UNIVERSIDADE D
COIMBRA

Sofia Onofre Ramos

POPULATION DIFFERENTIATION AND CONTRASTING DEMOGRAPHIS HISTORIES OF THE NATAL MULTIMAMMATE MOUSE (MASTOMYS NATALENSIS)

Dissertation within the scope of the Biological Resources, Territorial Development and Sustainability Masters, supervised by Dr. Luís das Neves Cunha and Dr. Isa-Rita Russo presented to the Department of Life Sciences of the Faculty of Sciences and Technologies of the University of Coimbra

January 2024

Population differentiation and contrasting demographic histories of the Natal multimammate mouse (*Mastomys natalensis*)

Sofia Onofre Ramos

Acknowledgments

I want to mention a special thank you to Luís Cunha. I highly appreciate all your guidance and assistance with this thesis. You considered my thoughts, opinions, and personal touches, and this thesis came to be under your diligent guidance.

To Isa Russo, I want to express my sincere gratitude for accepting me as your student. Your mentorship and the valuable knowledge you have shared were fundamental to the development of this document. Thank you for the opportunity to learn and grow under your guidance.

I extend my heartfelt gratitude to the research groups at Cardiff University and ENGEL at University of Coimbra for fostering an enriching environment that greatly contributed to my work. Special thanks go to Olivia and Rynhadt whose assistance in demography analysis and laboratory efforts was invaluable. A profound appreciation is also due to Luisa Fraga, whose unwavering support and companionship during the critical final phase of this thesis were indispensable. Without her, navigating this challenging period would have been significantly more difficult.

Abstract

Earth's ecosystems and biodiversity are intricately woven by complex factors, shaping the evolutionary history of populations. This study delves into the genetic dynamics of *Mastomys natalensis*, a key African rodent species with dual impacts as an agricultural pest and disease vector. Anthropogenic changes in habitats amplify selective pressures, prompting more frequent evolutionary events. Molecular ecology tools enable in-depth assessments of environmental factors and population parameters. The evolutionary forces shaping species are explored, emphasizing the deterministic and dynamic aspects of evolutionary biology.

The genetic mosaic of *Mastomys natalensis* is here explored using mitochondrial DNA (mtDNA) and microsatellite markers to illuminate the species' ancestry and evolutionary adaptations. While mtDNA phylogenies offer profound insights into this species' lineage and adaptive strategies, it's important to acknowledge inherent limitations such as sex bias and homoplasy. Conversely, the high variability of microsatellites plays a pivotal role in dissecting population structures and unravelling demographic histories, aiming to assess the genetic diversity influenced by mutation, gene flow, and natural selection across different populations.

The application of mtDNA and microsatellites in this research underscores their significance in decoding the population dynamics of *Mastomys natalensis*. Despite facing methodological challenges, these molecular tools are instrumental in refining taxonomic classifications, facilitating selection studies, and fostering a comprehensive grasp of evolutionary mechanisms. This inquiry contributes to the global initiative to study and conserve biodiversity and provides crucial insights for devising effective management and conservation strategies. Understanding the genetic and evolutionary nuances of *Mastomys natalensis* is paramount for mitigating its impact on agriculture and public health, highlighting the species' complex role within its ecosystem.

This work highlights the great genetic heterogeneity of the species by using extensive genomic studies to analyse the population architecture of *Mastomys natalensis*. Non-uniform genomic impacts are indicated by statistical analyses that show selective pressures with large departures from equilibrium and varying effects across different loci. Using STRUCTURE, AMOVA, and multivariate approaches, population genetic analyses revealed unique clustering patterns

emphasising the genetic architecture that differed between populations. With an emphasis on the geographic isolation of the Hluhluwe-iMfolozi Park—characterised by the Drakensberg Escarpment, the Indian Ocean, and particular habitat preferences—the study investigates potential causes driving this structuring. The species' restricted long-distance dispersion and affinities for soil humidity, slope orientation, and altitude could be the observed genetic isolation factors.

Structure in studied populations is suggested to have been greatly influenced by historical events during European settlements and urban human activities using the tool ABCtoolbox to create demographic stories. A bottleneck in the Hluhluwe cluster may be related to opportunities brought about by European colonists and changes in land use. Despite non-significant Bayesian Factor values, simulations reveal intricate population dynamics over the past 400 years.

Furthermore, mitochondrial DNA phylogenetic analyses indicate conserved traits consistent with other muroids, yet unusual tree topologies warrant caution when interpreting findings. Unusual differences in the Cricetidae family's division prompt inquiries on the source of samples and call for additional research. Interestingly, the sample (Mas12_4_H) branches differently from other *Mastomys natalensis* samples and aligns with *Mastomys erythroleucus*, requiring careful examination of sample identification and potential consequences for the distribution of the species. This study opens the door to a greater knowledge of the evolutionary trajectory and conservation implications of *Mastomys natalensis* by providing insightful information about the species' genetic quirks, historical impacts, and complex population dynamics.

Table of Contents

I. Introduction

Earth's ecosystems and their biodiversity form an elaborate tapestry woven by complex factors that, together, create its rich entanglements. With every knot between the yarns, the history of each population is shaped. Only through the unravelling of the aspects that built each species can we reveal what might influence today and how that ought to affect the future. Thus, understanding these evolutionary processes has been pivotal in management and conservation strategies. From endangered to invasive species, to successfully establish policies, a population's structure and its characteristics should be identified, as well as what shapes them (Alphey & Bonsall, 2018; Bowen et al., 2005; Le Roux & Wieczorek, 2009a). With the development of novel technologies, the future of environmental management seems to be fronted by genetic research (Haig, 1998). From whole genome sequencing and higher accessibility to genomic data and through a detailed systematics, these techniques have revolutionised the construction of environmental governance practices (Frankham, 1995; Hedrick, 2001; Stockwell et al., 2003). Recognizing the contribution of genetic resources to ecosystem services further promotes this development, with worldwide funding efforts to study and protect this pillar of biodiversity (Allendorf et al., 2013; Coker & Coker, 2019; Leroy et al., 2018).

Anthropogenic habitat change can lead to higher selective pressures and mutation rates, promoting more frequent evolutionary events (Stockwell et al., 2003). To identify these and other shaping forces, molecular ecology provides a plethora of tools, as it allows an in-depth assessment of how environmental factors influence genetic variation and diversity (Allendorf et al., 2010; Coker & Coker, 2019). Being able to quantitatively appraise parameters like variability, gene flow, migration, heterozygosity, inbreeding, parental lineages, effective population sizes or fitness allows us to confidently determine these relationships and their effects on ecosystems (Bažok, 2022; Coker & Coker, 2019; Stockwell et al., 2003). Aligning this knowledge with the history and landscape when targeting pests with the goal of population control, genetic-based research is successfully outlining collaborative and integrative management strategies worldwide (Alphey & Bonsall, 2018; Combs et al., 2019; Karsten et al., 2013). As we further discover the harmful effects of chemical toxicants-based approaches to pest management, both to ecosystems and human populations, these resources can inform decisions and offer sustainable and possibly more effective strategies to combat these threats (Arias et al., 2019; Backus & Gross, 2016; Bažok, 2022; Van Hooft et al., 2008).

1.1. Genetic Dynamics of Evolutionary History

A species' evolutionary history can be shaped by habitat changes, fragmentation, and destruction, which, in recent history, is often of anthropogenic origin (Colangelo et al., 2013; Kneidinger, 2010; Russo et al., 2016). Population genetics aims to identify levels of variation among and within populations, as well as the parameters influencing its structure, bridging ecology with evolutionary studies (Hamilton, 2021; Lowe et al., 2017; O'Neill et al., 2003). As discussed above, detailing evolutionary relationships between a species and its surrounding environment can not only allow for a better understanding of ecological history but further reveal its present and future impacts (Cutter, 2013; Manel et al., 2003; Russo et al., 2016; Stockwell et al., 2003).

Population genetic patterns can be evaluated from a broad range of perspectives, and the use of different methodologies and analyses can allow for the uncovering of different patterns in adaptation, evolutionary history, and population dynamics (Freeland, 2020; O'Neill et al., 2003; Wan et al., 2004). Historically, these studies began with classical, or Mendelian genetics, focussed on morphology, and, more recently, molecular ecological methodologies have become the leading role in the area (Casillas & Barbadilla, 2017; Crow, 1987; Hamilton, 2021). Genetic markers as a tool allow the quantification of genetic diversity with some ease, as well as the understanding of interactions and inheritance patterns (Allendorf et al., 2013; Casillas & Barbadilla, 2017; Freeland, 2020). Molecular changes at the DNA level arise from a single base (point mutations) to whole segments, with possible rearrangements, inversions, insertions and deletions (indels), as well as duplications, that can lead to copy number variations and transformations at scales as large as chromosome level (Le Roux & Wieczorek, 2009a; Loewe & Hill, 2010). The impact of these mutations can vary widely, as they can be detrimental, beneficial or neutral; effects that are dependent on gene expression, along with environmental factors and timescales (Loewe & Hill, 2010). The accumulation of each mutation in a population, and whether it's fixated or lost, is dependent on varied parameters such as natural selection and matting patterns, inbreeding, admixture, population size and structure, genetic drift, migration, environmental factors, amongst others (Otto & Whitlock, 1997; Uecker & Hermisson, 2011; Whitlock, 2003). These evolutionary forces can be deterministic, with both non-random or stochastic outcomes (Hamilton, 2021; Jenkins & Stekel, 2010; Loewe & Hill, 2010). As the essence of a process is often non-linear, and these forces coexist in nature, evolutionary biology has progressed to build dynamic models, allowing us to "replay the tape of life", as we discover how and what makes it unfold (Blount et al., 2018).

1.1.1 Unveiling Genetic Signatures Using Mitochondrial DNA and Microsatellite Markers

The genetic mosaic of each species is, as such, intrinsically connected to the characteristics of each population and the evolutionary forces moulding it. Knowledge of current populations and their structure can be applied to evolutionary models in order to identify a species' life history. Phylogeny, population structure and possible hypotheses for demographic history scenarios can be evaluated with DNA (Cutter, 2013; Fahey et al., 2014).

The Value of Phylogenies and the Use of Whole Mitochondrial Genomes

Defining phylogenies not only deepens the grasp on ancestry but allows for the comprehension of evolutionary adaptations and traits (Cutter, 2013). These analyses can be paired with historical, topographical and landscape information, to reveal evolutionary forces shaping biodiversity (Manel et al., 2003; Russo et al., 2016). Phylogeny can be assessed using different approaches, from behavioural, morphological, and fossil characteristics to molecular information, but here it is explored with mitochondrial (mtDNA) data (Bybee et al., 2008; Zrzavý et al., 2018).

Distinguished by high mutation rates, that can often become fixed and accumulate, as well as rapid base substitution, mtDNA evolutionary rates can be 5 – 10 times higher than in nuclear genomes (Allio et al., 2017; Boore, 1999a; Neiman & Taylor, 2009). Gene arrangements tend to be preserved amid generations, ranging between 15-20 kb, the small genome displays a structure that includes 13 protein-coding genes, 2 ribosomal RNAs, and 22 transfer RNAs (Boore, 1999b). Additionally, non-coding 'control regions' can be highly variable due to high substitution rates and have a central role in transcription and replication regulation (Ferris et al., 1983; González-Ittig et al., 2009). Other particularities, like being almost exclusively inherited trough the maternal line, generally without recombination except for few exceptions (Harrison, 1989; Rokas et al., 2003), having a compact genome, and the low possibility for heteroplasmy, make it an almost ideal genetic marker (Castro et al., 2014; Harrison, 1989). Defining phylogenies with this tool has additional advantages, as isolation, manipulation and sequencing are efficient, accessible, and affordable and, as it is commonly used in studies, it often comes with more well-supported and in-depth background knowledge (Castro et al., 2014; Harrison, 1989; Wan et al., 2004).

On the other hand, it is important to consider the possible detriments of mtDNA-based phylogenies in alternative to other data, such as nuclear DNA and morphology, as some of the mentioned advantages come with downsides. One of the main aspects to contemplate is the sex bias associated with exclusive maternal lineage, an aspect that can influence contrasting results when evaluated (Bowen et al., 2005; Dávalos & Russell, 2014). With rare recombination events, parentage inference can be challenging as genes are mainly inherited as a single linkage block,

risking an inaccurate representation of evolutionary events (Moore, 1995). Furthermore, homoplasy is often found within species due to short-lived mutation hotspots and can often hinder explorations attempts with deeper phylogenies; this happens due to saturation and false parallels created by its rapid substitution rates (Galtier et al., 2006; Rubinoff & Holland, 2005). When treating highly variable sites, not considering functionality, influence in fitness, and evolutionary knowledge, can construct false positives for selective pressures that are merited to neutrality (Castro et al., 2014; Hamilton, 2021). Selecting and manipulating mtDNA data requires the consideration of all factors discussed in order to properly establish lineages and more effectively identify the evolutionary pressures shaping species (Casillas & Barbadilla, 2017; Castro et al., 2014).

Even though, much like other data sources, the mtDNA genome has features that can limit its potential use in evolutionary analyses, it is an extremely effective phylogenetic tool time and time again. As Rubinoff & Holland, (2005) point out: "Mitochondrial DNA is no less a part of an organism, and therefore an organism's evolutionary history". The previously discussed characteristics of these types of data facilitate the identification of cladistic and reticulation events, as well as hybridisation and introgression events, that often reveal unique inter-species relations and other selective forces. This can be done through different methods and with various data subsects. Protein-codding genes have successfully been used to re-think and enhance taxonomic classifications (Carapelli et al., 2007; Deng et al., 2023; Foster et al., 2017) and determine selection and mutation mechanisms (Castellana et al., 2011; Kumar, 1996; Nabholz et al., 2013; Popadin et al., 2013).

Population structure and demography and the Value of Microsatellite Markers

Microsatellites, also known as Short Tandem Repeats (STRs), comprise motifs of 1-6 bases and can be highly variable between individuals, representing a large portion of non-coding DNA (Wan et al., 2004). Widely distributed in animal genomes, these markers are distinguished by high polymorphism levels and mutation rates (Hedrick, 2001; Selkoe & Toonen, 2006). Additionally, co-dominance inheritance and abundance throughout the genome, as well as their variability between and within species, have bought them to the forefront of genetic research. The particularities mentioned offer significant advantages for population genetic studies, as they can identify slight variations even with closely related individuals (Hedrick, 2001; Selkoe & Toonen, 2006) (Barker, 2002; Feng et al., 2002; Manel et al., 2003). Additionally, one of its main distinctive characteristics is the occurrence of several mutational events in short amounts of time due to high mutation rates, facilitating identification and dating of evolutionary events (Cavalli-Sforza, 1998; Schlotterer, 2000). Microsatellite data scoring and statistical analyses is simplified

with often easily distinguishable homozygous and heterozygous individuals (O'Neill et al., 2003; Wan et al., 2004). Co-dominance also allows for the calculation of frequencies between populations and, thus, the identification of their differences (Freeland, 2020). Considering these particularities, when accessing variation in sub-population analyses, notably when there is fracturing, or the populations are smaller, microsatellites can help reveal structuring patterns (O'Neill et al., 2003).

Even with the above-mentioned advantages in the use of microsatellite data, there are important challenges to consider. The mutation mechanisms of these markers are highly debated, yet this information is essential to consider when working on allele frequency analyses (Selkoe & Toonen, 2006; Wan et al., 2004). Mutations can alter the primer region segment in composition and length, which can negatively affect the PCR product (Selkoe & Toonen, 2006; Wan et al., 2004). Phenomena like slip-strand mispairing during replication, which is the most common mutation mechanism in STRs, or even simple repeats, can also lead to errors in the scoring process (Richard & Pâques, 2000; Schlotterer, 2000; Wan et al., 2004). Other risks are, for example, undetectable homoplasy, where sequences have different evolutionary stories yet are identical (Selkoe & Toonen, 2006). The problems listed can often be mitigated using mutation models, with the stepwise mutation model (SMM) (Kimura & Ohta, 1978) being the most common. SMM considers small stepwise alterations in microsatellite repeat lengths to simulate the natural mutation process (Hamilton, 2021; Luikart & England, 1999). While SMM encompasses a great number of changes, not all mutations are of stepwise origin, and alternative methods keep being developed to account for different scenarios (Freeland, 2020; Luikart & England, 1999; Richard & Pâques, 2000).

STR analyses are extremely cost-effective, especially in comparison to other alternatives, and while they comprise some downsides, controlling workable conditions can help resolve these issues (Wan et al., 2004). Species dynamics have been successfully assessed in the past with the use of these markers, and they remain in the forefront of population genetic studies. Microsatellites have also help uncover evolutionary stories and identification varied pressures influencing structure (Wright & Bentzen, 1995). Recently, research carried out with this tool established defining lines for endangered species conservation (Abdul-Muneer, 2014; Bowen et al., 2005; Lacher et al., 2016; Schwartz et al., 2007) as well as pest and invasive species management (Arias et al., 2019; Bažok, 2022; Combs et al., 2019; Karsten et al., 2013; Le Roux & Wieczorek, 2009b). Here, microsatellites are used to infer population structure and demographic history.

1.2. Ecosystem Engineers and Pests: The Ecological Significance of Small Mammals with a Focus on the Multimammate Mouse (*Mastomys natalensis*)

Small mammals are some of the most widely distributed and important species in the biomes they occupy (Linzey & Kesner, 1997). The possible impacts these species have in ecosystems vary widely, and whether their influence is a net positive is dependent on several factors. Serving as a major food source for predators, dispersing seeds and fungi, pollinating, and even affecting vegetation and soil structure as ecosystem engineers, are some of the ecosystem services they provide (Lacher et al., 2016). They are major contributors to ecosystem complexity and stability, having a crucial role maintaining ecosystem functionality and biodiversity, making it an indispensable group and an important bioindicator (Avenant, 2011; Pearce & Venier, 2005). At the same time, these species are often pests and major disease vectors, for both humans and other animals. The presence of some species is repeatedly associated with major economic losses and impacting the wellbeing of local communities, as well as leading environmental health issues (Lacher et al., 2016; Singleton et al., 2010). Because of this, several small mammal species pose a global threat to food and health security (REF). These complex relationships are major barriers in the development of conservation and management plans. Rapid reproductive rates, resilience and plasticity make it very hard to define successful guidelines, and local governments often have little to no alternatives to the use of raticides. It is also extremely difficult to study the standard ecology of these groups as they tend to occupy habitats with extreme anthropogenic influence and change (Linzey & Kesner, 1997). To better mitigate the detriments small mammal species can pose, it is essential to understand their ecology and population characteristics, as well as their relationship with the surrounding environment and landscape.

The multimammate mouse: Mastomys natalensis (Smith, 1834)

The Natal multimammate mouse, *Mastomys natalensis* (Smith, 1834) is the rodent species with the largest distributional range in Africa which is sub-Saharan of nature. Often associated with human communities and anthropogenic activity, populations can be found in the surrounding area of human settlements and in cultivated areas (Brouat et al., 2007; Kneidinger, 2010; Meester et al., 1979). This species can spread throughout a broad ecosystem spectrum, excluding deserts and high mountainous areas (Colangelo et al., 2013), where they occupy burrows (Leirs et al., 1996a; Meester et al., 1979). Previous studies suggest preferences for areas with low altitude and more than 600 mm of rainfall per year (Venturi et al., 2004). This species displays swimming behaviour and is dependent on available water, with population peaks often happening after longer rainy seasons (Martin & Dickinson, 1985; Russo et al., 2016). While several reports show that there is a preference for areas with high vegetative cover (Leirs et al., 1996; Monadjem & Perrin, 1998 Venturi et al., 2015), areas with low vegetative cover can be colonized if there is an abundance

of food availability (Monadjem & Perrin, 2017). This can explain this species' high colonization of agricultural land, even with the high risk of predation that comes with lower vegetation (Monadjem & Perrin, 2017; Taylor & Green, 1976).

As most small mammals, they are nocturnal opportunistic feeders, consuming a variety of food depending on availability, with a clear seasonal effect, being primarily granivorous (Kneidinger, 2010; Mulungu et al., 2011; Odhiambo et al., 2008; Russo et al., 2016). Different reports support that feeding preferences vary throughout the year, preferring seeds, arthropods and grasses during the wet season, and other less nutritious plant materials during the dry season, during which they might even disappear completely from the habitat (Monadjem & Perrin, 2015; Mulungu et al., 2011). According to Mulungu et al. (2011), grain is the main food source in every season and habitats, with maize fields and woodland grains representing 80% of diet during harvesting season.

These feeding patterns help illustrate how it has become the most important agricultural pest species in southern and eastern Africa, commonly associated with high numbers of crop losses (Mulungu et al., 2011; Mwanjabe & Leirs, 1997). Reports indicate damages in farmers' fields of 48% with a rodent population of just 18 animals per ha (Mdangi et al., 2013), and losses can exceed 80% of the planting and seedling stages during severe outbreaks (Mwanjabe et al., 2002; Mwanjabe & Leirs, 1997). Additionally, this species is a major disease vector for both humans and animals. The spread of the carrying viruses is influenced by this specie's cryptic genetic structure, with pathogen presence and abundance varying between subgroups (Cuypers et al., 2020).

Mastomys natalensis is a pioneer species, often the first to recolonize areas that went through habitat destruction, for example, after fires or anthropogenic activity (Kneidinger, 2010; Rowe-Rowe, 1995). Features of secondary succession vegetation, with high biodiversity, thicket, weeds, grasses and shrubs, provide shelter and varied food choices, and allows for the capitalization of anthropogenic habitat alterations (Meester et al., 1979; Russo et al., 2016). As landscape develops towards a climax community, population numbers stagnate to pre-damage conditions and it is preceded by specialist species (Kneidinger, 2010; Meester et al., 1979).

As the name suggests, the Natal multimammate mouse is highly distinguishable by a high number of mammary glands, from 8 to 12 pairs (Kneidinger, 2010), which correlates with its high reproduction rates. Breeding behaviour happens throughout the year, with reduced rates in colder months, and litters can have up to 24 pups *in utero* (Kennis et al., 2008; Kneidinger, 2010). Individuals can start breeding at around 3 and a half months and multiple paternity per litter is common, further facilitating success (Kennis et al., 2008; Kneidinger, 2010). These adaptations, when in conjunction with preferred environmental factors, often create population outbursts .

The study species belongs to the *Rodentia* order, within the *Praomys* complex of the *Muridae* family, along with three other genera: *Praomys*, *Myomys* and *Hylomyscus* (Lecompte et al., 2002a). The genus *Mastomys* is highly cryptic, (Lecompte et al., 2005) cyt-b assessments showed intra-species variation between 1.1 and 3.3%, and between species from 8.1%. Previous phylogenetic analyses suggested that *M. natalensis* has recently differentiated from other congeneric lineages following the expansion of sub-Saharan savannah habitats during the Plio-Pleistocene (Colangelo et al., 2013; Lecompte et al., 2002b). Colangelo et al. (2013) found that, throughout the continent, patterns of genetic diversity show six main phylogroups, that diverged due to long-term barriers. The differentiation seems to have been climate-driven at the continental scale and not due to regional-specific physical or climatic barriers. Russo et al. (2016) showed that, while there is a pattern of isolation-by-distance (IBD) at smaller scales, with kin clustering, larger scales fail to present it, due to long-distance dispersal in *M. natalensis* being rare. The same study found that the breeding age will reflect the location, being that the earliest the individuals breed, the higher the tendency to breed closer to their natal site (REF, not Russo et al.). Additionally, the findings of four sub-clusters aligned with landscape characteristics, indicating that genetic patterns are affected by external factors other than IBD. It is also relevant that relatedness between individuals was not necessarily linked with their respective areas, with restricted dispersal influencing patterns. The main identified landscape characteristics positively influencing gene flow were eastern-facing slopes and thicker vegetation (seasonally available water), while topographic complexity (like ridges and steep slopes) and rivers limited gene flow. . The eastern facing slopes sampled for this study are cooler and have higher humidity and rainfall. These results align with others that indicate that this species tends to expand their home range with age but, when there is food available and suitable environmental conditions, travel distance is reduced (Mulungu et al., 2015).

