Chapter 1

General Introduction
1. General

The evolutionary history of species and populations is the product of processes occurring over two time scales: evolutionary time - being based on broad-scale changes under specific environmental conditions with associated selective pressures and ecological time - over which population processes (e.g., demographic changes, migration, local extinction and colonisation) occur (Martin and Simon, 1990; Carroll et al., 2007). Evolutionary biology aims at unraveling these interactions and assessing the importance of the processes. Understanding evolutionary processes is brought about through the study of closely related taxa (Martin and Simon, 1990). Genetic structure of a population can therefore be correlated with both biogeographical factors as well as ecological and demographic processes (Carisio et al., 2004). Our understanding of species formation is based on population level comparisons; by examining the variation among these populations, their historical associations and the processes of genetic restructuring that may have lead to speciation can often be revealed (Knowles 2004; Knowles and Maddison 2002; Wright, 1931).

Biological diversity can be meaningfully divided into least common evolutionary denominators, namely, the ‘species’ (Hendry et al., 2000). Delineating distinct species is often problematic but most biologists agree with Mayr (1957) that “the living world comprises more or less distinct entities that we call species”. Taxonomic groups have historically been identified using morphological criteria but over the last couple of decades molecular techniques have provided a powerful tool for evaluating the validity of taxonomic units (Avise and Walker, 1999). Avise and Walker (1999) used patterns of mitochondrial DNA (mtDNA) variation to argue that mtDNA discontinuities and traditional taxonomic designations tend to converge which in turn may reveal real biotic units.

Speciation has been described as the evolutionary process by which new biological species arise (Mayr, 1942). Geographic models of speciation in nature have been described based on the extent to which populations are geographically isolated: allopatric (physical barrier separates populations), peripatric (species are formed in isolated, small peripheral populations that are prevented from exchanging genes with the main population),
parapatric (zones of two diverging populations are separate but do overlap) and sympatric (population sharing a geographic location is forced by environmental factors to diverge). The large genetic and phenotypic diversity observed within species is necessary for evolution to create new reproductively isolated species (Härdling et al., 2009). Although most biologist believe that reproductive isolation is the driving force behind species, it has been shown that reproductive isolation alone is not sufficient to permit coexistence of two species at the same locality (Mayr, 1949). Species should also be different in their ecological requirements to avoid competition (Crombie, 1947). Speciation then means the evolution of reproductive isolation as well as of ecological differentiation between populations. To this day, processes and mechanisms involved in speciation are still much debated.

There are many species concepts such as the phylogenetic species concept (PSC), the biological species concept (BSC), the evolutionary species concept (ESC), the cohesion species concept (CSC), the ecological species concept (ESC), the genetics species concept (GSC) and many others (Cracraft, 1989; Dobzhansky, 1940; Mayr, 1940; Mayr, 1942; Simpson, 1961; Schluter, 1998; Schluter, 2001; Templeton, 1989; Wiley, 1978). Regardless of the species concept (reviewed by De Queiroz, 2007; Mayden, 1997) chosen, biologists are confronted with the question of how much difference (or amount of isolation) defines a species (Hendry et al., 2000). The BSC states that “species are groups of interbreeding populations which are reproductively isolated from other such groups” (Mayr, 1940). If 100% reproductive isolation is used as a criterion for applying the BSC, then the identification of species would be relatively straight-forward (Hendry et al., 2000). On the other hand, if the BSC was universally adopted, many of the current taxonomic species would no longer be recognised due to hybridisation and introgression (see Petit and Excoffier, 2009) in populations in the wild (Hendry et al., 2000; Niemiller et al., 2008; Nosil, 2008).

