8 and 4.9. The availability of sequences (cytochrome b or 16S rRNA) on GenBank determined the number of individuals that were included in the pairwise Ti:Tv ratio tables.

Table 4.3 Variable sites of 13 mitochondrial DNA cytochrome b haplotypes (489 base pairs) of *Aethomys namaquensis* (NH16, NH29, NH30), *A. granti* (GH01-GH03), *A. ineptus* (H01, H11, H12, H15, H16) and *A. chrysophilus* (H18, H19) from South Africa and Swaziland. Variable positions six and 489 correspond to position 14178 and 14661 of *Mus musculus* (Bibb *et al.* 1981). Dots (.) indicate identity to the base in the reference sequence NH16 and question marks (?) indicate undetermined bases. Numbers following populations correspond to population numbers in Figure 4.1.

Haplotype No.	Variable positions	No. of individuals	Populations
	1111111111111111111111111122222222222222222222223333333333333333333444444444444444444 1112333345667777888999000011222333444567778889990001112234445555777889901112222334445566889111222333444456788 6258103468169256824814623581236925845668147369259124036251379235803958459259124506258173647615703625813459510 9		
NH16	TCCACCAGCCATCATCTACTAGCTTATCTTGCCCCACTCTCCTCTTCCTAGATACTGACCATAACTACACTTCCGATCTTTATCGTTACATACCATCCCCATCTT?????	1	Gariep Nature Reserve (3)
NH29	.T..T..A..G.T.C.....A.....A.....C.....G.....A.....CG.....A.....C.A....T.....TG....CGCTC	2	Wathaba-Uitkomst, Machadodorp (1)
NH30	.T..T.....G.T.C.....A.....A.....C.....G.....A.....CG.....A.....C.A....T.....TG....CGCTC	1	Farm: Rietfontein, Springbok (2)
GH01	..T.TA...T....ATA..CTAACA.A.CCA..T..AC.CT.CTCCTT...C..T.AC..C....C.TCT.C.TA...C.CTCTACG.....TT.A.TTT.A.A.TTTTA	1	Farm: Droëkloof, Calvinia (4)
GH02	C.T.TA...T....ATA..CCAA.A.A.CCA..T..AC.CT.CTCCTT...TCGT.AT..C....C.TCT.C.TA...C.CTCTACG.T...TT.A.TTT.A.A.TTTCA	1	Farm: Droëkloof, Calvinia (4)
GH03	..T.TA...T.C..ATA..CTAACA.A.CCA..T..AC.CT.CTCCTT...C..T.AC..C....C.TCT.C.TA...C.CTCTACG.....TT.A.TTT.A.A.TTTTA	2	Farm: Rietpoort, Loxton (5)
H01	CT.C..C.A..C.CA.AGT...ACACATCCATTAAGACTCT...A.TA..AC...CA..T.CCTTA.TT.CCT.AGGTAC.CCTACATT.CCT.TAT.TTCATCCCTCCT	1	Vaalkop Dam Nature Reserve (7)
H11	CT.C..C.A..C.CA.AGT...ACACATCCATTAAGACTCT...A.TG..AC...CA.TT.CCTTA.TT.CCT.AGGTAC.CCTACATT.CTT.TAT.TTCATCC?????	1	Wathaba-Uitkomst, Machadodorp (8)
H12	CT.C..C.A..C.CA.AGT...ACACA.CCATTAAGACTCT...A.TG..AC...CA.TT.CCTTA.TT.CCT.AGGTAC.CCTACATT.CCT.TAT.TTCATCCCTCCT	1	Malolotja Nature Reserve, Swaziland (9)
H15	CT.C..C.A..C.CA.AGT...ACACATCCATTAA.ACTCT...A.TA..AC...CA.TT.CCTTA.TT.CCT.AGGTAC.CCTACATT.CCT.TAT.TTCATCCCTCCT	2	Farm: Koedoesberg, Pongola (10)
H16	????..C.A..C.CA.AGT...ACACATCCATTAAGACTCT...A.TG..AC...CA.TT.CCTTA.TT.CCT.AGGTAC.CCTACATT.CCT.TAT.TTCATCCCTCCT	1	Ashburton (11)
H18	CT.C..C.A..C.CA.AG....ACACATCCATTAAGACTCTT..A.TAC.AC..TCA..T.CCTTA.TT.CCT.AGATA..CCTACACT.CC..TAT.TTCATCCCTCCT	1	Pafuri, Kruger National Park (6)
H19	CT.C..C.A..C.CA.AG....ACACATCCAT.AAGACTCTT..A.TAC.AC..TCA..T.CCTTA.TT.CCT..GATA..CCTACACT.CC..TAT.TTCATCCCT???	1	Pafuri, Kruger National Park (6)

Table 4.4 Variable sites of nine mitochondrial DNA 16SrRNA haplotypes (526 base pairs) of *Aethomys namaquensis* (NH16, NH29, NH30), *A. granti* (GH01, GH03), *A. ineptus* (H01, H16) and *A. chrysophilus* (H18, H19) from South Africa and Swaziland. Variable positions 14 and 455 correspond to position 1978 and 2419 of *Mus musculus* (Bibb *et al.* 1981). Dots (.) indicate identity to the base in the reference sequence NH16 and dashes (-) indicate alignment gaps. Population numbers correspond to those in Figure 4.1.

Haplotype No.	Variable positions	No. of individuals	Populations (Sample size, Population No.)
	11111111112222222222222222222222222223333333333333333333444 11350011234888011122233344455556666677778801233467777777789992 45 45122849241134437802323723404561235706792796504760123456935694 85		
NH16	TTTCATAATTTAACGGAACGATTCCTTGAATTATTATTCTTCTATCCCAATTTTAAACTCTCCT	1	Gariep Nature Reserve (1,3)
NH29	.....C.......T...G.....A..A......A.................C....TAA.T..	2	Wathaba-Uitkomst, Machadodorp (2,1)
NH30	.....C.......T...G.....-..A......A.............G...C....TAA.T..	1	Farm: Rietfontein, Springbok (1,2)
GH01	..........C..T.A..TA..C.T.AA..CA.A....T.AACT.T....C.AC...AGT..T.	2	Farm: Droëkloof, Calvinia (2,4)
GH03	..........C..T.A..TA..C.T.AA..CA.A....T.AACT.TT...C.AC...AGT..T.	2	Farm: Rietpoort, Loxton (2,5)
H01	CCATTCGGACCTTAACT..AGCAATAAAGT.ACGA.CCTAA.C.CT.TGT.AACTTTTATATTC	5	Vaalkop Dam Nature Reserve (1,7), Wathaba-Uitkomst, Machadodorp (1,8), Malolotja Nature Reserve, Swaziland(1,9) Farm: Koedoesberg, Pongola (2,10)
H16	CCATTCGGACCTTAACT..AGCA-TAAAGT.ACGA.CCTAA.C.CT.TGT.AA.TTTTATATTC	1	Ashburton (11)
H18	CCATTCGGACCTTAACT..AG-CATAAA.C.ACGA.CCTAATC.CT.TG..AACTTTTATATT.	1	Pafuri, Kruger National Park (1,6)
H19	CCATTCGGACCTTAACT..AGCA-TAAA.C.ACGATCCTAA.C.CT.TG..AACTTTTATATT.	1	Pafuri, Kruger National Park (1,6)

Table 4.5 Variable sites and the average percentage base composition (A, G, C & T) in a 489 bp fragment of the cytochrome b gene (A) and 16S rRNA (B) within *Aethomys namaquensis, A. granti, A. chrysophilus* and *A. ineptus* from South Africa and Swaziland.

A

Cytochrome b:	Variable sites	A	G	C	T
First position	19	29.0	25.1	14.3	31.0
Second position	2	18.1	18.0	25.0	39.3
Third position	89	50.0	1.7	28.0	20.2
Overall	110	32.0	15.0	23.2	30.1

B

16S rRNA:	Variable sites	A	G	C	T
Overall	64	0.36	0.21	0.19	0.25

Table 4.6 Pairwise estimates of Kimura two-parameter sequence divergence (%) between 13 cytochrome b haplotypes of *Aethomys* from South Africa and Swaziland and other African murid rodents obtained from GenBank. Haplotypes NH16, NH29, NH30 represent *A. namaquensis*, haplotypes GH01-GH03 represents *A. granti*, haplotypes H01, H11-H16 represent *A. ineptus* and haplotypes H18 and H19 represent *A. chrysophilus* in the present study.

		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
1	NH16																													
2	NH29	4.47																												
3	NH30	4.22	0.20																											
4	*Aethomys namaquensis*	2.76	3.60	3.38																										
5	GH01	13.67	13.06	12.81	13.59																									
6	GH02	14.55	14.09	13.83	14.63	1.88																								
7	GH03	13.94	13.32	13.06	13.33	0.20	2.09																							
8	H01	19.00	18.13	17.86	17.88	16.12	16.39	15.85																						
9	H11	19.54	19.24	18.94	18.66	16.63	17.27	16.34	0.66																					
10	H12	19.02	18.13	17.86	17.88	16.12	16.39	15.85	0.62	0.43																				
11	H15	18.97	18.13	17.86	17.88	16.12	16.39	15.85	0.41	0.67	0.62																			
12	H16	19.47	18.75	18.46	18.52	16.13	16.70	15.84	0.43	0.23	0.21	0.43																		
13	H18	18.97	18.13	17.86	17.88	16.12	16.39	15.85	1.66	2.47	2.30	2.09	2.19																	
14	H19	18.36	18.49	18.20	17.38	16.08	16.66	15.80	2.18	2.95	2.86	2.63	2.76	0.43																
15	*Aethomys sp.*	18.62	18.68	16.40	18.43	15.32	15.59	15.06	5.37	6.08	5.60	5.37	5.61	5.60	6.10															
16	*Dasymys incomtus*	17.34	17.06	16.79	17.62	18.15	18.43	18.43	15.53	16.51	15.80	15.53	16.02	15.26	15.51	14.19														
17	*Rattus rattus*	21.40	21.07	20.78	19.95	20.54	20.82	20.26	17.33	17.60	17.33	17.60	17.97	17.33	17.09	17.60	17.02													
18	*Mus musculus*	19.80	20.26	19.97	18.88	18.59	19.42	18.86	15.53	15.06	15.27	15.01	14.92	14.75	14.97	16.06	15.00	18.06												
19	*Otomys sloggetti*	19.66	18.47	18.75	18.49	16.50	15.97	16.76	15.31	15.60	15.31	15.31	15.49	16.10	16.96	15.04	16.08	17.76	16.99											
20	*Otomys irroratus*	18.68	17.37	17.10	16.85	16.50	17.84	16.23	14.01	14.69	14.01	14.01	14.41	13.75	14.71	12.74	16.31	16.22	18.85	8.83										
21	*Arvicanthis niloticus*	18.78	19.50	19.22	18.40	18.43	19.26	18.15	16.51	16.99	16.78	16.51	17.05	15.70	15.65	16.73	16.17	19.20	18.17	17.98	16.08									
22	*Arvicanthis abyssinicus*	18.22	18.40	18.13	18.43	17.62	18.17	17.90	16.76	16.39	16.48	16.76	16.73	15.94	15.91	17.26	13.81	19.18	17.88	16.60	17.69	9.82								
23	*Arvicanthis somalicus*	19.37	18.96	19.24	19.54	16.01	17.35	16.28	16.76	17.22	17.31	16.76	17.34	15.94	15.91	16.98	16.64	19.42	19.26	16.87	16.33	9.08	10.26							
24	*Lemniscomys bellieri*	20.03	19.22	18.94	18.13	17.37	17.92	17.64	15.16	15.24	15.16	15.16	15.38	14.90	14.31	15.97	16.44	18.96	17.98	16.81	16.25	16.12	14.78	16.91						
25	*Acomys cahirinus*	23.58	23.35	23.05	22.74	23.95	23.05	23.65	20.58	21.18	20.58	20.58	20.91	20.58	20.46	20.58	21.82	19.09	23.45	19.43	19.47	23.53	25.00	25.36	22.89					
26	*Microtus arvalis*	24.89	24.42	24.12	23.80	23.45	23.45	23.16	21.58	21.18	21.29	21.01	21.74	21.86	22.11	21.29	20.46	20.12	22.13	19.94	20.78	24.38	22.78	23.13	24.68	21.93				
27	*Meriones unguiculatus*	23.18	23.23	23.53	22.92	20.94	20.38	20.66	24.83	25.24	25.14	24.83	24.82	23.92	24.00	23.03	22.71	27.22	23.90	23.60	24.10	24.50	25.28	23.82	27.87	25.67	20.56			
28	*Tatera kempi*	25.93	23.86	24.15	25.33	21.32	22.47	21.61	23.21	24.16	23.21	23.22	23.54	22.92	23.19	22.92	23.44	26.59	23.99	24.11	25.65	24.27	25.10	25.11	22.34	28.45	28.41	24.02		
29	*Gerbillus nigeriae*	23.99	22.71	23.01	22.98	20.96	21.52	21.24	18.35	19.24	18.35	17.81	18.16	19.18	19.29	19.45	19.09	25.38	19.95	20.72	21.86	23.53	21.49	20.71	23.53	25.14	25.36	19.42	21.19	

Table 4.7 Pairwise estimates of Kimura two-parameter sequence divergence (%) between nine 16Sr RNA haplotypes of *Aethomys* from South Africa and Swaziland and other African murid rodents obtained from GenBank. Haplotypes NH16, NH29, NH30 represent *A. namaquensis*, haplotypes GH01 and GH03 represent *A. granti*, haplotypes H01 and H16 represent *A. ineptus* and haplotypes H18 and H19 represent *A. chrysophilus* in the present study.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	NH16																	
2	NH29	2.20																
3	NH30	2.21	0.20															
4	*Aethomys namaquensis*	1.99	1.99	2.01														
5	GH01	5.12	5.35	5.36	5.19													
6	GH03	5.34	5.57	5.58	5.44	0.20												
7	H01	12.01	10.88	10.65	11.33	9.94	10.17											
8	H16	11.55	10.65	10.40	10.82	9.94	10.17	0.20										
9	H18	11.34	10.21	9.98	11.11	8.82	9.05	1.20	1.40									
10	H19	11.39	10.47	10.23	11.08	9.31	9.54	1.00	1.20	0.60								
11	*Aethomys chrysophilus*	12.33	11.05	10.82	11.34	10.99	11.25	2.23	2.00	2.21	1.98							
12	*Dasymys incomtus*	9.62	8.38	8.16	8.12	7.81	8.06	7.88	8.35	7.61	7.85	8.10						
13	*Rattus rattus*	14.92	14.20	13.96	14.26	13.45	13.69	16.17	15.95	15.70	15.79	17.03	15.57					
14	*Mus musculus*	12.47	11.54	11.10	12.85	11.06	11.29	12.76	12.76	12.29	12.13	13.11	11.58	15.14				
15	*Arvicanthis niloticus*	9.59	9.59	9.37	8.12	8.53	8.78	9.35	9.35	9.11	9.34	9.38	5.49	12.42	10.55			
16	*Arvicanthis abyssinicus*	9.86	10.36	10.13	9.36	9.29	9.54	9.35	9.35	9.10	9.08	8.63	6.19	14.00	11.07	1.55		
17	*Gerbillus nigeriae*	22.97	22.69	23.06	23.12	21.12	20.81	20.16	20.71	20.11	20.10	20.80	18.83	22.74	22.93	19.97	19.40	

Table 4.8 Pairwise estimates of transition:transversion ratio (Ti:Tv) between 13 cytochrome b haplotypes of *Aethomys* from South African and Swaziland and other African murid rodents (above the diagonal) and pairwise protein differences between the different haplotypes and other African murid rodents (below the diagonal) obtained from GenBank. Haplotypes NH16, NH29, NH30 represent *A. namaquensis*, haplotypes GH01-GH03 represent *A. granti*, haplotypes H01, H11-H16 represent *A. ineptus* and haplotypes H18 and H19 represent *A. chrysophilus* in the present study.