1.3 Aims

This study aims to explore the population genetics and evolutionary history of *Mastomys natalensis*, an African rodent recognized not only as an agricultural pest and disease vector but also as an important commensal species with humans. The research emphasizes on evaluating the genetic diversity and structural variances across different populations, utilizing microsatellite markers and examining the role of environmental factors in shaping population dynamics. It is hypothesized that considerable genetic variation exists within and between *Mastomys natalensis*' populations in South Africa, influenced by evolutionary mechanisms like mutation, gene flow, and natural selection. Employing both mitochondrial DNA and microsatellite markers, this study seeks to provide a thorough understanding of the phylogeny, population structure, and demographic history of this species. These objectives are integral to grasping this species' genetic and evolutionary dynamics, a crucial aspect for effective management and conservation strategies

in light of its dual impact as both a pest and a disease vector (Allendorf et al., 2010; Frankham, 1995).

II. Materials and methods

2.1. Study area and sampling

The samples selected for this study (N= 445) were collected by Russo et al. (2016) and colleagues, four different sampling zones localized within two South African biomes: grassland and savanna (Figure 1). In the grassland biome a total of 22 samples were collected, nine from Koppies Dam Nature Reserve, and 13 from Soetdoring Nature Reserve (samples were provided by Johan Watson, Free State Department of Tourism, Environmental and Economic Affairs). The following 423 samples were collected from the country's east, in savanna-dominated areas. A total of 163 samples were collected from the Kruger Natural Park (Duncan MacFadyen provided samples), and the remaining 260 samples were collected from the Hluhluwe-iMfolozi Park (Russo et al., 2016).

Live-trapping of the animals using Sherman traps manufactured by H.B. Sherman Traps Inc. in Florida, U.S.A received ethical approval from Cardiff University. All handling of the animals complied with the standards established by the American Society of Mammalogists (ASM; Animal Care and Use Committee, 1998; (http://www.mammalogy.org/committees/index. Animals collected by Russo et al. 2016 were taken in accordance with permits. Ear clips from the remaining animals were either frozen at -20°C, kept in 99% EtOH, . Some animals were sacrificed by halothane inhalation.

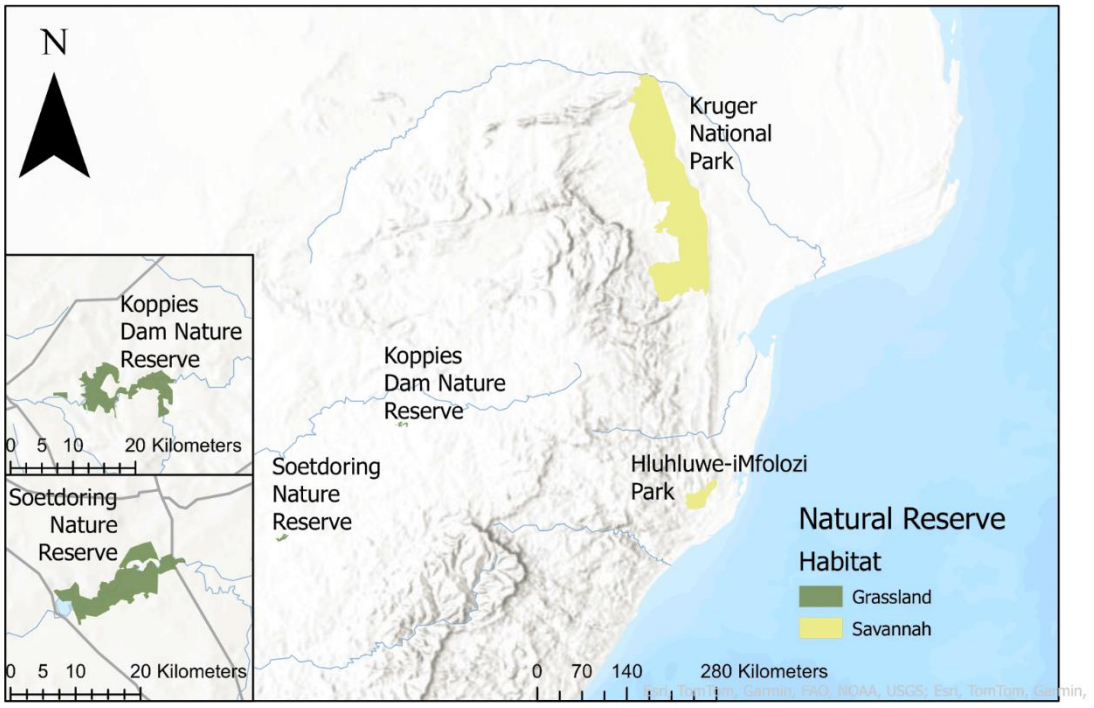


Figure 1. Map of eastern South Africa. Highlighted in green are the two grassland biome sampling regions: Koppies Dam Nature Reserve and Soetdoring Nature Reserve. In yellow we can see the two savannah biome sampling regions: Hluhluwe-iMfolozi Park and Kruger National Park.

2.2. Microsatellite analyses

Genotyping

Using a DNeasy® Tissue Kit (QIAGEN® Hilden, Germany), total genomic DNA was extracted from tissue samples collected form Soetdoring and Koppies Dam Nature Reserves following the manufacturer's instructions. All other samples were previously extracted. Ten microsatellite markers, initially identified from *Mastomys huberti* by Galan et al. (2004) and Loiseau et al. (2007), were used for genotyping, following the outlined PCR conditions. Denaturing was held at 95º C for 15 minutes; followed by 35 cycles of: 94º C for 30 seconds, 60º C (except for the Multiplex 1 which was run at 55º C) for one minute and 30 seconds, and 70º C during one minute; the final elongation stage was ran for 30 minutes at 60 º C. Amplifications were performed with a total volume of 10 µl, which contained 1µl of DNA (between 50 and 100 ng/µl), 0.2 µM of each forward and reverse primer, and 5 µl of QIAGEN Multiplex PCR Master Mix. The ten primer pairs were split into three multiplexes and one Singleplex with three dyes (FAM, HEX, and TAM): MH28, MH30, MH60, MH133, and MH141 encompassed Multiplex 1; MH05 and MH74 for Multiplex 2; MH105 and MH188 Multiplex 3; and MH80 the Singleplex. The characteristics

of the primers sent for fragment analyses and respective groupings are listed in Table 1. PCR Macrogen Inc., Korea (www.macrogen.com/eng/) processed the PCR products for commercial use.

Table 1. Isolated microsatellite loci characteristics. The repeat motif lists the STRs nucleotides, the indicated number is the of repeats. The dye indicates the used fluorescent dye. Allele size indicates the number of bp originally defined by Galan et al. (2004) and Loiseau et al. (2007). Markers MH60, MH105, MH30, MH141, MH133 and MH28 were identified by Galan et al. (2004), and the remaining MH188, MH5, MH74 and MH80 by Loiseau et al. (2007). Multiplex and singleplex divisions are as defined by (Russo et al., 2016).

	Primers	Repeat motif	Primer sequences 5'-3'	Size of cloned allele (bp)	Dye	GenBank accession code
Multiplex I	MH28	$(TC)_{28}$	GCCAAAAGCCTATAACCTTC CATCTTCCAAGTTCTATTTATGTG	186	6-FAM	AY583743
	MH30	$(TG)_{21}$	AAGGATTCTAACGTGCAGCC CCGGAGTTGAAGAGTTTCAG	249	HEX	AY583740.1
	MH60	$(TC)_{28}$	GCCAAAAGCCTATAACCTTC CATCTTCCAAGTTCTATTTATGTG	186	6-FAM	AY583743.1
	MH133	$(TC)_{35}$	CTCTTTTTAACACCTGTGTGG CTGCAGGTACTCACAACACG	338	NED	AY583745
	MH141	$(TG)_{22}$	TGGAAACAGCCTGTGCCAGC TTGGGCTCCCTCCTGGTTG	232	NED	AY583746
Multiplex II	MH05	$(GT)_{18}$	CACGTTCTGAAGGAAGTC CACAGATCACTCAGTGCAG	93	HEX	EF091896
	MH74	$(CA)_{10}$	CCTCTGCATCAGCACTTG GTGTTCTAGCATTGGACCTC	156	NED	EF091897
Multiplex III	MH105	$(TG)_{27}$	CAGCCATGATATGTCGCCAC AATGCCTGGATGGGGTGAG	249	HEX	AY583740
	MH188	$(AC)_{20}$	CAACCTTGGCTACTTGAGAC CATGGAGGTAGAATGGAACTG	104	NED	EF091893
Singleplex	MH80	$(CA)_7(GA)_{19}$	CGTAGTAGAAGAAACAAGTGG AAGGAATCTCCACAGGCT	252	NED	EF091898

Statistical analyses

For scoring the electropherograms from the sequenced DNA, GeneMarker v 1.91 (SoftGenetics LLC) was used. The MH60 was identified to have null allele and was removed from the data. With the remaining dataset, evaluation gene copy number, allelic statistics, heterozygosity and Theta θ(H) values were determined with Arlequin v3.5.2.2 (Excoffier & Lischer, 2010). Additionally, the following parameters were assessed through GenAlEx 6.51b2 (Peakall & Smouse, 2006, 2012): Shannon's mutual information index (I), the fixation index, Nei's genetic distance and identity, Fst, Rst and Gst. A principal coordinates analysis (PCoA) and Analysis of Molecular Variance (AMOVA) with 9999 permutations were also run. Hardy-Weinberg

Equilibrium (HWE) *p*-values were manually adjusted with both a Bonferroni and a False Discovery Rate (FDR) correction.

Population Structure

The most likely number of population clusters (*K*) was determined using Bayesian clustering in STRUCTURE v 2.3 (Pritchard et al., 2000). Following a burn-in of 50,000 iterations with *a priori* location data identifying each area, the analysis was run for 450,000 Markov chain Monte Carlo (MCMC) iterations with 20 independent iterations out for varying *K* values (20 for the complete dataset and 10 for each subsequent cluster). Structure outputs were processed with Structure Harvester (Earl & vonHoldt, 2012) and population structure was assessed through the Evanno method (Evanno et al., 2005), where the posterior probability (LnPr(X|K)) and rate of change (ΔK) were considered. The Clumpak (Kopelman et al., 2015) server was used to produce the visual outputs for the structure runs.

Demographic history

Recent demographic changes were inferred for the clusters identified in the previous analyses with the use of ABCtoolbox (Wegmann et al., 2010). Priors were defined broadly, with different ancestral (NA) and current (NC) effective population sizes tested (with a generation time of 1 year) to access for the most likely scenario. For each groups, n=20 representative samples were selected, and tree potential scenarios were tested: S1: a null scenario, where the population remains stable (NA=NC); S2: an expansion (NA<NC), and S3: a bottleneck (NA>NC). A contemporary timeframe was successfully defined for a period between 200 and 400ya, selected with the goal to identify recent demographic change, as this it corresponds to the colonial occupation in South-Africa. Scenarios and priors were tested and defined considering the best output summary statistics and posterior values. The scenario with the best *p*-value for each cluster was selected and compared with the other marginal densities to calculate the Bayes Factor (BF), this indicates the most likely scenario (validated for values of BF>3). All outputs are in the

2.3 Mitochondrial DNA analyses

Out of the 445 total samples collected, 5 were processed for mitochondrial analyses. A sample was selected from each park, with 2 from Hluhluwe-iMfolozi, as represented on Table 2.

Table 2. Summary of selected samples for mitochondrial DNA (mtDNA) analyses and respective collection sites.

Sample	**Location**
Mas_12_4_H	Hluhluwe-iMfolozi Park
Mas_15_H_3	Hluhluwe- iMfolozi Park

Mas_71_S	Soetdoring Nature Reserve
Mas527_Kr	Kruger National Park
MasKMC1_KD	Koppies Dam Nature Reserve

Assembly, annotation, and phylogeny

The raw genomic data was run through NOVOPlasty 4.3.1 (Dierckxsens et al., 2017) with a kmer size of 33 to assemble the complete mitochondrial DNA (mtDNA) genome and heteroplasmy was evaluated. A total of 40 additional complete mitogenomes from species within the same superfamily Muroidea (*Rodentia*), with 7 *Cricetidae* and 5 *Spalacidae* and the remaining *Muridae*, exept from one outgroup species, for rooting purposes (Chiroptera: Pteropodidae: *Pteropus vampyrus*), were chosen from GeneBank to construct phylogenetic threes. MITOS (Bernt et al., 2013) was used for sequence annotation and visualization. The resulting data was aligned and visualised per individual gene, edited, and trimmed using MEGA-X v.11.0.13 (Tamura et al., 2021). The complete sequences were concatenated using DAMBE v7.3.5 (Xia, 2018). Genetic distances on the complete genomes were calculated with MEGA-X. The substitution models for the *M. natalensis* samples were evaluated with JModelTest 2.1.10 (Darriba et al., 2012), and the best average one was chosen out of three total runs.

Maximum Likelihood (ML) and Bayesian Inference (BI) were used to assess phylogeny mitochondrial genomes. The ML tree was built with PAUP (Cummings, 2004) through an heuristic search of 100 repetitions and a tree-bisection-reconnection (TBR) branch-swapping algorithm with random stepwise addition. The model selected for this tree was *GTR + I + G*, as the best average one, since the software did not allow for partitioning. BI was tested with MrBayes (Ronquist et al., 2012) with a Markov Chain Monte Carlo (MCMC) method for two runs with 1,000,000 generations, with a burnin of 2500. For the BI tree, the model applied was selected considering the available options for the used software (selected models listed in Table 3). These analyses were run using all 13 mtDNA protein coding genes. For visualization, GenomeVx (Conant & Wolfe, 2008) was used for the complete genome image and FigTree (v1.4.4) to build the phylogenetic trees.

Table 3. JModelTest indicated models and chosen model for phylogeny analyses per gene. The applied models are the General Time Reversible (GTR) (Tavaré, 1986) and the Hasegawa-Kishino-Yano (HKY) (Hasegawa et al., 1985). The models account for gamma-distributed rate heterogeneity when + G (Yang, 1994), and for invariable sites when + I (Shoemaker & Fitch, 1989).

Gene	Best model	Model Used
ATP6	TPM3uf + *I* + *G*	HKY + *I* + *G*

ATP8	HKY + *G*	HKY + *G*
COB	GTR + *G*	GTR + *G*
COX1	GTR + *I* + *G*	GTR + *I* + *G*
COX2	TPM3uf + *I* + *G*	HKY + *I* + *G*
COX3	TIM1 + *I* + *G*	GTR + *I* + *G*
NAD1	GTR + *I* + *G*	GTR + *I* + *G*
NAD2	TIM2 + *G*	GTR + *G*
NAD3	TIM2 + *G*	GTR + *G*
NAD4	TIM3 + *G*	GTR + *I* + *G*
NAD4L	GTR + *I* + *G*	GTR + *I* + *G*
NAD5	GTR + *I* + *G*	GTR + *I* + *G*
NAD6	GTR	GTR + *I* + *G*

II Results

3.1 Microsatellite analyses

Statistical analyses

The complete sample data has a mean allele number of 69.556 ±37.71, with size ranging from 16 to 121. Hardy-Weinberg Equilibrium analyses showed highly significant deviation in most loci (excluding MH05 and MH80). The mean observed heterozygosity (Ho) is 0,683 ±0.311 and the mean expected heterozygosity (He) is 0.833 ±0.279. The mean fixation index (F) is 0.199 ±0.065. Table 4 shows selected analyses per locus.

Table 4. Genetic measures per loci for the complete dataset. N is the number of gene copies and Na the number of alleles. Ho and He are the observed and expected heterozygosity, respectively. HWE corresponds to the Hardy-Weinberg equilibrium's p values with a Bonferroni correction (*** indicates significant values of $p < 0.0001$). F corresponds to the fixation index (F_{st}), I is Shannon's information index, and θH indicates the Theta values from a He stepwise model.

Locus	N	Na	Ho	He	HWE	F	I	θH
MH28	884	78	0,828	0,951	***	0,130	3,256	279,52
MH30	886	89	0,847	0,950	***	0,109	3,150	334,87
MH133	870	60	0,759	0,948	***	0,200	3,193	206,24
MH141	882	73	0,871	0,951	***	0,084	3,209	304,18
MH05	886	103	0,761	0,921	0,037	0,174	2,876	149,86
MH74	862	11	0,088	0,129	***	0,316	0,353	0,17
MH105	882	70	0,868	0,946	***	0,082	3,128	379,24
MH188	870	15	0,202	0,612	***	0,669	1,109	3,77
MH80	850	127	0,929	0,959	0,023	0,031	3,472	650,08

The analysis of molecular variance (AMOVA) was calculated with the 4 sampling areas (Figure 2) with the evaluation of genetic differentiation by stepwise mutation (R_{st}) is illustrated in Figure 2. There is some variance within and among individuals, but the highest percentage is found among populations (53%).

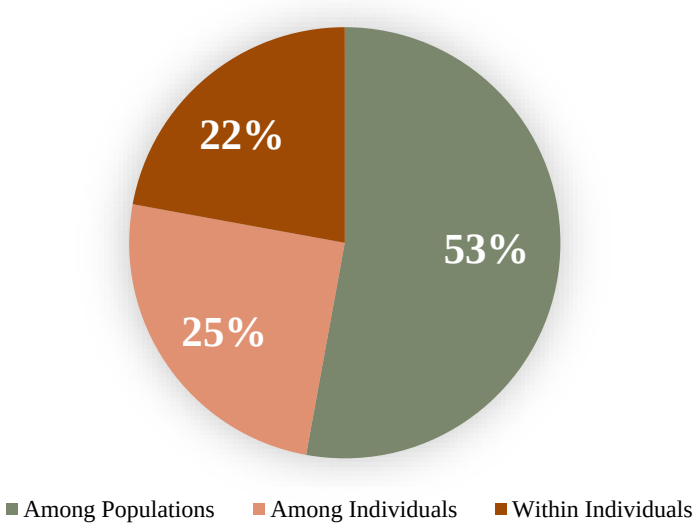


Figure 2. Analysis of molecular variance (AMOVA) in percentages. Results show estimated values among populations and among and within individuals.

Population structure

Analyses for the complete population data highly support a *K = 2* structuring between the samples, with a *K = 3* being the next best model. The Delta*K* (Δ*K*) values were the ones considered when selecting for *K*, instead of the mean probability of Ln(*K*), as these values increased throughout the runs forming an asymptote and indicating values that were not biologically representative. This was equally detected by (Russo et al., 2016) which might indicate that this analysis is not informative for small mammals or species with high genetic diversity.

The obtained q-matrixes (representing inferred genetic cluster per individual) from the STRUCTURE *K = 2* analyses are illustrated in Figure 3. The genetic segregation identifies samples from Soetdoring, Koppies Dam and Kruger as one cluster (named Cluster I), and samples from Hluhluwe-iMfolozi as a second group (Cluster II). There are a few individuals with ancestry from both clusters, indicating minimal admixture.

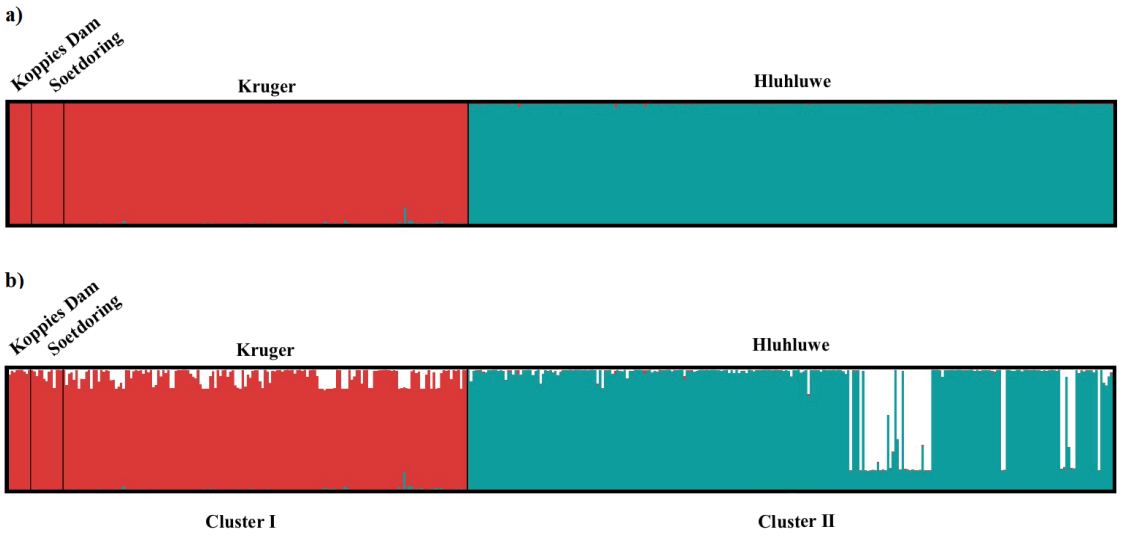


Figure 3. Complete dataset (N=445) genetic population Bayesian clustering analysis in STRUCTURE. Each bar represents a sample, and the colours identify the specific genetic clusters; with the length of the bar corresponding to the likelihood of being assigned to each grouping. a) *K = 2* : visualization of the individual membership coefficients value. In red, identified as Cluster I, are the samples from the Koppies Dam Nature Reserve (N=9) , Soetdoring Nature Reserve (N=13) and Kruger National Park (N=163). The Cluster II, in blue, encompasses the Hluhluwe-iMfolozi samples. b) *K = 3* : Plot for the second highest supported clustering by Δ*K* values. There is some admixture yet the third cluster does not show structured sub clustering.

The sub structuring was further evaluated for each cluster. For Cluster I, Δ*K* values indicate a K = 3 as the highest supported scenario, with K = 2 as the second best (Figure 4). For the Hluhluwe-iMfolozi Park samples, as Cluster II, K = 2 was the most supported analyses, with K = 4 as the second best (Figure 4).