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, 1997; Mayr and Ashlock, 1991). The problem is exacerbated by the numerous definitions that have been proposed to define species as mentioned above. For example, the BSC defines species as “interbreeding natural populations that are reproductively isolated form other such
populations” (Dobzhansky, 1940; Mayr, 1942), while the PSC (Cracraft, 1989) considers a species as “a cluster of organisms, diagnosably distinct from other such clusters”. The cohesion species concept (CSC) 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). Ecological speciation states that natural selection on traits between populations in different environments leads to the evolution of reproductive isolation and as a consequence species (Schluter, 2001). The genetic species concept defines a genetic species as a group of genetically compatible interbreeding natural populations that is genetically isolated from other groups (Baker and Bradley, 2006). The GSC differs from the BSC that the focus is rather on generic isolation than reproductive isolation (Baker and Baker, 2006). The BSC and its variations, is the most widely used in mammals and are thus followed in this study. Some of the variation observed in the present study was explained using the ESC and GSC.

Biologists should avoid the notion of discrete categories into which organisms should be allocated and rather emphasise the level of variation within and among groups of organisms (Hendry et al., 2000). Species descriptions should thus be based on the geographic distribution of genetic diversity (Buth and Mayden, 1981; Johns and Avise, 1998). In addition, multidisciplinary approaches should preferentially be used in the verification/identification of valid biological species such as a combination of cytogenetics, classical morphology and morphometrics together with multi-gene molecular data (Huchon et al., 2005; Taylor et al., 2009).

Similarly, several problems exist in defining subspecies because of different criteria that have been proposed (Moritz et al., 1987; O’Brien and Mayr, 1991; Ryder, 1986). 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) (reviewed by Fraser and Bernatchez 2001; Moritz, 1994b; Ryder, 1986; Ryder et al., 1988). The role of ESUs and MUs in biological classification has generally been limited to recognising intraspecific diversity, which is useful in the conservation and management of endangered and exploited species (Moritz, 1994a).

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 BSC, this subspecies definition may be useful in cases where it is difficult to test for reproductive isolation (Lidicker, 1962). Diversification at specific and subspecific levels is driven by a combination of intrinsic and extrinsic processes (deLong, 1967; Martens, 1997; Rogers and Bernatchez, 2007).

Detailed paleo-climatic records, extraordinary fossil discoveries and advanced analysis of extant fossil data have focused on the possible role that changes in African climate may have had in the evolutionary history of African mammalian fauna (deMenocal, 2004). Biological traits may evolve under the influence of a variety of selective forces such as environmental constraints (Ryan and Brenowitz, 1985). Large scale shifts in climate have altered the ecological composition of the landscape which, in turn, manifested specific faunal adaptation or speciation pressures leading to genetic selection (deMenocal, 2004) reflecting a long history of responses to habitat changes (Riddle, 1996). Small mammals would be greatly affected by environmental variables and climatic changes, such as those seen across the African continent and thus represent good models to understand the evolutionary past and to make predictions about potential future changes in the face of natural and anthropogenic environmental changes.

More specifically, the present study focuses on the southern African subregion which consists of a wide range of biomes. Although other geographical features like altitude and precipitation may play a role in the diversification between *M. namaquensis* lineages, I used biomes/bioregions for comparison with previous work (Chimimba, 2001a). The relatively moist, mostly winter-rainfall region, includes the Fynbos biome in the west. The drier Succulent Karoo biome forms the smallest of the world’s six floristic kingdoms (Cox, 2001) and is distributed across the sandy lowlands of the south-western Cape (Mucina and Rutherford, 2006). It also occurs in the Richtersveld, Namaqualand and the Little Karoo. Savanna from the summer-rainfall region on the north and east of the subregion represents the southern extension of the largest biome in Africa (Mucina and Rutherford, 2006). The summer-rainfall Grassland biome occurs on the cooler, elevated interior of South Africa (Low and Rebele, 1996). The mostly summer-rainfall Nama-Karoo biome is possibly the least species-rich and is confined to the western parts of the subregion. Desert occupies the
north-western parts of southern Africa. Thicket represents an unusual structural, floristic and evolutionary ancient biome on the subcontinent (Low and Rebelo, 1996). Afrotemperate forest in southern Africa are highly distinctive and are characterised by their small and patchy occurrence over the wetter parts of both the winter and summer-rainfall areas of the region (Mucina and Rutherford, 2006). A biome map is presented in Chapter 2 (see Fig. 2.1 for details). Both biomes and bioregions were considered for this study. A biome is therefore the largest land community unit recognised at a continental or sub-continental scale. Bioregions are focussed on plant diversity within biomes (i.e. a finer scale of resolution of each biome). Since some of the lineages corresponded to bioregions at a finer scale, I included both biomes and bioregions to explain the relationships between lineages and biomes/bioregions.