		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
1	NH16		18/1	17/1	12/0	39/17	42/17	40/17	45/30	46/30	45/30	45/30	44/29	45/30	43/30	44/30	41/28	48/35	43/34	50/27	47/27	43/31	43/29	46/30	49/29	49/41	49/46	45/45	49/49	49/43
2	NH29	2		1/0	16/1	38/20	42/20	39/20	46/32	45/31	46/32	46/32	46/31	46/32	44/32	48/32	43/31	50/39	49/37	50/29	46/29	50/33	47/32	49/32	49/33	56/41	55/46	51/46	50/49	52/43
3	NH30	1	1		15/1	37/20	41/20	38/20	45/32	44/31	45/32	45/32	45/31	45/32	43/32	47/32	42/31	50/39	48/37	51/29	45/29	49/33	46/32	50/32	48/33	55/41	54/46	52/46	51/49	53/43
4	*Aethomys namaquensis*	0	2	1		41/29	45/19	4019	46/31	44/30	46/31	46/31	46/30	46/31	41/31	48/31	46/30	50/38	45/36	51/28	45/28	47/32	48/31	52/31	46/32	53/42	52/47	49/47	54/50	52/44
5	GH01	4	5	4	4		9/0	1/0	46/24	44/23	46/24	46/24	44/23	46/24	44/23	43/24	47/31	50/37	43/37	39/33	39/33	48/31	46/30	40/30	46/29	58/41	48/50	43/46	45/45	44/45
6	GH02	4	5	4	4	0		10/0	47/24	46/23	47/24	4724	46/23	47/24	46/23	44/24	48/31	50/37	46/37	37/33	44/33	51/31	48/30	45/30	48/29	55/41	48/50	41/46	49/45	46/45
7	GH03	4	5	4	4	0	0		45/24	43/23	45/24	45/24	43/23	45/24	43/23	42/24	48/31	50/37	44/37	40/33	38/33	47/31	47/30	41/30	47/29	57/41	47/50	42/46	46/45	45/45
8	H01	7	9	8	7	8	8	8		3/0	3/0	2/0	2/0	8/0	10/0	23/2	52/15	50/31	41/27	42/25	37/25	52/19	52/20	52/20	47/19	52/35	49/42	60/42	57/39	44/35
9	H11	7	9	8	7	8	8	8	0		2/0	3/0	1/0	11/0	13/0	24/2	51/14	50/27	37/24	40/23	37/23	50/17	48/17	50/18	44/17	47/35	43/40	57/39	53/39	43/33
10	H12	7	9	8	7	8	8	8	0	0		3/0	1/0	11/0	13/0	24/2	53/15	50/31	40/27	42/25	37/25	53/19	51/20	54/20	47/19	52/35	48/42	61/42	57/39	44/35
11	H15	7	9	8	7	8	8	8	0	0	0		2/0	10/0	12/0	23/2	52/15	50/31	39/27	42/25	37/25	52/19	52/20	52/20	47/19	52/35	47/42	80/42	57/39	42/35
12	H16	7	9	8	7	8	8	8	0	0	0	0		10/0	12/0	23/2	51/15	50/31	37/26	41/24	37/24	52/18	50/19	52/19	48/18	52/33	47/41	58/40	55/38	42/33
13	H18	7	9	8	7	8	8	8	0	0	0	0	0		2/0	24/2	51/15	50/31	38/27	45/25	36/25	49/19	49/20	49/20	46/19	52/35	50/42	57/42	56/39	47/35
14	H19	7	9	8	7	8	8	8	0	0	0	0	0	0		25/2	49/15	50/30	36/27	45/25	37/25	46/19	47/19	46/20	42/18	48/35	47/42	55/40	53/39	44/35
15	*Aethomys sp.*	6	6	7	6	7	7	7	3	3	3	3	2	3	3		47/15	50/31	43/27	41/25	32/25	51/21	52/22	51/22	50/19	52/35	48/42	54/42	56/39	48/35
16	*Desymys incomtus*	7	7	6	7	7	7	7	8	8	8	8	8	8	8	8		50/34	36/28	44/26	43/28	48/22	40/21	47/25	49/22	48/44	46/41	52/43	55/42	42/40
17	*Rattus rattus*	9	11	10	9	12	12	12	9	9	9	9	9	9	9	10	10		42/36	39/38	37/34	48/34	47/35	46/37	49/32	42/40	41/45	54/57	58/50	47/58
18	*Mus musculus*	10	11	11	10	12	12	12	7	7	7	7	7	7	7	9	7	7		38/36	43/38	48/30	46/31	51/31	51/26	48/50	50/43	56/43	55/44	47/38
19	*Otomys sloggetti*	9	8	9	9	9	9	9	9	9	9	9	9	9	9	8	11	11	13		34/6	51/26	45/27	46/27	43/30	47/36	46/39	55/43	48/52	46/42
20	*Otomys irroratus*	6	7	6	6	7	7	7	7	7	7	7	7	7	7	6	9	8	11	5		44/26	49/27	44/27	39/32	49/34	49/39	53/47	55/50	50/42
21	*Arvicanthis niloticus*	9	10	9	9	9	9	9	6	6	6	6	6	6	6	7	7	11	9	10	7		39/5	36/5	46/24	52/46	63/37	58/43	56/44	52/46
22	*Arvicanthis abyssinicus*	8	9	8	8	8	8	8	6	6	6	6	6	6	6	7	8	13	11	8	7	6		38/8	40/25	56/47	55/40	56/48	56/47	44/47
23	*Arvicanthis somalicus*	9	8	9	9	9	9	9	8	8	8	8	8	8	8	8	8	11	9	8	7	6	8		48/25	59/45	58/38	53/46	56/47	45/43
24	*Lemniscomys bellieri*	11	11	10	11	11	11	11	9	9	9	9	9	9	9	10	8	11	8	10	11	11	10	12		59/36	64/37	67/45	55/38	52/46
25	*Acomys cahirinus*	15	15	14	15	16	16	16	16	16	16	16	15	16	16	17	15	11	15	16	15	18	20	18	15		53/39	60/45	71/42	61/42
26	*Microtus arvalis*	14	16	15	14	16	16	16	13	13	13	13	13	13	13	15	14	12	13	14	15	14	14	15	13	17		51/36	70/43	59/45
27	*Meriones unguiculatus*	18	16	17	18	17	17	17	19	19	19	19	16	19	19	17	18	22	20	17	18	20	19	18	19	23	24		65/33	57/25
28	*Tatera kempi*	27	25	26	27	26	26	26	26	26	26	26	26	26	26	26	25	28	24	25	27	26	27	25	23	30	26	20		51/38
29	*Gerbillus nigeriae*	17	15	16	17	17	17	17	18	18	18	18	17	18	18	17	17	19	17	17	17	17	18	15	18	22	21	8	17	

Table 4.9 Pairwise estimates of transition:transversion ratio (Ti:Tv) between 13 16Sr RNA haplotypes of *Aethomys* from South Africa and Swaziland and other African murid rodents obtained from GenBank. Haplotypes NH16, NH29, NH30 represent *A. namaquensis*, haplotypes GH01 and GH03 represent *A. granti*, haplotypes H01 and H16 represent *A. ineptus* and haplotypes H18 and H19 represent *A. chrysophilus* in the present study.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	NH16																	
2	NH29	6/5																
3	NH30	7/4	1/0															
4	*Aethomys namaquensis*	6/3	7/2	8/1														
5	GH01	16/9	18/8	19/7	17/6													
6	GH03	17/9	19/8	20/7	18/6	1/0												
7	H01	29/27	27/24	26/24	24/24	23/24	24/24											
8	H16	28/26	26/24	25/24	23/23	24/23	25/23	1/0										
9	H18	28/25	26/22	25/22	25/22	20/22	21/22	4/2	5/2									
10	H19	27/26	25/24	24/24	24/23	21/23	22/23	3/2	4/2	1/2								
11	*Aethomys chrysophilus*	24/28	22/25	21/25	23/25	22/25	23/25	7/3	6/3	5/5	4/5							
12	*Dasymys incomtus*	22/19	21/15	20/15	19/16	16/18	17/18	15/19	16/20	13/20	12/22	14/21						
13	*Rattus rattus*	28/39	26/38	26/37	22/36	23/38	24/38	28/44	28/43	28/42	27/43	25/43	23/40					
14	*Mus musculus*	29/29	24/30	23/29	24/30	22/30	23/30	28/31	28/31	27/30	25/31	24/31	20/29	22/46				
15	*Arvicanthis niloticus*	19/22	20/21	19/21	14/21	14/23	15/23	18/22	18/22	17/22	16/24	17/23	10/14	16/35	17/28			
16	*Arvicanthis abyssinicus*	21/21	22/22	21/22	18/22	16/24	17/24	17/23	17/23	16/23	14/25	13/24	12/15	20/37	20/27	6/1		
17	*Gerbillus nigeriae*	44/46	44/45	45/45	45/45	41/43	40/43	42/38	43/39	41/39	41/39	44/38	40/35	38/50	41/49	41/38	40/37	

Phylogenetic analyses

The results of saturation analysis are summarised in Figure 4.2. First and second position changes accumulated slowly. The rate of transversional substitutions was lower than that of transitions (Ti:Tv ratio = 2:1). Transversions at all positions and transitions at the first and second codon positions were not saturated. Transitions at third codon positions showed saturation and were, therefore, removed from subsequent analyses to reduce the weight of this type of substitution. Transitional saturation became more apparent at distances above approximately 15%, a value consistent with the pattern observed in other mammalian taxa (Irwin *et al.* 1991; Smith & Patton 1993).

Kimura two-parameter sequence divergence estimates were used to summarize relationships among the 13 and nine *Aethomys* haplotypes for the cytochrome b or 16S rRNA genes, respectively, using neighbour-joining trees (Figs. 4.3A and 4.4A). Data for other African murid taxa were also included in these analyses. Haplotypes NH16, NH29 and NH30 represented *A. namaquensis*, haplotypes GH01, GH02 and GH03 represented *A. granti*, haplotypes H18 and H19 represented *A. chrysophilus* and haplotypes H01, H11 and H12, H15 and H16 represented *A. ineptus*. The average separation between *A. namaquensis* and *A. granti* was estimated at 13.66% (range 12.81-14.55%) while the separation between *A. namaquensis* and *A. chrysophilus/A. ineptus* was estimated at 18.42% (range 17.38-19.54%). The estimated average separation between *A. granti* and *A. chrysophilus/A. ineptus* was 16.12% (range 15.06-17.27%). The separation between *A. chrysophilus* and *A. ineptus* varied from 1.66 to 2.86%.

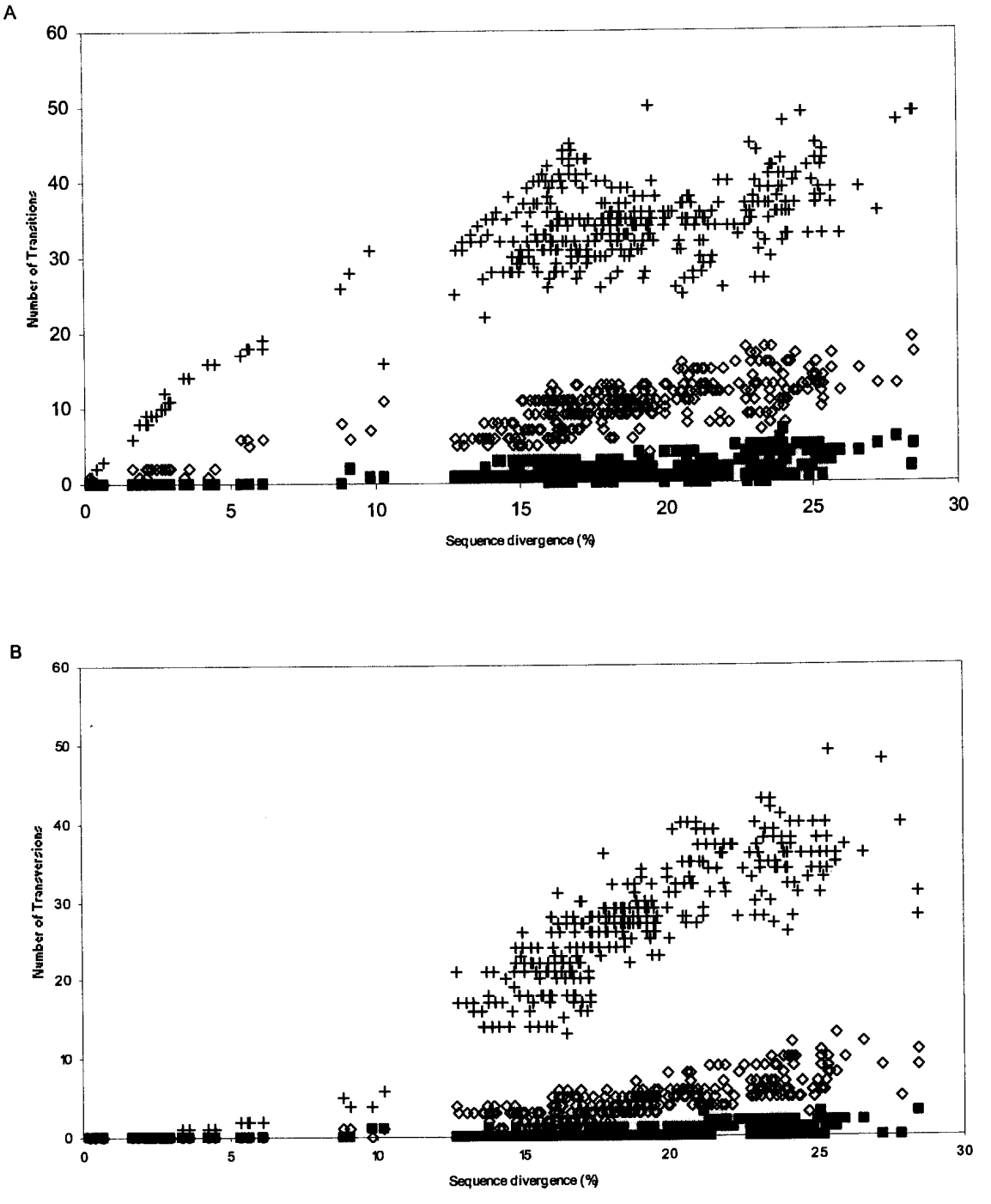


Fig. 4.2 The number of transitions (A) and transversions (B) plotted against cytochrome b Kimura two-parameter sequence divergence (%) estimates between 29 taxa of the rodent family *Muridae* from the present study and GenBank. (◊) First codon positions, (■) second codon positions and (+) third codon positions.

