The development of biological tools to aid in the genetic investigation of the black (*Diceros bicornis*) and white (*Ceratotherium simum*) rhinoceros mitochondrial genomes

A thesis submitted in fulfilment of the degree of Master of Science in Biochemistry of Rhodes University

July 2015

Michelle Parsons

Supervisor: Dr B. Wilhelmi

Co-Supervisors: Dr A. Edkins and Dr D. Parker

Department of Biochemistry and Microbiology

RHODES UNIVERSITY
Where leaders learn

ABSTRACT

The black (*Diceros bicornis)* and white (*Ceratotherium simum*) rhinoceros are found in South Africa. A decline in the populations of these species has resulted due to human activities such as habitat fragmentation and poaching. This has contributed to the loss of genetic diversity amongst the black and white rhinoceros. Conservation and anti-poaching efforts are needed to help maintain genetic diversity. These efforts could be improved through the development of non-invasive techniques to examine DNA from threatened animals. The aim of this research was to develop a molecular technique which would allow for the identification of the black and white rhinoceros and to develop a molecular technique which would allow for intraspecies genetic variation to be examined.

DNA extractions were performed on matched faecal and tissue samples that were collected from two regions in South Africa. Polymerase chain reaction (PCR) primer sets were designed to investigate several regions of the rhinoceros mitochondrial genome. PCR optimisation was completed for the target regions. Sequencing was conducted on all final PCR products. The *cytochrome c oxidase subunit 1* (*COIi*) gene allowed for the rhinoceros family to be identified. This region was digested with the *Hin*dIII restriction enzyme, which allowed for the specific identification of either the black or white rhinoceros. A subsequent region of the *cytochrome c oxidase subunit 1* (*COIii*) as well as the D-loop, hypervariable regions (HV1 and HV2), *cytochrome b* (*cytb*) and *16s rRNA* regions were investigated. These regions displayed potential for establishing geographic origin for black rhinoceros samples, whereas the D-loop and HV2 show potential for the white rhinoceros. The white rhinoceros displayed sequence variation in the HV2 and *cytb* region, while variation was observed in the *COIi* and HV1 for the black rhinoceros.

All investigated target regions allowed for the rhinoceros family to be identified. The *COI* (*COIi* and *COIii*), HV2 and *cytb* regions allowed for the subspecies of rhinoceros to be identified, however the D-loop was not able to identify the white rhinoceros species. The *16s rRNA* and HV1 regions allowed for the correct subspecies of rhinoceros to be identified, however as the primers were only compatible for the black rhinoceros therefore a subsequent investigation is required for the white rhinoceros.

The establishment of this novel PCR based technique to identify white and black rhinoceros will allow for efficient species identification in wildlife forensic cases. A biological method

was established to study intraspecies variation for the white and black rhinoceros; however the investigated target regions did not yield sufficient genetic variation. The core techniques developed in this study will be valuable for future studies that wish to investigate genetic variation in mammal species.

DECLARATION

I acknowledge that this is original work completed by myself, the undersigned, and is submitted for the degree of Master of Science of Rhodes University.

_____

Michelle Parsons

RESEARCH OUTPUTS

Parsons, M., Edkins, A.L., Parker, D. & Wilhelmi, B. Development of genetic techniques for the identification of individual black rhinoceros (*Diceros bicornis*): towards the development of rhinoceros DNA barcode for the use in wildlife forensics and conservation. Oral presentation presented at: the Eastern Cape symposium for the South African Society of Biochemistry and Molecular Biology, 2013 Oct 24; Fort Hare, Alice, South Africa.

Parsons, M., Edkins, A.L., Parker, D. & Wilhelmi, B. A Rhinoceros DNA Barcode for the use in Wildlife Forensics. Poster session presented at: the 24th biennial congress of the South African Society of Biochemistry and Molecular Biology; 2014 Jun 6-9; Rawsonville, South Africa.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Chapter Five

Chapter Six

LIST OF TABLES

LIST OF FIGURES

x

LIST OF ABBREVIATIONS

B:	Blood sample
Bp:	Base pairs
BLAST:	Basic Local Alignment Search Tool
CITES:	United nations convention on international trade in endangered species of wild fauna and flora
COI:	*Cytochrome c oxidase subunit 1*gene
COIi:	First region of the *cytochrome c oxidase subunit 1* gene examined
COIii:	Second region of the *cytochrome c oxidase subunit 1* gene examined
Cytb:	*Cytochrome b* gene
D-loop:	Control region found in a non-coding area of the mitochondrial genome
DNA:	Deoxyribonucleic acid
F:	Faecal sample
*Hin*dIII:	Restriction enzyme (5' AAGCTT 3')
HV1:	Hypervariable region 1
HV2:	Hypervariable region 2
IUNC:	World conservation union
KG:	Kilogram
M99:	Etorphine
MtDNA:	Mitochondrial DNA
NCBI:	National Centre for Biotechnology Information
ND:	NADH dehydrogenase
PCR:	Polymerase chain reaction

PCR-RFLP: Restriction fragment length polymorphism

RAPD: Random amplification of polymorphic DNA

RNA: Ribonucleic acid

rRNA: Ribosomal ribonucleic acid

SNP: Single nucleotide polymorphism

STR: Short tandem repeat

T: Tissue sample

Tm: Melting temperature

tRNA: Transfer RNA

UV: Ultra violet

Explanation of Sample Naming:

black (I, II):   black = black rhinoceros

                 I = Individual's unique number

                 II = Sample type (Blood [B], Tissue [T] or Faecal [F])

                 E.g. black (4, F) = black rhinoceros, termed 4 for the investigation, faecal
                 sample

BINOMIAL NOMENCLATURE

Genus and Subspecies for the Rhinocerotidae family

1. *Diceros bicornis*   Black
    a) *Diceros bicornis bicornis*   Southern black
    b) *Diceros bicornis brucii*   North-eastern black
    c) *Diceros bicornis chobiensis*   Chobe black
    d) *Diceros bicornis ladoensis*   Ugandan black rhinoceros
    e) *Diceros bicornis longipes*   Western black rhinoceros
    f) *Diceros bicornis michaeli*   Eastern black rhinoceros
    g) *Diceros bicornis minor*   South-central black rhinoceros
    h) *Diceros bicornis occidentalis*   South-western black

2. *Ceratotherium simum*   White
    a) *Ceratotherium simum cottoni*   Northern white
    b) *Ceratotherium simum simum*   Southern white

3. *Rhinoceros unicornis*   Indian
4. *Rhinoceros sondaicus*   Javan
5. *Dicerorhinus sumatrensis*   Sumatran

Other

Alligator mississippiensis	American alligator
Alligator sinensis	Chinese alligator
Argopecten purpuratus	Peruvian scallop
Balaenoptera acutorostrata	North Pacific minke whale
Canis lupus	Wolf
Canis lupus familiaris	Dog
Capricornis crispus	Japanese serow
Cervus nippon	Sika deer

Crassostrea gigas	Pacific oyster
Crocuta crocuta	Spotted hyaena
Gorilla beringei	Eastern gorilla
Helarctos malayanus	Malayan sun bears
Mus musculus	Lab mice
Otis tarda	Great Bustard
Pan paniscus	Bonobos
Papio hamadryas hamadryas	Hamadryas baboons
Rangifer tarandus caribou	Reindeer
Sepiella japonica	Cuttlefish
Theropithecus gelada	Gelada baboon
Ursus arctos	Grizzly bear

ACCESSION NUMBERS

AF187825.1   *Diceros bicornis* mitochondrial D-loop, partial sequence

EU107377.1   *Diceros bicornis* isolate CT *cytochrome b* gene, complete cds; mitochondrial

FJ905814.1   *Diceros bicornis* complete mitochondrial DNA sequence

JF718874.1   *Ceratotherium simum cytochrome b* (*cytb*) gene, complete cds; mitochondrial

Y07726.1   *Ceratotherium simum* complete mitochondrial DNA sequence

CHAPTER ONE
REVIEW OF LITERATURE

1.1   Introduction

The illegal trade of flora, fauna and the associated by-products is estimated to be worth approximately US $20 billion per year according to the International Policing Organisation (Alacs *et al.*, 2009). The illegal trading of wildlife items is appealing to criminals as there are substantial profits, while the consequences of trading in wildlife are less than if prosecuted for weapons or drugs (Alacs *et al.*, 2009). Identification of wildlife products which are illegally traded is the first step in the prevention of illegal activity (Alacs *et al.*, 2009). Molecular techniques are useful in wildlife forensics (Alacs *et al.*, 2009; Idaghdour *et al.*, 2003) and have the potential to save populations from local and global extinction. Molecular markers can accurately allow for species identification. DNA can be extracted from degraded or highly processed products such as dried or cooked meats, egg shells, animal hairs, ivory, bone and horns.

1.2   South African rhinoceros

The rhinoceros family can be separated into five species namely, the white (*Ceratotherium simum*), black (*Diceros bicornis)*, Indian (*Rhinoceros unicornis*), Javan (*Rhinoceros sondaicus)* and Sumatran (*Dicerorhinus sumatrensis*) rhinoceroses (Nielsen *et al.,* 2008). The black and white rhinoceros are found in South Africa.

1.2.1   The black rhinoceros

There has been much debate on how to categorise the subspecies of the black rhinoceros. There have been eight subspecies identified (Groves, 2011). These are the Southern black rhinoceros (*Diceros bicornis bicornis)*, North-eastern black rhinoceros (*Diceros bicornis brucii*), Chobe black rhinoceros (*Diceros bicornis chobiensis*), Ugandan black rhinoceros (*Diceros bicornis ladoensis*), Western black rhinoceros (*Diceros bicornis longipes*), Eastern black rhinoceros (*Diceros bicornis michaeli*), South-central black rhinoceros (*Diceros bicornis minor*) and South-western black rhinoceros (*Diceros bicornis occidentalis*) (Groves, 2011). The Southern, North-eastern, and Western black rhinoceros are considered to be extinct, with the Western black rhinoceros being classified as extinct as recently as 2011. The four subspecies that are more commonly recognised are the Western, Eastern, South-western and South-central black rhinoceros (Emslie & Brooks, 1999). According to the

World Wildlife Fund (2013a), the Southern-central, South-western and East African black rhinoceros are all classified as critically endangered.

The black rhinoceros is often referred to as the hook-lipped rhinoceros; because of the upper lip being a triangular shape (Stuart & Stuart, 2006). The natural colour of the black rhinoceros is a dark grey, although due to their frequent bathing these animals are often covered in mud. The black rhinoceros has two horns, with the front horn being longer than the second. The horns are attached to the animal's skin (Stuart & Stuart, 2006). The black rhinoceros is a solitary animal, however individuals may congregate at waterholes and areas where the soil is mineral rich. The cow and bull will briefly have an encounter for mating, and the newly born calf will stay with the cow for approximately three years (Stuart & Stuart, 2006). The black rhinoceros feeds on trees, grass and shrubs and is therefore found in regions with this type of habitat (Stuart & Stuart, 2006).

1.2.2   The white rhinoceros

The white rhinoceros has two subspecies, namely the Northern (*Ceratotherium simum cottoni*) and Southern (*Ceratotherium simum simum*) white rhinoceros (Emslie & Brooks, 1999). The Southern white rhinoceros faced extinction at the end of the 19th century. In 1885 it was estimated that only twenty individuals remained (Emslie & Brooks, 1999). However, this species of rhinoceros is now estimated to have a population of over 20 000 individuals worldwide. This subspecies is now classified as "near threatened", and as shown in Figure 1, this is a huge improvement from the previous category of almost extinct in the wild. The Northern white rhinoceros population status is critical and is classified as critically endangered or extinct in the wild. In 1960, there was an estimated 2 250 individuals, and as of 2013, there are only four known surviving Northern white rhinoceros (World Wildlife Fund, 2013b). Field protection is currently being used for this subspecies. This includes the use of fenced areas in order to concentrate wild populations so that individuals can be intensively monitored. Fenced areas include conservancies, sanctuaries, protection zones and conservation areas (Emslie, 2012). Monitoring of rhinoceros populations has become important to allow for decisions to be made in terms of the management of the different populations, as some populations can experience rapid growth. This allows for rhinoceroses to be translocated and for new populations to be formed (Emslie, 2012).
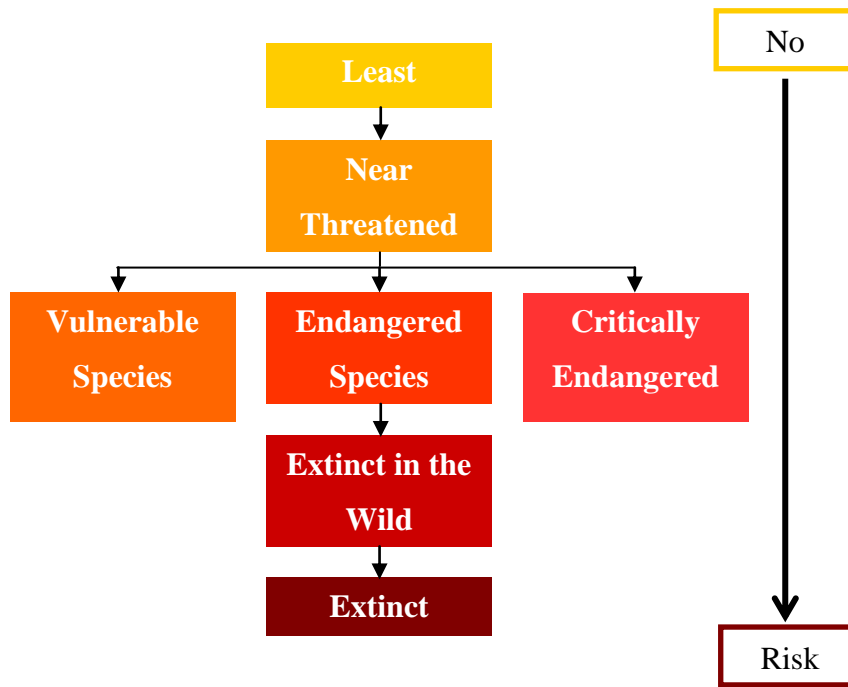
Figure 1: Conservation statuses which can be assigned to lower taxa or species. The risk increases for species as their status moves from top to bottom. Species that fall in the least concerned category are not at risk as there are abundant individuals and populations for these species. Species that fall into categories that are more critical, such as vulnerable, endangered and critically endangered species are at more risk, as they may face extinction if the populations continue to decrease in size. This chart is used to make conservation decisions as all species fall into one of the above categories. If species face no risk, conservation would not be deemed a priority. However, if species are critically endangered, conservation efforts would be deemed a priority as there species could face extinction in the future. The black rhinoceros is currently holds the critically endangered status and the white rhinoceros holds the status of near threatened (Adapted from International Union for Conservation of Nature and Natural Resources, 2013).

The white rhinoceros is referred to as the square-lipped rhinoceros, and is generally larger than the black rhinoceros (Stuart & Stuart, 2006). This rhinoceros also has two horns, with the front horn being longer than the second. The white rhinoceros is generally found in savannah areas, which is an ecosystem dominated by short grass and trees. This rhinoceros prefers fibrous plants while the black rhinoceros prefers more coarse woody material (Stuart & Stuart, 2006). The white rhinoceros is more social than the black. Usually, there is a territorial bull and then his subordinate bulls which are also accompanied by the calves and cows (Stuart & Stuart, 2006).