Uplifts affected African topography during the late Neogene (8 to 2 MYA) of which the most important topographic structure is the East African Rift System (Sepulchre et al., 2006; Axelrod and Raven, 1978). The western branch of the East African Rift System started to develop during the middle-late Miocene. This event initiated the formation of the central Tanganyika Basin at about 12 to 10 MYA (Cohan et al., 1993). More recent uplifts between 5 and 2 MYA were active in the Tanganyika and Malawi rifts (Ebinger et al., 1993). The major Tanzanian escarpments were present at 3 MYA (Foster et al., 1997). The Karoo plateau in South Africa has been raised during the past 5 million years. These uplifts have not only contributed to changes in climate but have also contributed to two major episodes of rapid speciation in South Africa (Axelrod and Raven, 1978). Uplifts gave rise to strong aridification and paleoenvironmental changes (Sepulchre et al., 2006).

Other extrinsic barriers that may influence genetic diversity are physical barriers like rivers and mountains. It has been suggested that the Orange River/Holgat River may act as a barrier to dispersal and consequently result in the genetic differentiation of insect populations (Sole et al., 2005). More specifically, rivers have also been cited as factors affecting the distribution of African mammal species (Quéroïl et al., 2003; Telfer et al., 2003). Likewise, from a topographic perspective, the Drakensberg mountain range, may act as a barrier to gene flow between populations of small mammals on the eastern and western sides of the escarpment (Chimimba, 2001a).
2. African Rock Rats

2.1 General Background

African rock rats of the genera *Aethomys* Thomas, 1915a, and *Micaelamys* Ellerman, 1941, are long-tailed murid rodents of medium to large size (Chimimba and Bennett, 2005; Meester et al., 1986). The name *Aethomys* is derived from the Greek word *aithos*, meaning sunburnt and *mys*, meaning mouse (De Graaff, 1981; Rosevear, 1969). These genera represent a diverse group of rodents endemic to East, Central and southern Africa, with a marginal extension into West Africa (Chimimba and Bennett, 2005; Musser and Carleton, 2005). Originally, the genera *Aethomys* and *Micaelamys* were considered subgenera within the genus *Aethomys* (Chimimba and Bennett, 2005). Although *A. namaquensis* Smith, 1834, and *A. granti* Wroughton, 1908, have traditionally been allocated to the subgenus *Micaelamys*, recent molecular studies (Castiglia et al., 2003; Ducroz et al., 2001; Russo, 2003) and data on albumin fixation (Watts and Baverstock, 1995) have provided evidence for the paraphyly of the genus. As a consequence, the subgenera *Micaelamys* and *Aethomys* have been elevated to full generic rank (Chimimba and Bennett, 2005). The genus *Micaelamys* includes *M. namaquensis* and *M. granti*. The genus *Aethomys* is currently considered to include nine species, namely *A. silindensis* Roberts, 1938, *A. chrysophilus* De Winton, 1897, *A. ineptus* Thomas and Wroughton, 1908, *A. nyikae* Thomas, 1897, *A. bocagei* Thomas, 1904, *A. hindei* Thomas, 1902, *A. stannarius* Thomas, 1913, *A. thomasi* De Winton, 1897, and *A. kaiser* Noack, 1887, (Chimimba et al., 1999; Musser and Carleton, 2005).

Although there is little biological data available, members of these genera are nocturnal, terrestrial and to some extent arboreal, living either in pairs or in small family units (Chimimba and Bennett, 2005). They utilise a wide variety of covered habitats, ranging from rock crevices, rocky terrains and outcrops, excavated burrows, tree trunks, grassland with some form of cover, to open savanna woodland (Chimimba and Bennett, 2005). 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 (Chimimba and Bennett, 2005). Little is known about their feeding biology, but all species currently included in the two genera are considered to be
omnivorous with a preference for seed, fruit and grain (Chimimba and Bennett, 2005; Watson, 1987; Woodall and Mackie, 1987). *Micaelamys namaquensis* (formerly known as *A. 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 (Chimimba and Bennett, 2005).