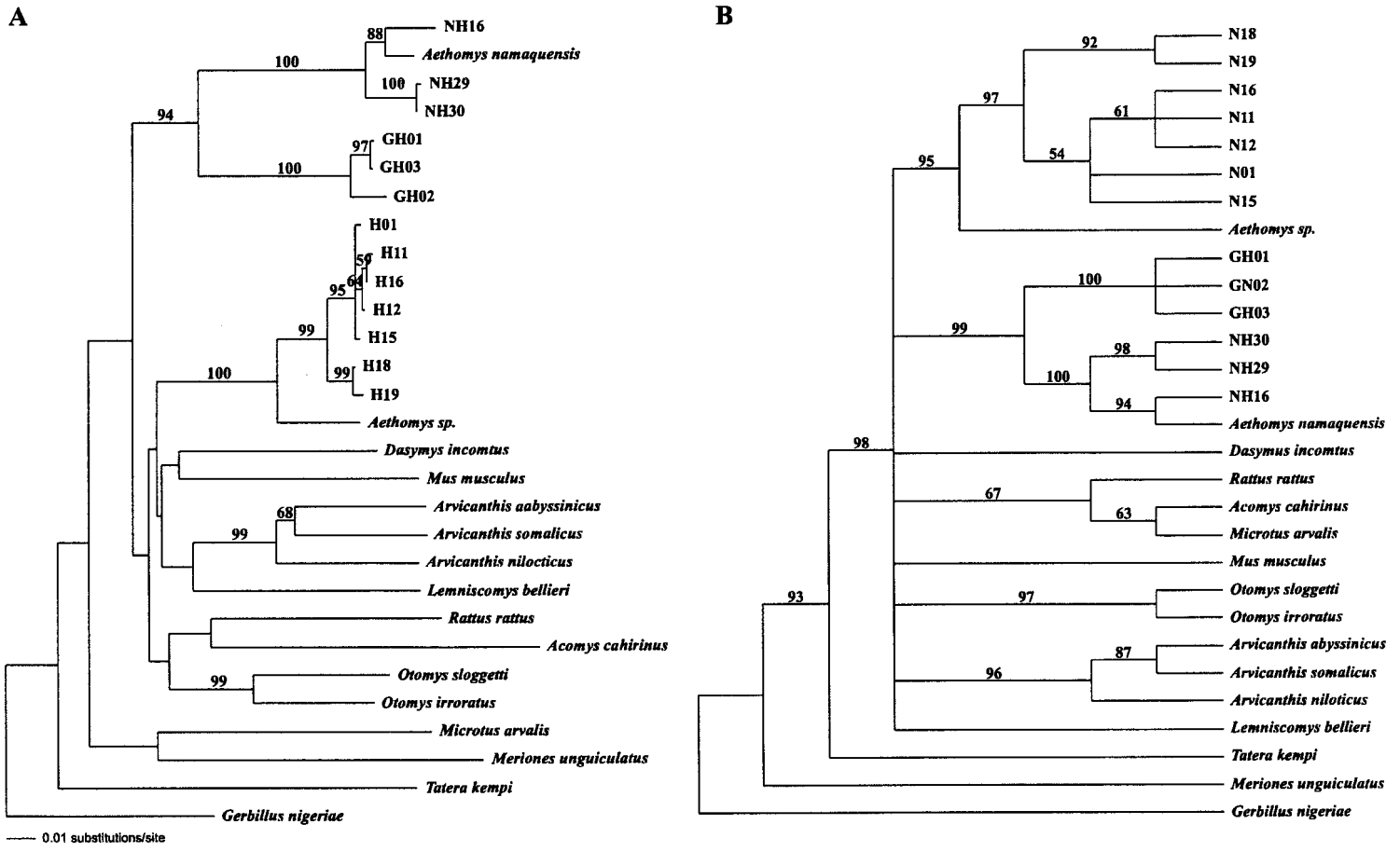


Fig. 4.3 Neighbour-joining phylogram (A) based on percentage nucleotide sequence divergence and a maximum parsimony strict consensus tree of four equally parsimonious trees (B) based on mitochondrial DNA cytochrome b gene sequences of *Aethomys* from South Africa and Swaziland and other African murid taxa obtained from GenBank. Haplotypes NH16, NH29, NH30 represent *A. namaquensis*, haplotypes GH01-GH03 represents *A. granti*, haplotypes H01, H11-H16 represent *A. ineptus* and haplotypes H18 and H19 represent *A. chrysophilus* in the present study. Bootstrap confidence limits (% occurrence in 1000 replicates) for internal branches are given at each node. *Gerbillus nigeriae* was used as an outgroup.

Kimura two-parameter sequence divergence of the 16S rRNA gene was, on average, lower than the cytochrome b gene sequence divergence values. The average separation between *A. namaquensis* and *A. granti* was 5.37% (range 5.12-5.58%), while the average sequence divergence between *A. namaquensis* and *A. chrysophilus/A. ineptus* was 10.88% (range 9.98-12.01%). The average separation between *A. granti* and *A. chrysophilus/A. ineptus* was 9.92% (range 8.82-11.25%). Separation between *A. chrysophilus* and *A. ineptus* was estimated at 1.51% (1.00-2.23%). These results supported the previously proposed close affinity between *A. namaquensis* and *A. granti* and between the two cryptic species, *A. chrysophilus* and *A. ineptus* (Chimimba *et al.* 1997).

The parsimony analysis with only cytochrome b data (saturated characters excluded) gave the same topology as the neighbour-joining phylogram (Figs. 4.3A & 4.3B). The following parsimony statistics were obtained: 165 parsimony informative characters; number of trees = 16; tree length = 484; CI = 0.43; RI = 0.63; RC = 0.27; g_1 = -0.78 (P < 0.01; see Hillis & Huelsenbeck, 1992). In an attempt to reduce the number of trees generated, reweighting according to the consistency index (CI) yielded the following statistics: number of trees = 4; tree length = 209; CI = 0.52; RI = 0.68; RC = 0.35; g_1 = -8.63 (P < 0.01), and a consensus tree was subsequently generated. The partition homogeneity did not reject the congruence of the different data sets (P < 0.05). The parsimony analysis (Fig. 4.4B) comprising the cytochrome b (saturated characters excluded) and 16S rRNA data sets gave the same topology as the neighbour-joining tree (Fig. 4.3A) with the following parsimony statistics: 211 parsimony informative characters; number of trees = 1; tree length = 439; CI = 0.61; RI = 0.74; RC = 0.45; g_1 = -0.71 (P < 0.01).

A parsimony analysis was also conducted including the cytochrome b, 16S rRNA and morphological characters yielding the following parsimony statistics: 178 parsimony informative characters; number of trees = 1; tree length = 202; CI = 0.89; RI = 0.94; RC = 0.84; g_1 = -0.76 (P < 0.01) (Fig. 4.5). These nodes showed high bootstrap support. The latter combined analysis only included sequences of *Aethomys* that were generated in the present study and suggested the presence of two well-supported clades that included: 1) *A. chrysophilus* and *A. ineptus* and 2) *A. namaquensis* and *A. granti*. This subdivision of *Aethomys* was also evident from the other two analyses but these analyses did not support the monophyly of *Aethomys*, suggesting that this genus may be paraphyletic.

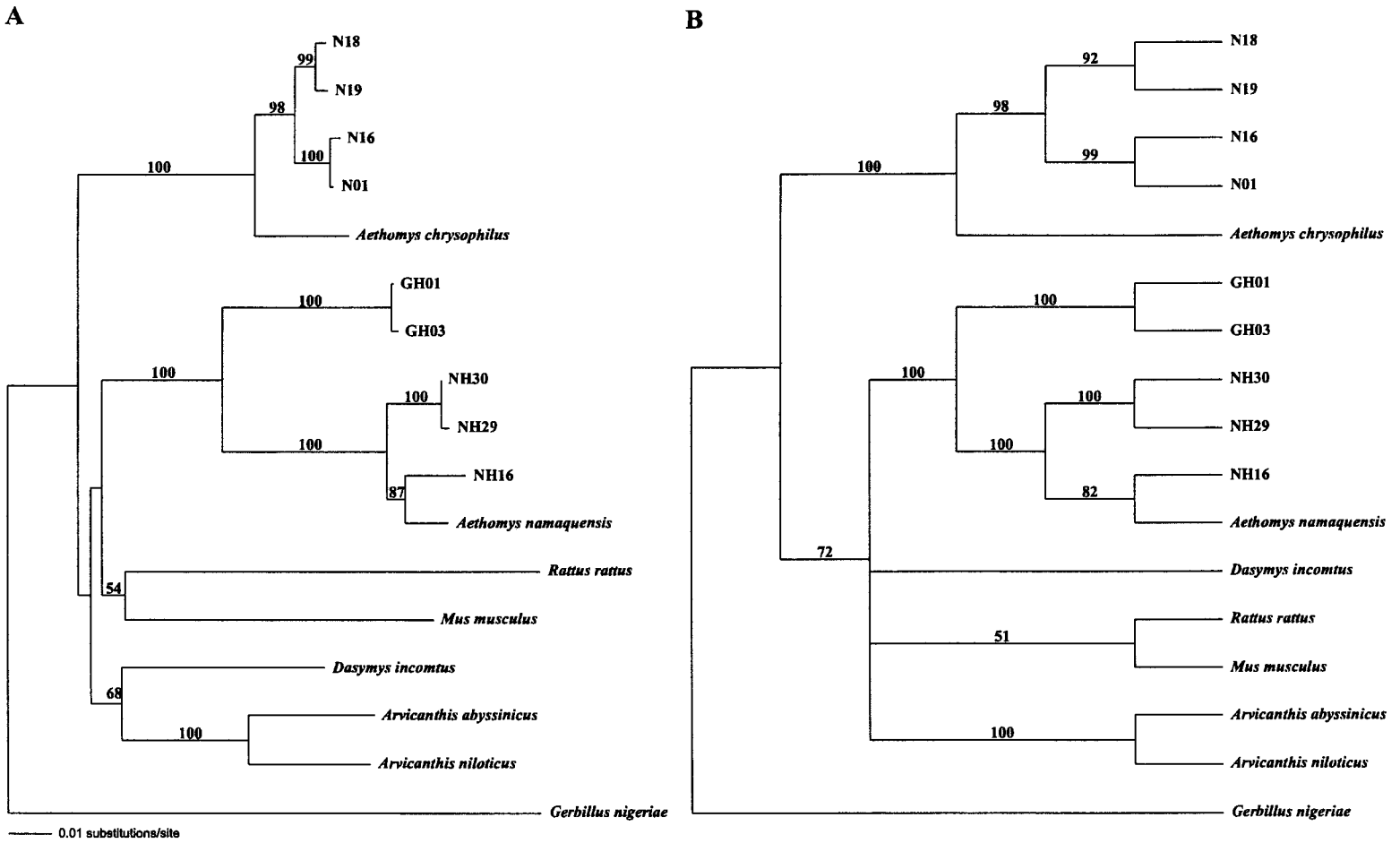


Fig. 4.4 Neighbour-joining phylogram (A) based on percentage nucleotide sequence divergence and a maximum parsimony tree (B) for combined mitochondrial DNA cytochrome b and 16S rRNA sequences of *Aethomys* from South Africa and Swaziland and other African murid taxa obtained from GenBank. Haplotypes NH16, NH29, NH30 represent *A. namaquensis*, haplotypes GH01, GH03 represent *A. granti*, haplotypes H01 and H16 represent *A. ineptus* and haplotypes H18 and H19 represent *A. chrysophilus* in the present study. Numbers above branches indicate bootstrap support (1000 replicates) for branches that occur on at least 50% of the time. *Gerbillus nigeriae* was used as an outgroup.

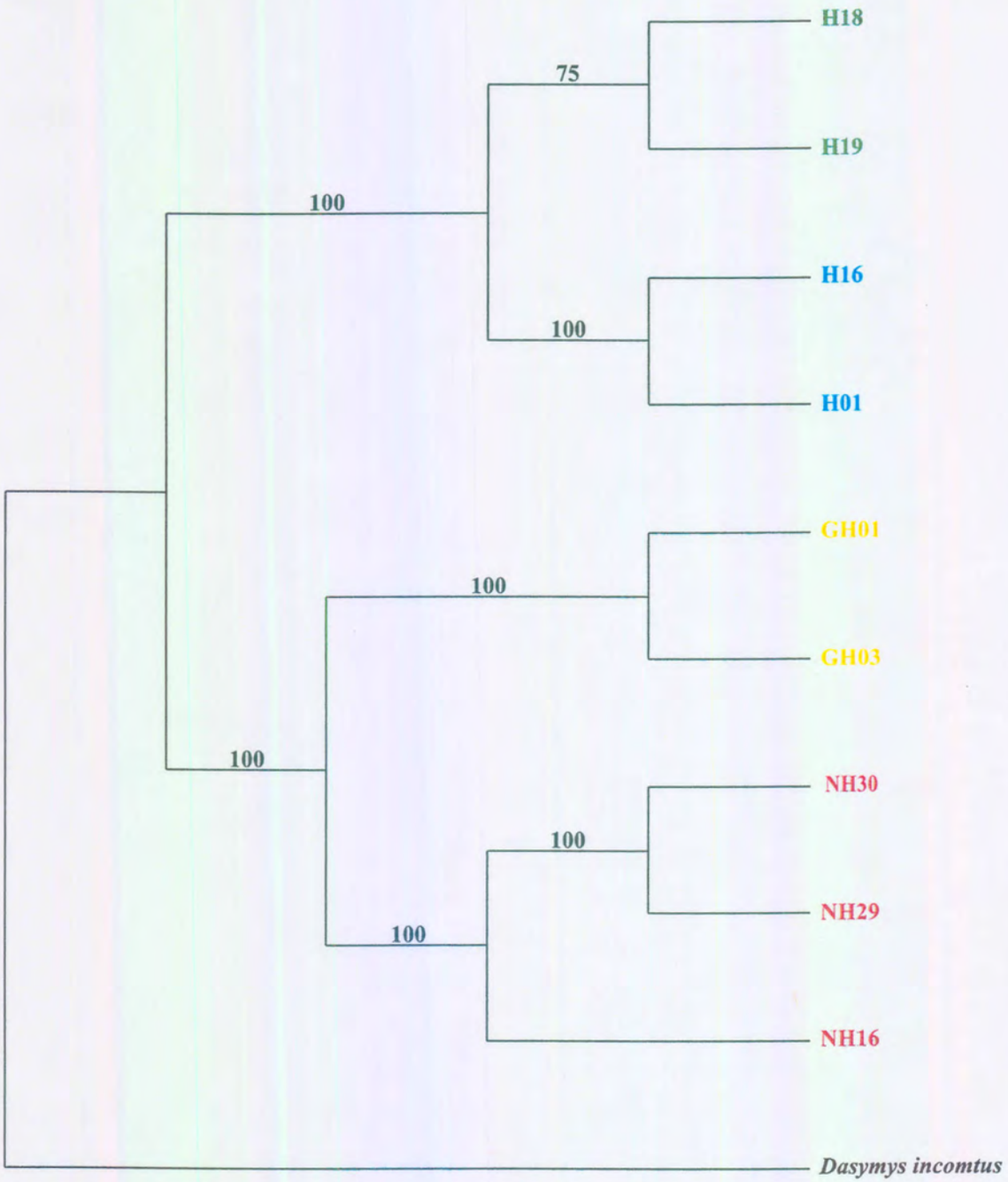


Fig. 4.5 A maximum parsimony tree derived from cytochrome b and 16S rRNA sequences of *Aethomys* from South Africa and Swaziland, and morphological data obtained from Chimimba (1997). Haplotypes NH16, NH29 and NH30 represent *A. namaquensis*, haplotypes GH01, GH03 represent *A. granti*, haplotypes H18-H19 represent *A. chrysophilus* and haplotypes H01, H16 represent *A. ineptus*. Different colours correspond to the colours used in Figure 4.1. The bootstrap confidence limits (% occurrence in 1000 replicates) for internal branches are given at each node. *Dasymys incomtus* was used as an outgroup.