As the white rhinoceros is bigger in size compared to the black rhinoceros, its horns are bigger. The horns from white rhinoceros can weigh up to 6 kg (Emslie & Brooks, 1999). The rhinoceros horn can regrow if dehorned (Emslie & Brooks, 1999). The rate of horn growth

depends on a variety of factors which includes: sex, age, species and the environment. According to Pienaar *et al.* (1991), on average a horn can grow 50 mm per year with growth ranging for 55-66 mm per year for the rhinoceroses sampled in the Kruger National Park.

1.3   Demand for rhinoceros horn

Rhinoceros horn is made out of keratin (Emslie & Brooks, 1999). All higher vertebrates have this protein. Keratin is chemically unreactive and mechanically robust (Coulombe & Omary, 2002). Appendages such as feathers, hair, nails and horns are composed of 85% keratin (Emslie & Brooks, 1999). Rhinoceros horn is used mainly in Chinese traditional medicine, and as ornaments in the Middle East (Emslie & Brooks, 1999). Rhinoceros horn is thought to serve as an aphrodisiac in the East. It has also been used for medicinal purposes as it is thought to help with influenza, fevers, leukaemia, hepatitis and to treat burns (Nowell *et al.,* 1992). In the 1980's, it was estimated that China used 600 to 700 kg of rhinoceros horn per year (Martin, 1990). Recently, rhinoceros horn was believed to cure cancer in Vietnam (Milliken & Shaw, 2012).

The use of rhinoceros horn can be dated back to 618 AD in items such as carvings, cups and bowls (Martin & Martin, 1982). Rhinoceros horn has also been used extensively in Oman and Yemen for dagger handles. It was estimated that three tonnes of rhinoceros horn was imported annually into Yemen from 1972 to 1978, at that time this accounted for approximately 40% of the world's rhinoceros horn mass (Martin, 1984).

In 1977, the trade of rhinoceroses and their by-products became prohibited by being included in the United Nations Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) appendix I (Leader-Williams, 1992). However, CITES permits were obtainable which allowed for white rhinoceros to be legally hunted as trophies (Emslie & Brooks, 1999).

1.4   The impact of poaching on rhinoceros

According to Emslie & Brooks (1999) the black rhinoceros had the highest population size for rhinoceros in the 19th century, with an estimate of several thousand individuals in Africa. The black rhinoceros was subsequently hunted when the Europeans populated the African continent. In 1933, only two breeding populations of ~110 black rhinoceros were found in Southern Africa. However by 1960, there were a reported 100 000 black rhinoceros present in Africa. Due to the recent increased in the demand for rhinoceros horn from the East, the

numbers have again rapidly declined (Emslie & Brooks, 1999). Unfenced areas were targeted in the 1980's due to the wide roaming nature of the black rhinoceros. There was a lack of funds and support to allow for protection of these animals. Protection was provided to the few remaining individuals who had evaded poachers whilst living in the unfenced areas. This protection involved translocating the animals to secure locations (Emslie & Brooks, 1999). A net increase of black rhinoceros was observed in Namibia and South Africa from 1980 to 1997, because of investments made in law enforcement and monitoring to help with conservation efforts (Emslie & Brooks, 1999).

The black rhinoceros is considered to be one of the most endangered species on the African continent (Van Coeverden de Groot *et al.,* 2011). Throughout the past several centuries' poaching, disease and habitat destruction as a result of human activities has led to a reduction in population sizes for multiple large African mammals (Cardillo *et al*., 2005). Due to an increased demand for rhinoceros horn, larger numbers of rhinoceros have been poached with the majority of cases being fatal. This has led to a 97% decrease in the black rhinoceros population in the last fifty years (Emslie & Brooks, 1999). Poaching of the black rhinoceros still occurs despite conservation efforts (Appendix 1 CITES). The number of poached rhinoceros (this includes black and white) continues to increase in South Africa. Figure 2 indicates that the number of reported poaching cases has tripled since 2010 (Department of Environmental Affairs, 2014). A census was conducted in 2012 and 2013, which estimated that the black rhinoceros population had 5 055 individuals and the white population has 20 405 individuals (Save the Rhino International, 2015).
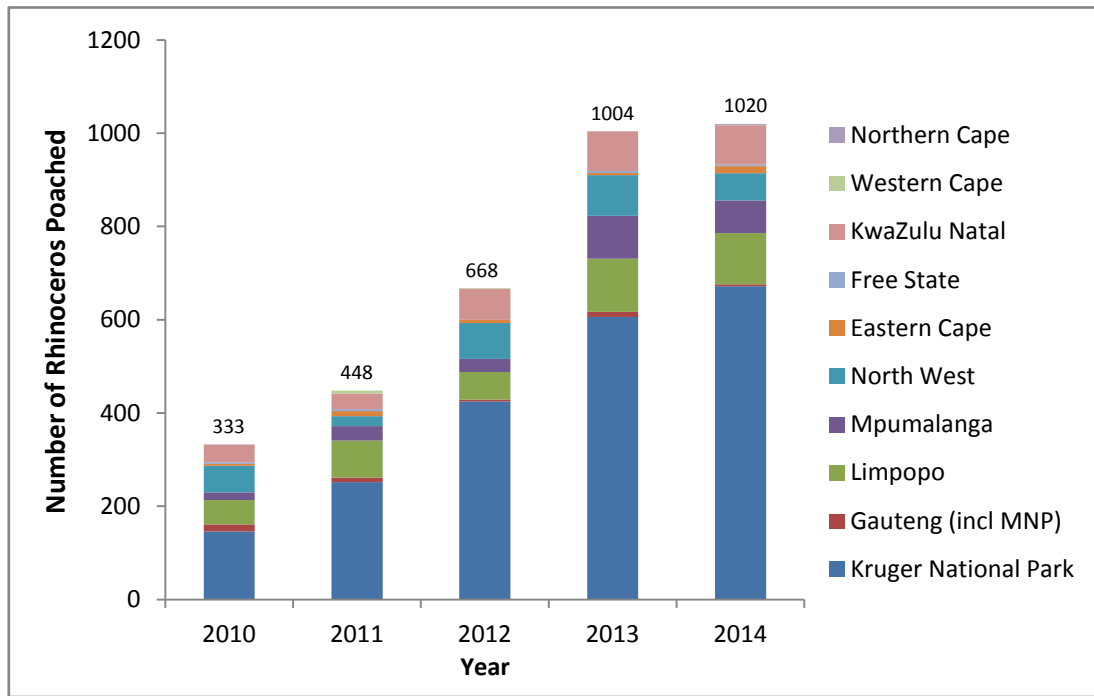
Figure 2: The statistics for rhinoceros poaching in South Africa as of the 20th of November 2014. MNP: Marakele National Park. All data was adapted from the Department of Environmental Affairs (https://www.environment. gov.za/mediarelease/rhinopoaching _wildlifetrade).

Internal threats also need to be considered with regards to rhinoceros conservation. During the 1960's to 1970's northern Africa experienced a 70% decrease in the Northern white rhinoceros population; as during the civil wars no one focused on the poaching problem. According to Emslie & Brooks (1999), Chad, Somalia, Sudan, Rwanda, Uganda, Angola, Namibia, Mozambique, the Democratic Republic of the Congo and Central African Republic all had a decrease in rhinoceros populations since the 1960's as these country's politics were unstable. Countries such as Angola even traded rhinoceros horn for weapons, and Potgieter (1995) has suggested that whilst South Africa was under the Apartheid regime, horn smuggling was ignored or disregarded.

Prohibiting the trade of rhinoceros horn has proven to be a controversial issue, as some believe this will create a larger demand and black market trading will increase (Emslie & Brooks, 1999). The prohibition of the rhinoceros horn trade will cause a negative cycle as the price of rhinoceros horn will increase and be more of an incentive to sell the illegal substance. Alternatively the prohibition could help with the conservation of the species. As a result of the trade ban, most horns have been acquired through the illegal poaching of rhinoceros. Emslie & Brooks (1999) suggest that if rhinoceros horns were harvested and provided to the

regions of demand, the amount of illegal poaching would decrease as the demand would be addressed (Emslie & Brooks, 1999).

1.5   Wildlife conservation

All animals contribute to ecosystems as well as ecosystem services. For example, large mammals help with nutrient cycling and the pollination of plants (Frankham *et al.,* 2002). Animals can also be used for aesthetic purposes. Zoos, nature reserves, game reserves and ecotourism all contribute to the economic value of a country, and therefore the IUCN (world conservation union) recognise the need to conserve biology with respect to ecosystems, genetics and species diversity (Frankham *et al.,* 2002). Protecting the habitats as well as the animals and plants that inhabit these regions is referred to as wildlife conservation (Mills, 2012). Human activity has had negative impacts on wildlife (*i.e.* industrialisation and poaching), and therefore wildlife conservation aims to protect existing wildlife for future generations (Mills, 2012). Numerous approaches can be undertaken with regards to wildlife conservation; of which one is genetic conservation.

1.5.1   Genetic conservation

Genetic conservation can be defined as the "application of genetics to preserve species as dynamic entities capable of coping with environmental change" (Frankham *et al.,* 2002). In 2002, a total of 9% of mammals were classified as critically endangered and 16% were classified as endangered (Frankham *et al.,* 2002). The loss of genetic diversity can lead to a decrease in animal fitness (Garnier *et al*., 2001). A defect in animal fitness is commonly observed as a result of inbreeding within the population group (Saccheri *et al*., 1999). Decreased fitness can also affect fertility, which can lead to survival issues (Frankham *et al.,* 2002). Strategy implementation, unit management and population groups that are genetically distinct strongly influence the outcomes of species conservation (Anderson-Lederer *et al*., 2012).

Populations that are genetically different from other groups have the potential of containing genetic variation which is specific to the group. These genetic variations may be as a result of local adaptation to the group's habitat (Anderson-Lederer *et al*., 2012). Conservation genetics focuses on problems such as interbreeding depression, loss of genetic diversity, decrease in gene flow, genetic drift replacing natural selection as the chief evolutionary progression, purging of deleterious mutations and adaption to captivity with regards to genetics (Frankham *et al.,* 2002). Outbreeding depression may be caused by mixing these distinct groups with

other individuals due to genetic modification of unique genetic traits. Alternatively, genetic drift between population groups can lead to differences in the group's genetics. Population groups which are diminishing and the fragmentation of populations' habitat can lead to a genetic drift in a species (Anderson-Lederer *et al*., 2012).

Small populations are susceptible to environmental, demographic and genetic factors (Gilpin & Soulé, 1986). This is especially true for population groups that are isolated. These factors have the potential to threaten the species to a point of extinction (Gilpin & Soulé, 1986). Small populations of wildlife are often forced to utilize inbreeding due to the lack of individuals, this leads to the decrease in genetic diversity. Species are at a greater risk when there is reduced genetic diversity. If there is a need for change (*i.e.* environmental disaster, drought etc.) populations with a higher degree of genetic diversity will have a better chance of survival (Frankham *et al.,* 2002).

According to Anderson-Lederer *et al*. (2012), low genetic variation is not always caused by human influence on habitat fragmentation or poaching. An hypothesis has been put forward that low levels of autosomal and mitochondrial DNA (mtDNA) variation are a result of small populations adaptations that are low and demographic separation formed over a long period (Anderson-Lederer *et al*., 2012). Miller & Waits (2003) conducted an investigation where they found that when compared to other American grizzly bears (*Ursus arctos*), the Grizzly bears found in the Yellowstone National Park (United States of America) had lower autosomal and mtDNA genetic variation. There was no indication of genetic bottlenecking in this geographical area (Miller & Waits, 2003).

Genetic conservation uses multiple markers to assess the level of genetic variation within a species (Frankham *et al.,* 2002). The genetic makeup of a eukaryote is comprised of the nuclear genome (incorporating autosomal and sex chromosomes) and a range of extranuclear genomes. This range of extranuclear genomes could include the chloroplast, kinetoplast or mtDNA genomes depending on the organism (Liu *et al.,* 2007). The nuclear and mitochondrial genomes are commonly used in forensic and genetic analyses (Liu *et al.,* 2007), however most wildlife forensic techniques use mtDNA analysis.

1.6 Mitochondrial DNA

MtDNA is useful for determining population history, structure and for identifying individuals (Moritz & Cicero, 2004). MtDNA is preferred as a source of genetic material as it tends to be

more abundant in cells (500-2 000 copies per cell) when compared to nuclear DNA (2 copies per cell) and it is more resistant to degradation than nuclear DNA (Moritz & Cicero, 2004).

MtDNA is vital for mitochondrial function maintenance (Kraytsberg *et al.,* 2004). MtDNA is a double stranded, histone free, circular molecule which is approximately 16 400 bp in size (Ashley *et al*., 1990). Due to its size, shape and abundance it is easily purified (Anderson *et al.,* 1982). Mitochondria are localised in the cytoplasm of almost all cells. The mitochondrial genome encodes 22 *tRNA*s and two *rRNA*s as well as approximately 80 protein subunits (13 polypeptides) which are all involved in oxidative phosphorylation (Budowle *et al*., 1999; Ashley *et al*., 1990). Mammalian mitochondria share similarities, such as the position and order of the genes (Linacre & Tobe, 2009). The general mtDNA structure is observed in Figure 3.
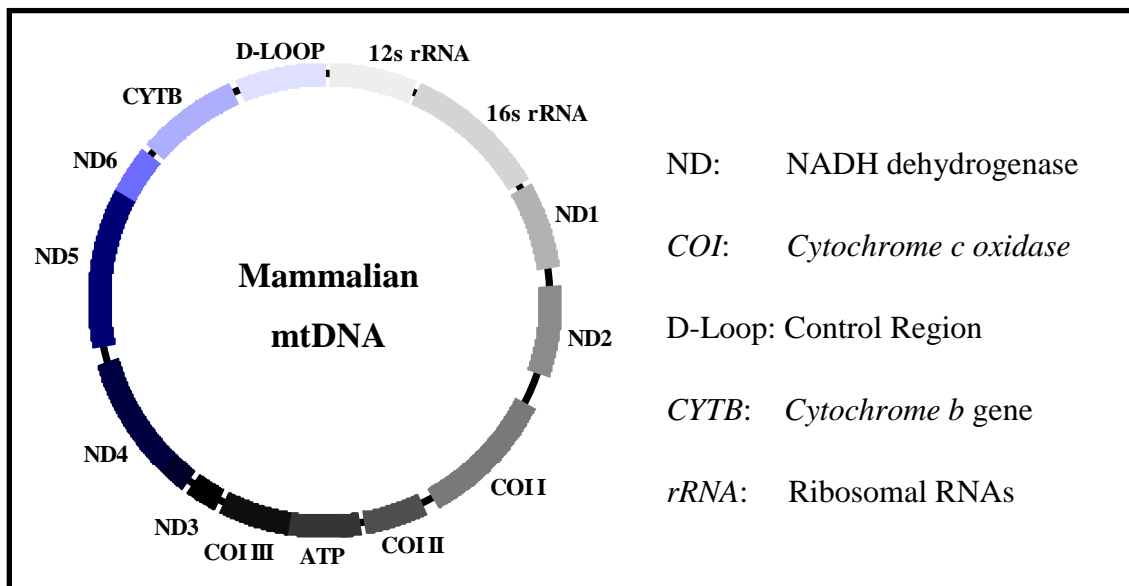
Figure 3: A schematic representation of the mammalian mitochondrial genome. This genome encompasses double stranded DNA in a circular loop. The 37 genes incorporated in this genome vary in size and produce the required RNA molecules or proteins for the mitochondria. The main genes used for species identification include *COI*, *12S rRNA* and *cytb* (adapted from Lincare & Tobe, 2009).