Some members may breed throughout the year with an average of three offspring per litter, while others show no evidence of breeding during colder months (Chimimba and Bennett, 2005; Rautenbach, 1978). They have a short generation time (0.16 to 0.33 years) producing as many as four litters by a single female (C.T. Chimimba pers. comm.). They also 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 crops and stored grain (De Graaff, 1981; Smithers, 1971; Wilson, 1970, 1975).

2.2 Systematics

Taxonomically, the genus *Aethomys* has undergone a number of nomenclatural changes (Rosevear, 1969). Thomas (1915a) proposed *Aethomys* as a subgenus of *Rattus* and later elevated it to 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, necessitating 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 and Robinson, 1986; 1987). For example, *A. chrysophilus*, traditionally considered a single species, was shown to contain two electrophoretically distinct cytotypes (*2n* = 50 and “slow” haemoglobin; *2n* = 44 and “fast” haemoglobin) (Baker et al., 1988; Gordon and Rautenbach, 1980; Gordon and Watson, 1986; Visser and Robinson, 1986, 1987). The two cytotypes have also been shown to differ in gross sperm and bacular morphology (Breed et al., 1988). Given the lack of evidence of hybrids, these studies strongly suggested the presence of two sibling species.
Subsequently, Chimimba et al. (1999) undertook an analysis of both morphometric and morphological patterns of inter- and intraspecific variation within the genus (*Aethomys senso lato*) across a more comprehensive geographical coverage in southern Africa than has previously been considered. Chimimba (2005) also undertook a phylogenetic appraisal of the then 11 recognised species. These studies 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. Subsequently, Russo et al. (2006) examined mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) genetic variation in these cryptic murid rodent species. Phylogenetic and phylogeographic analyses showed reciprocal monophyly between populations of the two species in southern Africa, but no support for monophyly of *A. chrysophilus* from southern and eastern Africa (Russo et al., 2006). This suggested that the analysis of mtDNA can be used to distinguish these sister species in southern Africa (Russo et al., 2006).

The above mentioned phylogenetic analysis (Chimimba, 1997, 1998) 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, 1998; Chimimba et al., 1999). 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. Other molecular studies have also suggested that the genus *Aethomys* may be paraphyletic (Castiglia et al., 2003; Lecompte et al., 2008; Russo, 2003). Worth noting is that *M. namaquensis* (formerly known as *A. namaquensis*) was previously placed in the genus *Thallomys* (Ellerman, 1941).

Earlier reports (e.g., Meester et al., 1964; Roberts, 1951) recognised 16 subspecies within *A. namaquensis*. These subspecies designations were made with little or no assessment of patterns of geographic variation over the entire distributitional range of the species in southern Africa. The intraspecific designations were also based on non-statistical comparisons of type material and/or small, geographically restricted samples (Smithers, 1971; Smithers and Wilson, 1979).
Prior to an intraspecific morphometric study within *A. namaquensis* (Chimimba, 2001a), the nature and extent of geographic variation within the species remained unknown. The morphometric analysis of intraspecific variation within *A. namaquensis* suggested the recognition of only four subspecies: *A. n. namaquensis*, Smith, 1834, *A. n. lehocla* Smith, 1836, *A. n. monticularis* Jameson, 1909 and *A. n. alborarius* Peters, 1852 (Chimimba, 2001a). The morphological discontinuities of these suggested subspecies broadly coincided with the major biomes of southern Africa (Chimimba, 2001a). 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, *monticularis* with Grassland, *alborarius* with Savanna and *lehocla* with Nama-Karoo (Chimimba 2001a. Morphometric variation within *A. chrysophilus* on the other hand, suggested the recognition of two subspecies, *A. c. chrysophilus* Thomas and Wroughton, 1908, and *A. c. imago* Thomas, 1927, whose distributions coincided with an altitudinal limit in the eastern parts of southern Africa (Chimimba, 2000). Geographic variation within *A. ineptus* and *A. granti* were 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, 2001b; Chimimba et al., 1998).