No significant differences in the rates of substitution between pairwise comparisons of lineages relative to an outgroup were detected at the 0.064% ($P \geq 0.0006$) and at the 0.138% ($P \geq 0.0014$) level for cytochrome b and 16S rRNA, respectively, indicating that the different lineages evolve at the same rate compared to the outgroup. The same approach as in analyses of *A. namaquensis* and *A. chrysophilus/A. ineptus* was followed to date the different lineages (Chapter 2 & 3). There are 32 third position transversions between *Mus* and *Rattus* (Bibb *et al.* 1981; Gadaleta *et al.* 1989) in the 489 bp fragment of the cytochrome b gene. This resulted in a divergence estimate of 19.63% between *Mus* and *Rattus* corresponding to the rate of change in third position transversions of 1.65% per Myr. Based on this calibration, the divergence time separating *A. namaquensis* from *A. chrysophilus*, *A. ineptus* and *A. granti* was estimated at 10.42 (range 10.04-10.78), 10.41 (range 9.67-10.78) and 6.02 (range 5.20-6.32) Myr ago, respectively.

The divergence time separating *A. granti* from *A. chrysophilus* and *A. ineptus* was estimated to have occurred 7.99 (range 7.81-8.82) and 8.03 (range 7.81-8.82) Myr ago, respectively. The divergence time separating *A. chrysophilus* and *A. ineptus* could not be estimated as no third position transversions were detected between these two cryptic species. When using the conventional substitution rate of 2% per Myr, the divergence time separating *A. namaquensis* and *A. granti* was estimated at 6.83 (range 6.41-7.28) Myr ago, while the separation between *A. namaquensis* and *A. chrysophilus/A. ineptus* was estimated at 9.21 (range 8.69-9.77) Myr ago. These dates suggest an early Pliocene and Miocene separation, respectively. The divergence time separating *A. granti* and *A. chrysophilus/A. ineptus* was estimated to have occurred 8.06 (7.53-8.64) Myr ago. Separation between *A. chrysophilus* and *A. ineptus* was estimated at 1.21 (830 000-1.43) Myr ago, suggesting an early Pleistocene separation. Both these methods yielded broadly similar estimates of times of divergence between the four southern African species of *Aethomys*.

The 16S rRNA gene showed 68 transitions and transversions (absolute number of differences) between *Mus* and *Rattus* (Bibb *et al.* 1981; Gadaleta *et al.* 1989) in the 526 bp fragment of the cytochrome b gene. This resulted in a divergence time of 12.93% between *Mus* and *Rattus* corresponding to the rate of change in all transitions and transversions of 1.10% per Myr. Based on this calibration, the divergence time separating *A. namaquensis* from *A. chrysophilus*, *A. ineptus* and *A. granti* was estimated at 8.58 (range 8.12-9.16),

8.93 (range 8.47-9.67) and 4.53 (4.30-4.67) Myr ago. The divergence time separating *A. granti* from *A. chrysophilus* and *A. ineptus* was estimated to have occurred 7.52 (range 7.23-7.78) and 8.21 (range 8.21-8.29) Myr ago, respectively. Although differences were detected between *A. chrysophilus* and *A. ineptus*, the time of separation between these two sibling species could not be estimated, as these differences were too low. The conventional substitution rate of 2% per Myr for the cytochrome b gene was adjusted for the 16S rRNA gene by using the Kimura two-parameter sequence divergence value between *Mus* and *Rattus*. This resulted in a divergence of 15.14% between *Mus* and *Rattus* corresponding to an overall rate of change of 1.26% per Myr. This overall rate of change is similar to the rate of 1.10% obtained using absolute number of differences, suggesting no differences in the method of calibration used.

Discussion

The phylogenetic analysis presented in this study offers several insights into the evolutionary relationships of southern African *Aethomys* and other African murid rodents. Of particular relevance is that the relationships between various combinations of species within each clade or group are broadly consistent with previously postulated relationships. The previously suggested sister-taxon relationship between the two cryptic species, *A. chrysophilus* and *A. ineptus* was apparent in this study. The close relationship between these species is also evident from the results of the molecular clock with a divergence time estimated at 1.12 Myr ago. Based on a marked change in sperm morphology, it was suggested that these two sibling species might be distantly related (Visser & Robinson 1987). Phillips *et al.* (1985), however, suggested that divergent sperm forms as is typical for these two cryptic species may evolve, even in closely related species that would act as a prezygotic reproductive isolating mechanism.

The close evolutionary relationship between *A. namaquensis* and *A. granti* was also evident from the present study. This affinity has also been observed in earlier classical taxonomic treatments of the genus (Ellerman *et al.* 1953), other studies based on cytogenetic analyses (Matthey 1954, 1958, 1964; Visser & Robinson 1986, 1987) and phenetic analyses (Chimimba *et al.* 1999; Chimimba 1997). It is not surprising that this

affinity led earlier taxonomic treatments to allocate these two species to the subgenus *Micaelamys*. Musser and Carleton (1993), however, considered characters separating *A. namaquensis* and *A. granti* from other species of *Aethomys* to be insignificant.

The phylogenetic analyses of the southern African species of *Aethomys* and other African murid rodents suggested the presence of two well-supported clades within *Aethomys* that included: 1) *A. chrysophilus* and *A. ineptus*; 2) *A. namaquensis* and *A. granti*. One of the major findings in this study, however, is that *Aethomys* may not be monophyletic. Based on molecular data, Ducroz *et al.* (2001) also found *Aethomys* to be paraphyletic, with *A. namaquensis* as the most basal member of a clade that combine arvicanthine rodents and other African murine taxa except *Mastomys*. The combined analysis (cytochrome b, 16S rRNA and morphological characters) could not test monophyly adequately due to the inclusion of a single closely related outgroup, *D. incomtus* and/or because the molecular support is relatively weak in comparison with the morphological data (Chimimba 1997). If sequence divergence values were taken into account, *D. incomtus* was more closely related to *A. ineptus* and *A. chrysophilus* than these two species were to *A. namaquensis*. These results may support the previous allocation of *A. namaquensis* to the genus *Thallomys* (Ellerman 1941).

The poor resolution that was obtained for higher-level relationships made it impossible to ascertain the phylogenetic affinities between *Aethomys* and the other African murine taxa. This is possibly due to the rapid radiation of rodents entering sub-Saharan Africa from Eurasia, possibly through the Middle East (Ducroz *et al.* 2001). It is possible that more samples and other DNA markers (e.g. nuclear genes) may assist in resolving these relationships.

The lack of any close relationship between fossil species of *Aethomys* found in East and South Africa has precluded interpretation of the biogeography of the genus (Chimimba 1997). Despite the need for additional palaeontological data and analyses based on other DNA markers, this study may nevertheless, allow a biogeographical hypothesis to be developed.

Aethomys chrysophilus represents a widely distributed species in southern as well as East and Central Africa. Chimimba (1997) postulated that the similarity between *A.*

chrysophilus (*senso lato*) and one of the two oldest known fossil species of *Aethomys*, *A. modernis* (Denys 1990a, b) may suggest that a form ancestral to *A. chrysophilus* may have given rise to all extant species within the genus *Aethomys*. From the molecular data in the present study, the divergence time separating *A. chrysophilus* from *A. namaquensis* and *A. granti* was estimated at 9.21 and 8.06 Myr ago, respectively, suggesting a separation during the Miocene. Baker *et al.* (1988) suggested that the acrocentric morphology of chromosome 5, 14, 15 and 20, evident in *A. ineptus*, may be primitive for the ancestor of the Muridae and Sigmodontidae rodent lineages.

Aethomys namaquensis is also a widely distributed species and extends from southern Africa, southern Angola and southern Malawi to Zambia, while *A. granti* is restricted to south-central South Africa (Musser & Carleton 1993). It is interesting to note that the teeth of *A. granti* are characteristic and sufficiently robust to facilitate preservation, suggesting that a lack of preservation or misidentification is unlikely (De Graaf 1981). The lack of palaeontological data within or near the present distributional range of *A. granti*, however, suggested that this species may be of more recent origin, perhaps emanating from an *A. namaquensis*-like form (Chimimba 1997). This is also evident from the molecular clock data in the present study that suggests that *A. granti* and *A. namaquensis* diverged more recently from one another (6.83 Myr ago) compared to the divergence between these two species and the two cryptic species *A. chrysophilus* and *A. ineptus* (8.01-9.21 Myr ago).

Chapter 5

Conclusions

There is a general lack of congruence between the present molecular dataset and previous morphological investigations. From the present sample sizes, there is only good support, from a molecular point of view, for the subspecies *A. n. lehocla*. In general, it seems that lineage B (*A. n. lehocla*) is not only restricted to the Nama Karoo but it also occurs in other Karoo microhabitats or in areas with invasions of Karoo elements. Two other genetically well-defined lineages were identified in the Grassland (lineage C) and Savanna (lineage A) biomes, as well as a further six unique lineages that are not geographically well represented. These may correspond to previously described subspecies. Further studies are needed to investigate if these lineages may represent biological species, since sequence divergences between lineages are substantial and with better sampling there appears to be several areas of overlap or contact where breeding integrity can be tested with nuclear markers.

Ecological speciation seems to be the most likely explanation for the origin of diversity within *A. namaquensis*, since the geographically well defined lineages corresponded broadly in their distribution to vegetation types as described by Low & Rebelo (1996) and Acocks (1975). The previously documented mtDNA phylogeographic break detected in the rock hyrax and red rock rabbits (Prinsloo & Robinson 1992; Prinsloo 1993; Matthee & Robinson 1996) was also evident in *A. namaquensis* (lineage C). The diversity and differentiation within *A. namaquensis* is more complex than the four previously proposed subspecies (Chimimba 2001a). Further geographic sampling as well as linkage to type material of described subspecies is needed to resolve the identity of the unique lineages and to gain more insight into the phylogeography and mode of speciation in this group of rodents.

Molecular techniques, specifically the analysis of the mtDNA cytochrome b gene, have been useful in the identification of sibling species. A phylogenetic analysis that included the two cryptic species, *A. chrysophilus* and *A. ineptus*, revealed two groups that correspond to the two species. Intraspecific mtDNA variation revealed that *A. ineptus* has a pattern that reflects phylogenetic continuity. The two morphologically defined subspecies detected within *A. chrysophilus* were not evident from the present study possibly due to inadequate sampling. More samples covering the distributional range of *A. chrysophilus* may reveal the presence of the two subspecies. Distributional data from Linzey *et al.* (2003) and from the present study suggested that the *A. chrysophilus* occupies

the low elevations of the Limpopo River drainage while *A. ineptus* occupies the remainder of South Africa at higher elevations but expands into lower elevation habitats in the southern portion of its range.

Phylogenetic relationships among species of *Aethomys* suggested the presence of two clades that included: 1) *A. chrysophilus* and *A. ineptus* and 2) *A. namaquensis* and *A. granti*, supporting previously postulated relationships between species of *Aethomys*. The present study suggests that *Aethomys* may be paraphyletic. This paraphyletic relationship between the species within *Aethomys* was also detected by Ducroz *et al.* (2001). This suggests that the taxonomic status of *A. namaquensis* and *A. granti* needs further investigation more especially by including other species currently recognised within *Aethomys*.

REFERENCES

Acocks JPH (1975) *Veld types of South Africa*. DJK Killick, Botanical Research Institute, Department of Agriculture Technical Services, South Africa.

Adams EN (1972) Consensus techniques and the comparison of taxonomic trees. *Systematic Zoology*, **21**, 390-397.

Allard MW, Honeycutt RL (1992) Nucleotide sequence variation in the mitochondrial 12S rRNA gene and the phylogeny of African mole-rats (Rodentia, Bathyergidae). *Molecular Biology and Evolution*, **9**, 27-40.

Allen J (1877) The influence of physical conditions in the genesis of species. *Radical Review*, **10**, 1-108.

Amos B, Hoelzel AR (1991) Long-term preservation of whale skin for DNA analysis. *Report of the International Whaling Commission*, **13**, 99-104.

Archie JW (1989) Homoplasy excess ratios: new indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. *Systematic Zoology*, **38**, 253-269.

Avery DM (1981) Holocene micromammalian faunas from the northern Cape Province, South Africa. *South African Journal of Science*, **77**, 265-273.

Avery DM (1982) Micromammals as palaeoenvironmental indicators and an interpretation of the late Quaternary in the southern Cape Province, South Africa. *Annals of the South African Museum*, **85**, 183-374.

Avery DM (1985) A preliminary assessment of the micromammalian remains from Gladysvale cave, South Africa. *Palaeontologia Africana*, **32**, 1-10.

Avise JC (1989) Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution*, **43**, 1192-1208.

Avise JC (1992) Molecular population structure and the biogeographic history of the regional fauna: a case in history with lessons for conservation biology. *Oikos*, **63**, 62-76.

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

Avise JC, Arnold RM, Ball E, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987a) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Avise JC, Bowen BW, Lamb T (1989) DNA fingerprinting from hypervariable mitochondrial genotypes. *Molecular Biology and Evolution*, **6**, 258-269.

Avise JC, Giblin-Davidson C, Laerm J, Patton JC, Lansman RA (1979b) Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Sciences of the USA*, **76**, 6694-6698.

Avise JC, Lansman RA (1983) Polymorphism of mitochondrial DNA in populations of higher animals. In: *Evolution of genes and proteins* (eds. Nei M and Koehn RK), pp. 147-164. Sunderland, Massachusetts.

Avise JC, Shapira JF, Daniel SW, Aquadro CF, Lansman RA (1983) Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. *Molecular Biology and Evolution*, **1**, 38-56.

Axelrod DI, Raven PH (1978) Late Cretaceous and Tertiary vegetation history of Africa. In: *Biogeography and ecology of southern Africa* (ed. Werger MJA), pp. 77-130. Dr W Junk Publishers, The Hague.

Baker BH, Wohlenberg J (1971) Structure and evolution of the Kenya rift valley. *Nature*, **229**, 538-542.

Baker RJ, Qumsiyeh MB, Rautenbach IL (1988) Evidence for eight tandem and five centric fusions in the evolution of the karyotype of *Aethomys namaquensis* A. Smith (Rodentia: Muridae). *Genetica*, **76**, 1661-1669.

Bell DM, Hamilton MJ, Edwards CW, Wiggins LE, Martínez RM, Strauss RE, Bradley RD, Baker RJ (2001) Patterns of karyotypic megaevolution in *Reithrontomys*: evidence from cytochrome-*b* phylogenetic hypothesis. *Journal of Mammalogy*, **82**, 81-91.

Bergmann C (1847) Ueber die Verhältnisse der Wärmeokonomie der Thiere zu ihrer Grösse. *Göttinger studien*, **3**, 595-708.

Bernatchez L (2001) The evolutionary history of Brown trout (Salmo Trutta L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, **55**, 351-379.

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell*, **26**, 167-180.

Black CG, Krishtalka L (1986) Rodents, bats and insectivores from Plio-Pleistocene sediments to the east of Lake Turkana, Kenya. *Contributions to Science*, **372**, 1-15.

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

Bothma JduP (1966) Color variation in Hyracoidea from southern Africa. *Journal of Mammalogy*, **47**, 687-693.

Bradbury JW, Vehrencamp SL (1998) *Principles of animal communication*. Sunderland, Massachusetts: Sinauer.

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

Bradley RD, Tiemann-Boege I, Kilpatrick WC, Schmidly DJ (2000) Taxonomic status of *Peromyscus boylii sacarensis*: inferences from DNA sequences of the mitochondrial cytochrome-*b* gene. *Journal of Mammalogy*, **81**, 875-884.

Brain CK (1985) Temperature-induced environmental changes in Africa as evolutionary stimuli. In: *Species and speciation* (ed. Vrba ES), pp. 45-52. Transvaal Museum Monograph no. 4, Transvaal Museum, Pretoria.

Breed WG (1997) Unusual chromatin structural organization in the sperm head of a murid rodent from southern Africa: the red veld rat, *Aethomys chrysophilus* type B. *Journal of Reproduction and Fertility*, **111**, 221-228.