MtDNA consists of one heavy chain (purine rich) and one light strand (pyrimidine rich) (Budowle *et al*., 1999). Repair mechanisms in mtDNA are reported to be lacking and this, in conjunction with the low fidelity polymerase, allows for mtDNA to have a mutation rate which is higher than the nuclear genome. This can be as high as 10 fold in comparison to single copy nuclear DNA (Budowle *et al*., 1999; Belay & Mori, 2006).

MtDNA can be advantageous over nuclear DNA in certain forensic instances. For example, when extracted DNA is degraded or low in concentration, nuclear DNA markers may be lost whereas mtDNA markers may still be present due to the abundance of mtDNA copies found in each cell. Therefore mtDNA would be a preferred source of DNA (Budowle *et al.,* 1999). However, when examining mtDNA issues such as heteroplasmy, recombination and paternal leakage should be considered (Budowle *et al.,* 1999).

Individuals are generally considered as homoplasmic (*i.e.* having identical copies of mtDNA) although heteroplasmy can be observed (Bendall *et al.,* 1997). Heteroplasmy is present when more than one type of mtDNA is present in an individual. Three forms of heteroplasmy may occur. The first form occurs when an individual has two sources of mtDNA. These sources will be present in different tissues. The second form occurs when two sources of mtDNA are present in a single source of tissue and homoplasmic in another source of tissue. The third form will be present when an individual contains more than one source of mtDNA in a particular tissue (Carracedo *et al.,* 2000). According to Budowle *et al.* (1999), the first form of heteromplasmy is considered to be the least frequent, and that when examining the HV1 and HV2 regions heteroplasmy is not likely to be observed. However, if this form did occur, the sequence of DNA would most likely differ by a single nucleotide (Budowle *et al.,* 1999).

Recombination with regards to human mtDNA has been observed. However mtDNA recombines with copies of itself and therefore remains unchanged from parent to offspring (Anderson-Lederer *et al.,* 2012). Paternal leakage refers to mtDNA that is inherited from the father, and due to the fact that sperm is destroyed in the fertilization process, this occurrence is rare. Laboratory mice (*Mus musculus*) have been reported to have an mtDNA paternal leakage rate, per generation, of between 1×10^{-5} to 5×10^{-5} (Gyllensten *et al.,* 1991). Human paternal leakage is often debated in court, in order to get evidence dismissed. However the occurrence is unknown (Budowle *et al.,* 1999).

Unlike nuclear DNA, which is inherited from both parents, mtDNA is usually inherited exlusively from the mother. Sperm contains mtDNA in the tail and neck regions; however the mitochondrial genome is subsequently destroyed after or during fertilization. Therefore, all maternal relatives and siblings share the same mtDNA (Anderson-Lederer *et al.,* 2012). MtDNA is therefore affected by population demographics, making mtDNA useful for determining familial lineages (Anderson-Lederer *et al.,* 2012). This has been proven to be

useful in the cases of missing individuals or other forensic work which concerns familial relationships as members from different generations can provide a sample of mtDNA.

According to Nabholz *et al*. (2008), mtDNA has been used extensively for determining the genetic history of a species and for population reconstruction because it is easily amplifiable and contains variation within and between species. MtDNA coding regions are often used for interspecific studies and non-coding regions are used for intraspecific or population studies (Glenn *et al*., 2002). MtDNA testing has been conducted on DNA extracted from numerous sources such as wigs, shoes, cell phones and clothes as the analysis can be conducted on low yields (Andréasson *et al*., 2002).

Regions which consist of high evolution rates can be used for human individualisation (Linacre & Tobe, 2009). The mitochondrial genome consists of coding and non-coding regions. MtDNA testing for humans involves PCR and/or the complete sequencing of the control region (D-loop) and hypervariable regions (HV1 and HV2). The mtDNA D-loop is a significant non-coding portion (approximately 1 100 bp) of the mitochondria genome (Linacre & Tobe, 2009; Parsons & Coble, 2001). The D-loop can be a beneficial region of mtDNA as it can be used for intraspecies phylogeny, population structures, as well as research (Hoelzel, 1993; Aranishi & Okimoto, 2005). Polymorphisms that occur in the D-loop have proven to be highly useful for determining familial relationships as well as showing considerable differences in sequences amongst non-related individuals (Fridman *et al*., 2011). HV1 and HV2 provide 610 bp of information and are extremely variable in populations (Parsons & Coble, 2001). According to Budowle *et al.* (1999), the HV1 region extends from 16 024-16 365 and HV2 extends from 73-340 in the mitochondrial genome. The evolution rate of HV1 and HV2 have a ten-fold increase in comparison with coding regions, therefore variation in the control region is more concentrated. However, the coding region is fifteen times greater than the control region and therefore has more total variation (Parsons & Coble, 2001).

The D-loop region has proven to be useful in human genetic studies. A pyrosequencing assay was developed by Andréasson *et al.* (2002) to identify sites that were polymorphic in the D-loop as well as the coding region. This was done in order for the mtDNA analysis to have an increased discrimination power. The assay resulted in accurate and sensitive results (Andréasson *et al.,* 2002). Bataille *et al.* (1999) developed an assay that utilised a single PCR that amplified the *cytb* and D-loop (which included the hypervariable regions) regions which

allowed for human identification as well as individualisation. The PCR products were electrophoresed on an agarose gel; if one band was present the assay indicated the sample originated from a non-human source. Subsequently the D-loop (including the hypervariable regions) was sequenced which allowed for individualisation of humans (Bataille *et al.,* 1999). The authors concluded that this assay produced results that were reliable and reproducible even when used on degraded DNA samples.

Greenberg & Aquardo (1983) investigated mtDNA variation in seven humans. The region was approximately 900 bp, and covered the D-loop and the heavy strand of replication regions. The authors found that the diversity amongst the investigated sequences was 1.77 (Greenberg & Aquardo, 1983). This was higher than expected in comparison with results published by Brown (1980). For the whole mtDNA genome, nucleotide variation for the human population was estimated to be between 0.003-0.004 (Brown, 1980; Ferris *et al*., 1981). This estimation was made based on the patterns that resulted from restriction fragment length polymorphism (RFLP). Greenberg & Aquardo (1983) concluded that this high variation was because of the non-coding region of DNA having a higher evolving rate than coding regions of DNA. Wilson & Cann (1983) also conducted an investigation into genetic variations found in human mtDNA genomes. The rate at which mutations occur in nuclear DNA and in coding mtDNA regions is much less than that of non-coding mtDNA regions (Wilson & Cann, 1983). A total of 112 humans were used for the investigation. This led to the identification of fourteen length variations. These length variations were caused by deletions or additions that ranged from 6 to 14 bp. Only three of these variations were found in the D-loop, the other eleven were found in seven other non-coding regions of the genome (Wilson & Cann, 1983).

A single nucleotide polymorphism (SNP) is a single base variant within a sequence and have been deemed the most prevalent variations found in DNA sequences (Kwok, 2002). SNPs can be found in coding and non-coding regions of DNA, however the majority of SNPs occurring in non-coding regions (Kwok, 2002). SNPs can be used for mtDNA analysis. A haplotype refers to SNPs that occur on specific chromosomes in the genome of eukaryotes (Gabriel *et al.,* 2002). Haplotypes can be either maternally or paternally inherited. A group of haplotypes that are similar and originate from the same ancestor are referred to as haplogroups. These haplogroups have the same SNP present in all haplotypes. MtDNA haplogroups are maternally inherited and differ for people that have originated from different geographic regions or populations due to the evolutionary nature of the mtDNA genome

(Achilli *et al.,* 2004). Caucasian individuals will usually have an average of eight nucleotide differences (Budowle *et al.,* 1999), while people of African ancestry have up to fifteen differences within the hypervariable regions (Vigilant *et al*., 1991). It has been suggested that examining SNPs may become the technique of choice for studies relating to conservation genetics, evolution and ecology. This is due to the fact that SNPs have shown potential for an increase in data quality, ease of interpreting results and the efficiency of genotyping (Morin *et al.,* 2004).

Fridman *et al.* (2011) investigated 26 SNPs in the coding region of mtDNA. Fifteen mother and child pairs were used for the investigation. The pairs were separated into different haplogroups with two having a European haplogroup, two having an Amerindian haplogroup, another two also showed the Amerindian haplogroup however they contained two different haplogroups, and seven showed African haplogroups (three African groups were formed due to different haplogroups). The sequences were aligned and the authors concluded that some haplogroups were not adequate genetic markers as they are conserved in some individuals and therefore these individuals cannot be separated (Fridman *et al*., 2011). By using 26 different SNPs (in different combinations) the authors were able to discriminate between individuals in thirteen of the fifteen pairs. This provided an 86.7% success rate, therefore showing that this method was efficient for discriminating individuals who had the sample D-loop haplogroups (Fridman *et al.,* 2011).

Meyer *et al.* (1999) conducted a study which investigated nucleotide substitution patterns in human mtDNA. The authors examined 1 229 HV1 regions and 385 HV2 regions. The sequences for the HV1 and HV2 regions were obtained from a collection that was available at: http://www.eva.mpg.de/hvrbase/ (Meyer *et al.,* 1999). The results obtained indicated that the HV1 region had a substitution rate that was approximately twice that of the HV2 region. Heterogeneity was found to be more prominent in the HV2 region when compared to that of the HV1 region (Meyer *et al.,* 1999). It is not clear if the substitution rate in rhinoceros HV1 is higher than that of HV2.

The *cytb* gene was used to assess the evolutionary rate of mtDNA (Nabholz *et al.,* 2008). A total of 1 696 mammalian species was used for the investigation. Sequences were obtained from the National Centre for Biotechnology Information (GenBank). The authors reported a two fold increase in mtDNA mutation rates when compared to the mutation rate of nuclear DNA (Nabholz *et al.,* 2008). The slowest mutation rate for a mammal species was 100

million years and the fastest every 1-2 million years. From this study, the authors were able to deduce that a decrease in mutation rates is caused by natural selection in species that have increased longevity (Nabholz *et al.,* 2008).

1.7 <u>Wildlife forensic science</u>

Wildlife forensic science has chiefly been used for identifying species, relationships, populations and individuals, and has developed alongside human forensic science, benefiting from techniques designed for humans (Ogden *et al.,* 2009). DNA evidence is becoming critical in wildlife forensics as it can help identity poached goods and fauna or flora which have been illegally traded (Ogden *et al.,* 2009).

The development of polymerase chain reaction (PCR) in 1983 (Saiki *et al.,* 1985) has proved to be vital for genetic work. This technique is commonly used in wildlife forensic science as it allows DNA to be amplified from trace amounts of sample (Linacre & Tobe, 2009). Another tool which is highly regarded for wildlife forensic science is restriction mapping (Wilson & Cann, 1983). Introns and repetitive DNA features that complicate the analysis of nuclear DNA are absent in mtDNA (Ashley *et al*., 1990). Therefore, analysis by utilizing restriction enzymes is a viable option. Restriction mapping incorporates several restriction enzymes, which aim to digest DNA sequences so that the resulting product will separate the different cleaved fragments of DNA (Wilson & Cann, 1983). Restriction mapping allows for estimations to be made with regards to the genetic relationship of individuals (Ashley *et al*., 1990). This is important in forensic science, as different individuals have unique maps of their DNA (with the exception of identical twins) (Aquadro & Greenberg, 1983).

MtDNA is more commonly used for species identification than nuclear DNA because it is easier to type from degraded or processed tissue (Alacs *et al.,* 2009). However, a limitation for mtDNA is that it is maternally inherited. Breeding efforts for the rhinoceros may result in the translocation of females to different populations; therefore separate populations may contain similar mtDNA. Paternal leakage which results in heteroplasmy may also be of concern as an organism may then contain two different mtDNA genomes (Alacs *et al.,* 2009).

MtDNA is not only used for identifying species but also for identifying geographically distinct populations by haplotype analysis. Haplotype analysis may include sequencing of particular regions in order to examine sequence variation, however the most commonly used technique to examine haplotypes is PCR-RFLP, which utilises restriction enzymes (Ogden *et al.,* 2009). Restriction enzymes will cleave sequences if the restriction site is present,

however different individuals may have SNPs present in some of the restriction sites. Therefore, PCR-RFLP can result in unique profiles being generated for different individuals (Ogden *et al.,* 2009). Cheaper and faster tests can be developed when examining a fragment of DNA for SNPs as long DNA fragments are not required (Ogden *et al.,* 2009).

The discriminatory power of mtDNA analysis can be increased by analysing polymorphisms not only in the region of the D-loop, but also coding regions. Andréasson *et al.* (2002) developed a technique to identity polymorphic sites in the D-loop and coding regions, which was achieved using a pyrosequencing (high throughput technology) assay where multiple PCR fragments were analysed. These fragments included eleven coding regions, four HV1 and six HV2 reactions (Andréasson *et al.,* 2002). These fragments covered the most discriminating regions of the mitochondrial genome. Control samples and casework samples were tested using pyrosequencing technology and they found that this method was sensitive and accurate, producing easily interpreted results. While this assay was initially developed for human samples, it has great potential application in wildlife forensics. While present wildlife forensic techniques have been limited to species identification, there is a need to develop techniques equivalent to the DNA barcode in humans, such that it becomes possible to link samples to individual animals (Andréasson *et al.,* 2002).

A genetics study conducted by Glenn *et al.* (2002) investigated a region of mtDNA which comprised of the 5' end of the *cytb* gene to the centre of the D-loop in 25 American alligators (*Alligator mississippiensis*). These sequences were compared to a single Chinese alligator (*Alligator sinensis*) due to it having a homologous sequence. The D-loop region examined was 624 bp in size. A sequence variation of 17.5% was found between the American and Chinese alligators. This was as expected due to the estimated divergence of the ancestry line. For the *cytb* region, two haplotypes were identified for the 25 American alligators, with 24 sharing one of the haplotypes and the other having the second haplotype (Glenn *et al.*, 2002). This second haplotype differed from the first haplotype by one base substitution. This substitution was silent. The D-loop was found to present three haplogroups for the 25 American alligators, of which 22 belonged to one haplogroup, two belonging to the second and one belonging to the third (Glenn *et al.*, 2002). The authors suggested that the American alligator species was severely affected by bottlenecking.

DNA barcoding uses short genetic markers which are useful in wildlife forensics as one can determine from which species a sample originated (Hebert *et al.,* 2003). The *cytochrome b*

(*cytb*), *cytochrome c oxidase subunit I* (*COI*) and *12S rRNA* genes encoded on the mitochondrial genome are commonly used for species identification. The *12S rRNA* is used when there is insufficient variation in the sequence for the *cytb* gene; however the *COI* gene is the main gene for barcoding, as it has sufficient sequence variation and information to distinguish between species (Linacre & Tobe, 2009).