A recent molecular study among 16 localities of *A. namaquensis* (currently known as *M. namaquensis*; Russo, 2003) confirmed that the species is polytypic but higher levels of variation than previously detected were revealed. This study showed some support for three of the four morphometrically-defined subspecies from the Chimimba (2001a) study: 1) a lineage found in the Limpopo valley and Botswana corresponding to the Savanna biome of southern Africa; 2) a lineage widely distributed across the Upper/Lower Karoo and 3) a lineage found across the Grassland biome of southern Africa. In addition, several unique and well-supported lineages defined by the molecular data were not concordant with the morphometrically-defined subspecies (Chimimba, 2001a). Furthermore, while some lineages showed considerable molecular sequence variation across the geographic area sampled, other lineages showed very little differentiation. These results lend support to earlier suggestions for the presence of a species complex within *M. namaquensis* from southern Africa.
Although the findings in the recent revision (Chimimba, 2000, 2001a, b, 2005) may be valid, these need to be independently tested using additional character sets before robust systematic conclusions can be drawn. 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 independently tests the findings of the morphologically based systematic revision by Chimimba et al. (1999) and intraspecific hypotheses (Chimimba, 2001a; Meester et al., 1964; Roberts, 1951; Smithers, 1971; Smithers and Wilson, 1979) by using molecular data following both phylogenetic and phylogeographic approaches. The main focus of the present study is on *M. namaquensis* from southern Africa.

2.3 Palaeontology

Fossils representing *A. namaquensis* Smith, 1834 (currently known as *M. namaquensis*; see Chimimba and Bennett 2005), and *A. chrysophilus (senso lato)* De Winton, 1897, have been described from South Africa (Avery, 1981, 1982, 1985; De Graaff, 1960; 1961; 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, were 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 and Krishtalka, 1986; Denys, 1987). There is, however, no close relationship between the East African species and those from South Africa, which would allow speculation on the origin and time of divergence of the genus *Micaelamys* (Denys, 1990a, b).
3. Molecular data and Phylogeny

The present investigation was largely based on the analysis of mtDNA data which are 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 itself and transcribes protein-coding genes within the organelle (Avise and Lansman, 1983; Moritz et al., 1987). Its gene content appears to be conserved, with two ribosomal RNA (rRNA), 22 transfer RNA (tRNA) 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, the history recorded in this molecule is not a complete characterisation of the intraspecific phylogeny of a species because relationships may be obscured by gender bias such as in levels of dispersal (Zhang and Hewitt, 1996).

Despite potential limitations of the cyt b gene, it has proved useful in addressing questions about relationships among and within species for a range of taxa (Ducroz et al., 1998; Fumagalli et al., 1999; Jansen van Vuuren and Robinson, 1997; Nicolas et al., 2008a Nicolas et al., 2008b; Ohdachi et al., 2001; Smith, 1998). More specifically, the gene has successfully been used to investigate systematic relationships in a number of murid rodents (e.g., Ducroz et al., 1998; Ducroz et al., 2001; Galewski et al., 2006; Patton and Smith, 1992; Russo et al., 2006; Smith and Patton, 1993, 1999; Verheyen et al., 1995, 1996), in addition to which its time scale calibrations (Smith and Patton, 1993) and rate of evolution are also well documented (Irwin et al., 1991). More recently, the cyt b gene has successfully been used to investigate the phylogeographic structure of the genus Acomys (Nicolas et al., 2009).