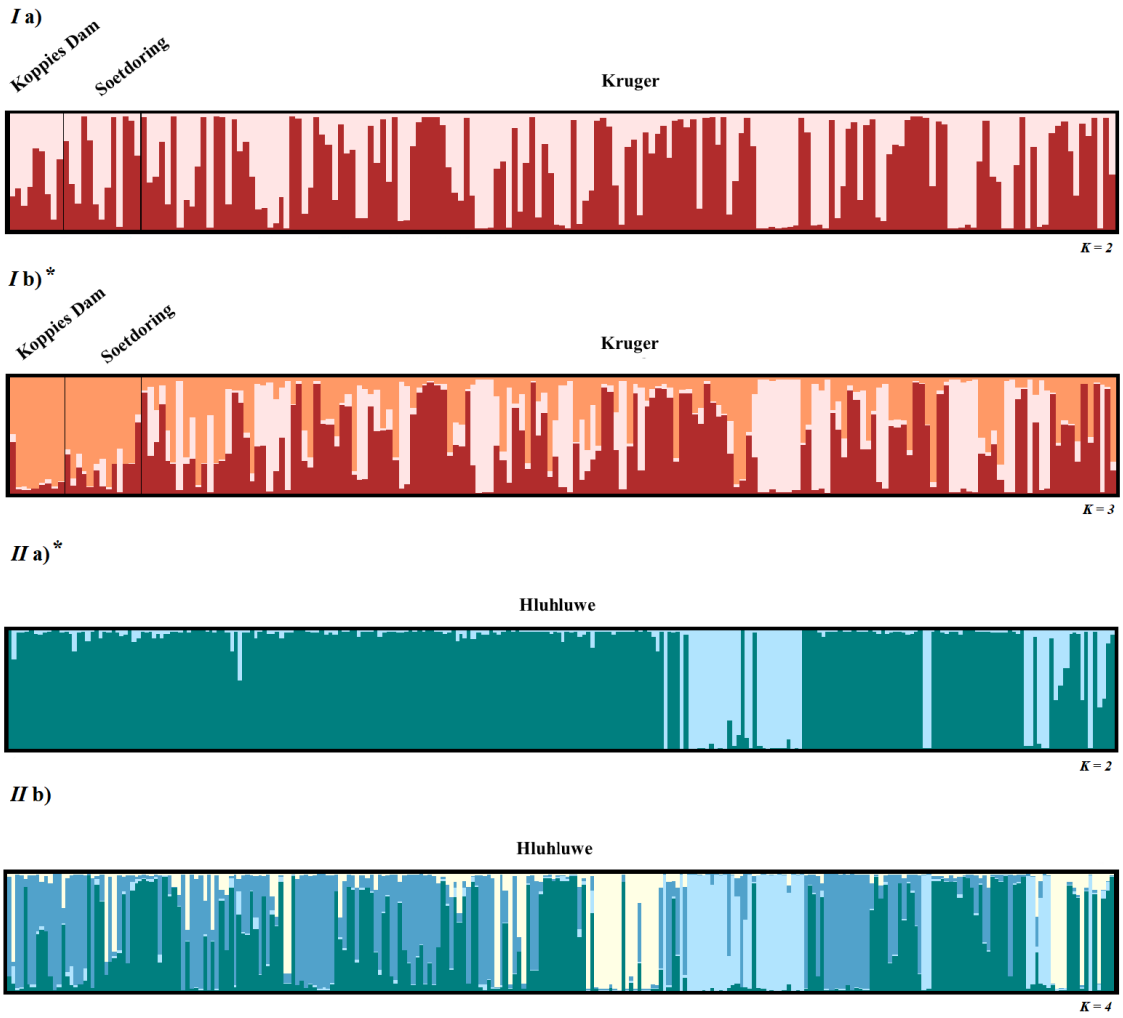


Figure 4. Mapping for the two best values for the STRUCTURE Bayesian clustering analysis of *Cluster I* and *II* population data. The highest supported scenarios are identified with a *.

Additional analyses

Considering the genetic structuring previously presented, the following analyses consider the defined clusters as two comparing populations, where Cluster I includes the combined Koppies Dam Nature Reserve, Soetdoring Nature Reserve and Kruger National Park samples, and Cluster II the Hluhluwe-iMfolozi Park samples.

The mean number of alleles was 60.11 ±35.635 in Cluster I, compared to 19.667 ±10.759 in Cluster II (with a total of 69.556 ±37.710), the values for each locus, of each cluster are displayed in Table 5. The Hardy-Weinberg Equilibrium (WHE in Table 5) analyses showed highly significant deviation in most loci (excluding MH30 and MH105 in cluster I, and MH05 and MH188 in cluster II). The mean observed heterozygosity (Ho) is 0,679 ±0.072 (with cluster I = 0.651 ±0.102, and cluster II = 0.706 ±0.106) and the mean total expected heterozygosity (He) is

0.833 ±0.279 (with cluster I = 0.814 ±0.305, and cluster II = 0.758 ±0.314), the values for locus of each cluster are displayed in Table 5. The fixation index (F_{st}) is 0.117, and Nei's unbiased genetic distance and identity values are 0.743 and 0.475, respectively. The evaluation of genetic differentiation by stepwise mutation (R_{st}) shows a value of 0.496. Analyses of θH values from a He stepwise model were all extremely significative except for Locus MH74 in both clusters. A principal coordinates analyses (PCoA) is illustrated in Figure 5, in this there is a clear differentiation of the previously demonstrated clusters.

Table 5. Genetic measures per loci for the complete dataset. N is the number of gene copies and Na the number of alleles. Ho and He are the observed and expected heterozygosity, respectively. HWE corresponds to the Hardy-Weinberg equilibrium's p values with a Bonferroni correction (*** indicates significant values of $p < 0.005$). F corresponds to the fixation index (F_{st}) and I is Shannon's information index.

Cluster	Locus	N	Na	Ho	He	HWE	F	I
I	**MH28**	368	59	0.809	0.936	***	0,105	2,565
	MH30	370	74	0.832	0.967	***	0,088	2,633
	MH133	360	58	0.688	0.962	***	0,280	3,352
	MH141	364	54	0.818	0.955	***	0,100	2,539
	MH05	368	101	0.690	0.978	0,019	0,266	3,142
	MH74	354	8	0.152	0.170	0,027	-0,024	0,355
	MH105	366	60	0.868	0.972	0,033	0,069	2,901
	MH188	356	11	0.101	0.401	0,058	0,209	0,318
	MH80	330	116	0.896	0.987	0,068	0,064	3,384
II	**MH28**	516	19	0.841	0.907	***	0,072	2,585
	MH30	516	15	0.856	0.902	***	0,049	2,437
	MH133	510	28	0.807	0.930	***	0,130	2,829
	MH141	518	19	0.907	0.904	***	-0,005	2,554
	MH05	518	19	0.810	0.885	***	0,083	2,361
	MH74	508	7	0.043	0.114	***	0,620	0,319
	MH105	516	27	0.868	0.928	***	0,063	2,900
	MH188	514	4	0.272	0.306	0,106	0,110	0,511
	MH80	520	39	0.950	0.944	1,991	-0,008	3,292

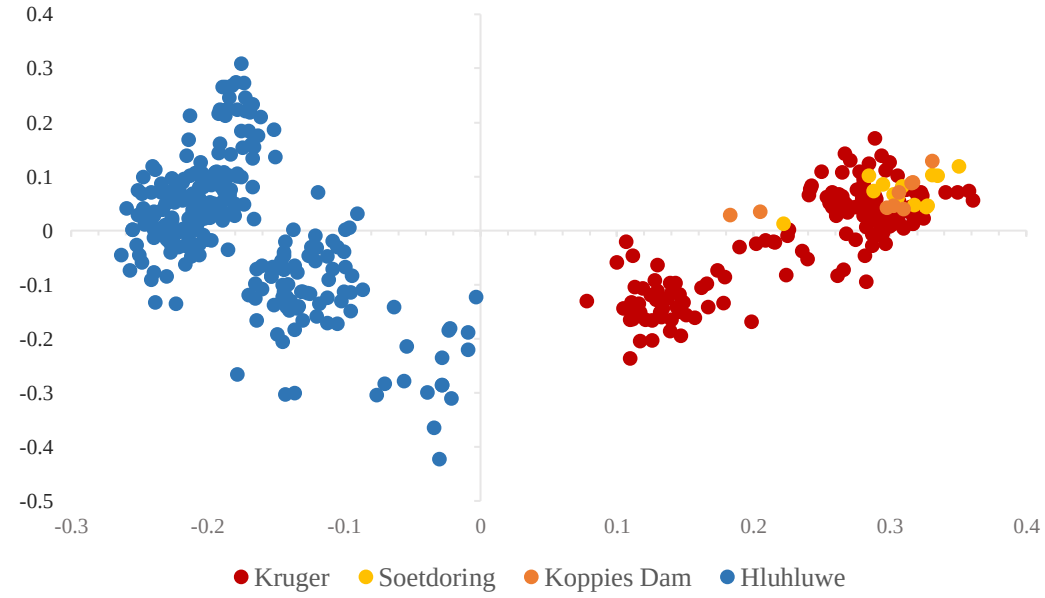


Figure 5. Principal Coordinates Analysis (PCoA) results based on genetic distances (using STR markers).

Demographic assessment

The null scenario (S1) can be rejected for both clusters as it was not supported either by marginal density or p-values. For Cluster I, an expansion scenario (S2) is the most supported, with a p-value of 0.956 and a BF of 6.34. On the other hand, the bottleneck scenario (S3) is the indicated one for Cluster II, with p-value of 0.764 and an extremely well supported BF of 113,3. The indicated expansion time for Cluster I is 268 ya, and the bottleneck time for Cluster II is 267 ya.

3.2 Phylogeny assessment with mtDNA

The full mtDNA, consisting of 15,709 basepairs was generated through NOVOPlasty. This comprised all 13 protein-coding genes (PCG), 2 rRNA sequences, 22 tRNA sequences and 2 origins of replication (OR) commonly found in mammal mitochondria (Boore, 1999a). The majority of PCGs and most tRNA genes are found on the forward (+) strand, which is significantly longer than the reverse (-) strand. The reverse strand only contains NAD6 and 8 different tRNA genes. The two largest intergenic overlaps were between ATP8 and ATP6 (43bp), and NAD5 and NAD6 (23bp), with 13 smaller others (≤9bp). These overlaps are prevalent in mtDNA genomes (Boore, 1999a). Table 6 provides a more comprehensive overview of each gene and characteristics.

The longest protein-coding gene, NAD5, had 1,839 base pairs, and the shortest, ATP8, had 204 base pairs. The shortest tRNA is TRNS1, which was 59 bp long, while the longest was TRNL2, which has 75 bp. The rRNA was 955bp and RRN1,563bp. Origins of replication are both in the

+ strand, OH has 243bp and it was located within a putative control region (where itis preceded by 262 bp and followed by 318 bp), and OL has 30 bp. In Figure 6 gene groupings and positioning were illustrated.

Table 6. An overview of every gene found in *Mastomys natalensis*' mitochondrial DNA. The Start and Stop column specify in which bp of the respective strand the gene begins and ends. Strand indicates whether it is located in the positive (+) or negative (-). Length is the number of basepairs per gene. Ovl/nc represents the overlap (negative number) and increment (positive number) of nucleotides in bp. The Start and Stop codons are identified in the last two columns, with incomplete ones in parenthesis.

Name	Start (bp)	Stop (bp)	Strand	Length (bp)	Ovl/nc	Start codon	Stop codon
trnS2	469	537	-	69	3		
trnD	541	607	+	67	1		
cox2	609	1292	+	684	3	ATG	TAA
trnK	1296	1359	+	64	1		
atp8	1361	1564	+	204	-43	ATG	TAA
atp6	1522	2202	+	681	-1	ATG	TAA
cox3	2202	2986	+	785	-1	ATG	TA(A)
trnG	2986	3052	+	67	0		
nad3	3053	3400	+	348	1	ATA	TAG
trnR	3402	3469	+	68	2		
nad4l	3472	3768	+	297	-7	ATG	TAA
nad4	3762	5139	+	1378	0	ATG	T(AA)
trnH	5140	5208	+	69	0		
trnS1	5209	5267	+	59	-1		
trnL1	5267	5337	+	71	-9		
nad5	5329	7167	+	1839	-23	ATA	TAA
nad6	7145	7663	-	519	0	ATG	TAA
trnE	7664	7732	-	69	5		
cob	7738	8881	+	1144	0	ATG	T(AA)
trnT	8882	8950	+	69	0		
trnP	8951	9017	-	67	262		
OH	9280	9522	+	243	318		
trnF	9841	9910	+	70	-1		
rrnS	9910	10864	+	955	2		
trnV	10867	10930	+	64	0		
rrnL	10931	12493	+	1563	-2		
trnL2	12492	12566	+	75	-3		
nad1	12564	13523	+	960	-2	ATA	TAG
trnI	13522	13590	+	69	-3		
trnQ	13588	13657	-	70	4		
trnM	13662	13730	+	69	0		
nad2	13731	14768	+	1038	-2	ATA	TAG
trnW	14767	14832	+	66	1		
trnA	14834	14902	-	69	1		
trnN	14904	14973	-	70	2		
OL	14976	15005	+	30	-1		
trnC	15005	15071	-	67	0		
trnY	15072	15139	-	68	1		
cox1	15141	471	+	1545	-3	ATG	TAA

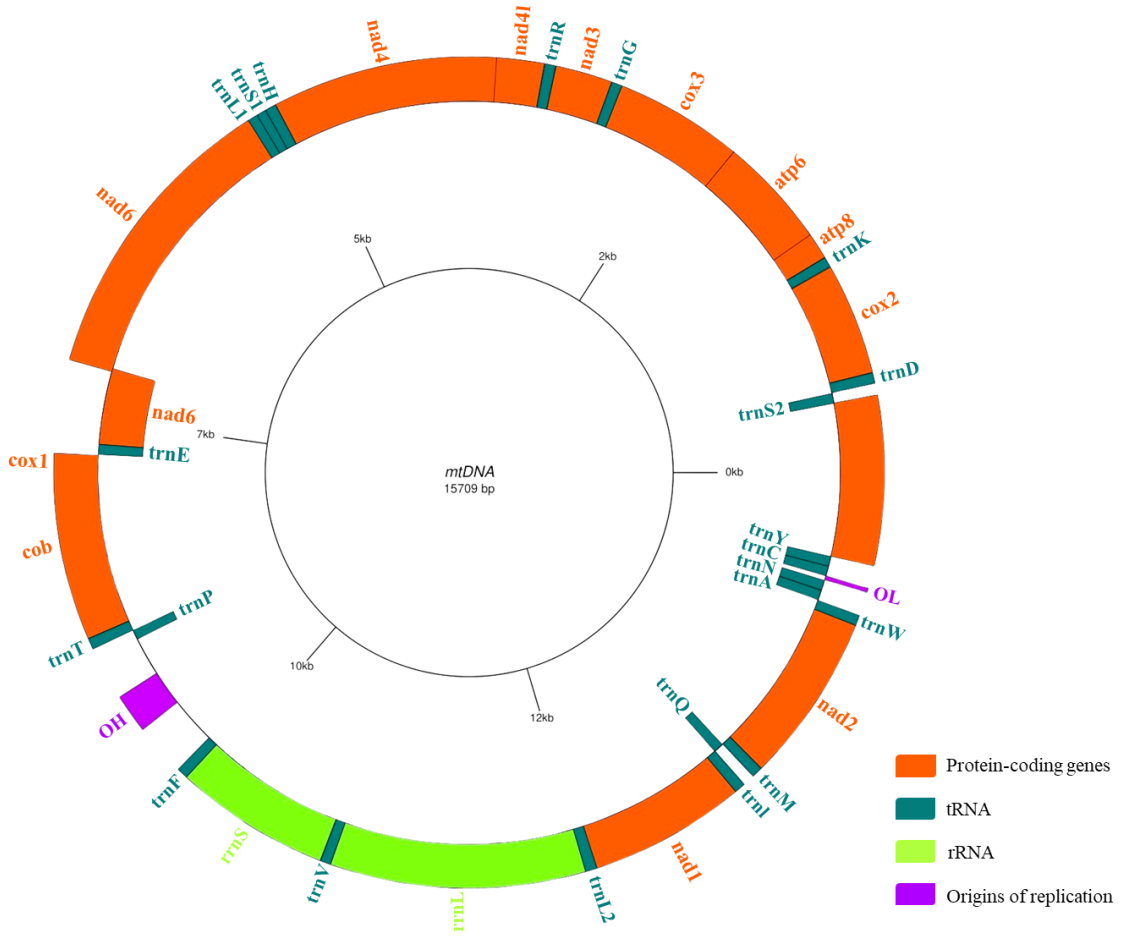


Figure 6. Visualization of the sequenced *Mastomys natalensis* mtDNA. The + strand is the outer ring and itis counterclockwise oriented. The inside ring is the – strand and can be read clockwise.

Phylogenetic analyses

Generated phylogenetic trees from ML and BI are represented in Figure 7 and 8, respectively. Both analyses placed the Mas_12_4_H sample, one of two Hluhluwe-iMfolozi Park samples, outside of the branch with the remaining samples. This sample was grouped with *Mastomys erythroleucus*, a central-north African species. The remaining samples were grouped together in a clade that incorporated *Mastomys coucha*. Both trees integrated *Cricetidae* individuals within the *Muridae* clades. Table 7 displays the variation between the 5 processed samples, where we can see that genetic difference between this sample and the remaining ones is over 9.5%.

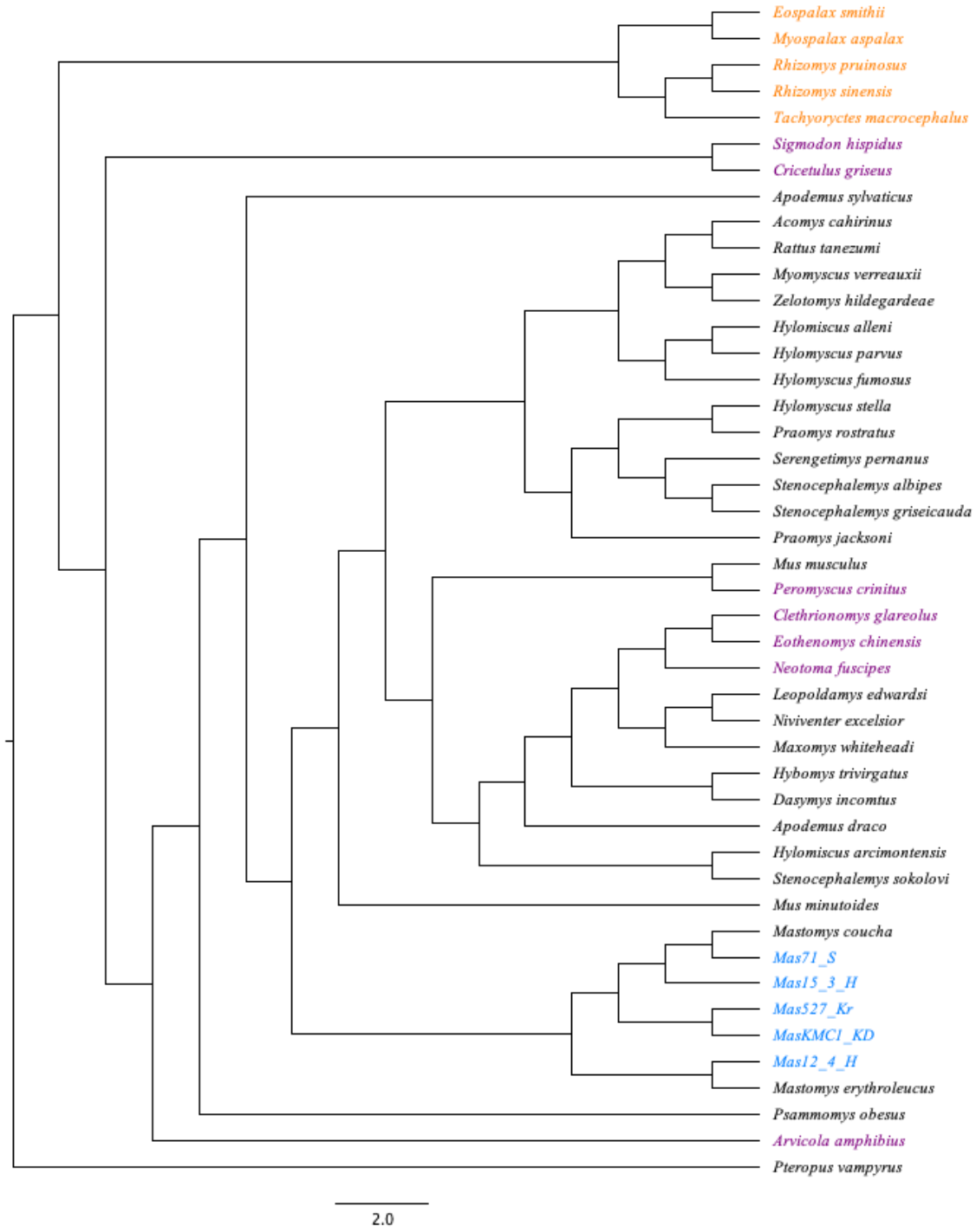


Figure 7. Visual representation of the Maximum Likelihood (ML) phylogeny built with PAUP. Species are coloured by family: *Muridae* in black (with analysed samples in blue), Spalacidae in orange and Cricetidae in purple. Species names are labelled on the right.

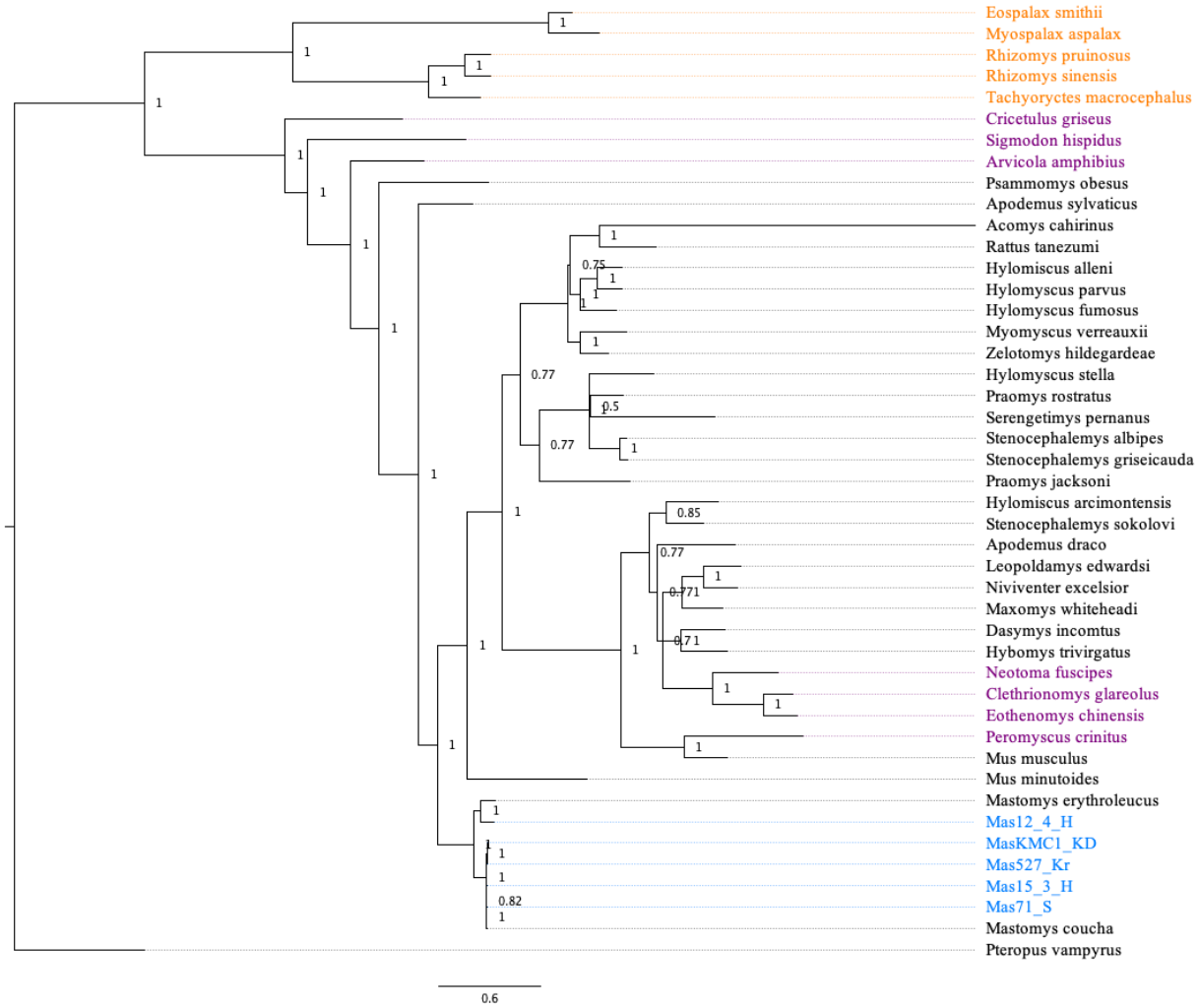


Figure 8. Visual representation of the Bayesian Inference (BI) phylogeny built with MrBayes. Probability values are indicated from 0-1. Species are coloured by family: *Muridae* in black (with analysed samples in blue), Spalacidae in orange and Cricetidae in purple. Species names are labelled on the right.