Breed WG, Cox GA, Leigh CM, Hawkins P (1988) Sperm head structure of a murid rodent from southern Africa: the red veld rat *Aethomys chrysophilus*. *Gamete Research*, **19**, 191-202.

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA*, **76**, 1967-1971.

Buth DG, Mayden RL (1981) Taxonomic status and relationships among populations of *Notropis pilsbryi* and *N. zonatus* (Cypriniformes: Cyprinidae) as shown by the glucosephosphate isomerase, lactate dehydrogenase, and phosphoglucomutase enzyme systems. *Copeia*, **1981**, 583-590.

Catzeflis FM, Aguilar J-P, Jaeger JJ (1992) Muriod rodents: phylogeny and evolution. *Trends in Ecology and Evolution*, **7**, 122-126.

Chimimba CT (1997) A systematic revision of southern African *Aethomys* Thomas, 1915 (Rodentia: Murdiae). PhD thesis, University of Pretoria, Pretoria.

Chimimba CT (1998) A taxonomic revision of synthesis of southern African *Aethomys* (Rodentia: Muridae) with a key to species. *Mammalia*, **62**, 427-437.

Chimimba CT (2000) Geographic variation in *Aethomys chrysophilus* (Rodentia: Muridae) from southern Africa. *Zeitschrift für Säugetierkunde*, **65**, 157-171.

Chimimba CT (2001a) Infraspecific morphometric variation in *Aethomys namaquensis* (Rodentia: Muridae) from southern Africa. *Journal of Zoology (London)*, **253**, 191-210.

Chimimba CT (2001b) Geographic variation in the Tete veld rat *Aethomys ineptus* (Rodentia: Muridae) from southern Africa. *Journal of Zoology (London)*, **254**, 77-89.

Chimimba CT, Dippenaar NJ, Robinson TJ (1998) Geographic variation in *Aethomys granti* (Rodentia: Muridae) from southern Africa. *Annals of the Transvaal Museum*, **36**, 405-412.

Chimimba CT, Dippenaar NJ, Robinson TJ (1999) Morphometric and morphological delineation of southern African species of *Aethomys* (Rodentia: Muridae). *Biological Journal of the Linnean Society*, **67**, 501-527.

Clark JD (1967) *Altas of African Prehistory*. University of Chicago Press, Chicago, London.

Coetzee JA (1978) Late Cainozoic palaeoenvironments of southern Africa. In: *Antarctic glacial history and world palaeoenvironments* (ed. Van Zinderen Bakker EM), pp. 114-127. AA Balkema, Rotterdam.

Cracraft (1989) Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: *Speciation and its consequences* (eds. Otte D and Endler JA), pp. 3-87. Sinauer Associates, Sunderland, Massachusetts.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 959-969.

Cummings MP, Otto SP, Wakeley J (1995) Sample properties of DNA sequence data in phylogenetic analysis. *Molecular Biology and Evolution*, **12**, 814-822.

Daniel JC, Blumstein DT (1998) A test of the acoustic adaptation hypothesis in four species of marmots. *Animal Behaviour*, **56**, 1517-1528.

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

De Graaff G (1960) A preliminary investigation of the mammalian microfauna in Pleistocene deposits of caves in the Transvaal system. *Palaeontologia Africana*, **7**, 79-118.

De Graaff G (1961) On the fossil mammalian microfauna collected at Kromdraai by Dapper in 1895. *South African Journal of Science*, **56**, 259-260.

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

De Graaff (1997) Namaqua rock mouse, *Aethomys namaquensis*. In: *The complete book of southern African mammals* (eds. Mills G and Hes L), pp. 149. Struik Winchester, Cape Town.

Denys C (1987) Micromammals from the West Natron Pleistocene deposits (Tanzania): biostratigraphy and paleoecology. *Scientific Geology Bulletin*, **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 novellas especes d'*Aethomys* (Rodentia: Muridae) a Langebaanweg (Pliocene, Afrique du Sud): implications phylogenetiques et paleoecologiques. *Annals of Paleontology (Vertebrates and Invertebrates)*, **76**, *41-69*.

De Queiroz A (1993) For consensus (sometimes). *Systematic Biology*, **42**, 368-372.

De Queiroz A, Donoghue MJ, Kim J (1995) Separate versus combined analysis of phylogenetic evidence. *Annual Review of Ecology and Systematics*, **26**, 657-681.

De Winton WE (1897) On collections of rodents made by Mr. J. F. Darling in Mashunaland and Mr. F. C. Selous in Matabeleland, with short field notes by the collectors. *Proceedings of the Zoological Society of London*, **1896**, 798-808.

Dippenaar NJ, Rautenbach IL (1986) Morphometrics and karyology of the southern African species of the genus *Acomys* I, Geoffroy Saint-Hilaire, 1838 (Rodentia: Muridae). *Annals of the Transvaal Museum*, **34**, 129-183.

Dobzhansky TH (1940) Speciation as a stage in evolutionary divergence. *American Naturalist*, 74, 312-321.

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. *Molecular Phylogenetics and Evolution*, **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. *Journal of Mammalian Evolution*, **8**, 173-206.

Ellerman JR (1941) *The families and genera of living rodents*. British Museum (Natural History), London.

Ellerman JR, Morrison-Scott TCS, Hayman RW (1953) *Southern African mammals 1758 to 1951: a reclassification*. British Museum (Natural History), London.

Ellison GTH, Taylor PJ, Nix HA, Bronner GN, McMahon JP (1993) Climatic adaptation of body size among pouched mice (*Saccostomus campestris*: Cricetidae) in the southern African subregion. *Global Ecology and Biogeography Letters*, **3**, 1-8.

Esposti MD, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A (1993) Mitochondrial cytochrome *b*: evolution and structure of the protein. *Biochimica et Biophysica Acta*, **1143**, 243-271.

Excoffier L, Langaney A (1989) Origin and differentiation of human mitochondrial DNA. *American Journal of Human Genetics*, **44**, 73-85.

Excoffier L, Smouse P (1994) Using allele frequencies and geographic subdivision to reconstruct gene genealogies within a species, molecular variance parsimony. *Genetics*, **136**, 343-359.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Farris JS (1969) A successive approximations approach to character weighting. *Systematic Zoology*, **18** ,374-385.

Farris JS (1989) The retention index and rescaled consistency index. *Cladistics*, **5**, 417-419.

Farris JS, Kluge AG, Eckardt MJ (1970) A numerical approach to phylogenetic systematics. *Systematic Zoology*, **19**, 172-179.

Farris JS, Källersjo M, Kluge Ag, Bult C (1995) Constructing a significance test for incongruence. *Systematic Biology*, **44**, 570-572.

Felsenstein J (1973) Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology*, **22**, 240-249.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368-376.

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

Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. *Molecular Ecology*, **11**, 1245-1262.

Frye M, Hedges SB (1995) Monophyly of the order Rodentia inferred from mitochondrial DNA sequences of the genes for 12S rRNA, 16S rRNA and tRNA-valine. *Molecular Biology and Evolution*, **12**, 168-176.

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

Fumagalli L, Taberlet P, Stewart DT, Gielly L, Hausser J, Vogel P (1999) Molecular phylogeny and evolution of Sorex shrews (Soricidae: Insectivora) inferred from mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution*, **11**, 222-235.

Gadaleta G, Pepe G, De Candia G, Quagliariello C, Sbisa E, Saccone C (1989) The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analyses between vertebrates. *Journal of Molecular Ecology*, **28**, 497-516.

Gaggiotti OE (1996) Population genetic models of source-sink metapopulations. *Theoretical Population Biology*, **50**, 178-208.

Gannon WL, Lowlor TE (1989) Variation of the chip vocalization of three species of Townsend chimpmunks (genus *Eutamias*). *Journal of Mammalogy*, **70**, 740-753.

Gear JHS, Davis DHS, Pitchford RJ (1966) The susceptibility of rodents to schistom infection with special reference to *Schistosoma haematobium*. *Bulletin of the World Health Organization*, **35**, 213-221.

Girman DJ, Kat PW, Mills MGL, Ginsberg JR, Borner M, Wilson V, Fanshawe JH, Fitzgibbon C, Lau LM, Wayne RK (1993) Molecular genetic and morphological analysis of the African wild dog (*Lycaon pictus*). *Journal of Hereditary*, **84**, 450-459.

Gordon DH, Rautenbach IL (1980) Species complexes in medically important rodents: chromosome studies of *Aethomys*, *Tatera* and *Saccostomus* (Rodentia: Muridae, Cricetidae). *South African Journal of Science*, **76**, 559-561.

Gordon DH, Watson CRB (1986) Identification of cryptic species of rodents (*Mastomys*, *Aethomys*, *Saccostomus*) in the Kruger National Park. *South African Journal of Zoology*, **21**, 95-99.

Grant SW, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Hereditary*, **89**, 415-426.

Gu X, Zhang J (1997) A simple method for estimating the parameters of substitution rate variation among sites. *Molecular Biology and Evolution*, **14**, 1106-1113.

Hackett SJ (1996) Molecular Phylogenetics and Biogeography of Tanagers in the Genus *Ramphocelus* (Aves). *Molecular Phylogenetics and Evolution*, **5**, 368-382.

Halanych KM (1996) Testing hypotheses of chaetognath origins: long branches revealed by 18s DNA. *Systematic Biology*, **45**, 223-246.

Halanych KM, Demboski JR, Jansen van Vuuren B, Klein DR, Cook J (1999) Cytochrome *b* phylogeny of North American hares and jackrabbits (*Lepus*, Lagomorpha) and the effects of saturation in outgroup taxa. *Molecular Phylogenetics and Evolution*, **11**, 213-221.

Halanych KM, Robinson TJ (1999) Multiple substitutions affect the phylogenetic utility of cytochrome *b* and 12S rDNA data: examining a rapid radiation in leporid (Lagomorpha) evolution. *Journal of Molecular Evolution*, **48**, 369-379.

Hallet AF, McNeill D, Meyer KF (1970) A serological survey of small mammals for plague in southern Africa. *South African Medical Journal*, **44**, 831-837.

Harpending H (1994) Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591-600.

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.

Hayes JP, Harrison RG (1992) Variation in mitochondrial DNA and the biogeographic history of woodrats (*Neotoma*) of the eastern United States. *Systematic Biology*, **41**, 331-344.

Hendey QB (1981) Paleoecology of the late tertiary fossil occurrences in "E" Quarry, Langebaanweg, South Africa and a re-introduction of their geological context. *Annals of the South African Museum*, **84**, 1-104.

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Hereditary*, **83**, 189-195.

Howell N (1989) Evolutionary conservation of protein regions in the proton-motive cytochrome b and their possible roles in redox catalysis. *Journal of Molecular Evolution*, **29**, 157-169.

Huelsenbeck JP, Bull JJ, Cunningham CW (1996) Combining data in phylogenetic analysis. *Trends in Ecology and Evolution*, **11**, 152-158.

Hutterer R, Joger U (1982) Klensauger aus dem hochland von Adamaoua, Kamerun. *Bonner Zoologische Beiträge*, **33**, 119-132.

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution*, **32**, 128-144.

Jaarola M, Tegelström H (1995) Colonization history of north European field voles (Microtus agrestis) revealed by mitochondrial DNA. *Molecular Ecology*, **4**, 299-310.

Jacobs LL, Downs WR (1994) The evolution of murine rodents in Asia. In: *Rodents and Lagomorph families of Asian origin and diversification* (eds. Tomida Y, Li D and Setoguchi T), pp. 149-156. National Science Museum Monographs, Tokyo.

Jaeger JJ (1976) Les rongeurs (Mammalia: Rodentia) du Pleistocene inferieur d'Olduvai Bed I (Tanzanie), 1e parti: les Murides. In: *Fossil vertebrates of Africa* (eds. Savage RJG and Coryndon SC), pp. 58-120. Academic Press, New York and London.

Jaeger JJ (1979) Les faunes de rongeurs et de lagomorphes du Pliocene et du Pleistocene d'Afrique orientale. *Bulletin of the Society of Geology, France*, **7**, 301-308.

Jameson HL (1909) On a collection of mammals from South Africa. *Annals and Magazine of Natural History*, series 4, **8**, 455-474.

Jansen van Vuuren B (1995) Mitochondrial variation and geographic population structure in the yellow mongoose (*Cynictis Penicillata*). MSc thesis, University of Pretoria, Pretoria.

Jansen van Vuuren B, Robinson T J (1997) Genetic population structure in the yellow mongoose, Cynictis penicillata. *Molecular Ecology*, **6**, 1147-1153.

Johnson SA, Van Tets IG, Nicolson SW (1999) Sugar preferences and xylose metabolism of a mammal pollinator, the Namaqua rock mouse (*Aethomys namaquensis*). *Physiological and Biochemical Zoology*, **72**, 438-444.

Kim I, Phillips CJ, Monjeau JA, Birney EC, Noack K, Pumo DE, Sikes RS, Dole JA (1998) Habitat islands, genetic diversity, and gene flow in a Patagonian rodent. *Molecular Ecology*, **7**, 667-678.

Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111-120.

King LC (1963) *South African Scenery: A text book of geomorphology*, Third edition. Hafner Publishers, Co., New York.

Kluge AG, Farris JS (1969) Quantitative phylogenetics and the evolution of anurans. *Systematic Zoology*, **18**, 1-32.

Knowles LL, Futuyma DJ, Eanes WF, Rannala B (1999) Insight into speciation from historical demography in the phytophagous beetle genus *Ophraella*. *Evolution*, **53**, 1846-1856.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paäbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the USA*, **86**, 6196-6200.

Kryger U (2002) Genetic variation among South African hares (*Lepus* spec.) as inferred from mitochondrial DNA and microsatellites. PhD thesis, University of Pretoria, Pretoria.

Lansman RA, Avise JC, Aquadro CF, JF Shapira, SW Daniel (1983a) Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Preomyscus maniculatus*. *Evolution*, **37**, 1-16.

Le Roux A, Jackson TP, Cherry MI (2002) Differences in alarm vocalizations of sympatric populations of the whistling rats, *Parotomys brantsii* and *P. littledadei* (Rodentia: Muridae). *Journal of Zoology, London*, **257**, 189-194.

Li W-H, Gouy M, Sharp PM, O'Huigin C, Yang Y-W (1990) Molecular phylogeny of Rodentia, Lagomorpha Primates Artiodactyla and Carnivora and molecular clocks. *Proceedings of the National Academy of Sciences of the USA*, **87**, 6703-6707.

Lidicker WZ (1962) The nature of subspecies boundaries in a desert rodent and its implications for subspecies taxonomy. *Systematic Zoology*, **11**, 160-171.

Linzey AV, Kesner MH (1997) Small mammals of a woodland-savannah ecosystem in Zimbabwe, I. Density and habitat occupancy patterns. *Journal of Zoology (London)*, **243**, 137-152.

Linzey AV, Kesner MH, Chimimba CT, Newberry (2003) Distribution of veld rat sibling species *Aethomys chrysophilus* and *Aethomys ineptus* (Rodentia: Muridae) in southern Africa. *African Zoology*, **38**, 169-174.

Low AB, Rebelo AG (eds.) (1996) *Vegetation of South Africa, Lesotho and Swaziland.* Department of Environmental Affairs and Tourism, Pretoria.

Maddison WP, Maddison DR (1992) *MacClade: Analysis of Phylogeny and Character Evolution, version 3*. Sinauer, Sunderland, Massachusetts.

Maddison DR, Ruvolo M, Swofford DL (1992) Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Systematic Biology*, **41**, 111-124.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-220.