The Recombination Activating Gene 1 (RAG1) gene was also used in the elucidation of phylogenetic relationships within southern African M. namaquensis in the present study. The protein encoded by the RAG1 gene is involved in the rearrangement and recombination of the genes of immunoglobulin and T cell receptor molecules during the process of V-D-J recombination (Wenhui et al., 2001). The cellular expression is restricted to lymphocytes during their developmental stages and the RAG1 gene is therefore essential to the generation of mature B and T lymphocytes, cell types that are important
components of the adaptive immune system. The RAG1 protein is fairly large such as the mouse RAG1 gene that contains 1040 amino acids (Abbas et al., 2003). The enzymatic activity of the RAG1 protein is largely concentrated in a core region between residues 384-1008 (Abbas et al., 2003). This core region contains three acidic residues (D\textsubscript{600}, D\textsubscript{708} and E\textsubscript{962}) in what is called the DDE motif (Janeway et al., 2005). Previous studies have shown that this nuclear gene is useful for phylogenetic studies of rodents (Steppan et al., 2004a, b; Steppan et al., 2005; Suzuki et al., 2004).

The cyt \textit{b} gene was used in all analyses in Chapter 2 and this was followed by a combined approach in Chapter 3 (i.e., a combination of a fragment of RAG1 and cyt \textit{b}); these genes were used to infer the phylogeny of \textit{M. namaquensis}. This approach was adopted for Chapter 3 as phylogenetic analyses of several short stretches from different genes show a better performance than analyses based on nearby sites from a single gene fragment (Cummings et al., 1995). In combining these genes, it was considered that the cyt \textit{b} gene with its faster mutation rate would possibly resolve relationships near the terminal nodes relative to the slower nuclear RAG1 gene which is considered useful for resolving relationships at deeper nodes (Suzuki et al., 2004).

4. Phylogeography

The analysis of the cyt \textit{b} gene in this study was also augmented by a phylogeographic approach which represents a “mtDNA bridge between population genetics and systematics” (Avise et al., 1987). It is concerned with patterns and processes governing geographic distributions of genealogical lineages among and within closely related species (Avise et al., 1987). Phylogeographic differentiation represents an interplay between vicariance and dispersal processes (Avise et al., 1987). Intraspecific population structure is therefore determined by the species’ potential for gene flow and the environmental influences on that potential (Avise et al., 1987). Phylogeographic patterns can be expected to range from total phylogenetic discontinuity between populations due to zoogeographic barriers and limited dispersal abilities, to genetic continuity in species with high dispersal potential and/or in unfragmented environments (Avise, 2000; Avise et al., 1987).
In the first decade of phylogeographic research, analyses largely involved the qualitative description of geographic patterns (Avise, 2000), yet valuable insights were gained regarding the structuring of intraspecific variation (see review by Avise, 2000), especially in response to Quaternary climate changes (see amongst others Hewitt, 2000; Hewitt, 2004). The subsequent introduction of statistical approaches now enables testing of hypotheses to investigate the processes underlying the patterns of diversity (Knowles, 2004; Knowles and Maddison 2002). One of the first of these approaches was nested clade analysis (Templeton et al., 1987; Templeton and Singh, 1993). Nested clade phylogeographic analysis (NCPA) attempts to distinguish between different historical processes that might have influenced the geographic distribution of haplotypes relative to higher level clades (Petit, 2008), however, the approach is a subject of considerable debate (Garrick et al., 2008; Petit, 2008). Recently, Petit (2008) suggested that the method should no longer be used until it has been more thoroughly and critically evaluated, while Garrick et al. (2008) suggested that the method could still be useful in generating “plausible hypotheses” as corroborating evidence or lack thereof. In reaction to previous debates on the effectiveness of NCPA (e.g., Knowles and Maddison, 2002; Templeton, 2004), a new software program (ANeCA) was developed that automates the complex NCPA methodology (Panchal, 2007). While I acknowledge the potential limitations of the approach, I adopted it along with other measures of population history (e.g., mismatch distribution analysis as implemented in Arlequin, version 3.0 (Excoffier et al., 2005)) to estimate population expansion and MIGRATE, version 2.4 that was used to estimate effective population sizes and past migration rates between \( n \) number of populations assuming a migration matrix model (Beerli and Felsenstein, 2001).