Table 7. The number of base substitutions per site from between sequences are shown. Standard error estimate(s) are shown above the diagonal. Analyses were conducted using the Maximum Composite Likelihood model. The rate variation among sites was modelled with a gamma distribution. All ambiguous positions were removed for each sequence pair. The final dataset a total of 11258 positions. Evolutionary analyses were conducted in MEGA.

	Mas12_4_H	Mas15_3_H	Mas71_S	Mas527_Kr	MasKMC1_KD
Mas12_4_H		0,004487	0,004473	0,004567	0,004559
Mas15_3_H	0,096212		0,000396	0,000765	0,000702
Mas71_S	0,095883	0,001781		0,000764	0,000719
Mas527_Kr	0,097347	0,006176	0,006176		0,000716
MasKMC1_KD	0,097210	0,005365	0,005544	0,005364	

III. Discussion

Population structuring of Mastomys natalensis

Statistical outputs strongly indicate selective pressure, with significant values for *HWE*, I, *F*-statistics, R_{st} and Nei's genetic distances; yet this is not universal throughout different loci, which suggests that these pressures do not have a uniform effect in the genome. The results obtained through both statistical and STRUCTURE analyses clearly illustrate the high genetic variability the study species has. While previous population genetic analyses indicated how variable *M. natalensis* populations can be (Colangelo et al., 2013; Russo et al., 2016), only now has such contrasting clustering been identified. AMOVA (Figure 2) analyses show a clear a split between the populations, and the PCoA maps (Figure 5) the sampling areas with the same clustering then identified through the STRUCTURE analyses. When evaluating diversity considering the clusters identified, the high levels of variance among individuals found (SX) is in accordance with the findings by Van Hooft et al. (2008), since there is high kin-clustering at these small scales. Influence on genetic structure is further supported by the high deviation from equilibrium.

While landscape factors and geographical distance was not considered in the analysis, the geographical location of the Hluhlue-iMfolozi Park might be the primary factor for such segregation. The area is mainly surrounded by the Drakensberg Escarpment, as well as by the Indian ocean, being about 30 km east. While the escarpment could allow for a passage through the coastline, this area has the St. Lucia Lake and is followed by wetlands. One of the main determinants of distribution in this species is altitude, being rarely present at high elevations and displaying a preference for easter-facing slopes (Colangelo et al., 2013; Russo et al., 2016; Venturi et al., 2004). Another strong habitat preference for *M. natalensis* is soil humidity, as they are commonly associated with wetlands and higher levels of rainfall (Kneidinger et al., 2014; Venturi et al., 2004). Movement restriction has previously been identified as a primary force shaping genetic structuring in rodents (Flores-Manzanero et al., 2019; Garrido-Garduño et al., 2016; Mapelli et al., 2020). Additionally, this species has a high tendency to remain in its natal area if breeding starts at an early age, and rarely displays long-distance dispersal (Leirs et al., 1996a; Van Hooft et al., 2008). This highly specific landscape, in conjunction with the previously mentioned reproductive and migration patterns, might encompass the perfect conditions that lead to the isolation of this population. In the subclustering analyses of Cluster I, the results do not reflect strongly divided populations, as we can observe individuals from Krugger with similar genetic patterns. Considering the habitat difference, it could be expected stronger differentiations, which might have failed to reflect these differences due to the small population size for the grassland habitat, in comparison with the remaining sampling areas. For the second cluster, the observed patterns reflect findings by Russo et al. (2016).

Establishing demographic stories with ABCtoolbox

Gryseels et al. (2016), while analysing population structure of urban and rural *Mastomys natalensis* and the possible factors shaping it, found that even with the presence of landscape and environmental determinants, as well as gene flow between the populations, subclustering seemed to be shaped by urban human activity. As a highly commensal species, that is a major agricultural pest, this specie is moulded with anthropogenic changes (Brouat et al., 2007; Mwanjabe et al., 2002; Odhiambo et al., 2008). The contemporary evolution analyses strongly indicates that population structure was affected during the early colonial periods of this area (Texp = 268; Tbot = 267). The expansion of the Cluster I data could be correlated with the land use changes and opportunities that were created with the European settlements and agricultural changes during this time. The bottleneck found for the Hluhluwe cluster could reflect the subclustering found for this population (Figure 4. *II* b). It is important to denote that, even though BF values were not as significant for these analyses, a scenario with a slightly wider timeframe (100-400ya) indicated a bottleneck (S3) for Cluster I and an expansion (S2) for Cluster II. The results obtained signal that reshaping of population structure through this period can be more complex than initially expected. It seems that both Clusters went through a bottleneck prior to the evaluated timeframe, segregating these populations; yet major contemporary events happened in the past 400 years. It is important to consider that, even with the use of microsatellite data, the simulated scenarios might be too recent to be reflected here.

Phylogeny and mtDNA particularities

The presented mitochondrial size and composition are parallel to other Muroids , and the outputs successfully identified all common sequences to mammalian mtDNA (Boore, 1999b; Chen et al., 2012; Deng et al., 2023; Hao et al., 2011; Lecompte et al., 2008). One of the particularities found in both trees is the separation of the Cricetidae family with integration of *Peromyscus crinitus*, Clethrionomys glareolus, *Neotoma fuscipes* and *Eothenomys chinensis* within Muridae samples. It is not clear why this happened as this does not align with previous studies (DeBry & Sagel, 2001; Flynn et al., 1985). Future analyses should not only consider different tree-building algorithms but also identify the specific origins of these samples.

In the analysis of Mastomys natalensis samples, all were closely related except for one, Mas12_4_H. This outlier, despite being analyzed for both mitochondrial and nuclear DNA and aligning with its species genetically, was found to cluster with Mastomys erythroleucus, a species prevalent in West Africa (Brouat et al., 2009). This unusual grouping, diverging from its expected

clade and aligning with a different species, suggests a potential introgression event, reminiscent of findings in other small mammal studies (BRUNET et al., 2002; Jahner et al., 2021; Lawson et al., 2021). The concept of introgression, particularly in small mammals, involves the exchange of genetic material between distinct species through hybridization and backcrossing (REfs). This phenomenon is well-documented across various species (Giev examnples), providing valuable insights into the evolutionary dynamics and species boundaries that define them. In the context of this specific sample of *Mastomys natalensis* and its unexpected genetic similarity with *Mastomys erythroleucus*, such instances underscore the complexity of mammalian genetics and evolutionary history. For instance, similar genetic intertwining has been observed in North American rodents, where environmental and historical factors facilitated gene flow between distinct species (REFs, examples). This case underscores the complexity of mammalian genetics and the need for further investigation to understand these evolutionary dynamics.

IV. Conclusion

Based on the comprehensive examination of population genetics, evolutionary history, and phylogenetic relationships within Mastomys natalensis populations across South Africa, this study brings some light on the complex interplay of genetic diversity, demographic history, and environmental factors shaping this species. The analyses employing microsatellite markers and mitochondrial DNA revealed significant genetic variation within and among populations, underlining the influence of evolutionary forces such as gene flow, mutation, and natural selection. These findings contribute to our understanding of Mastomys natalensis as both a pest and a commensal species, offering insights crucial for developing effective management and conservation strategies. The study underscores the importance of integrating genetic data with ecological and historical perspectives to address the challenges posed by species that impact agricultural productivity and public health.

V. Supplementary material

Table S1. Genetic measures per loci per cluster and HWE adjustments.

Cluster	Locus	DF	ChiSq	Prob	p-adj FDR	q	Bonferroni
I	MH05	630,0000	1456,7460	0,0000	0,0000	0,0083	0,0000
	MH105	378,0000	405,7130	0,1570	0,1762	0,0444	2,8194
	MH133	780,0000	1750,5920	0,0000	0,0000	0,0056	0,0000
	MH141	171,0000	371,3120	0,0000	0,0000	0,0222	0,0000
	MH188	21,0000	362,9880	0,0000	0,0000	0,0111	0,0000
	MH28	171,0000	413,4420	0,0000	0,0000	0,0167	0,0000
	MH30	231,0000	265,8070	0,0580	0,0691	0,0417	1,0369
	MH74	10,0000	47,9540	0,0000	0,0000	0,0333	0,0000
	MH80	780,0000	1085,1330	0,0000	0,0000	0,0278	0,0000
II	MH05	171,0000	183,3380	0,2460	0,2603	0,0472	4,4253
	MH105	351,0000	488,2650	0,0000	0,0000	0,0361	0,0000
	MH133	378,0000	560,6310	0,0000	0,0000	0,0306	0,0000
	MH141	171,0000	469,4730	0,0000	0,0000	0,0139	0,0000
	MH188	6,0000	4,2610	0,6410	0,6414	0,0500	11,5447
	MH28	171,0000	334,9390	0,0000	0,0000	0,0250	0,0000
	MH30	105,0000	281,3160	0,0000	0,0000	0,0194	0,0000
	MH74	21,0000	721,5280	0,0000	0,0000	0,0028	0,0000
	MH80	741,0000	868,3190	0,0010	0,0010	0,0389	0,0146

S3. G-stats per loci

Locus	Fis	Fst	Gis	Gst	G'stN	G'stH	G''st	Dest
MH28	-0,008	0,061	0,018	0,043	0,056	0,49	0,497	0,467
MH30	0,115	0,071	0,14	0,05	0,066	0,604	0,611	0,583
MH133	0,422	0,041	0,443	0,014	0,019	0,289	0,293	0,279
MH141	0,122	0,063	0,147	0,042	0,055	0,607	0,612	0,59
MH05	0,27	0,083	0,294	0,059	0,078	0,689	0,695	0,669
MH74	-0,512	0,209	-0,493	0,201	0,251	0,307	0,35	0,132
MH105	0,045	0,054	0,07	0,034	0,045	0,467	0,473	0,448
MH188	0,053	0,625	0,079	0,616	0,681	0,797	0,832	0,472
MH80	0,048	0,051	0,087	0,021	0,028	0,456	0,46	0,444

Table S3. Summary of resulting values from Structure Harvester's analysis of the complete dataset. *K* denotes the given number of clusters in the respective analyses, LnP(*K*) is the log likelihood. A total of 20 *K* values were tested. Highlighted in green is *K = 2*, with the highest *Delta K* (Δ*K*), rate of change of the log likelihood (Ln'(*K*)) and absolute value of the second derivative of the log likelihood (|Ln''(*K*)|). Highlighted in yellow is *K = 4*, the next best Δ*K* value.

K	Reps	Mean LnP(*K*)	Stdev LnP(*K*)	Ln'(*K*)	|Ln''(*K*)|	Delta *K*
1	19	-20703.831579	0.245068	—	—	—
2	20	-17940.355000	0.256443	2763.476579	2329.711053	9084.702927
3	19	-17506.589474	93.742104	433.765526	178.407632	1.903175
4	19	-17251.231579	10.804220	255.357895	65.781871	6.088535
5	18	-17061.655556	66.766267	189.576023	26.662573	0.399342
6	19	-16898.742105	58.324545	162.913450	55.818713	0.957036
7	19	-16791.647368	64.253330	107.094737	64.526316	1.004249
8	19	-16749.078947	90.550830	42.568421	90.057895	0.994556
9	19	-16616.452632	59.787600	132.626316	88.857895	1.486226
10	19	-16572.684211	92.214823	43.768421	23.710526	0.257123
11	19	-16552.626316	116.086595	20.057895	82.752632	0.712853
12	19	-16449.815789	87.862767	102.810526	100.021053	1.138378
13	19	-16447.026316	160.161144	2.789474	11.221053	0.070061
14	19	-16433.015789	157.019310	14.010526	21.684211	0.138099
15	19	-16397.321053	137.892102	35.694737	52.089474	0.377755
16	19	-16413.715789	265.871574	-16.394737	108.357895	0.407557
17	19	-16538.468421	509.502143	-124.752632	20.273684	0.039791
18	19	-16683.494737	490.873048	-145.026316	415.457895	0.846365
19	19	-17243.978947	1382.935992	-560.484211	783.789474	0.566758
20	19	-17020.673684	999.059682	223.305263	—	—

Table S4. Summary of resulting values from Structure Harvester's analysis of the Cluster I dataset (Soetdoring, Koppies Dam and Kruger). *K* denotes the given number of clusters in the respective analyses, LnP(*K*) is the log likelihood. A total of 10 *K* values were tested. Highlighted in green is *K = 3*, with the highest *Delta K* (Δ*K*), rate of change of the log likelihood (Ln'(*K*)) and absolute value of the second derivative of the log likelihood (|Ln''(*K*)|). Highlighted in yellow is *K = 2*, the next best Δ*K* value.

K	Reps	Mean LnP(*K*)	Stdev LnP(*K*)	Ln'(*K*)	|Ln''(*K*)|	Delta *K*
1	19	-7489.063158	0.624687	—	—	—
2	20	-7367.180000	18.987908	121.883158	47.228421	2.487289
3	19	-7198.068421	9.622950	169.111579	148.953684	15.479004
4	19	-7177.910526	48.094016	20.157895	1.347368	0.028015
5	18	-7159.100000	30.857147	18.810526	37.810526	1.225341
6	19	-7102.478947	56.621340	56.621053	4.115789	0.072690
7	19	-7049.973684	104.356642	52.505263	44.936842	0.430608
8	19	-7042.405263	83.213417	7.568421	80.094737	0.962522
9	19	-7114.931579	183.037859	-72.526316	1.442105	0.007879
10	19	-7188.900000	180.641739	-73.968421	—	—

Table S5. Summary of resulting values from Structure Harvester's analysis of the Cluster II (Hluhluwe-iMfolozi) dataset. *K* denotes the given number of clusters in the respective analyses, LnP(*K*) is the log likelihood. A total of 10 *K* values were tested. Highlighted in green is *K = 2*, with the highest *Delta K* (Δ*K*), rate of change of the log likelihood (Ln'(*K*)) and absolute value of the second derivative of the log likelihood (|Ln''(*K*)|). Highlighted in yellow is *K = 4*, the next best Δ*K* value.

K	Reps	Mean LnP(*K*)	Stdev LnP(*K*)	Ln'(*K*)	|Ln''(*K*)|	Delta *K*
1	19	-10336.210526	0.387147	—	—	—
2	20	-9877.775000	0.472257	458.435526	278.497368	589.716176
3	19	-9697.836842	7.790608	179.938158	13.296053	1.706677
4	19	-9531.194737	11.923109	166.642105	59.636257	5.001737
5	18	-9424.188889	11.577610	107.005848	40.638012	3.510052
6	19	-9357.821053	18.672368	66.367836	0.052047	0.002787
7	19	-9291.505263	19.702974	66.315789	79.894737	4.054958
8	19	-9305.084211	26.224125	-13.578947	11.684211	0.445552
9	19	-9330.347368	121.253332	-25.263158	82.668421	0.681783
10	19	-9272.942105	82.898555	57.405263	—	—

Section 1: ABCtoolbox data

Cluster	Scen	Marginal density	P value	Marginal density	P value
C1	Null	100576	0.044	588635	0.644
C1	Exp	637455	0.956	275098	0.496
C1	Bot	95799.4	0.612	154205	0.912
C2	Null	311578	0.84	197267	0.456
C2	Exp	12952.9	0.016	368769	0.912
C2	Bot	1.46764e+06	0.764	564630	0.252
		Scen:	200 – 400		100 - 400

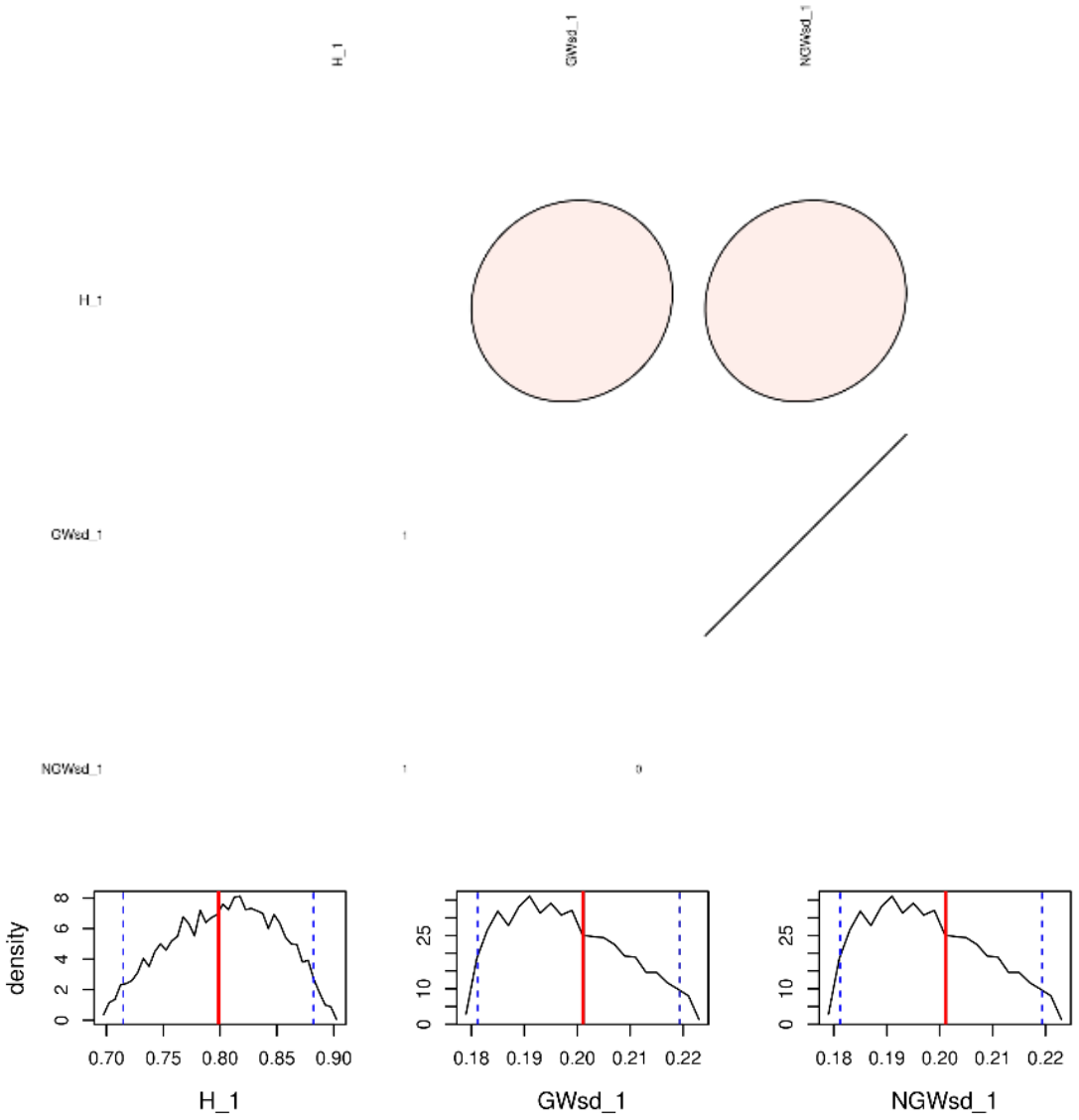
Results For Priors = 200 – 400 generations

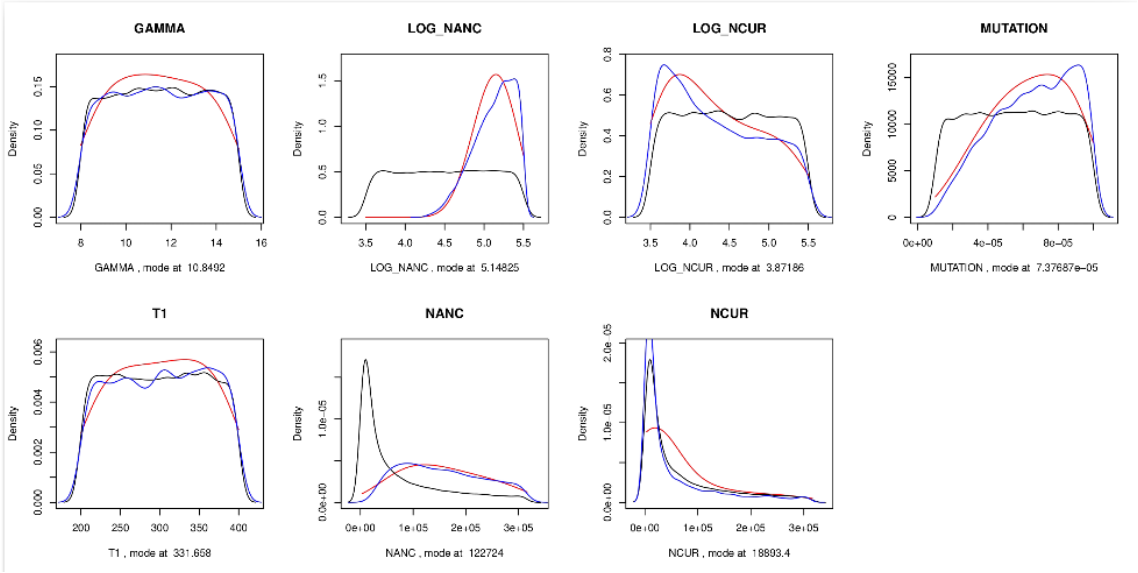
- **Cluster I (Koppies, Soetdoring and Kruger)**

S1 : Null

- marginal density: 100576 (p-value 0.044)