Dalton and Kotze (2011) investigated if the *COI* gene could provide accurate results for species determination in a forensic wildlife case. They used universal primers (HCO2198 and LCO1490) to amplify a 650 bp region of the *COI* gene from unknown samples (Dalton & Kotze, 2011). The authors were able to determine that two out of the three cases were not poaching incidents. However, the study did highlight the need for reliable reference sequences for correct species identification (Dalton & Kotze, 2011).

1.8   <u>Wildlife forensic science and the rhinoceros</u>

MtDNA has been found to be plentiful in sources such as tissues such as nails and hair. As a consequence, rhinoceros horn is rich with mtDNA and therefore mtDNA represents an obvious target for the identification of these samples (Willerslev *et al.,* 2009). The complete mitochondrial genome for a number of extant and extinct rhinoceros species have been sequenced and are available in public databases such as GenBank (Willerslev *et al.,* 2009).

Genetic differences in species for large mammals are characteristically less than 2% (Anderson-Lederer *et al.,* 2012). Conservation for the black rhinoceros is reliant on different population groups being managed effectively (Brown & Houlden, 2000). Identifying different genetic groups can help maintain genetic variation as well as help with the conservation of a population (Anderson-Lederer *et al.,* 2012). O'Ryan *et al.* (1994) used mtDNA to evaluate differences in black rhinoceros genetics. Skin biopsy samples were collected from black rhinoceros located in Namibia, KwaZulu-Natal and the Eastern Cape in South Africa. Sixteen restriction enzymes where used to create mtDNA restriction maps for different black rhinoceros, this resulted in the separation of the rhinoceros into different geographic groups**,** with restrictions maps for rhinoceros from the same geographical regions being identical (O'Ryan *et al.,* 1994).

Ferguson & Swart (1997) investigated genetic differences in southern African black rhinoceros populations (South-central and Southern black rhinoceros). The authors investigated thirty loci that were located in coding regions of mtDNA. Of these 30 loci, only 6 (Esterase-2, General Protein-3, General Protein-5, Glucose-6- phosphate dehydrogenase,

Hemoglobin-2 and Phosphoglucomutase) were found to be polymorphic. Therefore these 6 loci were used for the geographic analysis.

A total of 215 individuals were investigated. Of these, 21 Southern black rhinoceros were from Namibia (Etosha National Park), 135 South-central black rhinoceros were from Zimbabwe (Zambezi Valley), 25 South-central black rhinoceros came from Hluhluwe-Mfolozi Park and 34 South-central black rhinoceros came from Mkhuzi (Ferguson & Swart, 1997). The authors concluded their investigation by deeming the separate populations as isolated groups (Ferguson & Swart, 1997). The authors suggested that the black rhinoceros should not be translocated from populations in the East to populations in the West. The authors concluded that the genetic management for the black rhinoceros should not be deemed a priority as genetic variation occurred between populations with the exception of rhinoceros from the examined Zimbabwean population.

Brown & Houlden (2001) looked at genetic variation in the D-loop for the South-central and Eastern black rhinoceros. Tissue, faecal and blood samples were used for this investigation. A 450 bp region of the D-loop was examined. Five different haplotypes were identified in nine South-central black rhinoceros by mtDNA variation. This suggested that the lineages came from five different maternal ancestors. The haplotype diversity for the South-central black rhinoceros was found to be 0.86. No more than four nucleotide differences were found for each of the South-central black rhinoceros, therefore the sequence identity for the individuals was 99% (Brown & Houlden, 2001). Two Eastern black rhinoceros individuals were investigated. These individuals had different haplotypes, suggesting that they did not come from the same maternal lineage. The genetic variation between the two investigated subspecies of black rhinoceros was found to be 2.6%. The genetic variation between the South-central black rhinoceros individuals and the white rhinoceros was 14.7% and the variation for Eastern black rhinoceros individuals and the white rhinoceros was 14.2%. This study suggested that the two black rhinoceros subspecies had a divergence time of 0.93-1.3 million years (Brown & Houlden, 2001).

The *cytb* gene has also been used for species identification and phylogenetic studies. The amplification of *cytb* from highly processed and powdered samples has often however been problematic (Hsieh *et al*., 2003). The reason for this is most likely because the gene is 1 140 bp in mammals, which is considered to be large when amplifying from degraded sources.

Hsieh *et al*. (2003) were able to identify species from rhinoceros horn by utilizing a truncated version of the *cytb* gene.

Hsieh *et al.* (2003) developed a method for identifying rhinoceros DNA from highly degraded horn samples. For the study, horn samples were obtained from three white rhinoceros, three black rhinoceros, one Indian rhinoceros and one Holstein cow. Six unknown samples were investigated to determine if the samples originated from rhinoceros. The authors used a 402 bp region of the *cytb* gene (Hsieh *et al.,* 2003). This gene assisted in identifying not only what species the samples originated from but also the phylogenetic relations between the rhinoceros species. The authors obtained a maximum genetic distance value of 0.0333 among the black rhinoceros and 0.0176 among the white rhinoceros (Hsieh *et al.,* 2003). The genetic distance between the black and Indian rhinoceros was 0.1564. The phylogenetic study showed the Holstein cow to be an outlier and that the different species of rhinoceros came from a common ancestor (Hsieh *et al.,* 2003). Mixed samples were also investigated. At a ratio of 1:19 with Holstein cow DNA the c*ytb* gene was able to detect rhinoceros DNA. The authors concluded that all of the unknown samples originated from rhinoceros sources (Hsieh *et al.,* 2003).

A rhinoceros's diet can affect the composition of chemicals in its horn (Lee-Thorp *et al.,* 1992), and is being used for wildlife forensics. Rhinoceros fingerprinting is a forensic tool that examines the chemical composition of the rhinoceros horn. Individuals, within a species, that are located in the same geographical area will have a horn chemical composition which is different to other individuals that inhabit a different geographical area (Amin *et al.,* 2003). The digestive process allows for the absorption of different isotopes into the horn. Amin *et al.* (2003) described fingerprinting as on-going and that this technique may be a viable forensic tool which is both reliable and practical. An initial test is required to determine that the sample did originate from the rhinoceros family. This study aims to differentiate between rhinoceros species as well as to determine if mtDNA can be used for individual differentiation in order to aid forensic investigations.

An alternative approach to minimise the illegal trade of rhinoceros horn has been discussed for years. The hypothesis was introduced by Western (1982) in which the dehorning of rhinoceros would allow for the reduction of illegally traded rhinoceros horn. This approach would be considered a preventative measure in contrast to other forensic approaches, which could be considered as measures of repercussion. However wildlife studies have been

conducted to determine if this was a viable option. Dehorning rhinoceros was attempted in Namibia in 1989 for the first time due to limited anti-poaching funds. Not a single dehorned rhinoceros was poached (Lindeque, 1990). Berger & Cunningham (1994) later investigated the impact of dehorning female rhinoceros. Their study showed that dehorned mothers that lived with spotted hyaenas (*Crocuta crocuta*) had an infant mortality rate of 100% while horned mothers that lived with hyaenas had an infant survival rate of 100% (Berger & Cunningham, 1994). This demonstrated that the rhinoceros horn is vital for mothers to protect their offspring.

1.9   Faecal matter as a source of non-invasive DNA

Non-invasive sources of DNA allow for biological work to be completed on species without the effort and ethical concerns associated with the capturing of the animals (Taberlet *et al.,* 1999). The opioid which is commonly used for the immobilisation of large mammals is etorphine, which is more commonly known as M99 (Roussel & Patenaude, 1975). This drug is a chemical relative of morphine however it has a higher potency (~1 000 fold). This drug produces analgesia, which causes the animal to feel numb and therefore no pain (Roussel & Patenaude, 1975). M99 is commonly used with azaparone. Azaparone is a tranquiliser which helps with the side effects of M99s (*i.e.* nausea and dizziness) (Chui *et al.,* 1994).

Alibhai *et al.* (2001) investigated if immobilization affects black rhinoceros reproduction. Four reproductive parameters were investigated. These were the inter-calving interval, number of conceptions, calving rate and the number of calves born per year (Alibhai *et al.,* 2001). All four reproductive parameters appeared to be negatively affected by repeat immobilization (Alibhai *et al.,* 2001). The number of immobilisations and the interval at which immobilisations were administered were accounted for in these studies. The authors concluded that guidelines should be installed as to minimise the effect of immobilisation on female reproduction (Alibhai *et al.,* 2001). The suggested guidelines included that the immobilization of pregnant females should be avoided, the frequency of immobilization should my minimized, and through the duration of immobilization stress should be minimized/avoided (Alibhai *et al.,* 2001).

A study conducted by Heard *et al.* (1992) investigated the effect of M99 on a white rhinoceros female. The 28 year old female was immobilised twice (on consecutive days) while pregnant. Their findings concluded that hypoxemia, apparent hypertension and hypercapnia were associated with prolonged periods of resting and immobilization. Similar

results were found in a study conducted by Kock (1985) where the effects of M99 were tested on a white and Indian rhinoceros.

Kock *et al.* (1990) investigated the effects of immobilising black rhinoceros with combinations of M99, fentanyl and xylazine. Out of the 52 rhinoceros examined, 23 were stressed, and 29 showed no adverse effects. Although no fatalities were encountered, the fact that 44% of the tested rhinoceros were found to be stressed is problematic (Kock *et al.,* 1990). Wenger *et al.* (2007) investigated what the impact of adding butorphanol to an immobilisation mixture, which included M99, would have on cardiopulmonary effects on white rhinoceros. Darting from helicopters was used for immobilisation. When rhinoceros were immobilised with M99 (control and the mixture with butorphanol) severe hypoventilation and hypoxemia was observed (Wenger *et al.,* 2007). A non-invasive technique would be beneficial with regards to studying DNA obtained from female rhinoceros as the stress caused by immobilisation would be removed.

One of the sources of non-invasive DNA is faecal matter. Faecal sampling would allow for more animals to be sampled, lower costs for scientific research as one would not need the use of a helicopter or the required narcotics to immobilise individuals, and no stress could be caused to the animal. In a study by Ball *et al.* (2007), genomic DNA extracted from reindeer (*Rangifer tarandus caribou*) faecal matter yielded a concentration of 6.5-28.6 ng/µl. It should be noted that the amount of DNA extracted from the reindeer itself is unknown as the animals diet will contribute to the obtained genomic DNA. Their research resulted in a protocol which allowed for the determination of the amount of target DNA present in a sample before amplification (Ball *et al.,* 2007). The authors concluded that by characterizing DNA extracted from faecal samples, one can minimize the problems that can occur when working with this source of DNA (Ball *et al.,* 2007).

When working with faecal sample, analysis of shorter fragments of DNA (150-300 bp) is preferable. Peppin *et al.* (2010) were able to extract black rhinoceros DNA from blood, horn and faeces. Degraded DNA is not suitable for amplifying and analysing large regions of DNA, therefore Peppin *et al.* (2010) developed a reliable PCR-based method which allowed for sex determination for the white and black rhinoceros using a short fragment of DNA (approximately 100 bp). Their method used an unlabelled reverse primer and a fluorescently labelled forward primer, which allowed for the amplification of a zinc finger (structural motif) with either the X (ZFX) or Y (ZFY) chromosome. A size difference of 7 bp occurs between

ZFX and ZFY, therefore capillary electrophoresis was used to determine if sample originated from either a female or male rhinoceros (Peppin *et al.,* 2010).

Nielsen (2007) investigated whether DNA extracted from rhinoceros faeces could be a used as a competent substitute to DNA extracted from blood samples. She found that faecal samples did not have comparable DNA to that of blood samples. Through the use of RAPD (random amplification of polymorphic DNA) faecal samples were found to have lower molecular weights and to be more degraded in comparison to blood samples, which had a higher molecular weight and better quality DNA. Short tandem repeat (STR) analysis revealed that faecal samples did not compare to results obtained from blood samples. The sequences that resulted from faecal samples had no principal peaks. This was thought to result as a consequence of poor quality DNA that had been contaminated. Nielsen (2007) concluded that a different approach to RAPD should be undertaken when comparing blood and faecal samples, and concluded that techniques used for collection of faecal sample may cause problems with regards to the quality of extracted DNA.

Garnier *et al.* (2001) were able to obtain STRs from faecal samples for black rhinoceros, which allowed for genetic relationships to be assessed. The authors emphasized that specific precautions must be taken in order to ensure that good quality DNA could be extracted for faeces (Garnier *et al.,* 2001). These precautions included taking multiple samples of fresh dung and performing multiple extractions on collected samples (Garnier *et al.,* 2001). Sample collection was critical as the environment can lead to degradation such as damage from natural UV light, bacterial breakdown and physical destruction (Garnier *et al.,* 2001). Initially, it is necessary to compare faecal DNA to a positive control (*i.e.* horn, tissue or blood) determine the level of degradation.

Recent studies have examined nuclear DNA from faecal samples. These studies have highlighted problems such as false alleles and allelic dropout (Wehausen *et al*., 2004). False alleles are products created during the amplification procedure (Pompanon *et al.,* 2005). This can result when too many cycles were performed or the incorrect annealing temperature was used. Allelic dropout occurs when the target allele has not been amplified (Pompanon *et al.,* 2005). This may occur when degraded DNA was utilized, not enough template was used, too few cycles were employed or the incorrect annealing temperature was utilized. These analytical problems have led to the conclusion that examining nuclear DNA from faecal samples can cause unreproducible results (Wehausen *et al*., 2004).

The examination of mtDNA from faecal samples has proven to be uncomplicated in comparison to nuclear DNA due to the copy number ratio being in favour of mtDNA (Wehausen *et al*., 2004). However, caution should be taken as plant inhibitors (for herbivores) or bacteria that were obtained by consumed prey (carnivores) have resulted in amplification problems when using faecal matter as a source of DNA (Wehausen *et al*., 2004; Khanuja *et al*., 1999; Yamashiro *et al*., 2010). Wasser *et al.* (1997) were able to amplify mtDNA with sizes varying from 246 to 700 bp from faecal, tissue and blood. The group amplified a 246 bp gene from two Malayan sun bears (*Helarctos malayanus*). This region was digested with two restriction enzymes. The results for the three sources of DNA were identical. The two individual Malayan sun bears were found to have a 10 bp difference in the examined gene (Wasser *et al.,* 1997).

Immobilisation is routinely used for large mammals when DNA is obtained from invasive samples (Roussel & Patenaude, 1975) and has been found to affect black and white rhinoceros reproduction (Alibhai *et al.,* 2001). Faecal samples as a source of non-invasive DNA is cost and time efficient and no stress would be placed on the animals. However precautions, such as ensuring the faecal samples are collected when fresh and that no insect contamination has occurred, should be taken (Garnier *et al.,* 2001).