Although *M. namaquensis* is widely distributed in southern Africa, it is essentially restricted to rocky habitats (Chimimba and Bennett, 2005). Species occupying such habitat islands often display patterns of phylogenetic discontinuity. Geographically distant samples of the deer mouse (*Peromyscus maniculatus*) reflect the impact of physical barriers on dispersal (Lansman et al., 1983). Similarly, this pattern is also shown by southern African rock dwelling small mammals such as the rock hyrax (*Procavia capensis*; Prinsloo, 1993), the red rock rabbits (*Pronolagus sp.*; Matthee, 1993; Matthee and Robinson, 1996) and the Namaqua rock mouse (*M. namaquensis*; Russo, 2003). Nicolas et al. (2009) provided evidence of restricted gene flow with isolation by distance and a signal of population expansion was detected within several clades of *Acomys*, a rock-dwelling
small mammal. This species shows a strong phylogeographic structure (Nicolas et al., 2009).

In contrast, habitat uniformity of the open plains occupied by 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 and Robinson, 1997). 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 and Harrison, 1992). *Aethomys chrysophilus* and *A. ineptus* from southern Africa also reflect a lack of genetic divergence among samples over a broad geographic range (Russo et al., 2006).

Other life history characteristics of *M. namaquensis* that are predicted to impact on phylogeographic structuring include body size, age at weaning, age of sexual maturity, age of fecundity, time to first sexual activity and first reproduction, duration of gestation, litter size and interbirth interval. These life history traits relating to growth and reproduction vary greatly among species, populations and individuals within populations. Egron et al. (2001) demonstrated that the immediate environment (which includes water and food availability, nesting/breeding space ect.) plays a major role in shaping the life history characteristics of small mammals.

### 5. Aims of Study

Given the above background, the aims of the present study are as follow:

1. To assess the nature and extent mtDNA variation within *M. namaquensis* from southern Africa based on cyt *b* sequence data and to compare this variation to the previously described subspecies.

2. To assess phylogenetic relationships between lineages within *M. namaquensis* based on cyt *b* and RAG1 data.
3. To elucidate the phylogeographic pattern among *M. namaquensis* populations from the Eastern Kalahari Bushveld and to infer past processes underlying the current diversity.

6. **Research Questions**

Research questions were specific to each chapter. Questions 1 - 3 relate to Chapter 2, question 4 to Chapter 3 and question 5 to Chapter 4. The following specific research questions are addressed in the present study:

1. What is the nature and extent of mtDNA variation within *M. namaquensis* over a broad geographic area in southern Africa?

2. What is the phylogeographic structure within some of the *M. namaquensis* lineages?

3. What are the divergence dates of different *M. namaquensis* lineages?

4. What are the relationships among *M. namaquensis* lineages based on a combined phylogenetic analysis of the mitochondrial cyt b and the nuclear RAG1 genes?

5. What are the finer scale phylogeographic structure and underlying processes within *M. namaquensis* based on a case study of populations from the Eastern Kalahari Bushveld?

7. **Relevance of Study**

In addition to a significant contribution to small mammal systematics and biodiversity research in Africa, the findings of the present study may have implications in epidemiological and agricultural research associated with problem rodents. Some members of the genus *Aethomys* and *Micaelamys* 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 (De Graaff, 1981; Smithers, 1971;
Wilson, 1970, 1975). Consequently, the present study may assist health and agricultural authorities in gaining a better insight into these potentially problematic rodents.

In order to assist agricultural and health authorities it is important to assess the extent and nature of *M. namaquensis* diversity. High levels of genetic diversity within the species were evident in a previous study by Russo (2003). These levels of diversity should be taken into account when recommendations are made to the authorities given the balance between conservation of biodiversity and management of problem rodents are important. Furthermore, for the management of pest species it is essential to have some understanding of the life history traits of a species (i.e., movement patterns) with regard to the spread of diseases.

From an academic (scientific) point of view, small mammals such as these rodents are good models in evolutionary and population biology since they have short generation times, short gestation periods and large litter sizes (Chimimba and Bennett, 2005).