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

Mares MA, Lacher TE (1987) Ecological, morphological and behavioural convergence in rock-dwelling mammals. In: *Current Mammalogy* Volume 1 (ed. Genoways HH), pp. 307-348. Plenum Press, New York.

Martin Y, Gerlach G, Schlötterer C, Meyer A (2000) Molecular phylogeny of European muroid rodents based on complete cytochrome b sequences. *Molecular Phylogenetics and Evolution*, **16**, 37-47.

Matthee CA (1993) Mitochondrial DNA variability and geographic population structure in *Pronogalus ruepstirs* and *P. randensis* (Mammalia: Lagomorpha). MSc thesis, University of Pretoria, Pretoria.

Matthee CA, Robinson TJ (1996) Mitochondrial DNA differentiation among geographical populations of *Pronolagus rupestris*, Smith's red rock rabbit (Mammalia: Lagomorpha). *Heredity*, **76**, 514-523.

Matthee CA, Robinson TJ (1997) Mitochondrial DNA Phylogeography and Comparative Cytogenetics of the Springhare, *Pedetes capensis* (Mammalia: Rodentia). *Journal of Mammalian Evolution*, **4**, 53-73.

Matthey R (1954) Nouvelles recherches sur les chromosomes des Muridae. *Caryologia*, **6**, 1-44.

Matthey R (1958) Les chromosomes et al position systematique de quelques Murinea Africains. *Acta Tropica*, **15**, 97-117.

Matthey R (1964) Analyse caryologique de cing especes de Muridae Africains (Mammalia, Rodentia). *Mammalia*, **28**, 403-418.

Mayr E (1942) *Systematics and the origin of species*. Columbia University Press, New York.

Mayr E (1997) *Evolution and the biodiversity of life*. Cambridge, the Belknap Press of Harvard University Press.

Mayr E, Ashlock PD (1991) *Principles of systematic zoology*, 2nd edidtion. McGraw-Hill Inc., New York.

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

Meester J, Davis DHS, Coetzee CG (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.

Meester JAJ, Rautenbach IL, Dippenaar NJ, Baker CM (1986) Classification of southern African mammals. *Transvaal Museum Monograph*, **5**, 1-359.

Meier R, Kores P, Darwin S (1991) Homoplasy slope ratio: a better measurement of observed homoplasy in cladistic analyses. *Systematic Zoology*, **40**, 78-88.

Moritz (1994a) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution*, **9**, 373-375.

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, **18**, 269-292.

Morton ES (1975) Ecological sources of selection on avian sounds. *The American Naturalist*, **109**, 17-34.

Musser GG, Carleton, MD (1993) Family Muridae. In: *Mammal species of the world: a taxonomic and geographic reference,* Second edition (eds. Wilson DE and Reeder DM), pp. 501-755. Smithsonian Institution Press, in association with the American Society of Mammalogists, London and Washington D.C.

Nadler CF, Hoffmann RS (1970) Chromosomes of some Asian and South American squirrels (Rodentia: Sciuridae). *Experienta*, **26**, 1383-1386.

Nedbal MA, Honeycutt RL, Schlitter DA (1996) Higher level systematics of rodents (Mammalia, Rodentia): evidence from the mitochondrial 12S gene. *Journal of Mammalian Evolution*, **3**, 201-237.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, USA.

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics*, **97**, 145-163.

Neigel JE, Avise JC (1993) Application of a random-walk model to geographic distributions of animal mitochondrial DNA variation. *Genetics*, **135**, 1209-1220.

O'Brien SJ, Mayr E (1991) Bureaucratic mischief: recognizing endangered species and subspecies. *Science*, **251**, 1187-1188.

Ohdachi S, Dokuchaev NE, Hasegawa M, Masuda R (2001) Intraspecific phylogeny and geographical variation of six species of northeastern Asiatic Sorex shrews based on the mitochondrial cytochrome *b* sequences. *Molecular Ecology*, **10**, 2199-2213.

Paäbo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Research*, **16**, 1245-1255.

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) "*The Simple Fools Guide to PCR*, version 2.0". Department of Zoology and Kewalo Marine Laboratory, University of Hawaii Press, Honolulu.

Patton JL, Da Silva MNF, Malcolm JR (1996) Hierarchical genetic structure and gene flow in three sympatric species of Amazonian rodents. *Molecular Ecology*, **5**, 229-238.

Patton JL, Smith MF (1992) mtDNA phylogeny of Andean mice: a test of diversification across ecological gradients. *Evolution*, **46**, 174-183.

Pearson O (1975) An outbreak of mice in the coastal desert of Peru. *Mammalia*, **39**, 375-386.

Peppers L, Bradley RD (2000) Cryptic species in *Sigmodon hispidus*: evidence from DNA sequences. *Journal of Mammalogy*, **81**, 332-343.

Pereira L, Dupanloup I, Rosser ZH, Jobling MA, Barbujani G (2001) Y-Chromosome mismatch distributions in Europe. *Molecular Biology and Evolution*, **18**, 1259-1271.

Peters WCH (1852) Naturwissenschaftliche Reise nach Mossambique, Zoologie I. Saugetiere, Berlin: Georg Reimer.

Petit E, Excoffier L, Mayer F (1999) No evidence of bottleneck in the postglacial recolonization of Europe by the Noctule bat (*Nyctalus Nactula*). *Evolution*, **53**, 1247-1258.

Phillips DM, Shargi R, Dekel W (1985) Mammalian fertilization as seen with the scanning electron microscope. *American Journal of Anatomy*, **174**, 353-372.

Pocock TN (1987) Plio-Pleistocene mammalian microfauna in southern Africa – a preliminary report including description of two new fossil muriod genera (Mammalia: Rodentia). *Palaeontologia Africana*, **26**, 69-71.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817-818.

Prinsloo P (1993) Molecular and chromosomal phylogeny of the Hyracoidea. PhD thesis, University of Pretoria, Pretoria.

Prinsloo P, Robinson TJ (1992) Geographic mitochondrial DNA variation in the rock hyrax, *Procavia capensis*. *Molecular Biology and Evolution*, **9**, 447-456.

Pusey (1987) Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology and Evolution*, **2**, 295-299.

Rautenbach IL (1978) The mammals of the Transvaal. PhD thesis, University of Natal, Pietermaritzburg.

Roberts A (1913) The collection of mammals in the Transvaal Museum registered up to 31st March 1913, with descriptions of new species. *Annals of the Transvaal Museum*, **4**, 65-107.

Roberts A (1926) Some new South African mammals and some changes in nomenclature. *Annals of the Transvaal Museum*, **11**, 245-263.

Roberts A (1931) New forms of South African mammals. *Annals of the Transvaal Museum*, **14**, 221-236.

Roberts A (1946) Descriptions of numerous new subspecies of mammals. *Annals of the Transvaal Museum*, **20**, 303-328.

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

Robinson M, Gouy M, Gautier C, Mouchiroud D (1998) Sensitivity of the relative-rate test to taxonomic sampling. *Molecular Biology and Evolution*, **15**, 1091-1098.

Robinson-Rechavi M, Huchon D (2000) RRTree: Relative-rate test between groups of sequences on a phylogenetic tree. *Bioinformatics*, **16**, 296-297.

Robinson TJ, Skinner JD, Haim AS (1986) Close chromosomal congruence in two species of ground squirrel: *Xerus inauris* and *X. princeps* (Rodentia: Sciuridae). *South African Journal of Zoology*, **21**, 100-105.

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

Rosevear DR (1969) *The rodents of West Africa*. British Museum of Natural History, London.

Rozas J, Rozas R (1999) *DnaSP*, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174-175.

Rutherford MC, Westfall RH (1986) Biomes of southern Africa – an objective categorization. *Memoirs of the Botanical Survey of South Africa*, **54**, 1-98.

Ryan MJ, Brenowitz EA (1985) The role of body size, phylogeny, and ambient noise in the evolution of bird song. *The American Naturalist*, **126**, 87-100.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9-10.

Ryder OA, Shaw JH, Wemmer CM (1988) Species, subspecies and *ex situ* conservation. *International Zoo Yearbook*, **27**, 134-140.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.

Saitou N, Nei M (1987) The neighbour joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 509-514.

Sambrook, J, Fritsh EP, Maniatis T (1989) A Laboratory Manual. Second edition. Cold Spring Harbor Laboratory Press.

Santuccci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology*, **7**, 1163-1172.

Schluter D (2001) Ecology and the origin of species. *Trends in Ecology and Evolution*, 16, 372-380.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079-1089.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

She JX, Bonhomme F, Boursot P, Thaler L, Catzeflis F (1990) Molecular phylogenies in the genus Mus: comparative analysis of electrophoretic, scnDNA hybridization and mtDNA RFLP data. *Biological Journal of the Linnean Society*, **41**, 83-103.

Simonsen KL, Churchill GA, Aquadro CF (1995) Properties of statistical test of neutrality for DNA polymorphism data. *Genetics*, **141**, 413-429.

Skinner JD, Smithers RHN (1990) The mammals of the southern African subregion. University of Pretoria, Pretoria.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555-562.

Smith A (1834) African zoology: Mammalia. *South African Quarterly Journal*, **2**,160.

Smith A (1836) Report of the expedition for exploring Central Africa, from the Cape of Good Hope: 1934, under the superintendence of A Smith. Government Gazette Office, Cape Town.

Smith MF (1998) Phylogenetic relationships and geographic structure in Pocket Gophers in the genus *Thomomys*. *Molecular Phylogenetics and Evolution*, **9**, 1-14.

Smith MF, Patton JL (1993) The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe. *Biological Journal of the Linnean Society*, **50**, 149-177.

Smith MF, Patton JL (1999) Phylogenetic relationships and the radiation of Sigmodontine rodents in South America: evidence from cytochrome *b*. *Journal of Mammalian Evolution*, **6**, 89-128.

Smithers RHN (1971) The mammals of Botswana. *Memoirs of the National Museums of Rhodesia*, **4**, 1-340.

Smithers RHN, Wilson VJ (1979) Checklist and atlas of the mammals of Zimbabwe Rhodesia. *Museum memoirs of the National Museums and Monuments of Zimbabwe Rhodesia*, **9**, 1-193.

Sullivan J, Holsinger KE, Simon C (1995) Among-site rate variation and phylogenetic analysis of 12S rRNA in Sigmodontine rodents. *Molecular Biology and Evolution*, **12**, 988-1001.

Swanepoel R, Blackburn NK, Efstratiou S, Condy JB (1978) Studies of Rift Valley fever in some African murids (Rodentia: Muridae). *Journal of Hygiene of Cambridge*, **80**, 183-196.

Swofford DL (2002) PAUP*: *phylogenetic analysis using parsimony (* and other methods), Version 4*. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1989a) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

Tajima F (1989b) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597-601.

Tajima F (1993) Measurement of DNA polymorphism. In: *Mechanisms of molecular evolution. Introduction to molecular paleopopulation biology* (eds. Takahata N and Clark AG), pp. 37-59. Tokyo, Sunderland, MA:Japan Scientific Societies Press, Sinauer Associates, Inc.

Templeton AR (1989) The meaning of species and speciation: a genetic perspective. In: *Speciation and its consequences* (eds. Otte D and Endler JA), pp. 3-27. Sinauer Associates, Sunderland, Massachusetts.

Thomas O (1915a) New African rodents and insectivores, mostly collected by Dr Christy for the Congo Museum. *Annals and Magazine of Natural History*, series 8, **16**, 146-152.

Thomas O (1915b) List of mammals (exclusive of Ungulata) collected by Dr Christy for the Congo Museum, Tervuren. *Annals and Magazine of Natural History*, series 8, **16**, 465-481.

Thomas O (1927) On mammals from the Gobabis district, eastern Damaraland, South-West Africa, obtained during Captain Shortridge's fourth Percy Sladen and Kaffrarian Museum expidition, with field notes by the collecter. *Proceedings of the Zoological Society of London*, **1927**, 371-398.

Thomas O, Wroughton RC (1908) The Rudd exploration of South Africa. – X. List of mammals collected by Mr. Grant near Tette, Zambesia. *Proceedings of the Zoological Society of London*, **1908**, 535-553.

Thompson JD, Gibson TJ, Plewniak TF, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876-4882.

Taylor PJ, Meester J (1993) Morphometric variation in the yellow mongoose, Cynictis penicillata (Cuvier, 1829) (Carnivora: Viverridae) in southern Africa. *Durban Museum Novitates*, **18**, 37-71.

Van Zinderen Bakker EM (1957) A pollen analytical investigation of the Florisbad deposits (South Africa). *Proceedings of the Third Pan African Congress on Prehistory, Livingstone*, 1955, pp. 56-67.

Van Zinderen Bakker EM (1978) Quaternary vegetation changes in southern Africa. In: *Biogeography and ecology of southern Africa* (ed. Werger MJA), pp. 131-143. Dr. W Junk Publishers, The Hague.

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

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

Visser DS, Robinson TJ (1986) Cytosystematics of the South African *Aethomys* (Rodentia: Muridae). *South African Journal of Zoology*, **21**, 264-268.

Visser DS, Robinson TJ (1987) Systematic implications of spermatozoan and bacular morphology for the South African *Aethomys*. *Mammalia*, **51**, 447-454.

Wall JD (1999) Recombination and the power of statistical tests of neutrality. *Genetic Research, Cambridge*, **74**, 65-79.

Watson CBR (1987) The comparative ecology of two small mammal communities in the Kruger National Park. MSc thesis, University of Pretoria, Pretoria.

Weir BS (1996) *Genetic data analysis II: Methods for discrete population genetic data.* Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Wesselman H (1984) Omo micromammals. In: *Contributions to vertebrate evolution*, (eds.) MK Hecht and FS Szalay, pp. 219. S Karger, New York.

Wilson VJ (1970) Notes on the breeding and feeding habits of a pair of barn owls, *Tyto alba* (Scopoli) in Rhodesia. *Arnodia, Rhodesia*, **4**, 1-8.

Wilson VJ (1975) Mammals of the Wankie National Park, Rhodesia. *Memorials of the National Museum Monument of Rhodesia*, **5**, 1-147.

Withers PC, Louw GN, Henschel JR (1980) Energetics and water relationships of Namib desert rodents. *South African Journal of Zoology*, **15**, 131-137.

Woodall PF, Mackie RI (1987) Caecal size and function in the rock elephant shrew *Elephantulus myurus* (Insectivora: Microscelidae) and the Namaqua rock mouse *Aethomys namaquensis* (Rodentia: Muridae). *Comparative Biochemistry and Physiology*, **87A**, 311-314.

Yang Z (1996) Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution*, **144**, 1941-1950.

Yang Z, Goldman N, Friday A (1994) Comparrison of models from nucleotide substitution used in maximum-likelihood phylogenetic estimation. *Molecular Biology and Evolution*, **11**, 316-324.

Yang Z, Kumar S (1996) Approximate methods for estimating the pattern of nucleotide substitution and the variation of substitution rates among sites. *Molecular Biology and Evolution*, **13**, 650-659.