CHAPTER TWO
PROBLEM STATEMENT, AIMS AND OBJECTIVES

The black rhinoceros is a critically endangered species while the white rhinoceros is a near threatened species (Van Coeverden de Groot *et al.,* 2011). The black rhinoceros has been severely affected by poaching which has contributed to the 97% decrease in the population over the last fifty years (Emslie & Brooks, 1999). There is currently no accepted forensic method to identity the origin of retrieved poached or smuggled rhinoceros body parts to prosecute perpetrators. Most forensic analyses involving poached samples are limited to the identification of species present in the sample. However, this information is insufficient if it is to be used to prosecute perpetrators or to understand the networks by which these samples are trafficked. In these cases, it is necessary to identity both the geographic location of a sample and to link the sample to a particular individual animal. In addition, these techniques would also allow the identification of DNA from an individual animal by the analysis of blood, hair or tissue samples obtained from the belongings of a suspected poacher. The development of a non-invasive molecular technique which will allow for the identification of species, geographical location of rhinoceros populations as well as the development of a technique to allow individualisation of particular animals would support conservation and anti-poaching efforts. While the present study is based on the rhinoceros, it may inform future similar studies on other species.

The aim of this project was therefore to develop forensic techniques to study mtDNA from non-invasive samples to attempt the identification of specific individuals from the species *Diceros bicornis* and *Ceratotherium simum.* Such an approach would also allow for the identification of poached or smuggled goods (such as the horn).

The following objectives were set for this research project.
1. Isolate mitochondrial DNA from rhinoceros faeces, tissue and blood;
2. Develop a non-invasive mtDNA PCR-based technique for species identification between *Diceros bicornis* and *Ceratotherium simum* using *COI*; and
3. Develop a non-invasive mitochondrial DNA PCR based technique to examine genetic variation by investigating the D-loop, *cytb* and *16s rRNA* and hypervariable regions in the *Diceros bicornis* and *Ceratotherium simum.*

CHAPTER THREE

MATERIALS AND METHODS

3.1 <u>Materials</u>

The ZR faecal DNA MiniPrep™, quick-gDNA™ MiniPrep, geneJET gel extraction kit, geneRuler™ 50bp DNA ladder and O'Gene ruler 1kb plus DNA ladder were purchased from Inqaba Biotec (RSA). All PCR primers were synthesized by Inqaba Biotec (RSA). The KAPA HiFi hot start ready mix was purchased from Lasec (RSA). The restriction enzyme (*Hin*dIII) was purchased from Promega. All other reagents were purchased from Sigma Chemicals (USA) or Merck Chemicals (Germany), unless otherwise stated.

The black rhinoceros samples were collected from three locations that were situated in two provinces in South Africa, while the white rhinoceros samples originated from two locations in the same province. Eight black and seven white rhinoceros faecal samples were used for this investigation. Two tissue samples were used for each species of rhinoceros. Faecal sample was taken directly from the rectum to ensure that samples were fresh and originated from different individuals. Tissue samples were obtained from collected ear notch samples. These samples were provided by a qualified veterinarian (Dr Peter Brothers, Bothers Safaris) who was conducting routine immobilizations of the animals for private landowners. No request for immobilization was made for this investigation. Samples that came from the same individual were termed matched samples.

3.2 <u>Analysis of the rhinoceros mitochondrial DNA sequence *in silico*</u>

The NCBI database (http://www.ncbi.nlm.nih.gov/) was used to locate the mitochondrial genome sequences for the black (FJ905814.1) and white (Y07726.1) rhinoceros. The genome and its features were plotted using the bioinformatics programme BioEdit. Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) was used to perform all multiple alignments. The parameters for the multiple sequence alignment were set to default (dealign input sequences: no; mBed-like clustering guide-tree: yes; mBed-like clustering iteration: yes; number of combined iterations: 0; max guide tree iterations: 0 and max HMM iterations: 0). The multiple alignment diagrams were created using BioEdit (Hall, 1999). Bioinformatics programme, TotalLab (Newcastle, UK), was used to determine the size of the visualised PCR bands.

3.4 DNA extraction from tissue, blood and faecal samples

DNA was extracted from rhinoceros faeces using the ZR Faecal DNA MiniPrep™, and from rhinoceros tissue and blood using the Quick-gDNA™ MiniPrep. Faecal, tissue and blood samples were stored in 70% ethanol for the duration of the project. A total of 150 mg of faecal sample, 20 mg of tissue sample and 100 µl of blood samples was required for each DNA extraction. The instruction manuals for both extraction kits were followed with the exception of the elution buffer volume. Faecal DNA was eluted with 30 µl and tissue and blood DNA was eluted with 40 µl. The Nanodrop2000 spectrometer was used to quantify the extracted DNA. This was achieved by reading the absorbance of the DNA at a wavelength of 260 nm. The extracted DNA was stored at -20°C for the duration of this project.

3.5 Visualisation of PCR products on agarose gels

Agarose gels were constructed with 1x TAE buffer and were made to have a final agarose concentration of 1% (w/v), with the exception of gels utilized to visualise restriction digestion products which were made to a final concentration of 2% (w/v). Ethidium bromide (0.20 mg/mL) was added to the gels to allow for UV visualisation. Gels were set for 20 minutes and electrophoresed in 1x TAE buffer. The 1% gels were run at 70 V for 90 minutes and the 2% gels were run at 70 V for 140 minutes, respectively. A total of 4 µl of GeneRuler™ 50bp DNA Ladder or O'Gene Ruler 1kb Plus DNA Ladder was added to each gel. A total of 150 ng was loaded from each sample for all gels, unless otherwise stated. The agarose gels were visualised using the ChemiDoc (Uvitec®). The gel was viewed with ultraviolet light (λ= 300 nm) by utilizing the UVI Prochemi software.

3.6 Amplification of rhinoceros target regions located on the mitochondrial genome

Primers were designed to amplify denatured DNA from faeces. Factors listed by Bradley *et al.* (2007) were taken into consideration when designing the primers utilized in this study. These factors were; choosing a target size that would allow for amplification from degraded DNA extracted from faecal matter, choosing gene regions that excluded non-relevant species and allowed for the target species to be investigated, selection of gene regions that had sufficient reference information and lastly, selection of regions that would include the greatest amount of sequence variation (Bradley *et al.*, 2007).

Primer sets were designed to amplify seven target regions of the black rhinoceros mtDNA genome. These regions were the *COI* (two sets of primers were designed for this region), D-

loop, HV1, HV2, *cytb* and *16s rRNA*. The primer sets were designed to the following parameters: a nucleotide length of 18-30 bp, a cytosine and/or guanine content of 40-60% and a melting temperature (Tm) of 42℃-62℃ (Dieffenbach *et al.,* 1993). The melting temperature for each primer was determined using the following equation (Dieffenbach *et al.,* 1993):

$$Tm= 2(Adenine + Thymine) + 4(Cytosine\ and\ Guanine)$$

An online oligo analyser (http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/) was used to determine if there were secondary structures present in the designed primers. The designed primers can be seen in Table I.

A MiniOpticon™ System (Bio-Rad) was used for amplification of the KAPA Taq HiFi polymerase reactions. The initial temperature was 95℃ for 5 minutes, the heat denaturation temperature was 98℃ for 20 seconds, the extension temperature was 72℃ for 60 seconds and the final extension temperature was 72℃ for 5 minutes. An initial template concentration of approximately 150 ng for faecal samples and 60 ng for tissue samples was used (unless otherwise stated). Faecal sample amplification utilised 35 cycles for amplification and tissue samples utilized 20 (unless otherwise stated). A gradient PCR was performed on all primer sets in order to determine their optimal annealing temperature (Table I). All sequencing was conducted at Inqaba Biotec (RSA), using the appropriate primers form Table I.

3.7 Restriction digest of the *COIi* region

Amplified *COIi* PCR products were digested with *Hin*dIII. The single restriction digest used 10 U of the restriction enzyme and 2 µl of 10x reaction buffer. The PCR product (50 ng) was added to the mixture, which was brought to a final volume of 15 µl using ddH$_2$0. Reaction tubes were incubated at 37℃ for 6 hours. The template was replaced with ddH$_2$0 for the negative control. The samples were run on a 2% (w/v) TAE agarose gel (ethidium bromide 0.20 mg/mL) in 1x TAE buffer for 140 minutes. The gel was visualised using the ChemiDoc (Uvitec®), at a wavelength of 300 nm.

Table I: Primer sequences designed to amplify target regions of mtDNA for the black and white rhinoceros. This table includes the binding site position on the black rhinoceros mtDNA genome and the expected amplicon size for each region. The optimal annealing temperature was obtained by preforming a gradient PCR for each primer set.

	Forward (5'- 3')	Reverse (5'- 3')	Binding Position		Amplicon Size (bp)	Tm
			Forward	Reverse		
COIi	CATATGAGCCTTCTAATTCGCG	GTCGACAGAAGATGGTTAGGTC	5 450	5 795	347	62.0℃
COIii	GTCGACCACATATTTACTGTTGGC	CATATGATGCGAATGCTTCTC	6 221	6 782	562	60.0℃
D-loop	GTCGACCCCTGAACATTCCCCCTT	CATATGCAGTCTAGAAACCCC	15 451	15 897	447	60.0℃
HV1	GTCGACCTAAAATCGCCTATT	CATATGAGGCGAGTTTTGGGG	15 940	16 411	471	63.0℃
HV2	GTCGACGTTAATGTAGCT	CATATGAGCTGTTGCTTGTAG	1	580	580	63.0℃
16s rRNA	GTCGACATTTACCTAAAAGCC	CATATGGTTAGGATACCGCGG	1 498	2 028	531	62.0℃
Cytb	GTCGACTGACTAACATCCGTA	CATATGGTGGAAAGCGAAGAATCG	14 181	14 725	545	64.0℃

∗ Binding position: the position at which the primers bind to the black rhinoceros genome (white rhinoceros binding sites were similar to that of the black rhinoceros)
∗ Tm: optimal annealing temperature for designed primers

CHAPTER FOUR

RESULTS

4.1 The analysis of the mitochondrial genome for the black and white rhinoceros *in silico* to determine target regions of interest

The complete mitochondrial genomes for the black and white rhinoceros are available on the NCBI website (Accession numbers FJ905814.1 and Y07726.1, respectively). The black rhinoceros mtDNA genome was constructed using sample provided by the Zoological Museum, University of Copenhagen (Denmark) (Willerslev *et al.,* 2009). DNA extractions were performed on powdered nail obtained from a black rhinoceros specimen that was estimated to be 100 years old, however the geographic origin of the specimen is not available (Willerslev *et al.,* 2009). The Zoological department, University of Copenhagen (Denmark) provided striated muscle from a white rhinoceros specimen in order to construct the mtDNA genome (Xu & Arnason, 1997). The estimated age and geographic origin of this white rhinoceros specimen has not been published. The mitochondrial genomes for each species are represented schematically in Figure 4.

Seven regions of the mtDNA genome were chosen for investigation (Figure 4A). The first two regions came from the *COI* gene (*COIi* and *COIii*). The CBOL (Consortium for the Barcode of Life) aim to create a DNA barcode global reference library (Hanner, 2009). This library was to include databases of species, biogeographic information and specimens. The reasoning for a standardized barcoding region was to allow comparison between genetic studies as well as to ensure that collected data would have maximum usefulness for future studies. The CBOL barcode utilizes a 648 bp region of the *COI* gene, which is located at position 58-705 (relative the mitochondrial genome of the mouse) (Hanner, 2009). However, no relevant information relating to the rhinoceros has been released by the CBOL. *COIi* was chosen as a restriction site was present in the white rhinoceros reference sequence (Y07726.1) and absent in the black rhinoceros reference sequence (FJ905814.1). Therefor this site would allow for samples to be identified as either white or black rhinoceros. Literature is lacking with regards to intraspecies variation in mammalian *COI* genes, however studies conducted on non-mammalian species indicated that the *COI* gene showed potential for intraspecies variation (Li *et al.,* 2014). Therefore the *COIii* region was investigated to determine if this gene would allow for intraspecies variation.

The level of genetic variation within the D-loop of rhinoceros has been investigated (Anderson-Lederer *et al*., 2012; Brown & Houlden, 2001). Studies have been conducted on other species to determine if this region could show intraspecies variation (Belay & Mori, 2006; Glenn *et al*., 2002). This literature indicated that the D-loop had potential for inferring intraspecies variation for the black and white rhinoceros. Non-mammalian species indicated that the hypervariable regions showed potential for intraspecies variation (Aranishi & Okimoto, 2005), however there is a lack of literature for mammals. Primer sets were designed in order to investigate the level of intraspecies variation present in thee hypervariable regions. These findings would be beneficial for future mammalian studies with regards to using the hypervariable regions as a source of intraspecies variation.

The *cytb* gene is commonly used for species identification (Linacre *et al.,* 2009), however studies have shown intraspecies variation for this gene (Bradley & Baker, 2001). This literature indicated that the *cytb* region may have been a potential site of intraspecies variation. An investigation into intraspecies variation was conducted in a non-mammalian species for the *16s rRNA* gene (Marín *et al.,* 2013). This study indicated that this gene could be a potential site for intraspecies variation in mammalian species; therefore primers were designed to investigate this region in the rhinoceros.

The primers, with the exception of the *COIi* region, were initially designed for the black rhinoceros as sample was available for this species. White rhinoceros sample became available during the course of the investigation, therefore primer compatibility was examined. Five of the seven primer sets were compatible for the white rhinoceros (D-loop, HV2, *COIi*, *COIii*, and *cytb*) (Figure 4B).
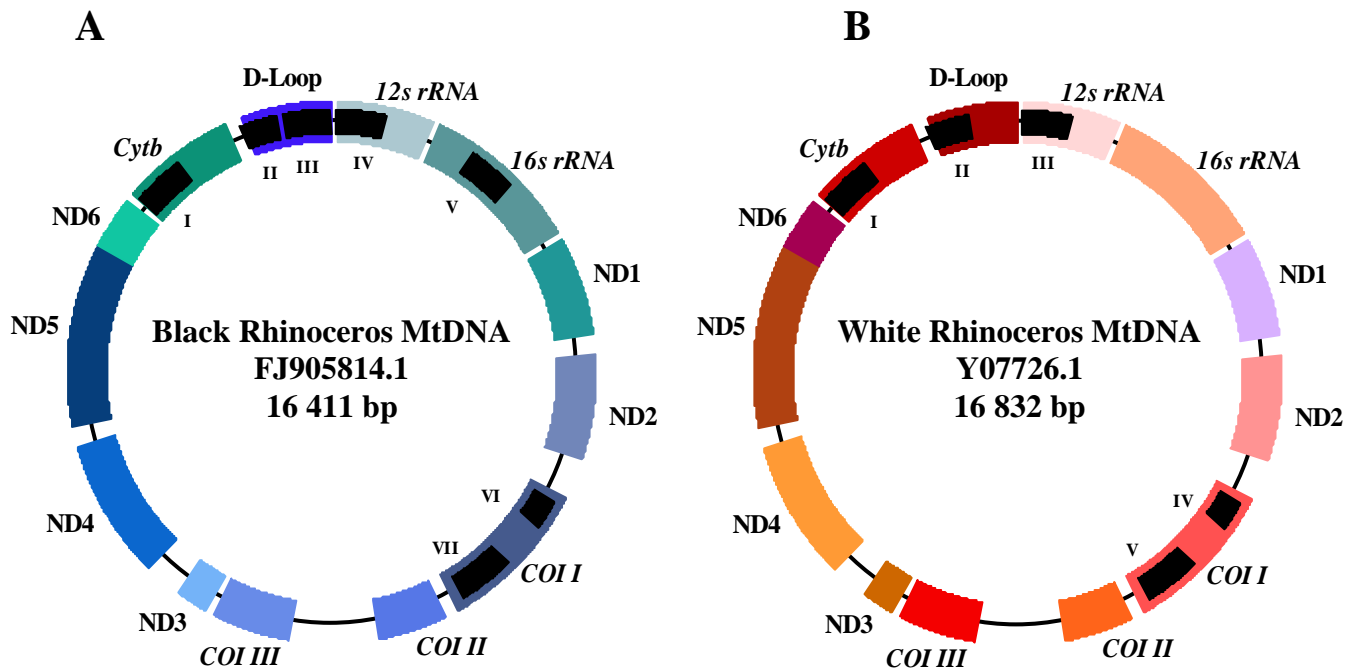
Figure 4: Schematic representation of the mitochondrial genomes for the black and white rhinoceros. This figure was adapted from Lincare & Tobe, 2009. (A) The black rhinoceros mitochondrial genome is 16 411 bp in size (Lincare & Tobe, 2009). Seven primer sets were designed to amplify six target regions. These regions are shaded black. The intended target regions are truncated as the primer sets were not designed to amplify the complete gene/region of interest. The reason for this is that smaller fragments of DNA have had greater amplification success when working with sample that has the potential of being degraded. I: *cytb*, II: D-loop, III: HV1, IV: HV2, V: *16s rRNA*, VI: *COIi* and VII: *COIii*. (B) The white rhinoceros mitochondrial genome is marginally larger than the black rhinoceros genome (Lincare & Tobe, 2009). Five of the seven primer sets were compatible for the white rhinoceros genome. These regions are highlighted black on the genome map. I: *cytb*, II: D-loop, III: HV2, IV: *COIi* and V: *COIii*.