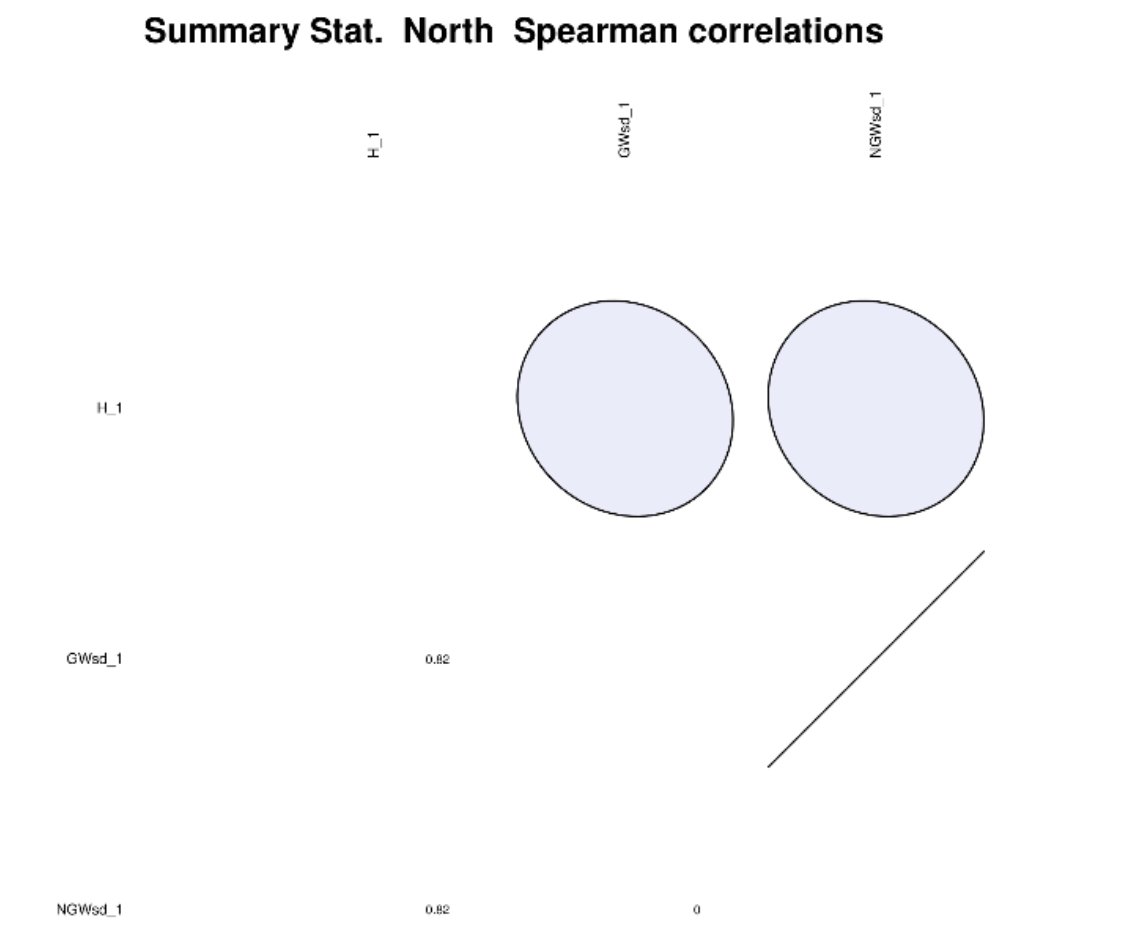
Summary Stat. North Spearman correlations

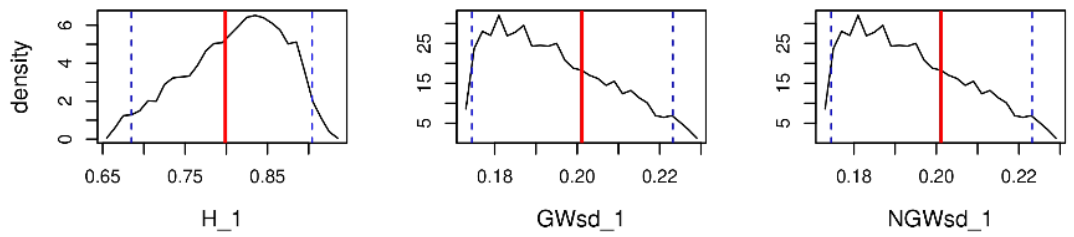


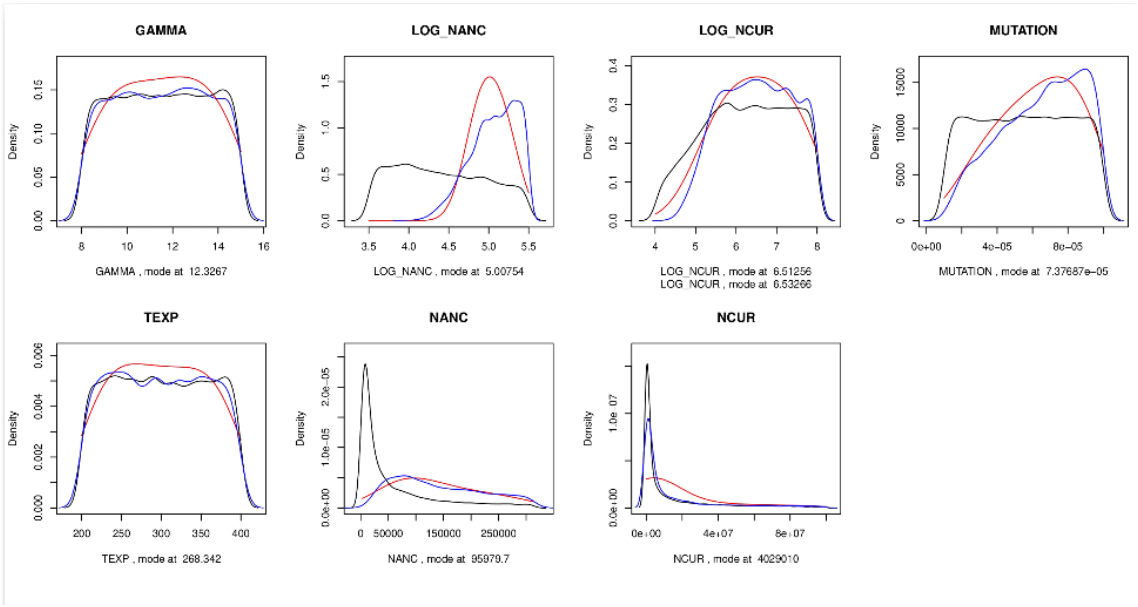


S2: Expansion

- marginal density: 637455 (p-value 0.956)



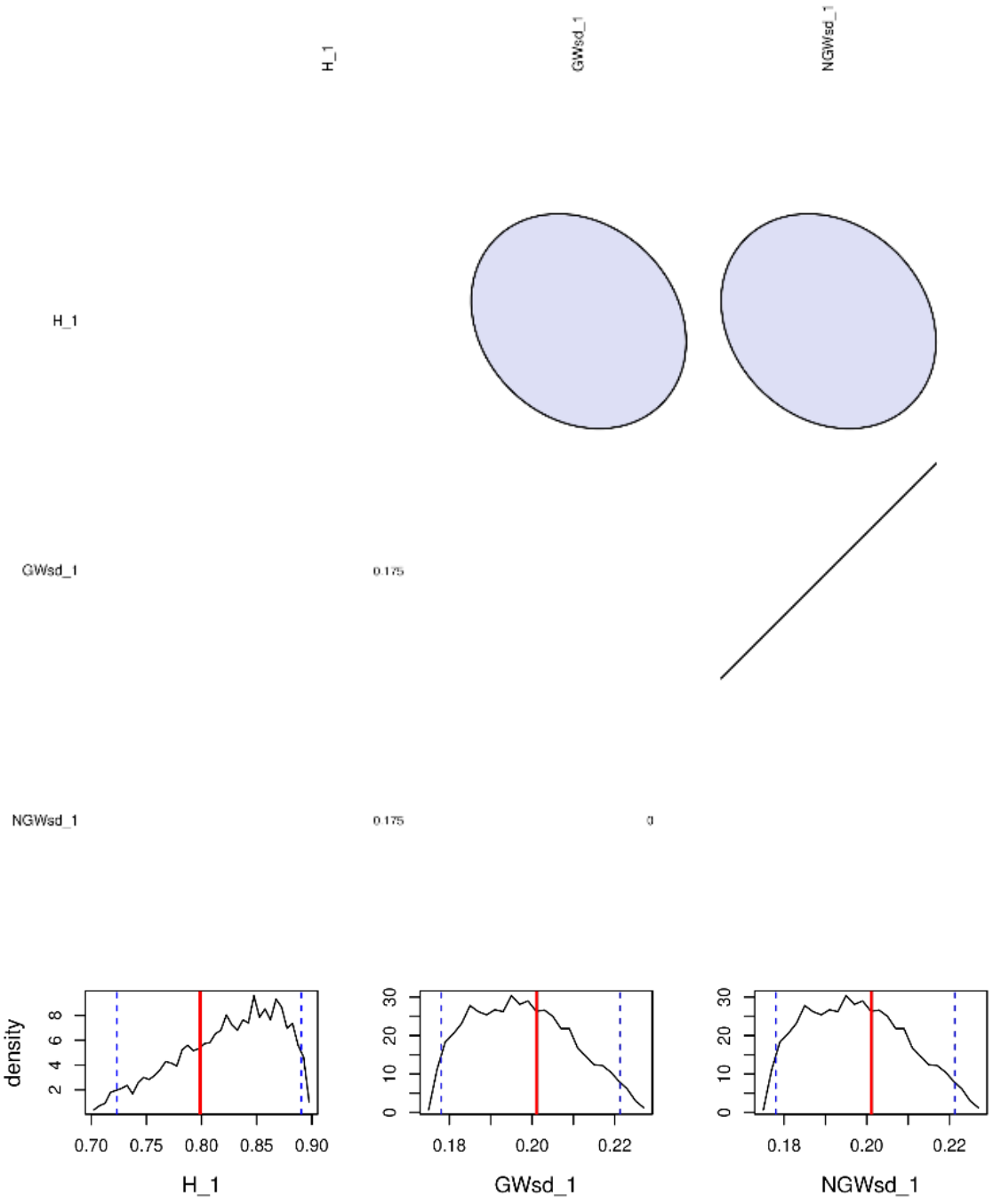


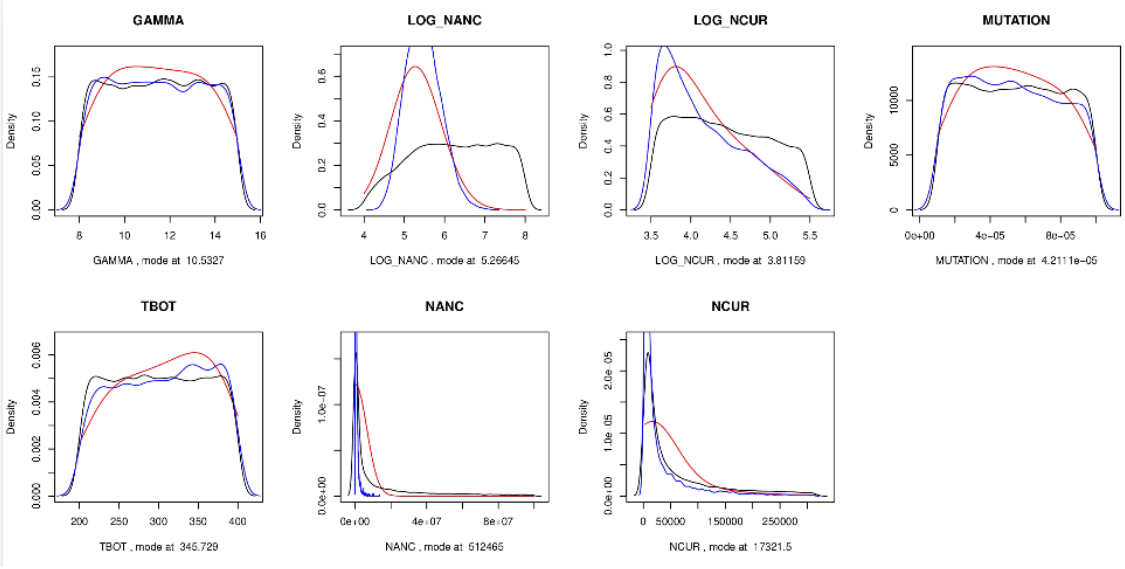


S3: Bottleneck

- marginal density: 95799.4 (p-value 0.612)

Summary Stat. North Spearman correlations



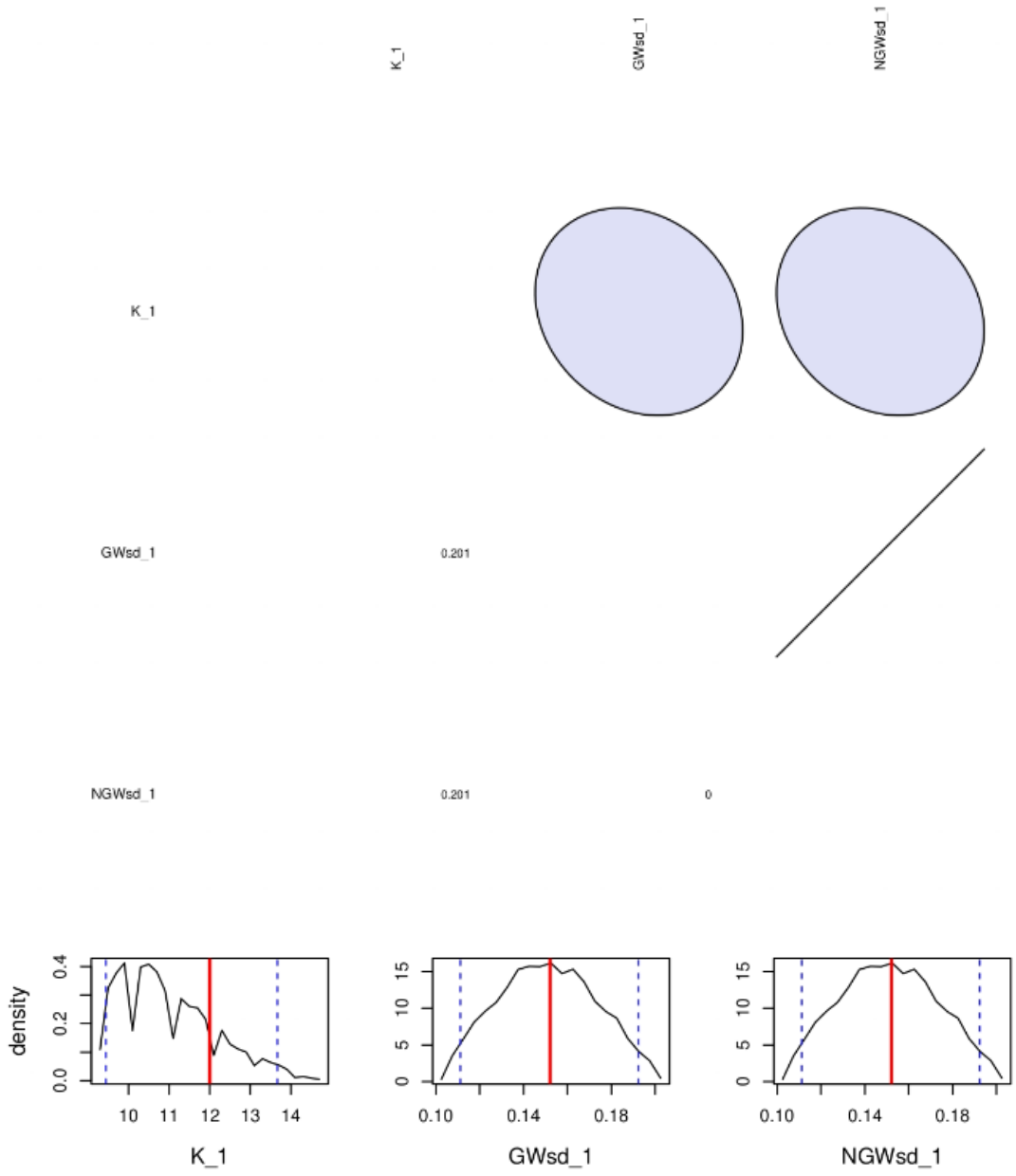


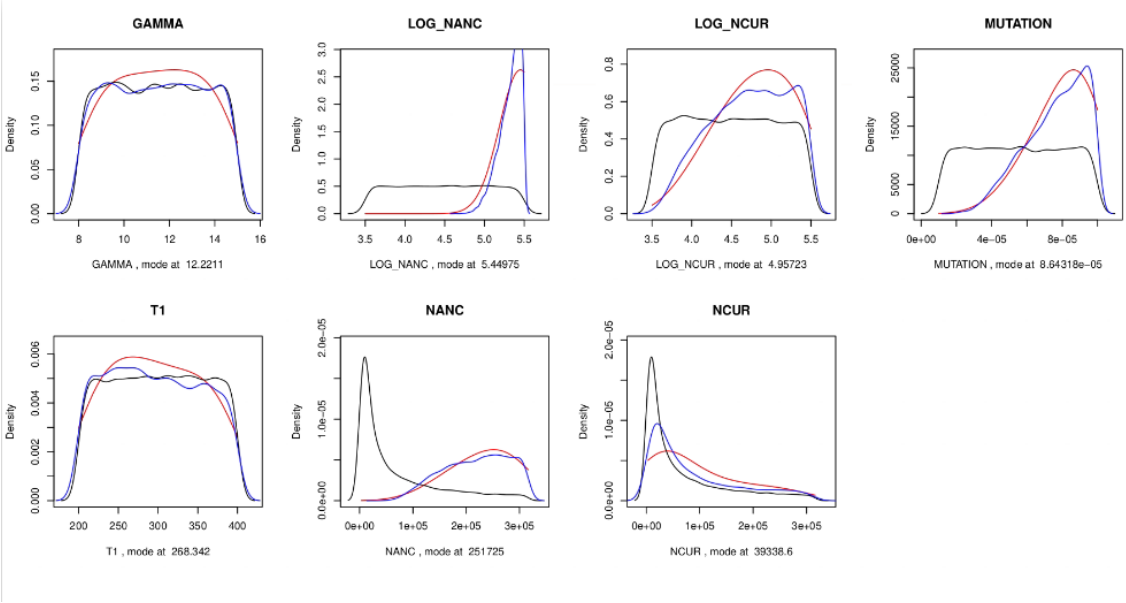
- **Cluster II (Hluhluwe)**

S1: Null

- marginal density: 311578 (p-value 0.84)

Summary Stat. North Spearman correlations

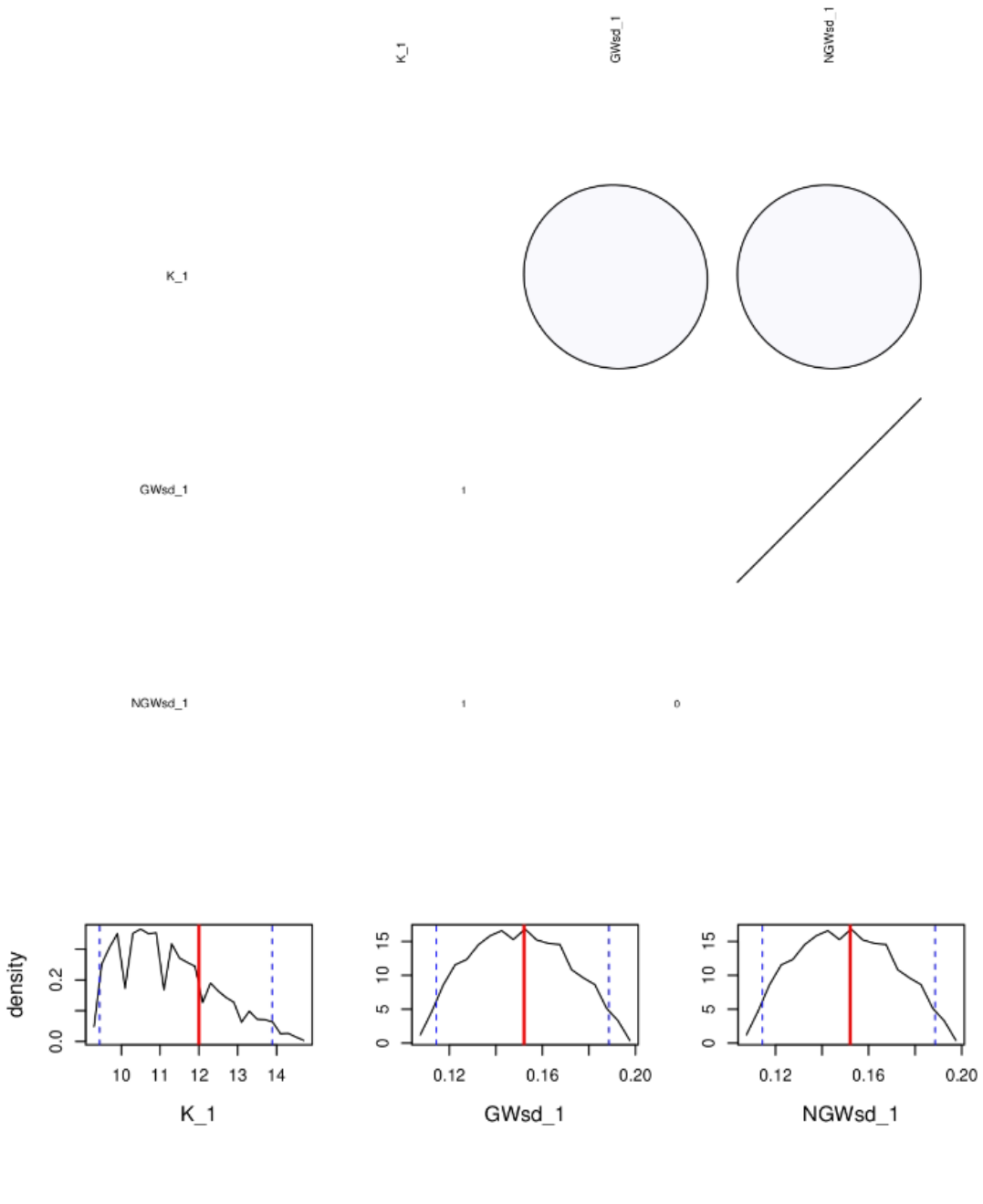


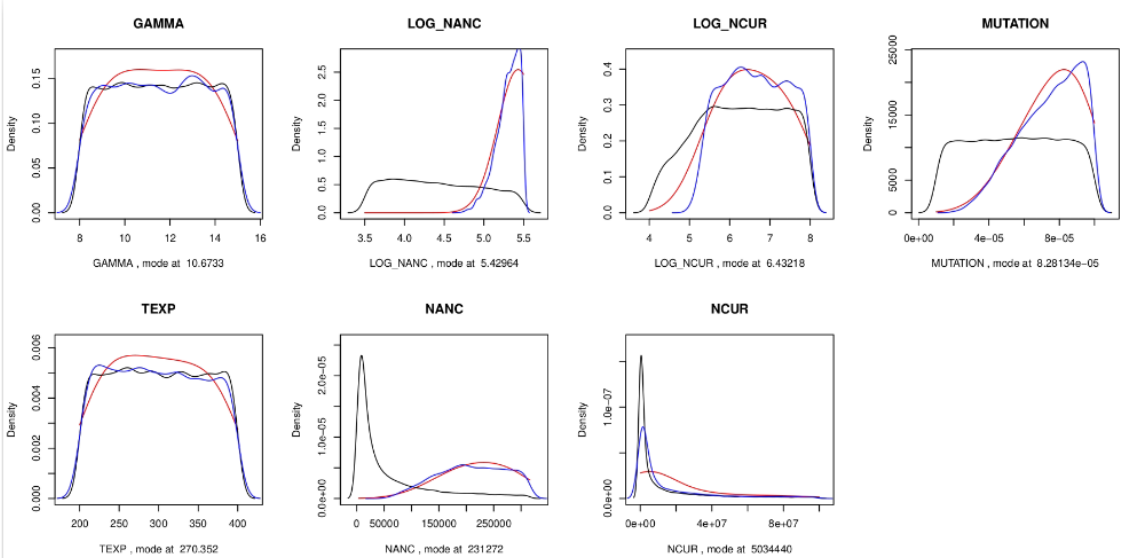


S2: Expansion

- marginal density: 12952.9 (p-value 0.016)