8. Thesis Outline

CHAPTER 2: The *Micaelamys namaquensis* (Rodentia: Muridae) species complex from southern Africa: Patterns of mitochondrial DNA versus morphological diversity

This chapter represents the main focus of the study which was to assess the nature and extent of the mtDNA variation, phylogeographic structuring within some lineages and estimated times of divergence for the lineages of interest identified within *M. namaquensis* from southern Africa. Phylogenetic analyses revealed 14 distinct lineages of which several show strong geographic association with particular vegetation types of southern Africa. The diversity and differentiation detected in the present study are much more complex than suggested by previous morphological assessments. An Early to Late Miocene time of divergence was suggested. *Micaelamys namaquensis* lineages show contrasting patterns of phylogeographic structure but the majority of the diversification events date to the Late Miocene and Pliocene that coincide with major periods of aridification in southern Africa.
CHAPTER 3: Phylogenetic relationships within *Micaelamys namaquensis* (Rodentia: Muridae) from southern Africa as inferred from mitochondrial and nuclear genes

This chapter extends the research reported in Chapter 2 by addressing the phylogenetic relationships among *M. namaquensis* lineages based on combined mitochondrial cyt b and nuclear RAG1 data. Incongruence was shown between the cyt b, the combined (cyt b and RAG1 genes) analyses and the independent RAG1 analysis possibly as a result of incomplete lineage sorting in the nuclear gene. The combined molecular data suggested that *Micaelamys* from southern Africa comprises 11 lineages of which six were well-supported with strong geographical associations to biomes/bioregions of southern Africa, while the remaining lineages were not associated with specific vegetation types.

CHAPTER 4: Phylogeography of *Micaelamys namaquensis* (Rodentia: Muridae) from the Eastern Kalahari Bushveld bioregion of South Africa

In order to refine our understanding of evolutionary and ecological processes underlying the remarkable diversity within *M. namaquensis*, this chapter presents a fine scale phylogeographic analysis of the mtDNA lineage from the Eastern Kalahari Bushveld bioregion. This bioregion has specifically been chosen since samples were distributed over a fairly small, well-defined geographic area. Secondly, since localities of sympatry have been identified in Chapter 2 for this region, it was thought that a fine scale phylogeographic analysis would give some insight in understanding the ecological processes underlying diversity. This bioregion was also represented by a fairly large sample size compared to some of the other biomes/bioregions. This fine scale analysis revealed a genetic pattern of phylogenetic continuity. Mismatch distribution analysis suggests that the lineage has experienced recent population growth following environmental changes associated with habitat modification over the past 3 000 to 10 000 years. Historical female gene flow does not appear to be equal amongst all localities.

CHAPTER 5: Thesis conclusion

The thesis concludes with a general synthesis of the major findings of this multidisciplinary study of *M. namaquensis* from southern Africa. In addition, some recommendations for future research are made.
9. General Notes

All chapters (except 1 and 5) were written as manuscripts that have been or will be submitted for publication in peer-reviewed journals. Chapter 2 has been submitted to BMC Evolutionary Biology. Given that it has been submitted for publication, I refer to the work being done by us (I.M. Russo, C.T. Chimimba and P. Bloomer). A reference list occurs at the end of each chapter rather than at the end of the thesis. Due to this format, there may be instances of duplication across chapters. Each chapter contains its own appendices. The introduction and the concluding chapter were tailored from the respective chapters, giving a general overview of this thesis and the overall concluding remarks and synthesis. For the purposes of this thesis the format of Molecular Phylogenetics and Evolution was followed and references in the text were listed in alphabetical rather than chronological order.
References


Dobzhansky, T.H., 1940. Speciation as a stage in evolutionary divergence. Amer. Nat. 74, 312-321.


Riddle, B.R., 1996. The molecular phylogeography bridge between deep and shallow history in continental biotas. TREE 11, 207-211.


Woodall, P.F., Mackie, R.I., 1987. Caecal size and function in the Rock elephant shrew
*Elephantulus myurus* (Insectivora: Microscelidae) and the Namaqua rock mouse


markers. TREE 11, 247-251.