According to Taberlet *et al.* (1999), faecal DNA may present problems with regards to molecular investigations as environmental factors can cause faecal DNA to be of low quality. The rationale behind this research was to develop and optimise molecular techniques in order to examine rhinoceros mtDNA. Therefore to ensure that the techniques produced reproducible results, faecal sample needed to be as fresh as possible. An invasive method for faecal sample collection was therefore used. This method of collection ensured that no insect contamination was present and that degradation caused by environmental factors was minimal. The collected samples remained in containers without fixation (*i.e.* 70% ethanol) for a minimum of one and a maximum of two days. Primers where therefore designed to amplify relatively short regions of DNA, with the maximum size being 580 bp and the minimum size being 349 bp as the source of DNA used in this study was likely to be degraded. This

collection method also ensured that faecal samples originated from different individuals and that tissue same could be linked to the same individual's faecal sample. The techniques developed in this research would be able to be applied to fresh, non-invasively, collected faecal samples.

4.2 <u>Isolation of DNA from rhinoceros faeces, tissue and blood</u>

The non-invasive source of DNA used for this study was faeces (F). This source of DNA is naturally abundant, easy to collect, cost effective, and no immobilisation of the animal is required. This source of DNA may be inefficient with regards to time as the collector would need to ensure the sample originated from the individual of interest. For this investigation, a positive control was required to ensure that the results obtained from the faecal DNA were reliable. Matched tissue (T) and blood (B) samples were investigated as control sources of DNA to determine which would be most suitable to serve as the positive control.

DNA extractions were performed on matched tissue, blood and faecal samples from black and white rhinoceroses. A Nanodrop2000 spectrophotometer was used to quantify the amount of DNA in each extraction. This was achieved by reading the individual samples absorbance at a wavelength of 260 nm. The average concentration of extracted DNA can be seen in Table II.

Table II: The average concentration of DNA extracted from matched blood, tissue and faecal samples for the black and white rhinoceros. Faecal sample was to serve as the non-invasive source of DNA and tissue/blood as the invasive source. The appropriate invasive source of DNA would be used at the positive control for all experimentation.

Sample Type	Blood		Tissue		Faeces	
Rhinoceros	black	white	black	white	black	white
	(N= 4)	(N= 4)	(N= 12)	(N= 8)	(N= 52)	(N= 30)
Average Concentration (ng)	31.51	26.85	66.81	33.48	29.33	43.64
Standard Deviation	5.12	11.41	36.64	16.26	16.71	19.74

∗ N: Number of independent extractions preformed on source of DNA
∗ Blood: One blood sample was provided for each species
∗ Tissue: Two tissue samples were provided for each species
∗ Faeces: Eight faecal samples were provided for the black and seven for the white rhinoceros

DNA extracted from black and white rhinoceros tissue samples had the highest concentration (66.81 ng and 33.48 ng respectively). DNA extracted from blood samples (31.51 ng and 26.85 respectively) had concentrations that were comparable to the extractions performed on faecal samples (29.33 ng and 43.64 respectively) (Table II).

Matched sources (blood, tissue and faecal sample that originated from the same individual) of black and white rhinoceros DNA were analysed by agarose gel electrophoresis in order to compare the quality of the extracted products (Figure 5). For the black rhinoceros, a concentration of 62 ng for the tissue samples, 59 ng for the blood samples and 50 ng for the faecal samples were loaded onto the gel (Figure 2A). For the white rhinoceros, a concentration of 58 ng for the tissue samples, 58 ng for the blood samples and 56 ng for the faecal samples were loaded (Figure 2B).
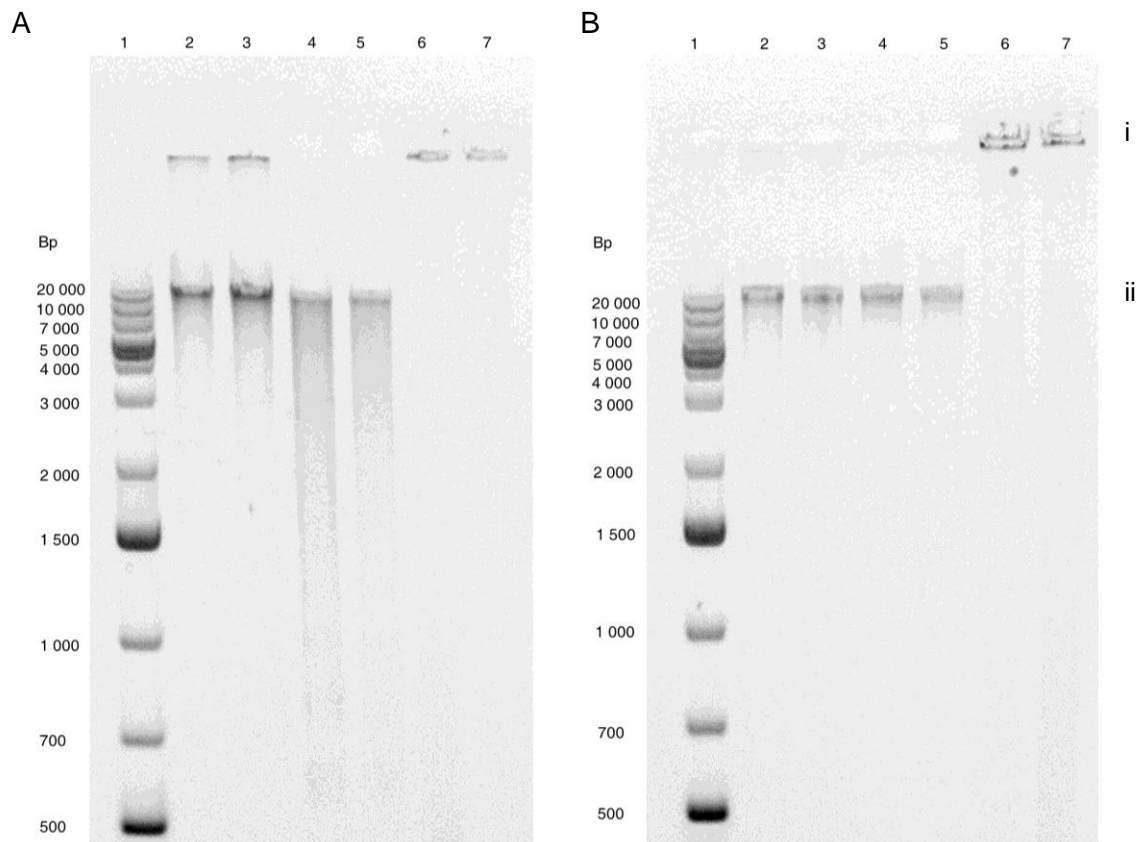
Figure 5: Duplicate DNA extractions for black and white rhinoceros blood, tissue and faecal samples. (A) DNA extractions performed on three sources of samples from the black rhinoceros. (B) DNA extractions performed on three sources of samples from the white rhinoceros. For both A and B: Lane 1: O'Gene Ruler 1kb Plus DNA Ladder, Lane 2: Tissue, Lane 3: Tissue, Lane 4: Blood, Lane 5: Blood, Lane 6: Faeces, Lane 7: Faeces. For both A and B: 1% (w/v) agarose gel electrophoresis was used. DNA that did not migrate past the well (i) and DNA extracted as expected (ii)

Defined bands (ii) of a high molecular weight were seen for DNA extracted from black and white rhinoceros blood and tissue samples in lanes 2-5 of Figure 5 (A and B). DNA was present for faecal sample extractions; however the samples were hindered in the wells of the

agarose gel (i). The smears observed in Figure 5A were most likely an indication of degraded DNA (Roux, 1995). Tissue sample DNA was used as the positive control for subsequent experimentation due to it having a higher average concentration than the DNA extracted from the blood sample.

4.3 <u>The development of a molecular technique to differentiate between the black and white rhinoceros using the *COI* gene</u>

4.3.1   Optimisation of PCR amplification for the COIi region

The black and white rhinoceros *COI* gene (FJ905814.1 and Y07726.1, respectively) was aligned, in order to identity a restriction enzyme in one of the species and absent in the other. A restriction site was identified for the *Hin*dIII enzyme in the white rhinoceros sequence and this sequence was absent for the black rhinoceros. A primer set was designed to amplify a 349 bp region of the *COI* gene, which included the restriction site. This region was termed *COIi*. The melting temperatures for the forward and reverse primer were calculated to be 64.0℃ and 66.0℃. A gradient PCR was conducted on three faecal samples to determine the optimal annealing temperature for amplification of the *COIi* region (Figure 6A).
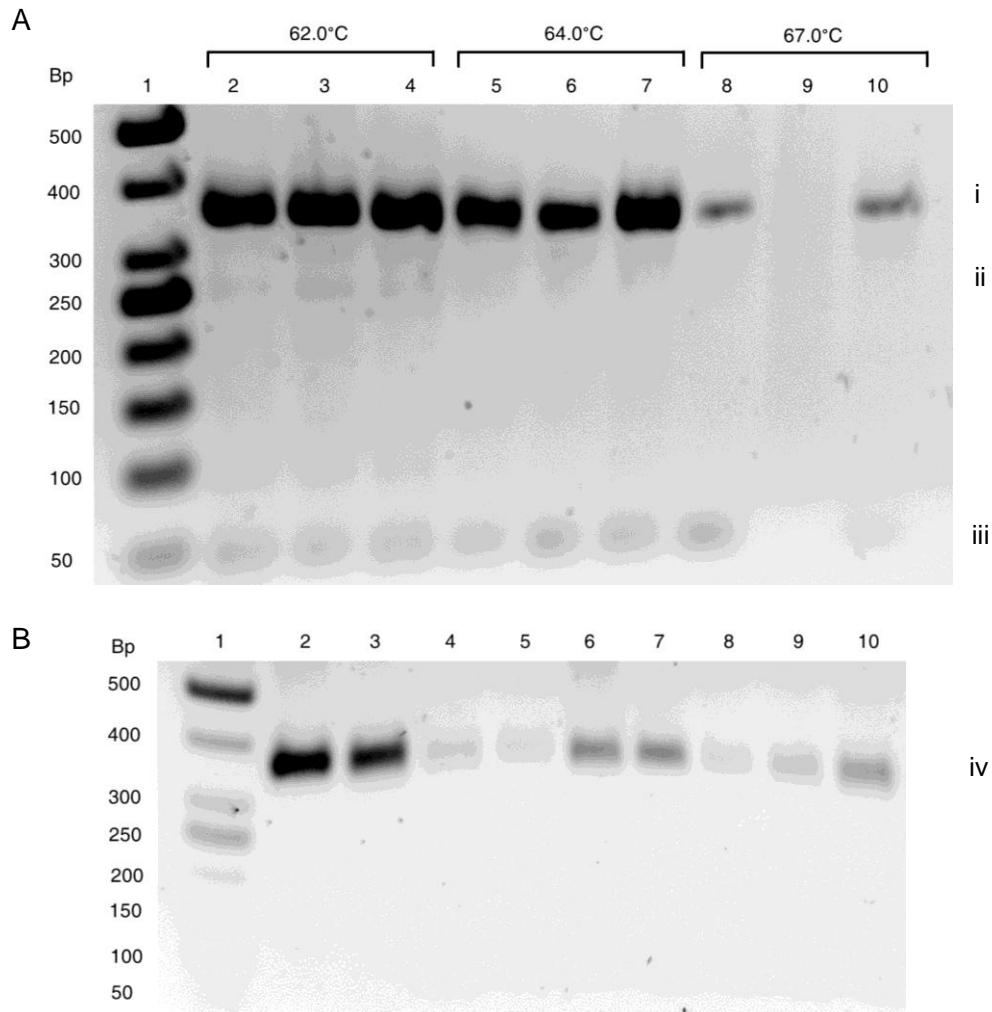
Figure 6: Amplified *COIi* region for the black rhinoceros. (A) A gradient PCR for the *COIi* region to determine the optimal annealing temperature for amplification. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, F), Lane 3: white (1, F), Lane 4: white (2, F), Lane 5: black (1, F), Lane 6: white (1, F), Lane 7: white (2, F), Lane 8: black (1, F), Lane 9: white (1, F), Lane 10: white (2, F). (B) The optimized PCR products for the black rhinoceros *COIi* region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (2, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (4, F), Lane 7: black (5, F), Lane 8: black (6, F), Lane 9: black (7, F), Lane 10: black (8, F). For both A and B: Agarose (1% w/v) gel electrophoresis was used. Target bands to determine optimal annealing temperature (i), non-specific bands (ii), primer dimer bands (iii) and optimised PCR product bands for the *COIi* region (iv).

The annealing temperatures 62.0℃, 64.0℃ and 67.0℃ were investigated in Figure 6A. Primer dimers of approximately 50 bp were observed (iii). Non-specific bands can also be seen in lane 3 and 4 (ii). This may have resulted from too much initial template being used in the PCR reaction, or too many cycles being performed in the amplification procedure (Roux, 1995). The annealing temperature of 62.0℃ (lane 2, 3 and 4) was selected for subsequent amplification due to the consistent size and density of the bands (i) at the expected size of the

34

predicted amplicon. To remove the non-specific bands and primer dimers the cycle number was decreased from 30 to 28. The optimised *COIi* PCR products for the black rhinoceros were represented in Figure 6B.