Summary Stat. North Spearman correlations

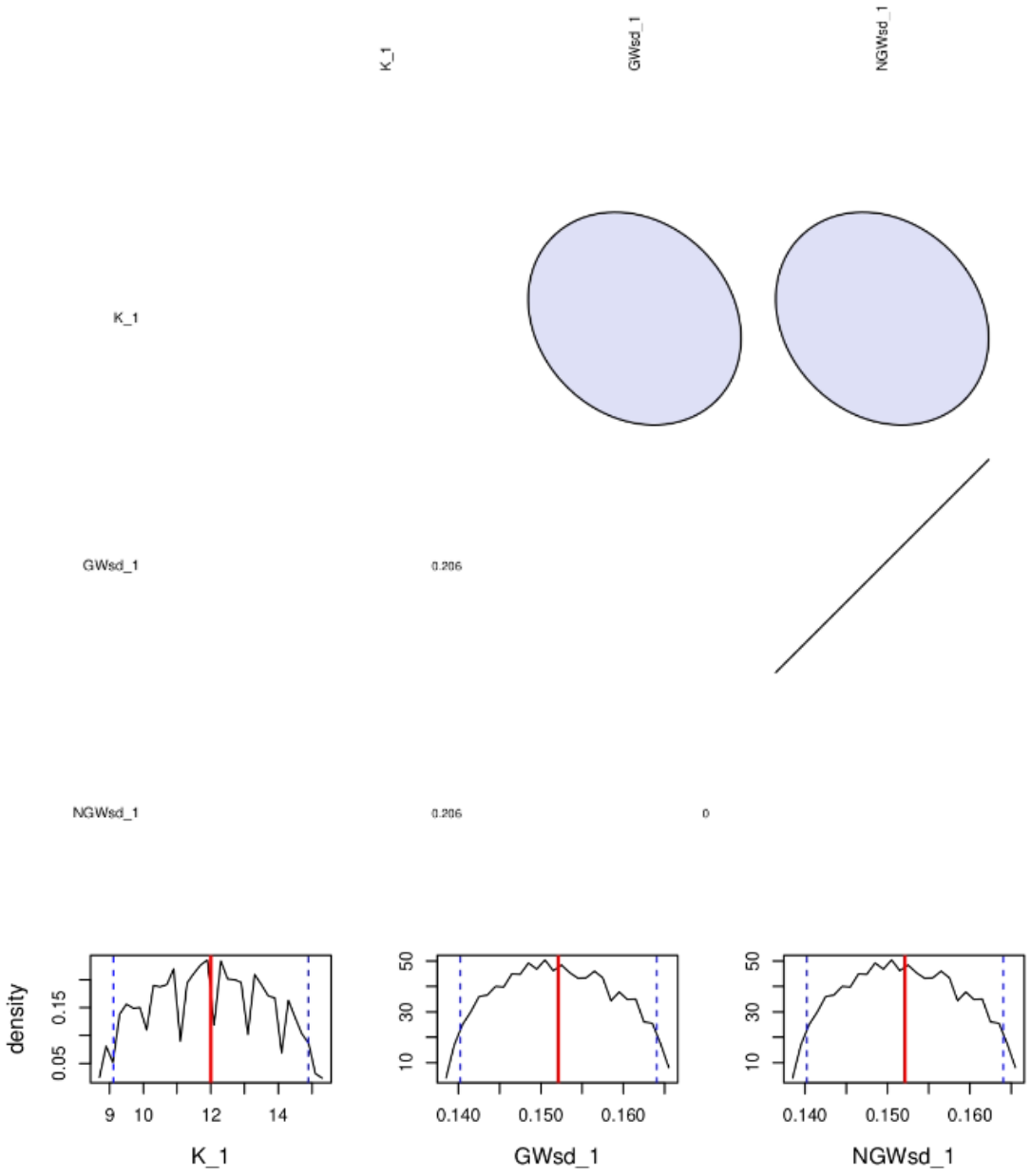


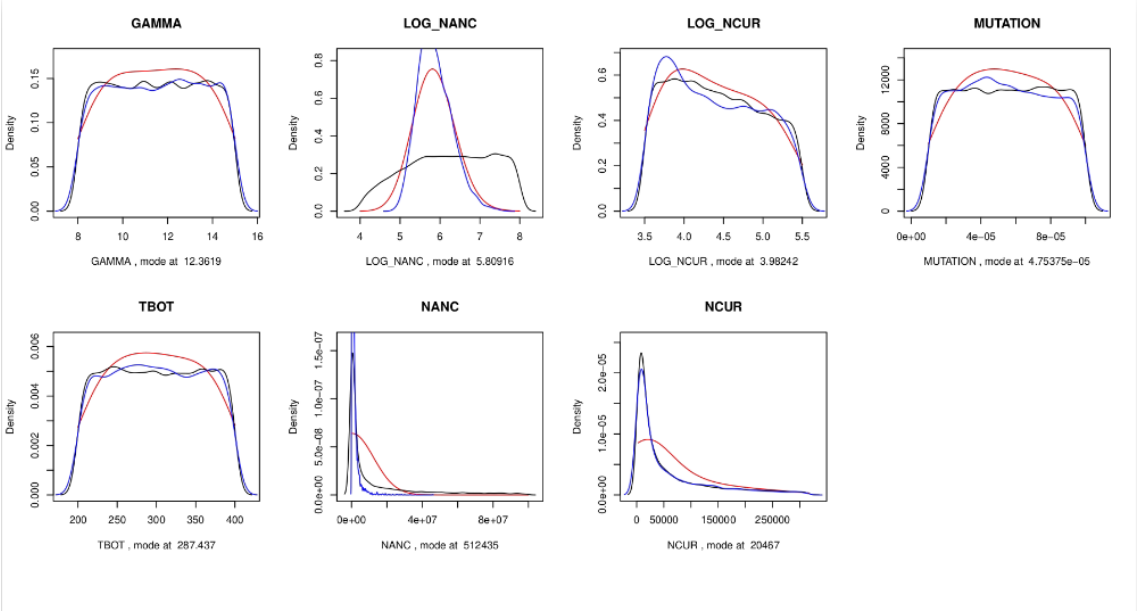


S3: bottleneck

- marginal density: 1.46764e+06 (p-value 0.764)

Summary Stat. North Spearman correlations





The following pages display the ABCtoolbox analyses outputs:

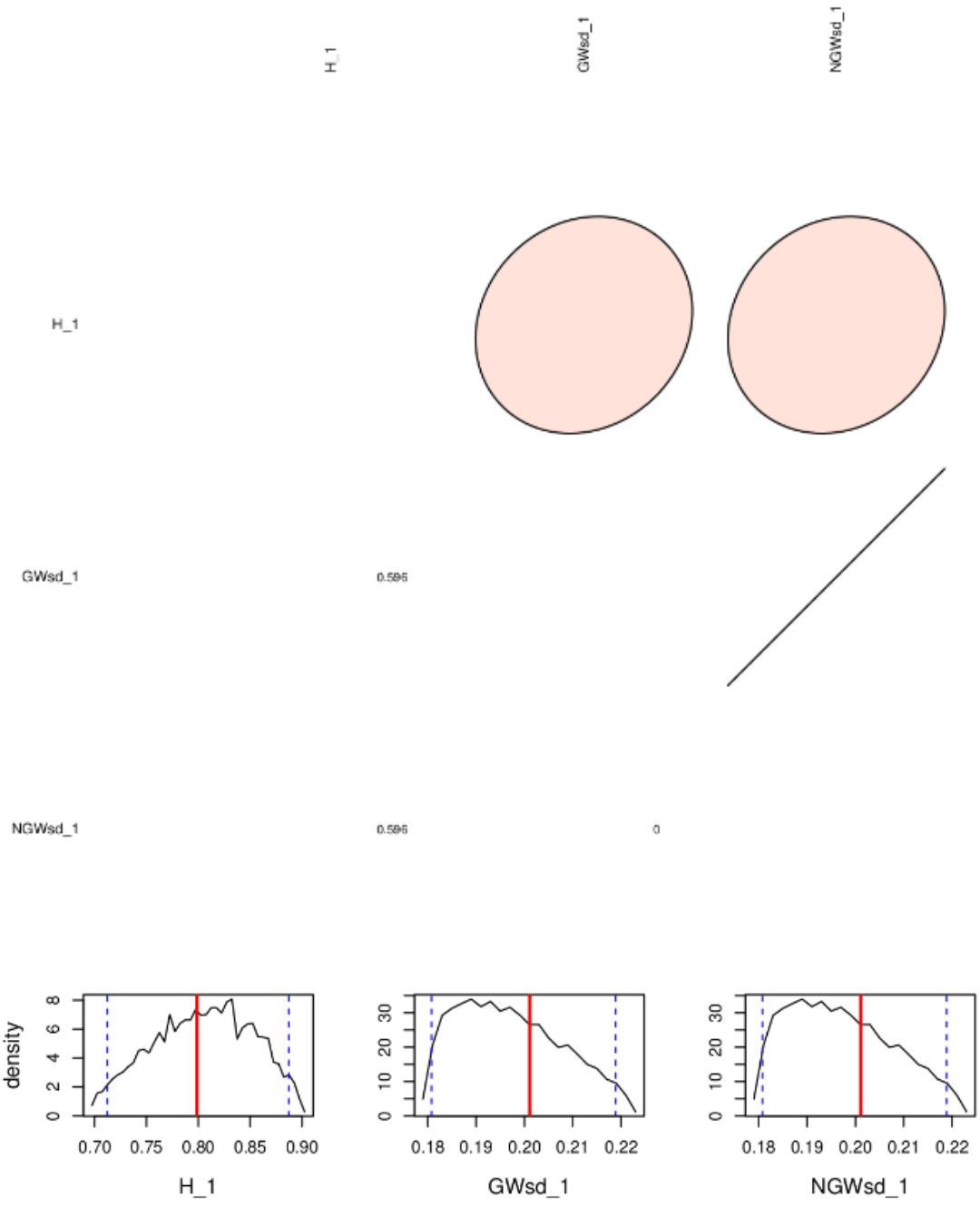
Results For Priors = 100 – 400 generations

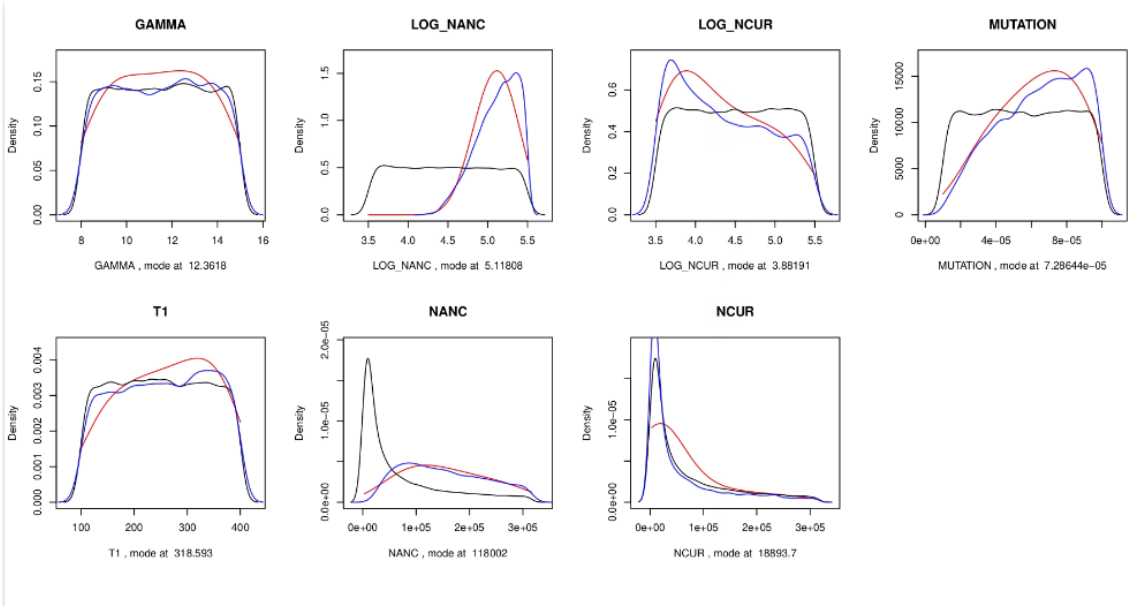
- **Cluster I (Koppies, Soetdoring and Kruger)**

S1: Null

- marginal density: 588635 (p-value 0.644)

Summary Stat. North Spearman correlations

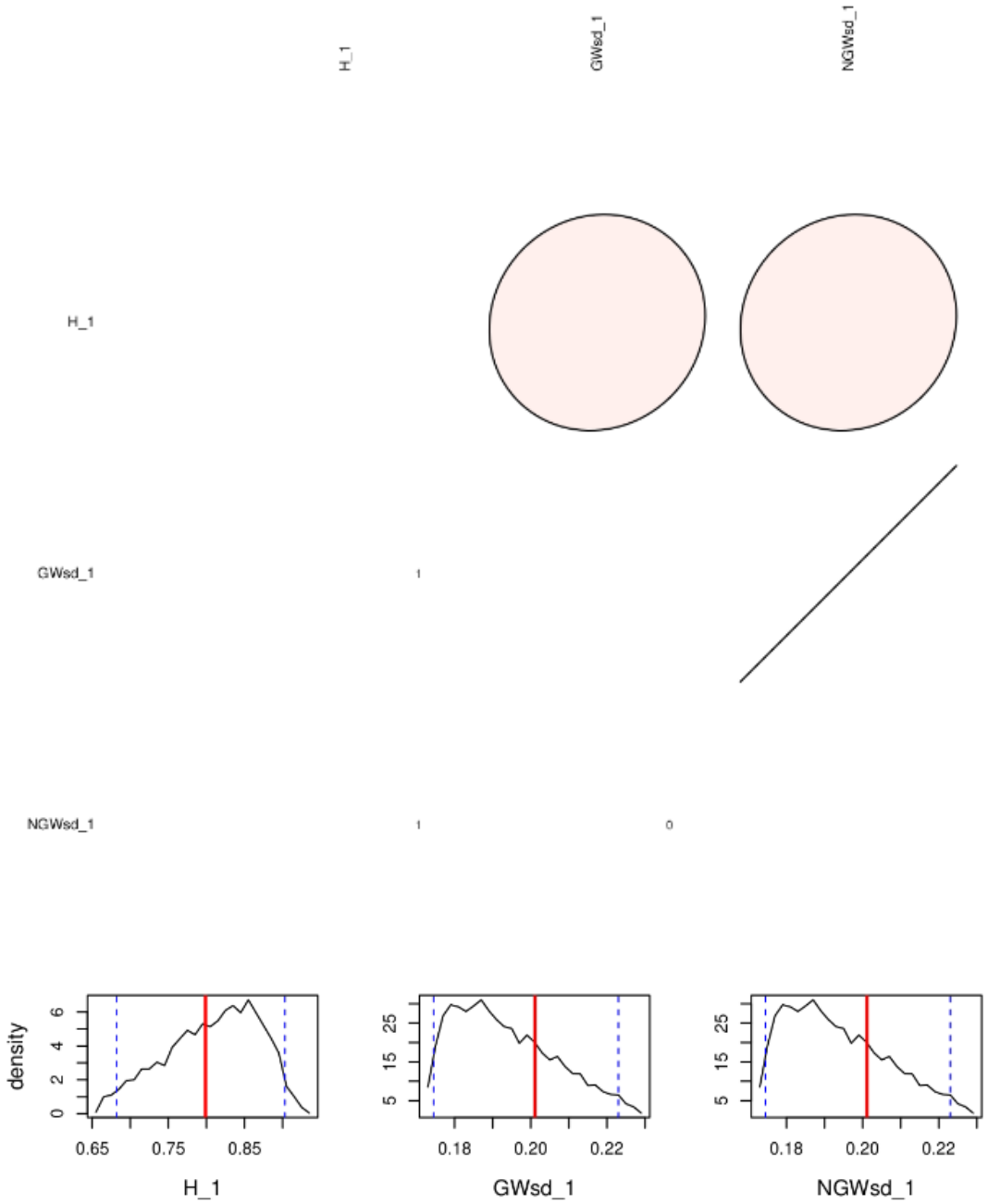


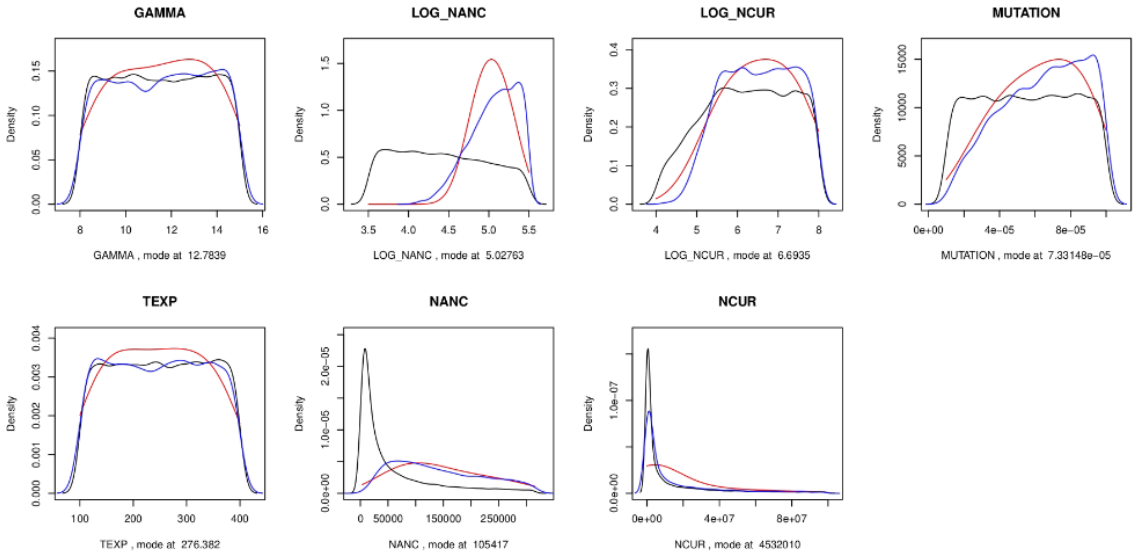


S2: expansion

- marginal density: 275098 (p-value 0.496)

Summary Stat. North Spearman correlations

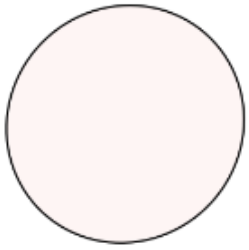
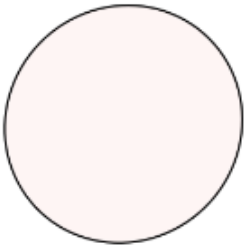
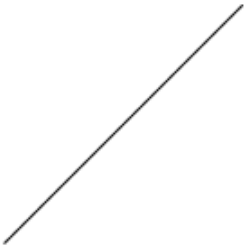


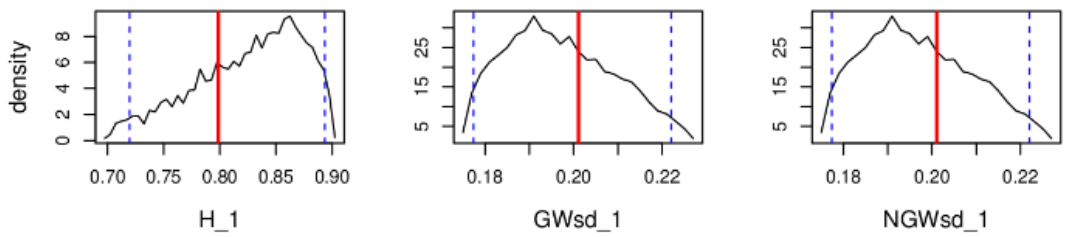


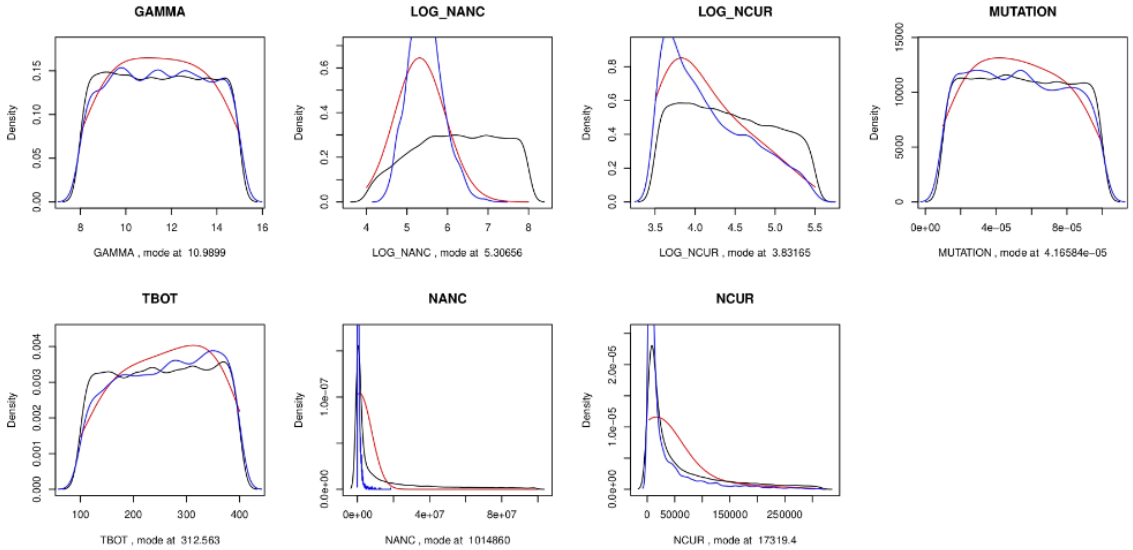
S3: bottleneck

- marginal density: 154205 (p-value 0.912)