Appendix 2.1

Sequence comparisons of 550 base pairs of the 5' end of the mitochondrial DNA cytochrome b gene for 37 haplotypes identified within *Aethomys namaquensis* from southern Africa. Dots (.) indicate identity with the haplotype NH01 sequence and question marks (?) indicate missing nucleotide data. *Mus* = *Mus musculus* and *Das* = *Dasymys incomtus*.

```
NH01  AAAATTATTAATCACTCATTTATTGACCTCCCAGCCCCATCCAACATCTCATCATGATGAAACTTCGGTTCACTTCTTGGAATCTGCTTAATAGTCCAAA
NH02  ...............................A.................................C.............................
NH03  ????????????????????????????????????????.....................................................
NH04  ........................T.....................................................................A......
NH05  ..........................................................................G.....................
NH06  ..........................................................................................
NH08  ...............................A..........................................................
NH31  ..................................................G.............T........C..................A......
NH36  ??????????????..................................G.............T........C..................A......
NH37  ..................................................G.............T.....G..C..................A......
NH34  ??????????????????????????????????????????????.....G.............T........C..................A......
NH26  ..................................................G.............T........C..................A......
NH27  ..................................................G.............T........C..................A......
NH28  ..................................................G.............T........C..................A......
NH29  ...............................A..................G.............T........C..................A......
NH30  ..................................................G.............T........C..................A......
NH23  ..................................................G.............T........C..................A......
NH14  ??????????????????????????????????????????????????????????????????????.....C..................
NH15  ...........C........C.....................................................T.....................
NH16  ...........C........C.....................................................T..C..................
NH07  ...........C........C.....................................................T.....................
NH10  ???????....C........C.....................................................T.....................
NH09  ..................C.....................................................T.....................
NH12  ........C..C........C.....................................................T..C..................
NH13  ........C..C........C.....................................................T..C..................
NH11  ??????????????????????????????????????????????????????????????????????.....T..................
NH32  ..........................................................................................
NH33  ..........................................................................................
NH35  ??????....................................................................................
NH18  ..............T.....................................................G.......................
NH20  ..............T.....................................................................
NH22  ..............T.....................................................................
NH19  ..................................................G.............C.....................
NH17  ...........C......................................................................
NH21  ..................C........T........G.....................................................
NH24  ...........C..................................G.....................................A......
NH25  ?????.....C..................A.............G.....................................A......
Mus   ...........C........C.....A..T.T...........T.............T..G..C.....A...G.....C..........
Das   ........C..C.....C..C..........T.............C.............T..C..C..GT.A...........CTA......
```

```
NH01  TTTTAACTGGCTTATTCTTAGCTATACACTACACCTCCGACACACTAACAGCATTCTCATCAGTAACTCACATCTGCCGAGACGTAAACTACGGCTGACT
NH02  ..........................................................................................
NH03  ..........................................................................................
NH04  ..........................................................................................
NH05  ..........................................................................................
NH06  ..........................................................................................
NH08  ..........................................................................................
NH31  ..........................................................C........T........T..T.........T.
NH36  ..........................................................C........T........T..T.........T.
NH37  ..........................................................C........T........T..T.........T.
NH34  ..........C...............................................C........T........T..T.........T.
NH26  ..........................................................C........T........T..T.........T.
NH27  ..........................................................C........T........T..T.........T.
NH28  ..........................................................C........T........T..T.........T.
NH29  ..........................................................C........T........T..T.........T.
NH30  ..........................................................C........T........T..T.........T.
NH23  ..........................................................C........T........T..T.........T.
NH14  ........................G.................................................T........T..T.........T.
NH15  ........................G.................................................T........T..T.........T.
NH16  ........................G.................................................T........T..T.........T.
NH07  ..........................................................................T........T..T.........T.
NH10  ..........................................................................T........T..T.........T.
NH09  ..........................................................................T........T..T.........T.
NH12  ..........C...............................................................T........T..T.........T.
NH13  ..........................................................................T........T..T.........T.
NH11  ..........................................................................T........T..T.........T.
NH32  ..............................T.....................T.....................T........T..T.........T.
NH33  ..............................T.....................T.....................T........T..T.........T.
NH35  ..............................T.....................T.....................T........T..T.........T.
NH18  .......C..................................................................T.....T..T...........T.
NH20  .......C..................................................................T.....T..T...........T.
NH22  .......C..................................................................T.....T..T...........T.
NH19  .......C..................................................................T........T..T.........T.
NH17  ..........................................................................T........T..T.........T.
NH21  ..........................................................C.....T........T..T.........T.
NH24  ..........................................................................T........T..T...........
NH25  ..........................................................................T........T..T...........
Mus   .CA.T..A..TC.T.......C..........A..A..T...A.....C..T..........A.....T..T...........T.....G.....
Das   ..A.C..A...C....TC.............A..A..T...A.....C.............C..T.................T.....A.....
```

```
NH01    AGTACGATATATACACGCAAATGGAGCCTCAATATTCTTCATCTGCCTATTTATACACGTAGGACGAGGTATATACTACGGATCATACACATTTCTAGAA
NH02    ...........................................................................................................
NH03    ...........................................................................................................
NH04    ...........................................................................................................
NH05    ...........................................................................................................
NH06    ...........................................................................................................
NH08    ...........................................................................................................
NH31    C..........................................................C..G.....................T.....
NH36    G..........................................................C..G.....................T.....
NH37    G..........................................................C..G.....................T.....
NH34    G..........................................................C..G.....................T.....
NH26    G........C..................................................C..G.....................T.....
NH27    G........C..................................................C..G.....................T.....
NH28    G..........................................................C..G.....................T.....
NH29    G..........................................................C..G.....................T.....
NH30    G..........................................................C..G.....................T.....
NH23    G..........................................................C..G.....................T.....
NH14    ........................G..................................................................T.....
NH15    ............G...........G..................................................................T.....
NH16    ........................G..................................................................T.....
NH07    .A.......C.............G..................................G...............................T.....
NH10    .........C.............G..................................G...............................T.....
NH09    .........C.............G..................................G...............................T.....
NH12    ........................G..................................G...............................T.....
NH13    ........................G..................................G...............................T.....
NH11    ........................G..................................G...............................T.....
NH32    ...........................................................G.....A.......................T.....
NH33    ...........................................................G...........................T.....
NH35    ...........................................................G...........................T.....
NH18    G..........................................................G...........................T.....
NH20    G..........................................................G...........................T.....
NH22    G............G.............................................G...........................T.....
NH19    ..........................................................C..G...........G.........T.....
NH17    G..........................................................................................T.....
NH21    G..........................................................................................T.....
NH24    ..................................T.....................................T.................C.....
NH25    G.................................T.....................................G..T.................C.....
Mus     .A.C..................C..............T..T..T...T....CC.T..T..C........CT....T..T........T......A.....
Das     ..................T..C..............T..T......T.....C.C...........G..G..............C..T..G..C.....
```

```
NH01    ACATGAAACATCGGAGTAATCCTTTTATTTACAGTAATAGCCACCGCATTTATAGGTTATGTACTCCCATGAGGACAAATATCATTTTGAGGAGCAACAG
NH02    ...........................................................................................................
NH03    ................G.....................................................G........................................
NH04    ................G...........................................................................................
NH05    ................G...........................................................................................
NH06    ..................................................C........................................................
NH08    ................G...........................................................................................
NH31    ....................T.....................................................T..................
NH36    ....................T.....................................................T..................
NH37    ....................T.....................................................T..................
NH34    ....................T.....................................................T..................
NH26    ....................T.....................................................T..................
NH27    ....................T.....................................................T..................
NH28    ....................T.....................................................T..................
NH29    ....................T.....................................................T..................
NH30    ....................T.....................................................T..................
NH23    ....................T.....................................................T..................
NH14    ..............G.....T..................T.....G...........................................
NH15    ..............G.....T..................T.....G...........................................
NH16    ..............G.....T..................T.....G...........................................
NH07    ..............G.....T........C.........T...................................................
NH10    ..............G.....T..................T...................................................
NH09    ..............G.....T..................T...................................................
NH12    ........T.....G.....T..................T.....................................G.............
NH13    ........T.....G.....T..................T.....................................G.............
NH11    ....................T..................T...................................................
NH32    ...........T........T..................T.....................G.............................G.............
NH33    ...........T........T..................T.....................C.............................G.............
NH35    ...........T........T..................T.....................C.....................................
NH18    ....................T.....................................................G.....................
NH20    ....................T...................................................................
NH22    ................G..T.....................................................G.....................
NH19    ........T..T........T..................T.....................G.....................G.............
NH17    ..............G.....T..C..............T...................................................
NH21    ..............G.....T..................T...................................................
NH24    ....................T..................T...................................................
NH25    ..............G.....T..................T........C..........................................
Mus     ..C........T......C.T..AC.G..CG....C........G...........C..C..C..T.....................C.....T..C....
Das     ..................C....CC....CG....C........T.....C.....C..C..C..A.....................C.............
```

```
NH01    TAATCACAAACTTACTATCGGCCATCCCATACATTGGAACTACTTTAGTAGAATGAATTTGAGGGGGATTCTCAGTAGATAAAGCTACCCTAACCCGATT
NH02    ...........................................................A..........................................
NH03    ...........................................................A..........................................
NH04    ...........................................................A..........................................
NH05    ...........................................................A..........................................
NH06    ...........................................................A..........................................
NH08    ...........................................................A..........................................
NH31    ....T.............T.................G.....................C..........................................
NH36    ....T.............T.................G.....................C.....T....................................
NH37    ....T.............T.................G.....................C.....T....................................
NH34    ....T.............T.................G.....................C.....T....................................
NH26    ....T.............T.....................................C..........................................
NH27    ....T.............T.................G.....................C..........................................
NH28    ....T.............T.................G.....................C..........................................
NH29    ....T.............T.................G.....................C..........................................
NH30    ....T.............T.................G.....................C..........................................
NH23    ....T.............T.................G.....................C.............................T............
NH14    ....T.............T.............C.....................C.....A.....................C..T............
NH15    ....T.............T.............C......G..............C.....A.....................C..T............
NH16    ....T.............T.............C.....????????????????????????????????????????????????????????
NH07    ....T.............C.............C.....................C.....A.....................C..............
NH10    ....T.............C.............C.....................C.....A.....................C..............
NH09    ....T.............C.............C.....................C.....A.....................C..............
NH12    ....T.............C.............C.....................C...........................C..............
NH13    ....T.............C.............C.....................C...........................C..............
NH11    ....T.............C.............C.....................C.....A.....................C..............
NH32    ....T.........T....T.....................................G...........A.....T.............C..............
NH33    ....T.........T....T.....................................G...........A.....T.............C..............
NH35    ....T.........T....T.....................................G...........A.....T.............C..............
NH18    ....T.............T...........T.........................C...........................G..............
NH20    ....T.............T...........T.........................C...........................G..............
NH22    ....T.............T...........T.........................C...........................G..............
NH19    ....T.............T.....................G...............C..........................................
NH17    ....T.............T..A.....................................C.....A.....................C..............
NH21    ....T.............T.....T.....................................C.....A.....................C..............
NH24    ....T.............T.....................C...............C..........................................
NH25    ....T.............T.....................C.............C..........................................
Mus     .T..T......C.C.....A...........T........A..CC....C.............A..C...........C.....C...T.G........
Das     ..........C....T..A..T...........C..C.....CC.............C.....C..C...........C.....C.........T.....
```

```
NH01    CTTTGCATTCCACTTCATTTTACCATTCATCATTACAGCTCTCGTAGTTG
NH02    ...................................................
NH03    ...................................................
NH04    ...................................................
NH05    ...................................................
NH06    ...................................................
NH08    ...............................................A...
NH31    ...C..............C.................C......A.C.
NH36    T..C..............C.................C......A.C.
NH37    T..C..............C.................C......A.C.
NH34    T..C..............C.................C......A.C.
NH26    T..C........T......C.................C......A.C.
NH27    T..C........T......C.................C......A.C.
NH28    T..C..............C.................C......A.C.
NH29    T..C..............C.................C......A.C.
NH30    ...C..............C.................C....??????
NH23    ...C..............C.................C......A.C.
NH14    ...C.....T............G.............C......A...
NH15    ...C.....T............G.............C......A...
NH16    ???????????????????????????????????????????????
NH07    ...C.....T.............................C......A...
NH10    ...C.....T.....................T........C...???????
NH09    ...C.....T.............................???????????
NH12    ...C.....T.....T.....................C......A...
NH13    ...C.....T.....T.....................C......A...
NH11    ...C.....T.....T.................??????????????
NH32    ...C.....T.....................???????????????
NH33    ...C.....T.............................C......A...
NH35    ...C.....T.............................C......A...
NH18    ...C..G................................C......A...
NH20    ...C..G................................C..........
NH22    ...C...................................C..........
NH19    ...C.....T.........C...................C......A...
NH17    ...C...................................C..T...A...
NH21    ...C.....T.............................C.??????????
NH24    ...C.....T.................T..T........C......A.C.
NH25    ...C.....T.................T..T........C......A.C.
Mus     ...C..T...........C........T..T..CG.G..C..A.C.A.C.
Das     ...C..............CC.T.................C..T...A...
```

Appendix 3.1

Sequence comparisons of 370 base pairs of the 5' end of the mitochondrial DNA cytochrome b gene for 21 haplotypes identified within *Aethomys ineptus* (H01-H16) and *A. chrysophilus* (H17-H21) from South Africa and Swaziland. Dots (.) indicate identity with the haplotype H01 sequence and question marks (?) indicate missing nucleotide data.

```
H01    ?????TAGGAGTTTGCTTAATAGTACAAATCATCACAGGTCTATTCTTAGCCATACATTA
H02    ???AC.....................................................
H03    CCTAC.....................................................
H04    CCTAT.....................................................
H05    CCTAC.....A...............................................
H06    CCTAC.....................................................
H07    CCTAC.....................................................
H08    CCTAC.....................................................
H09    ?????.....................................................
H10    CCTAC.....................................................
H11    CCTAC.....................................................
H12    CCTAC...................................C.................
H13    CCTAC.....................................................
H14    CCTAC...................................T.................
H15    CCTAC.....................................................
H16    CCTAC.....................................................
H17    CCTAC.....................................................
H18    CCTAC.......C.............................................
H19    CCTAC.......C.............................................
H20    CCTAC.....A.C.....G.......................................
H21    CCTAC.....................................................
```

```
H01    TACATCAGACACGACAACAGCATTTTCATCAGTAACCCATATCTGTCGAGACGTAAATTA
H02    ...........................................................
H03    ...........................................................
H04    ...........................................................
H05    ...........................................................
H06    ...........................................................
H07    ...........................................................
H08    ...................................G.......................
H09    ............A..............................................
H10    ...........................................................
H11    ...........................................................
H12    ...........................................................
H13    ...........................................................
H14    ...........................................................
H15    ............A..............................................
H16    ...........................................................
H17    ...........................................T...............
H18    ...........................................T...............
H19    C..........................................T...............
H20    ...........................................T...............
H21    ...........................................T...............
```

```
H01    TGGATGATTAATCCGATATATACACGCAAACGGAGCCTCAATATTCTTCATCTGCTTATT
H02    ...G.....................................................
H03    C..G.....................................................
H04    ...G.....................................................
H05    ...G.....................................................
H06    ...G.....................................................
H07    C........................................................
H08    .........................................................
H09    .........................................................
H10    ...G.....................................................T........
H11    ...G.....................................................T........
H12    ...G.....................................................T........
H13    .........................................................
H14    .........................................................
H15    .........................................................T........
H16    ...G.....................................................T........
H17    ...G...C.................T.....................................
H18    .......C.................T.....................................
H19    .......C.................T.....................................
H20    .......C.................T.....................................
H21    ...G...C.................T.....................................
```

```
H01    CCTTCATGTAGGACGAGGAATATACTATGGATCTTACACATTCCTAGAAACATGAAATAT
H02    ............................................................
H03    ............................................................
H04    ............................................................
H05    ............................................................
H06    ............................................................
H07    ............................................................
H08    ............................................................
H09    ............................................................
H10    ............................................................
H11    ............................................................
H12    ............................................................
H13    ............................................................
H14    ............................................................
H15    ............................................................
H16    ............................................................
H17    ..................................C.........................
H18    ............................................................
H19    ............................................................
H20    ............................................................
H21    ............................................................
```

```
H01    CGGAGTAGTGTTACTATTTACAGTCATAGCCACTGCATTCATAGGATATGTTCTTCCATG
H02    ............................................................
H03    ............................................................
H04    ............................................................
H05    ............................................................
H06    ..................T.........................................
H07    ............................................................
H08    ............................................................
H09    ............................................................
H10    .........A..................................................
H11    ............................................................
H12    ............................................................
H13    ...........................................................C........
H14    ...........................................................C........
H15    ............................................................
H16    ............................................................
H17    .........C...T..............................................
H18    .........A...T..............................................C........
H19    ...G.....A...T..............................................C........
H20    ............T..............................................C........
H21    .........C...T..............................................
```

```
H01    AGGACAAATATCATTCTGAGGAGCCACAGTAATTACAAATTTACTTTCAGCTATCCCATA
H02    ............................................................
H03    ............................................................
H04    ............................................................
H05    ............................................................
H06    ............................................................
H07    ............................................................
H08    ............................................................
H09    ..........................................C.................
H10    ...G........................................................
H11    ........................T...................................
H12    ............................................................
H13    ............................................................
H14    ............................................................
H15    ............................................................
H16    ............................................................
H17    ........................................C...................
H18    ........................................C...................
H19    ....................................??????????????????????????
H20    ........................................C...................
H21    ........................................C...................
```

```
H01     TATTGGCACA
H02     ..........
H03     ..........
H04     ..........
H05     ..........
H06     ..........
H07     ..........
H08     ..........
H09     ..........
H10     ..........
H11     ..........
H12     ..........
H13     ..........
H14     ..........
H15     ..........
H16     ..........
H17     ..........
H18     ..........
H19     ??????????
H20     ..........
H21     ..........
```

Appendix 4.1

Morhpological characters (Chimimba 1997) included in the combined analysis in the elucidation of phylogenetic relationships within southern African *Aethomys*.