TotalLab (Newcastle, UK) was used to determine the size of the bands present in Figure 6B (iv) relative to the molecular weight markers. The average band size was 348.45 bp (SD±4.79), which was similar to the size of 349 bp predicted by the *in silico* analysis. Non-specific bands and primer dimers were successfully removed in Figure 6B. The density of the bands at position (iv) were thought to differ due to an inconsistent initial template concentration being used. Due to the presence of the bands (iv), and the size of the bands being correct, the PCR reaction was considered optimized. The optimized protocol was performed on white rhinoceros tissue and faecal samples (data not shown). Black and white rhinoceros *COIi* products were sequenced by Inqaba (Biotech, SA).

4.3.2 Distinguishing between the black and white rhinoceros using a single restriction digest

A restriction digest, using the *Hin*dIII enzyme, was performed on several of the *COIi* PCR products for the black and white rhinoceros. This was done to determine if the two species could be experimentally separated with a single restriction digest (Figure 7).
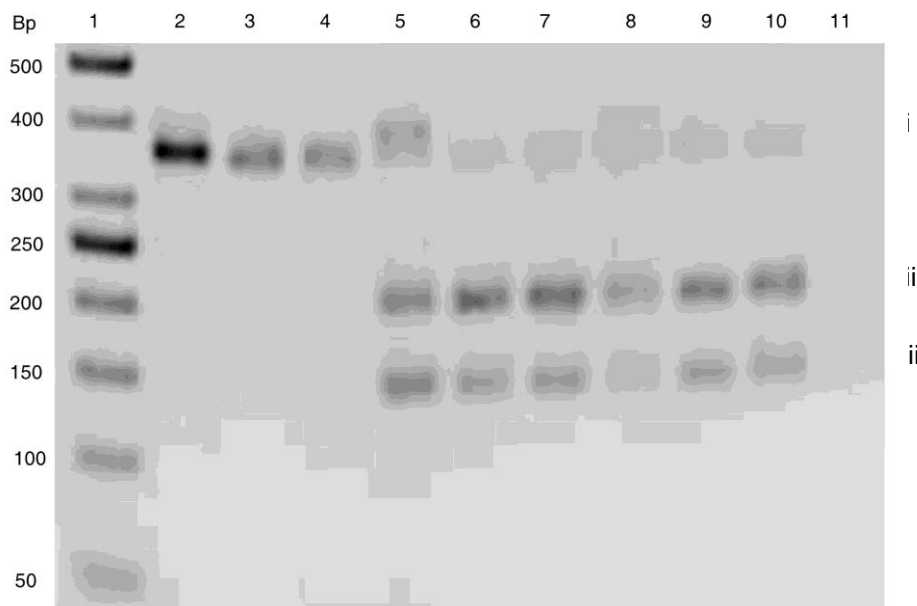
Figure 7: A restriction digestion of the black and white rhinoceros *COIi* region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (2, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: white (5, T), Lane 6: white (2, F), Lane 7: white (3, F), Lane 8: white (4, F), Lane 9: white (5, F), Lane 10: white (6, F), Lane 11: Negative Control. Agarose (2% w/v) gel electrophoresis was used. Undigested DNA for the *COIi* region (i), the larger band of digested DNA (ii) and the smaller band for the digested *COIi* DNA (iii).

Figure 7 represented a clear separation between the black (one band present at position i, lane 2-4) and white (multiple bands present at i, ii and iii, lane 5-10) rhinoceros. TotalLab (Newcastle, UK) was used to determine the size of the respective bands (i, ii and iii) relative to the molecular weight markers. The black rhinoceros bands had an average size of 352.05 bp (SD±5.07). This was expected as the black rhinoceros mtDNA sequence did not contain the restriction enzyme, and the *in silico* size of the region amplified was 349 bp. The bands present for the white rhinoceros in Figure 7 were calculated to have an average size of 367.04 (SD±12.56), 207.24 bp (SD±4.43) and 147.21 bp (SD±5.17), respectively. An *in silico* digestion was performed on the white rhinoceros *COIi* region, which produced bands of 201 bp and 146 bp. The bands present for the white rhinoceros in Figure 7 correlate with the *in silico* sizes, with the exception of the first band (367.04 bp at position i). This may have resulted due to an insufficient digestion time or a concentration that was too low for the restriction enzyme. No bands where present in the negative control indicating no contamination in KAPA HiFi hot start ready mix, the *Hin*dIII enzyme or the ddH$_2$0.

In order to confirm the experimental results, which indicated that the black and white rhinoceros could be separated with a single restriction digest, the sequences obtained from Inqaba (Biotech, SA) were aligned with two reference sequences (FJ905814.1 and Y07726.1), which was shown in Figure 8.
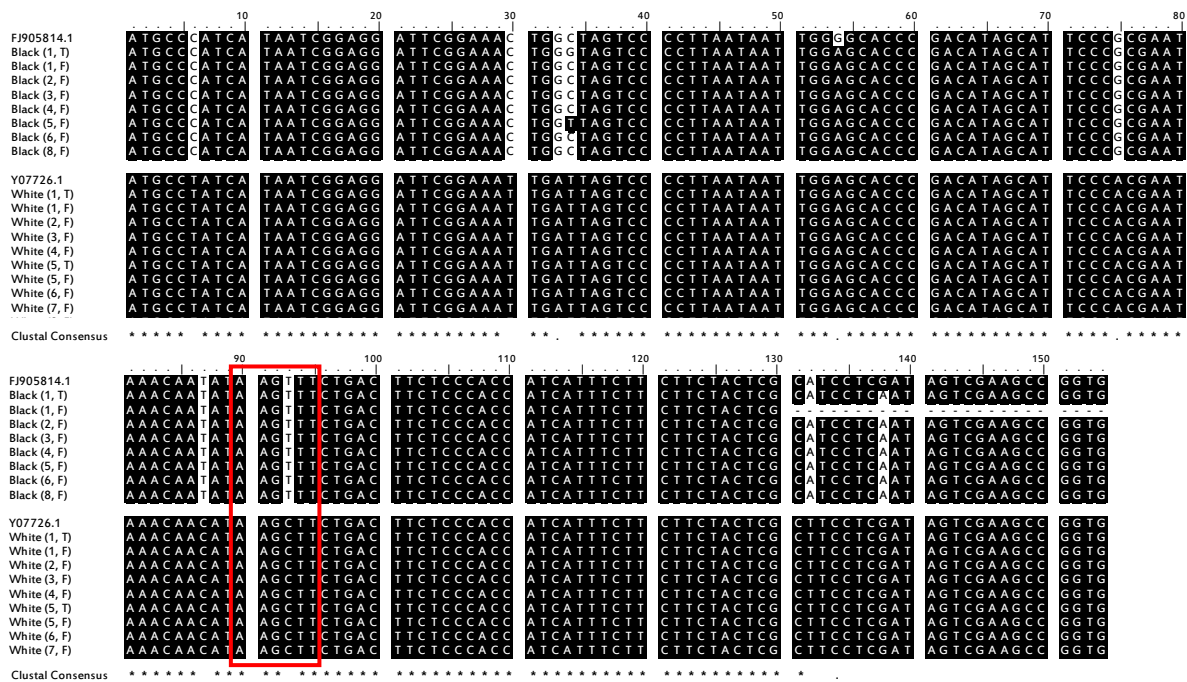
Figure 8: A multiple sequence alignment for the obtained black and white rhinoceros *COIi* sequences. Sequences obtained from one tissue and seven faecal samples were used for the black rhinoceros sequence alignment. Sequences obtained from two tissue and seven faecal samples were used for the white rhinoceros. The black and white rhinoceros sequence alignments included a reference sequence for each species (FJ905814.1 and Y07726.1 respectively). The *Hin*dIII restriction site was highlighted by the red box.

A total of 154 bp were used for the multiple alignments in Figure 8. This was due to sequence quality being lost at the 5' and 3' ends of the isolated DNA sample sequences. The black rhinoceros sequences resulted in a BLAST identity of 99% to the *Diceros bicornis* complete mitochondrial DNA sequence (FJ905814.1) with a maximum score of 523, a query cover of 97% and an E-value of $4e^{-145}$. The isolated black rhinoceros *COIi* sequences were highly conserved, with one nucleotide variation for black (5, F) at position 34. Sequence variation was observed between the reference sequence (FJ905814.1) and collected sample sequences at position 54 and 138. The white rhinoceros sequences resulted in a BLAST identity of 100% to the *Ceratotherium simum* complete mitochondrial DNA sequence (Y07726.1) with a maximum score of 499, a query cover of 100% and an E-value of $6e^{-138}$. No sequence variation was observed between the white rhinoceros reference sequence (Y07726.1) and isolated DNA sample sequences. Variation between the black and white rhinoceros shown occurred at positions 6, 30, 33, 34, 75, 87, 93, 132 and 138 (Figure 8). The sequence for the restriction enzyme, *Hin*dIII (AAGCTT) is at positions 89-95 for the white rhinoceros (red box). However as there is nucleotide difference at position 93, the black rhinoceros sequences

do not contain the restriction site. This confirms the experimentally obtained results from the restriction analysis.

4.4 Development of a PCR based technique to examine genetic variation in target regions found on the mitochondrial DNA genome for the black and white rhinoceros

4.4.1 Optimisation of PCR amplification for the COIii region

Primers were designed to amplify a second region of the *COI* (6222-6582) which was termed *COIii*. A total of 195 bp were not viable for sequence alignment for the *COIi* region as sequence quality was lost at the 5' and 3' ends of the isolated DNA sample sequences, therefore a primer set was designed to amplify a 562 bp region for the *COIii* region. The forward and reverse primers had melting temperatures of 70.0℃ and of 58.0℃, respectively. A gradient PCR with annealing temperatures of 60.0℃, 62.0℃ and 64.0℃, was performed to determine the optimal annealing temperature for amplification (Figure 9). Faecal sample PCR reactions included 35 cycles and tissue samples had a cycle number of 25.

Figure 9: Amplified *COIii* region for the black rhinoceros. (A) A gradient PCR for the *COIii* region to determine the optimal annealing temperature for amplification. Lane 1: O'Gene Ruler 1 kb Plus DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (1, T), Lane 7: black (1, F), Lane 8: black (2, T), Lane 9: black (3, F), Lane 10: black (1, T), Lane 11: black (1, F), Lane 12: black (2, F), Lane 13: black (3, F). (B) Optimized PCR products for the black rhinoceros *COIii* region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (4, F), Lane 7: black (5, F), Lane 8: black (6, F), Lane 9: black (7, F). For both A and B: Agarose (1% w/v) gel electrophoresis was used. Target bands to determine optimal annealing temperature (i), primer dimer bands (ii), target amplicon bands (iii) and primer dimer bands (iv).

Primer dimer bands (ii) were present for all faecal *COIii* PCR products (ii) in Figure 9A. No other non-specific bands were observed. The annealing temperature of 60.0℃ was chosen for subsequent amplification due to the primer dimer bands being the least dense for this temperature. The optimised PCR protocol was performed on all samples. The optimised *COIii* PCR products for the black rhinoceros are shown in Figure 9B. Although primer dimers (iv) were observed in Figure 9B, no other non-specific bands were present. TotalLab (Newcastle, UK) determined that the bands (iii) in Figure 9B had an average size of 560 bp (SD±6.12), which is similar to the *in silico* size of 562 bp. The difference in density of the bands (Figure 9B, iv) were as a result of the initial template concentration not being consistent. The optimised PCR protocol for the *COIii* region was performed on white rhinoceros sample (data not shown). The black and white rhinoceros *COIii* PCR products were sequenced by Inqaba (Biotech, SA).

A multiple alignment for the *COIii* region, which included a reference sequence for the black rhinoceros (FJ905814.1) and a reference sequence for the white rhinoceros (Y07726.1), was performed with the sequences obtained for the collected samples to determine if any genetic differences occurred between individuals (Figure 10). Bioinformatics programme BioEdit (Hall, 1999) was used to create sequence identity matrixes for all sequence alignments preformed in this investigation. These matrixes can be observed in the appendix (Table AI-AXII).

Figure 10: A multiple sequence alignment for the obtained black and white rhinoceros *COIii* sequences. The sequences included in the black rhinoceros sequence alignment included a reference sequence (FJ905814.1) and sequences obtained from one tissue and eight faecal samples. The white rhinoceros alignment included a reference sequence (Y07726.1) and sequences obtained from one tissue and six faecal samples. Conserved sequence variation is observed between the two species of rhinoceros. The sequence identity matrix for the *COIii* region can be seen in Table AIII for the black rhinoceros and Table AIV for the white rhinoceros.

The black rhinoceros *COIii* sequences resulted in a BLAST identity of 99% to the *Diceros bicornis* mitochondrion complete mitochondrial DNA sequence (FJ905814.1) with a maximum score of 948, a query cover of 98% and an E-value of 0.0. Sequences for the black rhinoceros samples were 100% conserved (sequence identity of 1.000). The reference sequence and the obtained black rhinoceros sequences had a sequence identity of 0.993 (variation at positions 20, 280 and 446), which indicates a sequence difference of 0.7%. The white rhinoceros sequences resulted in a BLAST identity of 100% to the *Ceratotherium simum* complete mitochondrial DNA sequence (Y07726.1) with a maximum score of 937, a query cover of 100% and an E-value of 0.0. Sequences for the white rhinoceros collected samples were 100% conserved (sequence identity of 1.000), and no variation occurred between the collected sample sequences and the reference sequence (sequence identity of 1.000). Variation was observed (Figure 12) between the black and white rhinoceros sequences at 26 different positions (sequence difference of 5.8%), namely 6, 20, 41, 62, 80, 83, 86, 95, 101, 104, 106, 122, 146, 167, 170, 251, 278, 299, 319, 335, 344, 359, 365, 437, 452 and 467. A total of 92 bp were lost due to sequence quality being lost at the 5' and 3' ends of the isolated DNA sample sequences.