Summary Stat. North Spearman correlations

	H_1	GWsd_1	NGWsd_1
H_1			
GWsd_1	1		
NGWsd_1	1	0	



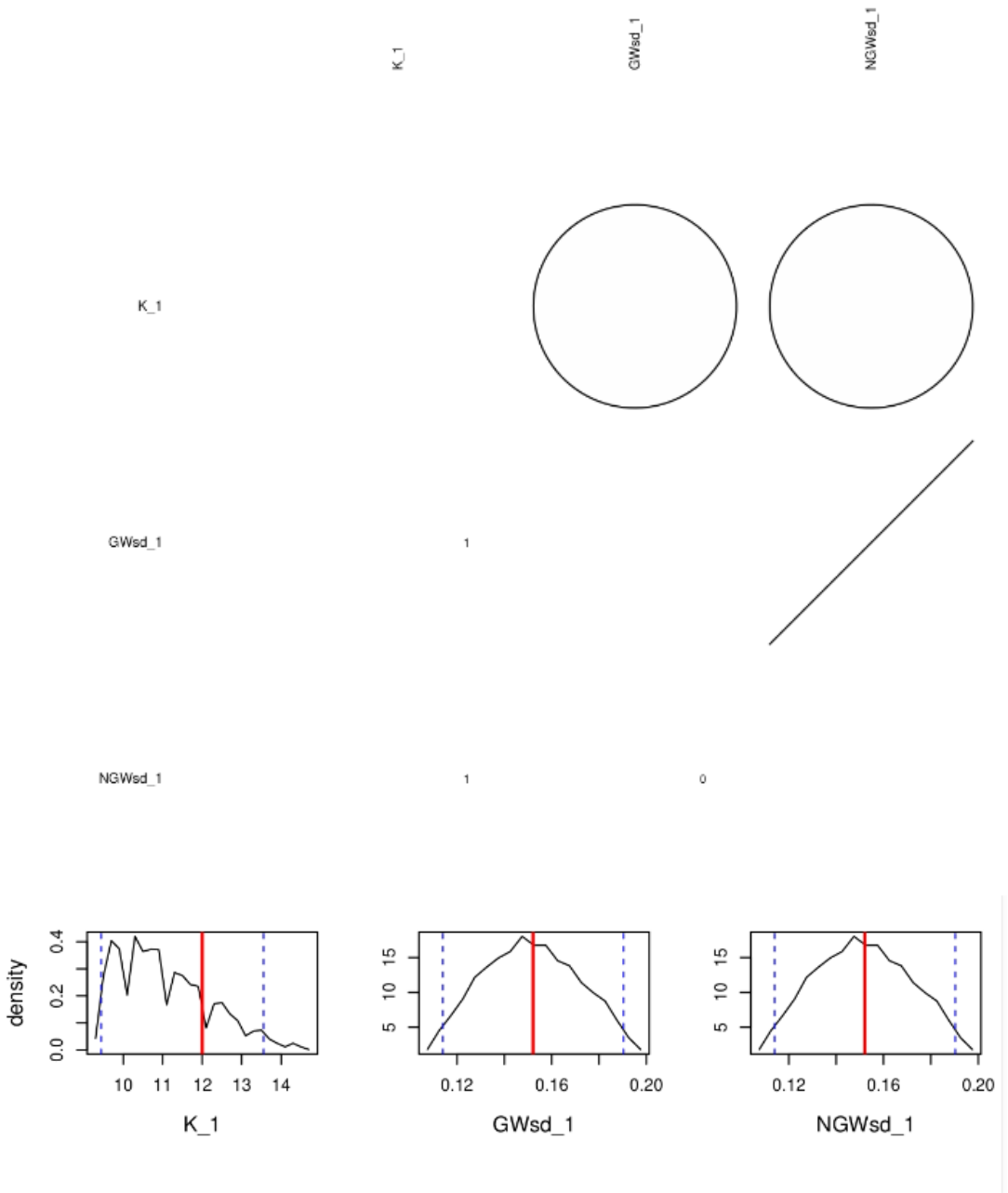


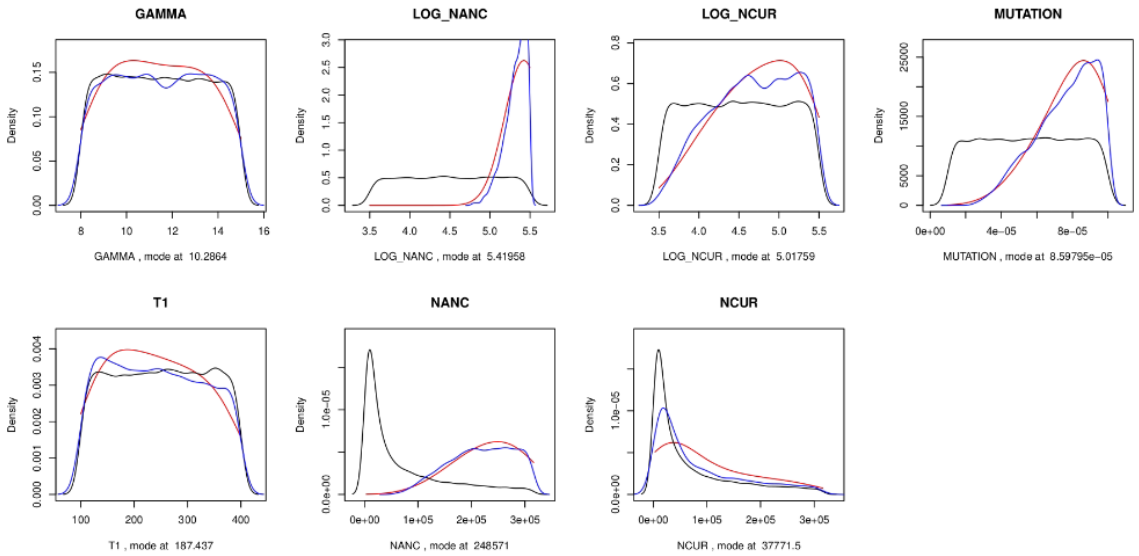
- **Cluster II (Hluhluwe)**

S1: Null

- marginal density: 197267 (p-value 0.456)

Summary Stat. North Spearman correlations

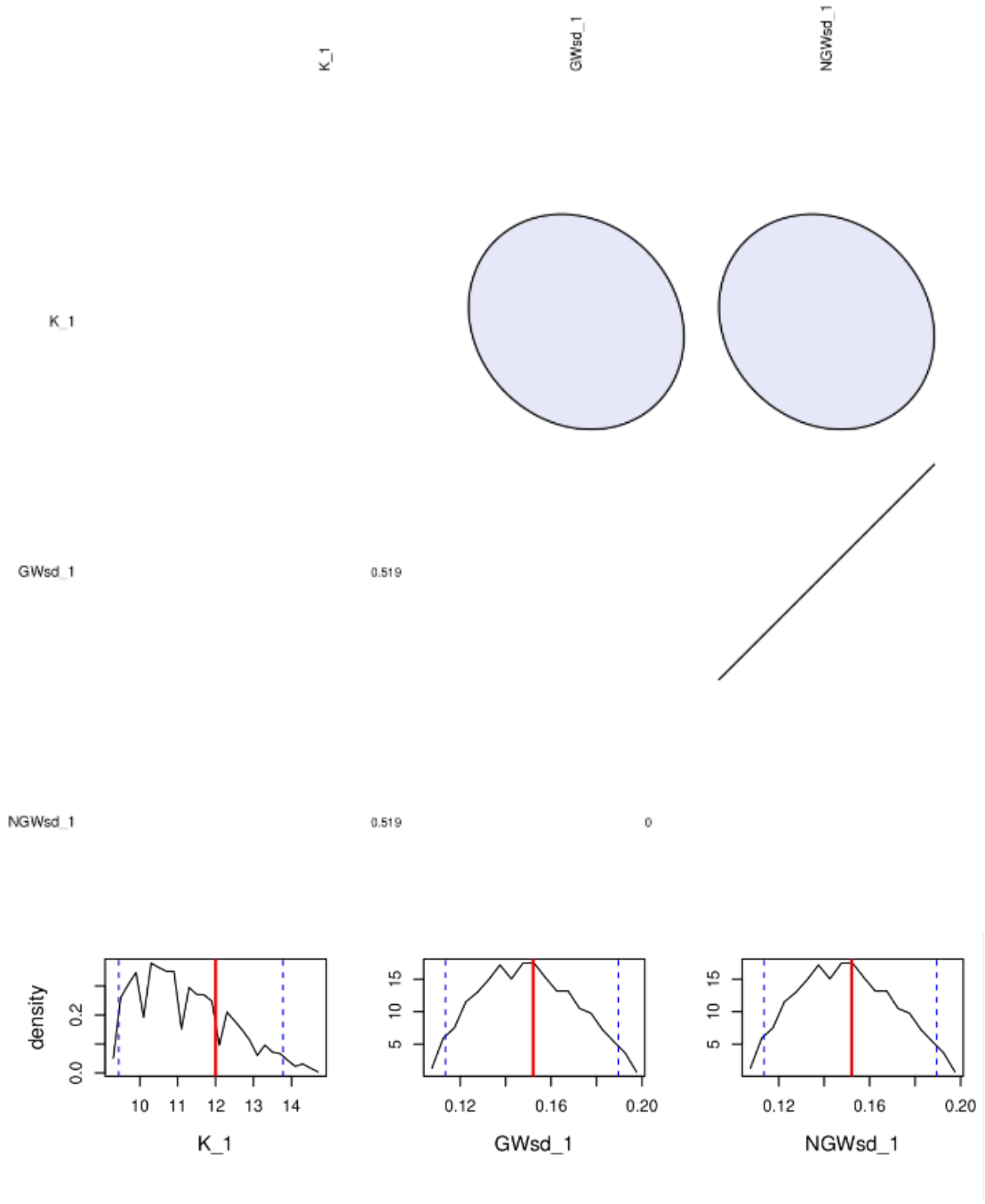


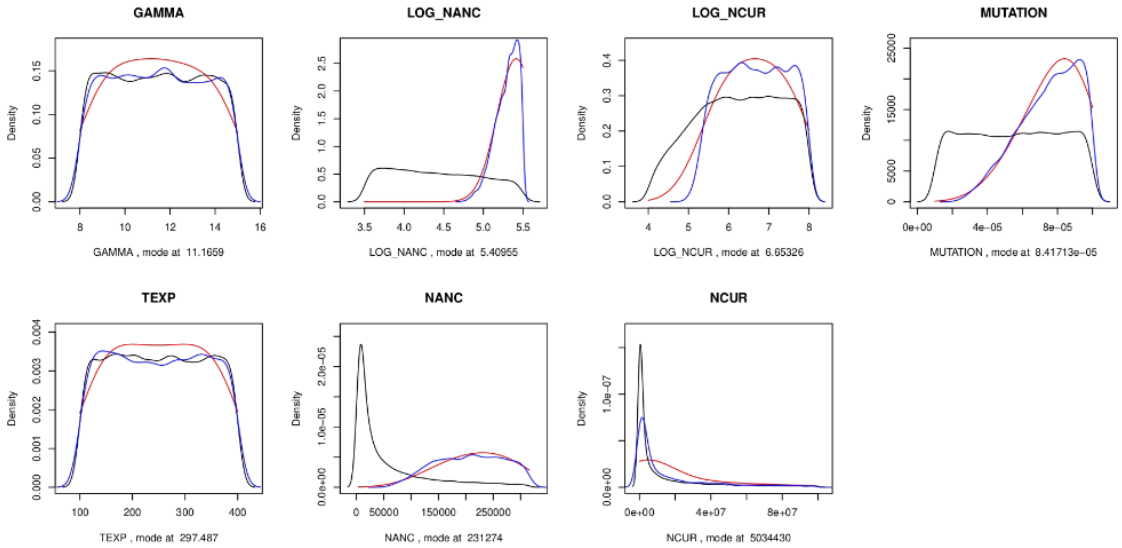


S2: Expansion

- marginal density: 368769 (p-value 0.912)

Summary Stat. North Spearman correlations

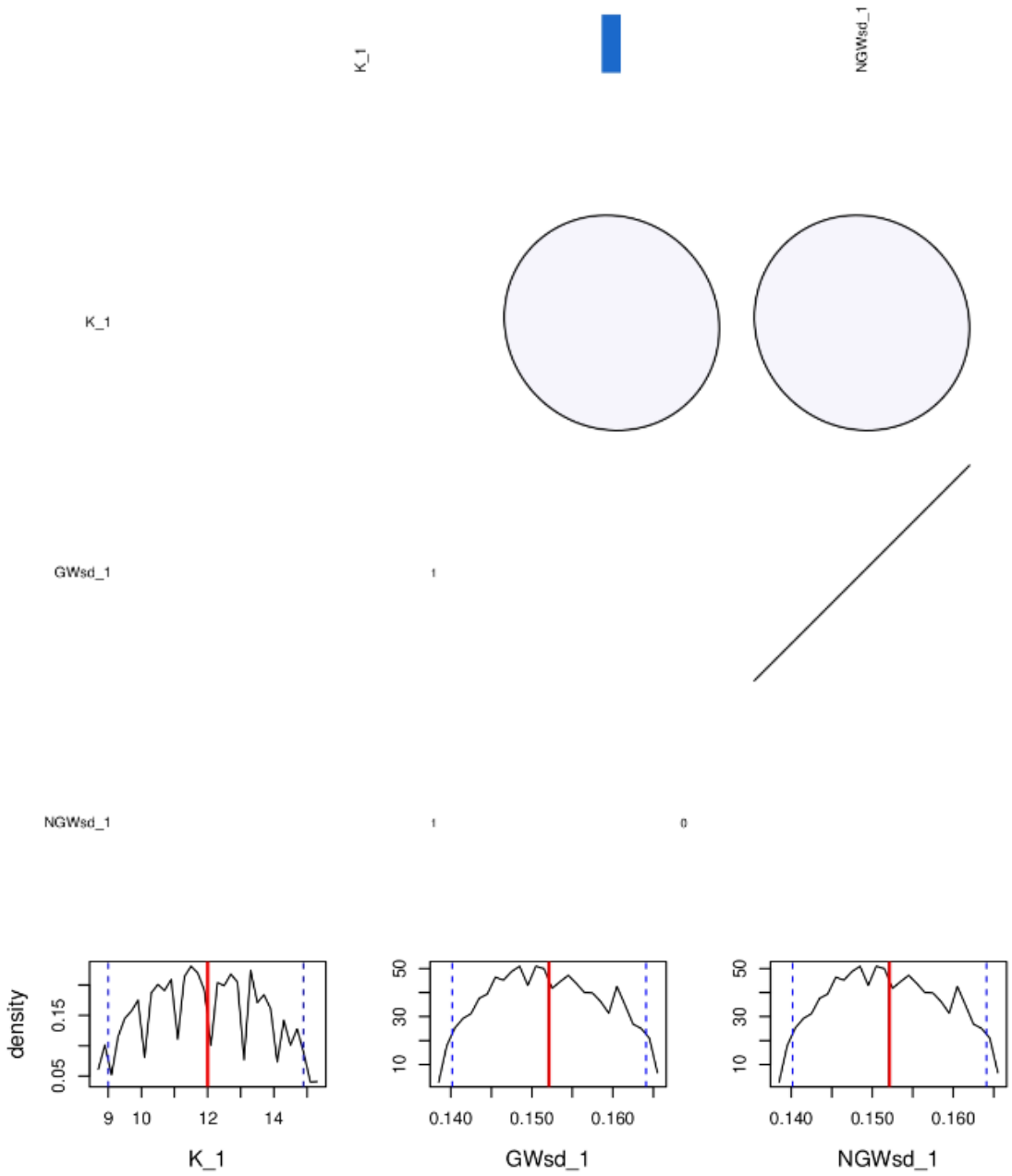


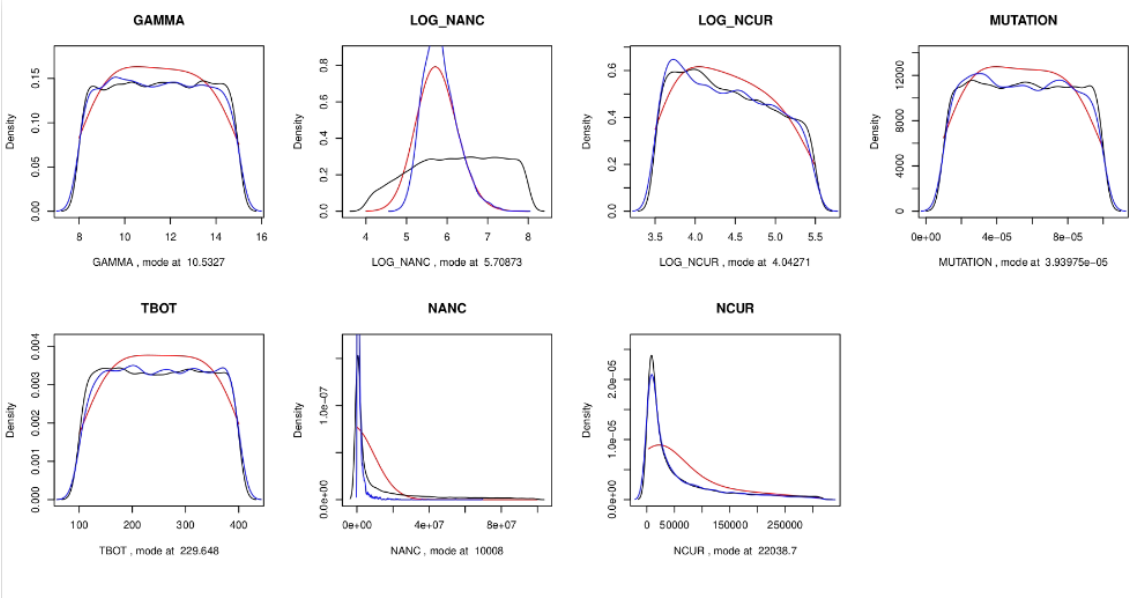


S3: bottleneck

- marginal density: 564630 (p-value 0.252)

Summary Stat. North Spearman correlations




GAMMA
Density
GAMMA , mode at 10.5327
LOG_NANC
Density
LOG_NANC , mode at 5.70873
LOG_NCUR
Density
LOG_NCUR , mode at 4.04271
MUTATION
Density
MUTATION , mode at 3.93975e-05
TBOT
Density
TBOT , mode at 229.648
NANC
Density
NANC , mode at 10008
NCUR
Density
NCUR , mode at 22038.7

References

Abdul-Muneer, P. M. (2014). Application of Microsatellite Markers in Conservation Genetics and Fisheries Management: Recent Advances in Population Structure Analysis and Conservation Strategies. *Genetics Research International*, *2014*, 1–11. https://doi.org/10.1155/2014/691759

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics 2010 11:10*, *11*(10), 697–709. https://doi.org/10.1038/nrg2844

Allendorf, F. William., Luikart, Gordon., & Aitken, S. N. (2013). *Conservation and the genetics of populations*. 602. https://books.google.com/books/about/Conservation_and_the_Genetics_of_Populat.html?id=MkFv5EEpGF4C

Allio, R., Donega, S., Galtier, N., & Nabholz, B. (2017). Large Variation in the Ratio of Mitochondrial to Nuclear Mutation Rate across Animals: Implications for Genetic Diversity and the Use of Mitochondrial DNA as a Molecular Marker. *Molecular Biology and Evolution*, *34*(11), 2762–2772. https://doi.org/10.1093/MOLBEV/MSX197

Alphey, N., & Bonsall, M. B. (2018). Genetics-based methods for agricultural insect pest management. *Agricultural and Forest Entomology*, *20*(2), 131–140. https://doi.org/10.1111/AFE.12241

Arias, O., Cordeiro, E., Corrêa, A. S., Domingues, F. A., Guidolin, A. S., & Omoto, C. (2019). Population genetic structure and demographic history of Spodoptera frugiperda (Lepidoptera: Noctuidae): implications for insect resistance management programs. *Pest Management Science*, *75*(11), 2948–2957. https://doi.org/10.1002/PS.5407

Avenant, N. (2011). The potential utility of rodents and other small mammals as indicators of ecosystem "integrity" of South African grasslands. *Wildlife Research*, *38*(7), 626–639. https://doi.org/10.1071/WR10223

Backus, G. A., & Gross, K. (2016). Genetic engineering to eradicate invasive mice on islands: modeling the efficiency and ecological impacts. *Ecosphere*, *7*(12), e01589. https://doi.org/10.1002/ECS2.1589

Barker, G. C. (2002). Microsatellite DNA: a tool for population genetic analysis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *96*(SUPPL. 1), S21–S24. https://doi.org/10.1016/S0035-9203(02)90047-7

Bažok, R. (2022). Integrated Pest Management of Field Crops. *Agriculture 2022, Vol. 12, Page 425*, *12*(3), 425. https://doi.org/10.3390/AGRICULTURE12030425

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., & Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319. https://doi.org/10.1016/J.YMPEV.2012.08.023

Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. *Science*, *362*(6415). https://doi.org/10.1126/SCIENCE.AAM5979/SUPPL_FILE/AAM5979_BLOUNT_SM.PDF

Boore, J. L. (1999a). Animal mitochondrial genomes. *Nucleic Acids Research*, *27*(8), 1767–1780. https://doi.org/10.1093/NAR/27.8.1767

Boore, J. L. (1999b). Animal mitochondrial genomes. *Nucleic Acids Research*, *27*(8), 1767–1780. https://doi.org/10.1093/NAR/27.8.1767

Bowen, B. W., Bass, A. L., Soares, L., & Toonen, R. J. (2005). Conservation implications of complex population structure: lessons from the loggerhead turtle (Caretta caretta). *Molecular Ecology*, *14*(8), 2389–2402. https://doi.org/10.1111/J.1365-294X.2005.02598.X

Brouat, C., Loiseau, A., Kane, M., Bâ, K., & Duplantier, J. M. (2007). Population genetic structure of two ecologically distinct multimammate rats: the commensal Mastomys natalensis and the wild Mastomys erythroleucus in southeastern Senegal. *Molecular Ecology*, *16*(14), 2985–2997. https://doi.org/10.1111/J.1365-294X.2007.03353.X

Bybee, S. M., Ogden, T. H., Branham, M. A., & Whiting, M. F. (2008). Molecules, morphology and fossils: a comprehensive approach to odonate phylogeny and the evolution of the odonate wing. *Cladistics*, *24*(4), 477–514. https://doi.org/10.1111/J.1096-0031.2007.00191.X

Carapelli, A., Liò, P., Nardi, F., Van Der Wath, E., & Frati, F. (2007). Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evolutionary Biology*, *7*(SUPPL. 2), 1–13. https://doi.org/10.1186/1471-2148-7-S2-S8/FIGURES/3

Casillas, S., & Barbadilla, A. (2017). Molecular Population Genetics. *Genetics*, *205*(3), 1003–1035. https://doi.org/10.1534/GENETICS.116.196493

Castellana, S., Vicario, S., & Saccone, C. (2011). Evolutionary Patterns of the Mitochondrial Genome in Metazoa: Exploring the Role of Mutation and Selection in Mitochondrial Protein–Coding Genes. *Genome Biology and Evolution*, *3*(1), 1067–1079. https://doi.org/10.1093/GBE/EVR040

Castro, J. A., Picornell, A., & Ramon, M. (2014). *Mitochondrial DNA: A tool for populational genetics studies*. https://www.researchgate.net/publication/12378146

Cavalli-Sforza, L. L. (1998). The DNA revolution in population genetics. *Trends in Genetics*, *14*(2), 60–65. https://doi.org/10.1016/S0168-9525(97)01327-9

Coker, O. M., & Coker, O. M. (2019). *Importance of genetics in conservation of biodiversity*. https://www.researchgate.net/publication/336124583

Colangelo, P., Verheyen, E., Leirs, H., Tatard, C., Denys, C., Dobigny, G., Duplantier, J. M., Brouat, C., Granjon, L., & Lecompte, E. (2013). A mitochondrial phylogeographic scenario for the most widespread African rodent, Mastomys natalensis. *Biological Journal of the Linnean Society*, *108*(4), 901–916. https://doi.org/10.1111/BIJ.12013

Combs, M., Byers, K., Himsworth, C., & Munshi-South, J. (2019). Harnessing Population Genetics for Pest Management: Theory and Application for Urban Rats. *Human–Wildlife Interactions*, *13*(2), 11. https://doi.org/https://doi.org/10.26077/0bcx-v339

Conant, G. C., & Wolfe, K. H. (2008). GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics*, *24*(6), 861–862. https://doi.org/10.1093/BIOINFORMATICS/BTM598

Crow, J. F. (1987). *POPULATION GENETICS HISTORY: A PERSONAL VIEW*.