	Character numbers																					
Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Aethomys namaquensis	1	0	1	1	0	1	0	2	2	2	2	1	2	1	1	1	2	0	0	0	0	0
A. granti	1	0	1	1	0	1	0	2	2	2	2	1	2	1	2	2	2	1	1	1	0	0
A. ineptus	2	2	2	2	1	0	0	1	2	1	2	1	2	2	2	2	2	0	0	1	0	1
A. chrysophilus	2	2	1	2	1	0	0	1	2	1	2	1	2	2	2	2	2	0	0	1	0	1
Dasymys incomtus	3	3	1	3	1	0	0	1	1	1	1	1	1	1	2	1	2	1	0	0	0	1

Appendix 4.2

Sequence comparisons of 489 base pairs of the 5' end of the mitochondrial DNA cytochrome b gene for 13 haplotypes identified within *Aethomys namaquensis* (NH16, NH29, NH30), *A. granti* (GH01-GH03), *A. ineptus* (H01, H11, H12, H15, H16) and *A. chrysophilus* (H18, H19) from South Africa and Swaziland. Dots (.) indicate identity with the haplotype NH16 sequence and question marks (?) indicate missing nucleotide data.

```
NH16    AAAATTATTAACCACTCATTCATTGACCTCCCAGCCCCATCCAACATCTCATCATGATGA
NH29    ..........T........T...........A...............G.........
NH30    ..........T........T.............................G.........
GH01    ..............T.....T........A................T............
GH02    .....C........T.....T........A................T............
GH03    ..............T.....T........A................T............
H01     .....C.....T.....C.............C..A.........................
H11     .....C.....T.....C.............C..A.........................
H12     .....C.....T.....C.............C..A.........................
H15     .....C.....T.....C.............C..A.........................
H16     ????????????????????.....C..A.........................
H18     .....C.....T.....C.............C..A.........................
H19     .....C.....T.....C.............C..A.........................

NH16    AACTTTGGCTCACTTCTTGGAATCTGCTTAATAGTCCAAATTTTAACTGGCTTATTCTTA
NH29    ........T.....C.................A.............................
NH30    ........T.....C.................A.............................
GH01    ..............AT.A.........C..T..A.A.....CA....A...C.........
GH02    ..............AT.A.........C..C..A.A......A....A...C.........
GH03    .....C........AT.A.........C..T..A.A.....CA....A...C.........
H01     .....C.....C..A..A...G.T...........A.....CA.C..A..TC........
H11     .....C.....C..A..A...G.T...........A.....CA.C..A..TC........
H12     .....C.....C..A..A...G.T...........A.....CA.C..A...C........
H15     .....C.....C..A..A...G.T...........A.....CA.C..A..TC........
H16     .....C.....C..A..A...G.T...........A.....CA.C..A..TC........
H18     .....C.....C..A..A...G.............A.....CA.C..A..TC........
H19     .....C.....C..A..A...G.............A.....CA.C..A..TC........

NH16    GCTATGCACTACACCTCCGACACACTAACAGCATTCTCATCAGTAACTCACATCTGTCGA
NH29    .....A.....................................................C............
NH30    .....A.....................................................C............
GH01    ..C..A........T.........AC.....................C..T.....C...
GH02    ..C..A........T.........AC.....................C..T.....C...
GH03    ..C..A........T.........AC.....................C..T.....C...
H01     ..C..A..T..T..A..A.....GAC.........T...........C..T.........
H11     ..C..A..T..T..A..A.....GAC.........T...........C..T.........
H12     ..C..A..T..T..A..A.....GAC.........T...........C..T.........
H15     ..C..A..T..T..A..A......AC.........T...........C..T.........
H16     ..C..A..T..T..A..A.....GAC.........T...........C..T.........
H18     ..C..A..T..T..A..A.....GAC.........T...........C..T..T......
H19     ..C..A..T.....A..A.....GAC.........T...........C..T..T......
```

```
NH16    GACGTTAATTACGGCTGATTAGTACGATATATACACGCAAATGGGGCCTCAATATTCTTC
NH29    ....................G.......................A..............
NH30    ....................G.......................A..............
GH01    ..T..C..C..T..T........C...........T........A.....C........
GH02    ..T..C..C..T..T........T.....C..G..T........A.....T........
GH03    ..T..C..C..T..T........C...........T........A.....C........
H01     .....A.....T..A......A.C.................C..A..............
H11     .....A.....T..G......A.C.................C..A..............
H12     .....A.....T..G......A.C.................C..A..............
H15     .....A.....T..A......A.C.................C..A..............
H16     .....A.....T..G......A.C.................C..A..............
H18     .....A.....T..A...C..A.C...........T.....C..A..............
H19     .....A.....T..A...C..A.C...........T.....C..A..............

NH16    ATCTGCCTATTTATACACGTAGGACGAGGTATATACTACGGATCATACACATTTTTAGAA
NH29    ............................C..G.............................
NH30    ............................C..G.............................
GH01    ........C...................C........T.....C..T......C.....
GH02    ........C...................C........T.....C..T......C.....
GH03    ........C...................C........T.....C..T......C.....
H01     ......T....CC.T..T...........A........T.....T........CC.....
H11     ..T...T....CC.T..T...........A........T.....T........CC.....
H12     ..T...T....CC.T..T...........A........T.....T........CC.....
H15     ..T...T....CC.T..T...........A........T.....T........CC.....
H16     ..T...T....CC.T..T...........A........T.....T........CC.....
H18     ......T....CC.T..T...........A........T.....T........CC.....
H19     ......T....CC.T..T...........A........T.....T........CC.....

NH16    ACATGAAACATCGGGGTAATTCTTTTATTTACAGTAATAGCTACCGCGTTTATAGGTTAT
NH29    ..............A.............................C.....A...........
NH30    ..............A.............................C.....A...........
GH01    ...........T..A........C.....C.....T.....C..T..A..C.....G...
GH02    ...........T..A........C.....C.....T.....C..T..A..C.....G...
GH03    ...........T..A........C.....C.....T.....C..T..A..C.....G...
H01     ........T.....A...G.GT.AC..........C.....C..T..A..C.....A...
H11     ........T.....A...G.GT.AC..........C.....C..T..A..C.....A...
H12     ........T.....A...G.GT.AC..........C.....C..T..A..C.....A...
H15     ........T.....A...G.GT.AC..........C.....C..T..A..C.....A...
H16     ........T.....A...G.GT.AC..........C.....C..T..A..C.....A...
H18     ........T.....A...G.AT.A...........C.....C..T..A..C.....A...
H19     ........T.........G.AT.A...........C.....C..T..A..C.....A...

NH16    GTACTCCCATGAGGACAAATATCATTTTGAGGAGCAACAGTAATTACAAACTTACTATCT
NH29    ........................T.....................................
NH30    ........................T.....................................
GH01    ..............................................T...T....A
GH02    .....T........................................T...T....A
GH03    ..............................................T...T....A
H01     ..T..T.....................C........C.............T.....T..A
H11     ..T..T.....................C........T.............T.....T..A
H12     ..T..T.....................C........C.............T.....T..A
H15     ..T..T.....................C........C.............T.....T..A
H16     ..T..T.....................C........C.............T.....T..A
H18     ..C..T.....................C........C.....................T..A
H19     ..C..T.....................C........C.....................T..A
```

```
NH16    GCCATCCCATACATCGGAACTACTTTA??????????????????????????????????
NH29    ..............T..G.........GTAGAATGAATCTGAGGGGGATTCTCAGTAGAT
NH30    ..............T..G.........GTAGAATGAATCTGAGGGGGATTCTCAGTAGAT
GH01    .....T.....T..T.....A..A...GTAGAATGAATTTGAGGTGGATTTTCAGTAGAT
GH02    .....T.....T..T.....A..A...GTAGAATGAATTTGAGGTGGATTTTCAGTAGAC
GH03    .....T.....T..T.....A..A...GTAGAATGAATTTGAGGTGGATTTTCAGTAGAT
H01     ..T........T..T..C..A.TCC..GTAGAATGAATCTGAGGTGGATTCTCAGTAGAC
H11     ..T........T..T..C..A.TCC..??????????????????????????????????
H12     ..T........T..T..C..A.TCC..GTAGAATGAATCTGAGGTGGATTCTCAGTAGAC
H15     ..T........T..T..C..A.TCC..GTAGAATGAATCTGAGGTGGATTCTCAGTAGAC
H16     ..T........T..T..C..A.TCC..GTAGAATGAATCTGAGGTGGATTCTCAGTAGAC
H18     ..T........T..T..C..A.TCC..GTAGAATGAATCTGAGGTGGATTCTCAGTAGAC
H19     ..T........T..T..C..A.TCC..GTAGAATGAATCTGAGGTGGA???????????
```

```
NH16    ?????????
NH29    AAAGCTACC
NH30    AAAGCTACC
GH01    AAAGCTACA
GH02    AAAGCTACA
GH03    AAAGCTACA
H01     AAAGCTACT
H11     ?????????
H12     AAAGCTACT
H15     AAAGCTACT
H16     AAAGCTACT
H18     AAAGCTACT
H19     ?????????
```

Appendix 4.3

Sequence comparisons of 526 base pairs of the 3' end of the mitochondrial DNA 16S rRNA gene for nine haplotypes identified within *Aethomys namaquensis* (NH16, NH29, NH30), *A. granti* (GH01, GH03), *A. chrysophilus* (H18, H19) and *A. ineptus* (H01, H16) from South Africa and Swaziland. Dots (.) indicate identity with the haplotype NH16 sequence and question marks (?) indicate missing nucleotide data. Dashes (-) indicate alignment gaps.

```
NH16    AGTATTAGAGGCATTGCCTGCCCAGTGACTTAAGTTAAACGGCCGCGGTATCCTGACCGT
NH29    ?...........................................................
NH30    ...........................................................
GH01    ...........................................................
GH03    ...........................................................
H01     ............CC.............A.................T........
H16     ............CC.............A.................T........
H18     ............CC.............A.................T........
H19     ???.........CC.............A.................T........

NH16    GCAAAGGTAGCATAATCACTTGTTCCTTAATTAGGGACTAGAATGAATGGCTAAACGAAG
NH29    ..........................................C............
NH30    ..........................................C............
GH01    ...........................................................
GH03    ...........................................................
H01     .....................................T.....C.....G....G.
H16     .....................................T.....C.....G....G.
H18     .....................................T.....C.....G....G.
H19     .....................................T.....C.....G....G.

NH16    GTTCAACTGTCTCTTATTCTTAATCAGTGAAATTGACCTTCCAGTGAAGAGGCTGGAATA
NH29    ...........................................................
NH30    ...........................................................
GH01    ..................C........................................
GH03    ..................C........................................
H01     .A.........C......C........................................
H16     .A.........C......C........................................
H18     .A.........C......C........................................
H19     .A.........C......C........................................

NH16    AAACAATAAGACGAGAAGACCCTGTGGAGCTTGAATAAACTGACTTAATTATTAAACAAA
NH29    ...T.............................G................A...
NH30    ...T.............................G................-...
GH01    ...T.........................A.....T.A.........C.......
GH03    ...T.........................A.....T.A.........C.......
H01     T.TA................A........C...T....AG........CA...A.-.
H16     T.TA................A........C...T....AG........CA...-...
H18     T.TA................A........C...T....AG........-C...A...
H19     T.TA................A........C...T....AG........CA...-...
```

```
NH16    ACTT-CCTAGTGGAATAAAATATATAA--TATAAGTCAT-ATATTTCGGTTGGGGTGACC
NH29    ...A-.................A....--.........-.....................
NH30    ...A-.................A....--.........-.....................
GH01    .T.A-....A.....C....A.A....--.......T..-.A....A.............
GH03    .T.A-....A.....C....A.A....--.......T..-.A....A.............
H01     .TAA-....A...GT.....ACG.A..--C.....CT.A-.A..................
H16     .TAA-....A...GT.....ACG.A..--C.....CT.A-.A..................
H18     .TAA-....A....C.....ACG.A..--C.....CT.A-.A....T.............
H19     .TAA-....A....C.....ACG.A.T--C.....CT.A-.A..................
```

```
NH16    TCGGAGAATAAAAAAACCTCCGAATGATTCTAGCATAG------ACC-------AACAAG
NH29    .....................................------...-------......
NH30    .....................................------...-------......
GH01    ........C......T.............T........------...-------......
GH03    ........C......T.............T...T....------...-------......
H01     ........C..............C....T........------..T-------......
H16     ........C..............C....T........------..T-------......
H18     ........C..............C....T........------..T-------......
H19     ........C..............C....T........------..T-------......
```

```
NH16    TCAAAACAAATTTTAAA-ATCTCATTGACCCAAATC-ATTTTGATCAACGGACCAAGTTA
NH29    ............C....-....T...........AA-........................
NH30    .....G......C....-....T...........AA-........................
GH01    ..........C.AC...-....A...........GT-........................
GH03    ..........C.AC...-....A...........GT-........................
H01     .....G...T.AACTT.-T...T...........AT-.A......................
H16     .....G...T.AA.TT.-T...T...........AT-.A......................
H18     .....G.....AACTT.-T...T...........AT-.A......................
H19     .....G.....AACTT.-T...T...........AT-.A......................
```

```
NH16    CCCCAGGGATAACAGCGCAATCCTATTCAAGAGTTCATATCGACAATTAGGGTTTACGAC
NH29    ...T.....................................................
NH30    ...T.....................................................
GH01    ........................T.................................
GH03    ........................T.................................
H01     ...T.....................T......C.........................
H16     ...T.....................T......C.........................
H18     ...T.....................T.................................
H19     ...T.....................T.................................
```

```
NH16    CTCGATGTTGGATCAGGACATCCCAATGGTGCAGAAGCTATTAATG
NH29    .............................................
NH30    .............................................
GH01    .............................................
GH03    .............................................
H01     .............................................
H16     .............................................
H18     .............................................
H19     .............................................
```