Cummings, M. P. (2004). PAUP* (Phylogenetic Analysis Using Parsimony (and Other Methods)). *Dictionary of Bioinformatics and Computational Biology*. https://doi.org/10.1002/9780471650126.DOB0522.PUB2

Cutter, A. D. (2013). Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution*, *69*(3), 1172–1185. https://doi.org/10.1016/J.YMPEV.2013.06.006

Cuypers, L. N., Baird, S. J. E., Hánová, A., Locus, T., Katakweba, A. S., Gryseels, S., Bryja, J., Leirs, H., & de Bellocq, J. G. (2020). Three arenaviruses in three subspecific natal multimammate mouse taxa in Tanzania: same host specificity, but different spatial genetic structure? *Virus Evolution*, *6*(2). https://doi.org/10.1093/VE/VEAA039

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods 2012 9:8*, *9*(8), 772–772. https://doi.org/10.1038/nmeth.2109

Dávalos, L. M., & Russell, A. L. (2014). Sex-biased dispersal produces high error rates in mitochondrial distance-based and tree-based species delimitation. *Journal of Mammalogy*, *95*(4), 781–791. https://doi.org/10.1644/14-MAMM-A-107/2/JMAMMAL-95-4-781-FIG6.JPEG

Deng, M., Zwick, A., Chen, Q., Liao, C.-Q., Wang, W., Wang, X., & Huang, G.-H. (2023). Phylogeny and classification of Endromidae (Lepidoptera: Bombycoidea) based on mitochondrial genomes. *Arthropod Systematics & Phylogeny*, *81*, 395–408.

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *45*(4), e18–e18. https://doi.org/10.1093/NAR/GKW955

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/S12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/J.1365-294X.2005.02553.X

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/J.1755-0998.2010.02847.X

Fahey, A. L., Ricklefs, R. E., & Dewoody, J. A. (2014). DNA-based approaches for evaluating historical demography in terrestrial vertebrates. *Biological Journal of the Linnean Society*, *112*(3), 367–386. https://doi.org/10.1111/BIJ.12259

Feng, Z., Zhang, Y., & Ling, H. (2002). [Microsatellite markers and applications in the barley genome]. *Yi Chuan = Hereditas*, *24*(6), 727–733. https://europepmc.org/article/med/15979979

Ferris, S. D., Sage, R. D., Prager, E. M., Ritte, U., & Wilson, A. C. (1983). MITOCHONDRIAL DNA EVOLUTION IN MICE. *Genetics*, *105*(3), 681–721. https://doi.org/10.1093/GENETICS/105.3.681

FigTree. (n.d.). Retrieved January 31, 2024, from http://tree.bio.ed.ac.uk/software/figtree/

Flores-Manzanero, A., Luna-Bárcenas, M. A., Rodney, |, Dyer, J., & Vázquez-Domínguez, E. (2019). Functional connectivity and home range inferred at a microgeographic landscape genetics scale in a desert-dwelling rodent. *Ecology and Evolution*, *9*(1), 437–453. https://doi.org/10.1002/ECE3.4762

Foster, P. G., De Oliveira, T. M. P., Bergo, E. S., Conn, J., Sant'Ana, D. C., Nagaki, S. S., Nihei, S., Lamas, C. E., Gonzalez, C., Moreira, C. C., & Sallum, M. A. M. (2017). Phylogeny of Anophelinae using mitochondrial protein coding genes. *Royal Society Open Science*, *4*(11). https://doi.org/10.1098/RSOS.170758

Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics*, *29*, 305–327. https://doi.org/10.1146/ANNUREV.GE.29.120195.001513

Freeland, J. R. (2020). *Molecular ecology*. John Wiley & Sons.

Galan, M., Van Hooft, W. F., Legrand, D., Berthier, K., Loiseau, A., Granjon, L., & Cosson, J. F. (2004). A multiplex panel of microsatellite markers for widespread sub-Saharan rodents of the genus Mastomys. *Molecular Ecology Notes*, *4*(3), 321–323. https://doi.org/10.1111/J.1471-8286.2004.00681.X

Galtier, N., Enard, D., Radondy, Y., Bazin, E., & Belkhir, K. (2006). Mutation hot spots in mammalian mitochondrial DNA. *Genome Research*, *16*(2), 215–222. https://doi.org/10.1101/GR.4305906

Garrido-Garduño, T., Téllez-Valdés, O., Manel, S., & Vázquez-Domínguez, E. (2016). Role of habitat heterogeneity and landscape connectivity in shaping gene flow and spatial population structure of a dominant rodent species in a tropical dry forest. *Journal of Zoology*, *298*(4), 293–302. https://doi.org/10.1111/JZO.12307

González-Ittig, R. E., Rossi-Fraire, H. J., Cantoni, G. E., Herrero, E. R., Benedetti, R., Gallardo, M. H., & Gardenal, C. N. (2009). Population genetic structure of long-tailed pygmy rice rats (Oligoryzomys longicaudatus) from Argentina and Chile based on the mitochondrial control region. *Https://Doi.Org/10.1139/Z09-115*, *88*(1), 23–35. https://doi.org/10.1139/Z09-115

Gryseels, S., Goüy de Bellocq, J., Makundi, R., Vanmechelen, K., Broeckhove, J., Mazoch, V., Šumbera, R., Zima, J., Leirs, H., & Baird, S. J. E. (2016). Genetic distinction between contiguous urban and rural multimammate mice in Tanzania despite gene flow. *Journal of Evolutionary Biology*, *29*(10), 1952–1967. https://doi.org/10.1111/JEB.12919

Haig, S. M. (1998). *MOLECULAR CONTRIBUTIONS TO CONSERVATION*. *79*(2), 413–425. https://doi.org/10.1890/0012-9658

Hamilton, M. B. (2021). *Population genetics*. John Wiley & Sons.

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology & Evolution*, *4*(1), 6–11. https://doi.org/10.1016/0169-5347(89)90006-2

Hasegawa, M., Kishino, H., & Yano, T. aki. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*(2), 160–174. https://doi.org/10.1007/BF02101694

Hedrick, P. W. (2001). Conservation genetics: Where are we now? *Trends in Ecology and Evolution*, *16*(11), 629–636. https://doi.org/10.1016/S0169-5347(01)02282-0

Jenkins, D. J., & Stekel, D. J. (2010). Stochasticity Versus Determinism: Consequences for Realistic Gene Regulatory Network Modelling and Evolution. *J Mol Evol*, *70*, 215–231. https://doi.org/https://doi.org/10.1007/s00239-010-9323-5

Karsten, M., van Vuuren, B. J., Barnaud, A., & Terblanche, J. S. (2013). Population Genetics of Ceratitis capitata in South Africa: Implications for Dispersal and Pest Management. *PLOS ONE*, *8*(1), e54281. https://doi.org/10.1371/JOURNAL.PONE.0054281

Kennis, J., Sluydts, V., Leirs, H., & Van Hooft, W. F. P. (2008). Polyandry and polygyny in an African rodent pest species, Mastomys natalensis. *Mammalia*, *72*(3), 150–160. https://doi.org/10.1515/MAMM.2008.025/MACHINEREADABLECITATION/RIS

Kimura, M., & Ohta, T. (1978). Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences*, *75*(6), 2868–2872. https://doi.org/10.1073/PNAS.75.6.2868

Kneidinger, C. M. (2010). *Mastomys Natalensis and Mastomys Coucha: Identification, Habitat Preferences and Population Genetics*. University of Johannesburg (South Africa).

Kneidinger, C. M., van Heerden, H., MacFadyen, D., van der Merwe, M., Avenant, N. L., & van der Bank, H. (2014). Species identification, habitat preferences and population genetics of Mastomys natalensis (A. Smith, 1834) and M. coucha (A. Smith, 1836) in an enclosed area, Kruger National Park, South Africa. *Navors. Nas. Mus., Bloemfontein*, *30*(Part 3).

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179. https://doi.org/10.1111/1755-0998.12387

Kumar, S. (1996). Patterns of Nucleotide Substitution in Mitochondrial Protein Coding Genes of Vertebrates. *Genetics*, *143*(1), 537–548. https://doi.org/10.1093/GENETICS/143.1.537

Lacher, T., Rogan, J. E., Smith, A., & Upham, N. S. (2016). Evolution, phylogeny, ecology and conservation of the Clade Glires: Lagomorpha and Rodentia. *Handbook of the Mammals of the World* , *6*, 15–26. https://paleobiodb.org/navigator/

Le Roux, J., & Wieczorek, A. M. (2009a). Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology*, *154*(1), 1–17. https://doi.org/10.1111/J.1744-7348.2008.00280.X

Le Roux, J., & Wieczorek, A. M. (2009b). Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology*, *154*(1), 1–17. https://doi.org/10.1111/J.1744-7348.2008.00280.X

Lecompte, E., Brouat, C., Duplantier, J. M., Galan, M., Granjon, L., Loiseau, A., Mouline, K., & Cosson, J. F. (2005). Molecular identification of four cryptic species of Mastomys (Rodentia, Murinae). *Biochemical Systematics and Ecology*, *33*(7), 681–689. https://doi.org/10.1016/J.BSE.2004.12.015

Lecompte, E., Granjon, L., & Denys, C. (2002a). The phylogeny of the Praomys complex (Rodentia: Muridae) and its phylogeographic implications. *Journal of Zoological Systematics and Evolutionary Research*, *40*(1), 8–25. https://doi.org/10.1046/J.1439-0469.2002.00172.X

Lecompte, E., Granjon, L., & Denys, C. (2002b). The phylogeny of the Praomys complex (Rodentia: Muridae) and its phylogeographic implications. *Journal of Zoological Systematics and Evolutionary Research*, *40*(1), 8–25. https://doi.org/10.1046/J.1439-0469.2002.00172.X

Leirs, H., Verheyen, W., & Verhagen, R. (1996a). Spatial patterns in Mastomys natalensis in Tanzania (Rodentia, Muridae). *Mammalia*, *60*(4), 545–555. https://doi.org/10.1515/MAMM.1996.60.4.545/MACHINEREADABLECITATION/RIS

Leirs, H., Verheyen, W., & Verhagen, R. (1996b). Spatial patterns in Mastomys natalensis in Tanzania (Rodentia, Muridae). *Mammalia*, *60*(4), 545–555. https://doi.org/10.1515/MAMM.1996.60.4.545/MACHINEREADABLECITATION/RIS

Leroy, G., Baumung, R., Boettcher, P., Besbes, B., From, T., & Hoffmann, I. (2018). Animal genetic resources diversity and ecosystem services. *Global Food Security*, *17*, 84–91. https://doi.org/10.1016/J.GFS.2018.04.003

Linzey, A. V., & Kesner, M. H. (1997). Small mammals of a woodland-savannah ecosystem in Zimbabwe. I. Density and habitat occupancy patterns. *Journal of Zoology*, *243*(1), 137–152. https://doi.org/10.1111/J.1469-7998.1997.TB05760.X

Loewe, L., & Hill, W. G. (2010). The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1544), 1153–1167. https://doi.org/10.1098/RSTB.2009.0317

Loiseau, A., Konečný, A., Galan, M., Bryja, J., Cosson, J. F., & Brouat, C. (2007). New polymorphic microsatellite loci for rodents of the genus Mastomys using PCR multiplexing, and cross-species amplification in Myomys and Praomys. *Molecular Ecology Notes*, *7*(4), 684–687. https://doi.org/10.1111/J.1471-8286.2006.01677.X

Lowe, W. H., Kovach, R. P., & Allendorf, F. W. (2017). Population Genetics and Demography Unite Ecology and Evolution. *Trends in Ecology & Evolution*, *32*(2), 141–152. https://doi.org/10.1016/J.TREE.2016.12.002

Luikart, G., & England, P. R. (1999). Statistical analysis of microsatellite DNA data. *Trends in Ecology and Evolution*, *14*(7), 253–256. https://doi.org/10.1016/S0169-5347(99)01632-8

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, *18*(4), 189–197. https://doi.org/10.1016/S0169-5347(03)00008-9

Mapelli, F. J., Boston, E. S. M., Fameli, A., Gómez Fernández, M. J., Kittlein, M. J., & Mirol, P. M. (2020). Fragmenting fragments: landscape genetics of a subterranean rodent (Mammalia, Ctenomyidae) living in a human-impacted wetland. *Landscape Ecology 2020 35:5*, *35*(5), 1089–1106. https://doi.org/10.1007/S10980-020-01001-Z

Martin, G. H. G., & Dickinson, N. M. (1985). Small mammal abundance in relation to microhabitat in a dry sub-humid grassland in Kenya. *African Journal of Ecology*, *23*(4), 223–234. https://doi.org/10.1111/J.1365-2028.1985.TB00953.X

Mdangi, M., Mulungu, L. S., Massawe, A. W., Eiseb, S. J., Tutjavi, V., Kirsten, F., Mahlaba, T., Malebane, P., von Maltitz, E., Monadjem, A., Dlamini, N., Makundi, R. H., & Belmain, S. R. (2013). Assessment of rodent damage to stored maize (Zea mays L.) on smallholder farms in Tanzania. *Http://Dx.Doi.Org/10.1080/09670874.2012.744495*, *59*(1), 55–62. https://doi.org/10.1080/09670874.2012.744495

Meester, J. A. J., Lloyd, C. N. V, & ROWEROWE, D. T. (1979). NOTE ON THE ECOLOGICAL ROLE OF PRAOMYS-NATALENSIS. *South African Journal of Science*, *75*(4), 183–184.

Monadjem, A., & Perrin, M. (2015). Population fluctuations and community structure of small mammals in a Swaziland grassland over a three-year period. *Http://Dx.Doi.Org/10.1080/15627020.2003.11657200*, *38*(1), 127–137. https://doi.org/10.1080/15627020.2003.11657200

Monadjem, A., & Perrin, M. R. (1998). Effects of supplemental food on the habitat selected by Mastomys natalensis (Muridae: Rodentia) in a subtropical grassland in Swaziland. *African Zoology*, *33*(3), 172–177. https://www.ajol.info/index.php/az/article/view/154743

Monadjem, A., & Perrin, M. R. (2017). Effects of supplemental food on the habitat selected by Mastomys natalensis (Muridae: Rodentia) in a subtropical grassland in Swaziland. *African Zoology*, *33*(3), 172–177. https://www.ajol.info/index.php/az/article/view/154743

Moore, W. S. (1995). INFERRING PHYLOGENIES FROM mtDNA VARIATION: MITOCHONDRIAL-GENE TREES VERSUS NUCLEAR-GENE TREES. *Evolution*, *49*(4), 718–726. https://doi.org/10.1111/J.1558-5646.1995.TB02308.X

Mulungu, L. S., Borremans, B., Ngowo, V., Mdangi, M. E., Katakweba, A. S., Tesha, P., Mrosso, F. P., Mchomvu, M., & Kilonzo, B. S. (2015). Comparative study of movement patterns of Mastomys natalensis in irrigated rice and fallow fields in eastern Tanzania. *African Journal of Ecology*, *53*(4), 473–479. https://doi.org/10.1111/AJE.12233

Mulungu, L. S., Mahlaba, T. A., Massawe, A. W., Kennis, J., Crauwels, D., Eiseb, S., Monadjem, A., Makundi, R. H., Katakweba, A. A. S., Leirs, H., & Belmain, S. R. (2011). Dietary differences of the multimammate mouse, Mastomys natalensis (Smith, 1834), across different habitats and seasons in Tanzania and Swaziland. *Wildlife Research*, *38*(7), 640–646. https://doi.org/10.1071/WR11028

Mwanjabe, P. S., & Leirs, H. (1997). AN EARLY WARNING SYSTEM FOR IPM-BASED RODENT CONTROL IN SMALLHOLDER FARMING SYSTEMS IN TANZANIA. *Belg. J. Zoo*, *127*(1), 49–58.

Mwanjabe, P. S., Sirima, F. B., & Lusingu, J. (2002). Crop losses due to outbreaks of Mastomys natalensis (Smith, 1834) Muridae, Rodentia, in the Lindi Region of Tanzania. *International Biodeterioration & Biodegradation*, *49*(2–3), 133–137. https://doi.org/10.1016/S0964-8305(01)00113-5

Nabholz, B., Ellegren, H., & Wolf, J. B. W. (2013). High Levels of Gene Expression Explain the Strong Evolutionary Constraint of Mitochondrial Protein-Coding Genes. *Molecular Biology and Evolution*, *30*(2), 272–284. https://doi.org/10.1093/MOLBEV/MSS238

Neiman, M., & Taylor, D. R. (2009). The causes of mutation accumulation in mitochondrial genomes. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1660), 1201–1209. https://doi.org/10.1098/RSPB.2008.1758

Odhiambo, R. O., Makundi, R. H., Leirs, H., & Verhagen, R. (2008). Dietary selection in Mastomys natalensis (Rodentia: Muridae) in the maize agro-ecosystems of central and southwestern Tanzania. *Mammalia*, *72*(3), 169–177. https://doi.org/10.1515/MAMM.2008.007/MACHINEREADABLECITATION/RIS

O'Neill, R., Snowdon, R., & Köhler, W. (2003). *Population Genetics: Aspects of Biodiversity*. 115–137. https://doi.org/10.1007/978-3-642-55819-1_7

Otto, S. P., & Whitlock, M. C. (1997). *The Probability of Fixation in Populations of Changing Size*. https://academic.oup.com/genetics/article/146/2/723/6054000

Peakall, R., & Smouse, P. E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288–295. https://doi.org/10.1111/J.1471-8286.2005.01155.X

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, *28*(19), 2537. https://doi.org/10.1093/BIOINFORMATICS/BTS460

Pearce, J., & Venier, L. (2005). Small mammals as bioindicators of sustainable boreal forest management. *Forest Ecology and Management*, *208*(1–3), 153–175. https://doi.org/10.1016/J.FORECO.2004.11.024

Popadin, K. Y., Nikolaev, S. I., Junier, T., Baranova, M., & Antonarakis, S. E. (2013). Purifying Selection in Mammalian Mitochondrial Protein-Coding Genes Is Highly Effective and Congruent with Evolution of Nuclear Genes. *Molecular Biology and Evolution*, *30*(2), 347–355. https://doi.org/10.1093/MOLBEV/MSS219

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/GENETICS/155.2.945

Richard, G. F., & Pâques, F. (2000). Mini- and microsatellite expansions: the recombination connection. *EMBO Reports*, *1*(2), 122–126. https://doi.org/10.1093/EMBO-REPORTS/KVD031

Rokas, A., Ladoukakis, E., & Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution*, *18*(8), 411–417. https://doi.org/10.1016/S0169-5347(03)00125-3

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, *61*(3), 539–542. https://doi.org/10.1093/SYSBIO/SYS029

Rowe-Rowe, D. T. (1995). Small-mammal recolonization of a fire-exclusion catchment after unscheduled burning. *South African Journal of Wildlife Research-24-Month Delayed Open Access*, *25*(4), 133–137.

Rubinoff, D., & Holland, B. S. (2005). Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. *Systematic Biology*, *54*(6), 952–961. https://doi.org/10.1080/10635150500234674

Russo, I. R. M., Sole, C. L., Barbato, M., Von Bramann, U., & Bruford, M. W. (2016). Landscape determinants of fine-scale genetic structure of a small rodent in a heterogeneous landscape (Hluhluwe-iMfolozi Park, South Africa). *Scientific Reports 2016 6:1*, *6*(1), 1–14. https://doi.org/10.1038/srep29168

Schlotterer, C. (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma*, *109*(6), 365–371. https://doi.org/10.1007/S004120000089/METRICS

Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, *22*(1), 25–33. https://doi.org/10.1016/j.tree.2006.08.009

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, *9*(5), 615–629. https://doi.org/10.1111/J.1461-0248.2006.00889.X

Shoemaker, J. S., & Fitch, W. M. (1989). Evidence from nuclear sequences that invariable sites should be considered when sequence divergence is calculated. *Molecular Biology and Evolution*, *6*(3), 270–289. https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A040550

Singleton, G. R., Belmain, S. R., Brown, P. R., & Hardy, B. (2010). *Rodent Outbreaks: Ecology and Impacts*. 289. https://www.researchgate.net/publication/233992246_Rodent_Outbreaks_Ecology_and_Impacts

Stockwell, C. A., Hendry, A. P., & Kinnison, M. T. (2003). Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, *18*(2), 94–101. https://doi.org/10.1016/S0169-5347(02)00044-7

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/MOLBEV/MSAB120

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Some Mathematical Question in Biology-DNA Sequence Analysis*, 57–86. https://doi.org/10.3106/1348-6160(2005)30

Taylor, K. D., & Green, M. G. (1976). The influence of rainfall on diet and reproduction in four African rodent species. *Journal of Zoology*, *180*(3), 367–389. https://doi.org/10.1111/J.1469-7998.1976.TB04683.X

Uecker, H., & Hermisson, J. (2011). On the Fixation Process of a Beneficial Mutation in a Variable Environment. *Genetics*, *188*(4), 915–930. https://doi.org/10.1534/GENETICS.110.124297

Van Hooft, P., Cosson, J. F., Vibe-Petersen, S., & Leirs, H. (2008). Dispersal in Mastomys natalensis mice: use of fine-scale genetic analyses for pest management. *Hereditas*, *145*(6), 262–273. https://doi.org/10.1111/J.1601-5223.2008.02089.X

Venturi, F. P., Chimimba, C. T., van Aarde, R. J., & Fairall, N. (2004). The distribution of two medically and agriculturally important cryptic rodent species, Mastomys natalensis and M. coucha (Rodentia: Muridae) in South Africa. *Http://Dx.Doi.Org/10.1080/15627020.2004.11657220*, *39*(2), 235–245. https://doi.org/10.1080/15627020.2004.11657220

Wan, Q. H., Wu, H., Fujihara, T., & Fang, S. G. (2004). Which genetic marker for which conservation genetics issue? *ELECTROPHORESIS*, *25*(14), 2165–2176. https://doi.org/10.1002/ELPS.200305922

Wegmann, D., Leuenberger, C., Neuenschwander, S., & Excoffier, L. (2010). ABCtoolbox: A versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, *11*(1), 1–7. https://doi.org/10.1186/1471-2105-11-116/FIGURES/4

Whitlock, M. C. (2003). Fixation Probability and Time in Subdivided Populations. *Genetics*, *164*(2), 767–779. https://doi.org/10.1093/GENETICS/164.2.767

Wright, J. M., & Bentzen, P. (1995). Microsatellites: genetic markers for the future. *Molecular Genetics in Fisheries*, 117–121. https://doi.org/10.1007/978-94-011-1218-5_7

Xia, X. (2018). DAMBE7: New and Improved Tools for Data Analysis in Molecular Biology and Evolution. *Molecular Biology and Evolution*, *35*(6), 1550–1552. https://doi.org/10.1093/MOLBEV/MSY073

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *Journal of Molecular Evolution*, *39*(3), 306–314. https://doi.org/10.1007/BF00160154

Zrzavý, J., Duda, P., Robovský, J., Okřinová, I., & Pavelková Řičánková, V. (2018). Phylogeny of the Caninae (Carnivora): Combining morphology, behaviour, genes and fossils. *Zoologica Scripta*, *47*(4), 373–389. https://doi.org/10.1111/ZSC.12293

Anexos