4.4.2   *Optimisation of PCR amplification for a region of the D-Loop*

A study conducted by Brown & Houlden (2000) examined sequence variation in the D-loop (control region) for two subspecies of the black rhinoceros (*D. b. michaeli* and *D. b. minor*). The selected region was able to identity several different mtDNA haplotypes. Therefore primers were designed to amplify a truncated version of the D-loop region investigated by Brown & Houlden (2000). Primers were designed to amplify a 447 bp region of the D-loop. The forward and reverse primers had melting temperatures of 76.0℃ and of 62.0℃, respectively. A gradient PCR was conducted to determine the optimal annealing temperature for amplification. Figure 11A shows the gradient PCR, which investigated the annealing temperatures of 60.0℃, 62.5℃ and 67.5℃.
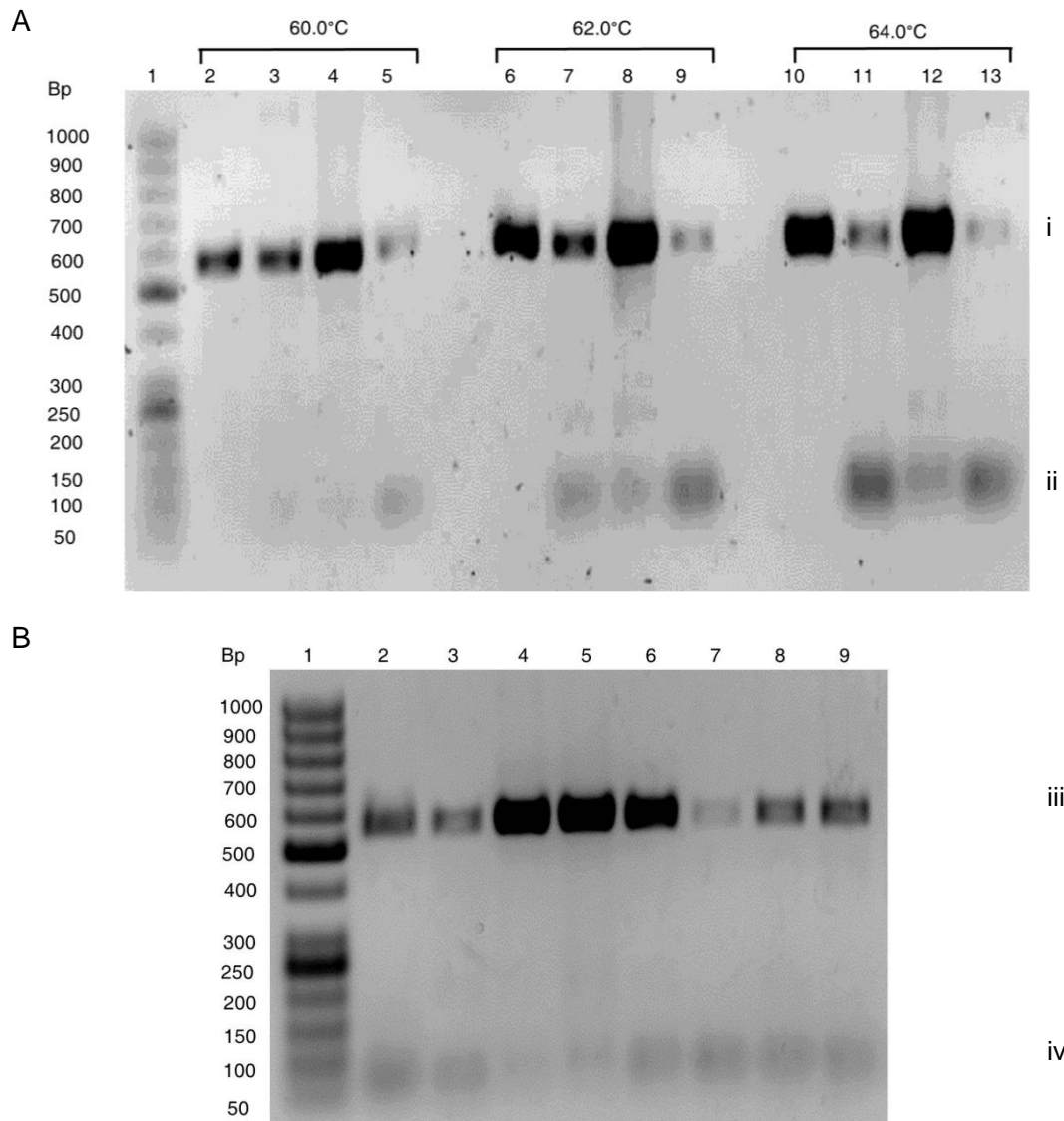
Figure 11: Amplified D-loop region for the black rhinoceros. (A) A gradient PCR for the D-Loop region to determine the optimal annealing temperature for amplification. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (1, T), Lane 6: black (1, F), Lane 7: black (2, F), Lane 8: black (1, T), Lane 9: black (1, F), Lane 10: black (2, F). (B) The optimized PCR products for the black rhinoceros D-loop region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, T), Lane 3: black c(1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (4, F), Lane 7: black (5, F). For both A and B: Agarose (1% w/v) gel electrophoresis was used. Non-specific bands (i, ii, iii, iv and vi), target bands to determine optimal annealing temperature (v) and optimised PCR product bands for the D-loop region (vii).

An annealing temperature of 60.0℃ was chosen for amplification as the three different samples were amplified. Non-specific bands were observed in lane 2 and 5 of Figure 11A (i, ii, iii, iv and vi). These lanes represent amplification of the tissue sample at annealing temperatures of 60.0℃ and 62.5℃, respectively. These non-specific bands may have been due to the initial concentration of sample being too high or too many cycles in the amplification process (Roux, 1995). Altering the initial template concentration was investigated and this is shown in Appendix 1 (page 83). The cycle number for amplification was also investigated and this is shown in Appendix 2 (page 84). A concentration of 130 ng and 30 cycles was optimal for faecal DNA amplification, and a concentration of 50 ng and 25 cycles was optimal for tissue DNA amplification. The optimised PCR products for the black rhinoceros are shown in Figure 11B.

TotalLab (Newcastle, UK) was used to determine the size of the bands present in Figure 11B (vii). The average size of the bands was 458 bp (SD±11.62), which was similar to the size predicted *in silico* of 447 bp. The optimised PCR protocol for the D-loop region was performed on white rhinoceros samples (data not shown). The optimized black and white rhinoceros D-Loop PCR products were sequenced by Inqaba (Biotech, SA).

A multiple alignment, which included a reference sequence for the black (FJ905814.1) and white rhinoceros (Y07726.1), was performed with the sequences obtained for the collected samples in order to determine if any genetic differences occurred in the D-loop of these individuals (Figure 12).

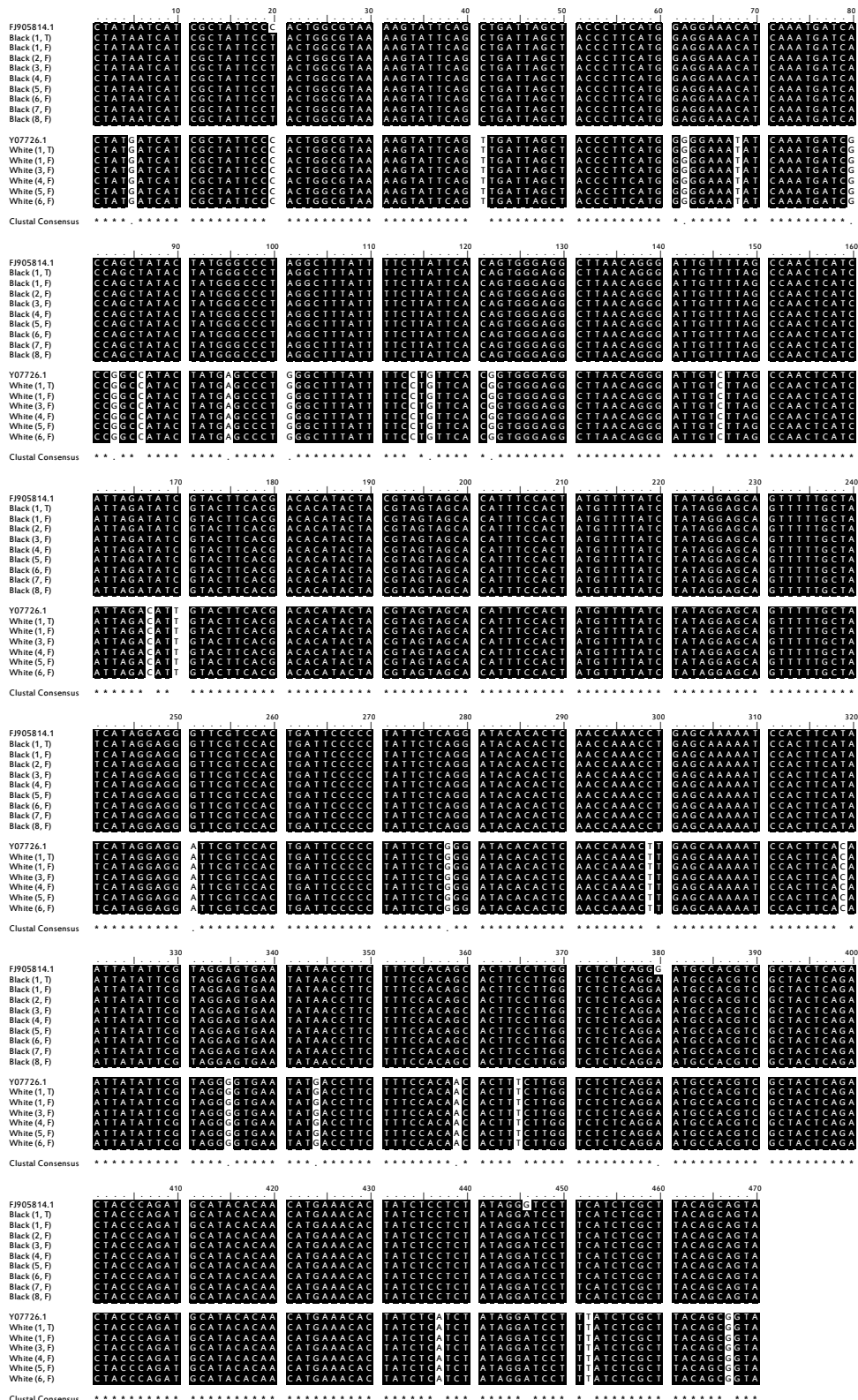**Figure 12: A multiple sequence alignment for the obtained black and white rhinoceros D-Loop sequences.** The black rhinoceros sequence alignment included sequences obtained from one tissue and seven faecal samples as well as a reference sequence (FJ905814.1). A reference sequence (Y07726.1) and sequences obtained from five faecal samples were used for the white rhinoceros sequence alignment. The sequence identity matrix for the D-loop can be seen in Table AV for the black rhinoceros and Table AVI for the white rhinoceros.

A total of 237 bp were lost due to sequence quality being lost at the 5' and 3' ends of the isolated DNA sample sequences. The black rhinoceros D-loop sequences resulted in a BLAST identity of 96% to the *Diceros bicornis* complete mitochondrial DNA sequence (FJ905814.1) with a maximum score of 815, a query cover of 97% and an E-value of 0.0. The white rhinoceros sequences resulted in a BLAST identity of 99% to the *Diceros bicornis* mitochondrial D-loop partial sequence (AF187825.1), with a maximum score of 695, a query cover of 97% and an E-value of 0.0. The black rhinoceros reference sequence and isolated DNA sample sequences has a sequence identity of 0.945 (variation at positions 18, 21, 26, 31,

33, 45, 126, 129, 130, 138, 139, 140, 161, 166, 185, 262 and 265), while the white rhinoceros had a sequence identity of 0.845 (variation at positions 7, 9, 13, 17, 18, 23, 27, 29, 33, 34, 35, 44, 55, 66, 75, 94, 96, 97, 99, 102, 105, 112, 116, 119, 123, 126, 129, 135, 139, 141, 146, 148, 149, 150, 152, 154, 157, 160, 166, 171, 184, 185, 196, 211, 224, 262, 265 and 299). This indicates a sequence difference between the isolated DNA sample sequences and the reference sequence of 5.5% for the black and 15.5% for the white rhinoceros. The isolated DNA sample sequences for the black and white rhinoceros both had sequence identities of 1.000 as no variation was observed.

4.4.3   *Optimisation of PCR amplification for the HV1 and HV2 regions*

The hyervariable regions (HV1 and HV2) were investigated to identify possible sequence variation. Primers were designed to amplify a 471 bp region of the HV1. The forward and reverse primers had melting temperatures of 60.0℃ and of 64.0℃, respectively. A gradient PCR, with the annealing temperatures of 59.0℃, 63.0℃ and 64.0℃, was performed to determine the optimal annealing temperature for amplification (Figure 13A). Amplified faecal DNA samples had a cycle number of 35 and amplified tissue DNA samples had a cycle number of 25.
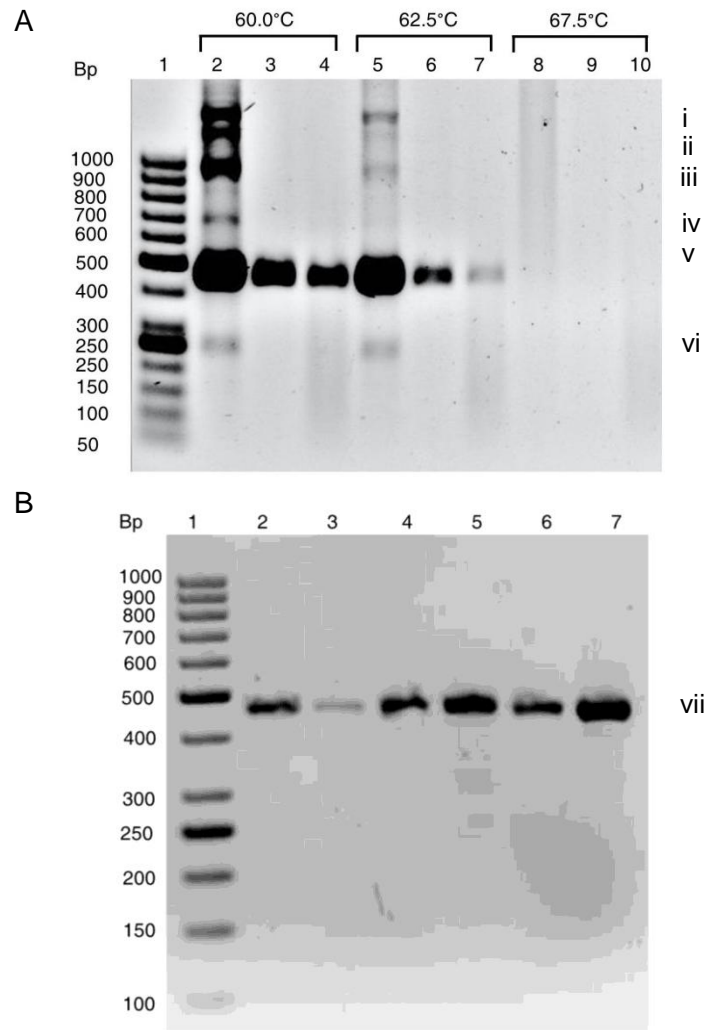
Figure 13: Amplified HV1 region for the black rhinoceros. (A) A gradient PCR for the HV1 region to determine the optimal annealing temperature for amplification. Lane 1: O'Gene Ruler 1 kb Plus DNA Ladder, Lane 2: black (1, F), Lane 3: black (2, F), Lane 4: black (1, T), Lane 5: black (1, F), Lane 6: black (2, F), Lane 7: black (1, T), Lane 8: black (1, F), Lane 9: black (2, F), Lane 10: black (1, T). (B) Optimized PCR products for the black rhinoceros HV1 region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (4, F), Lane 7: black (5, F), Lane 8: black (6, F), Lane 9: black (7, F), Lane 10: black (8, F). For both A and B: Agarose (1% w/v) gel electrophoresis was used. Non-specific bands (i, ii, iii, and v), target bands to determine optimal annealing temperature (iv), primer dimer bands (vi) and optimised PCR product bands for the HV1 region (vii).

Non-specific bands were present in Figure 13A in lanes 4, 6, 8 and 9. Primer dimers where present for all of the samples. The annealing temperature of 63.0℃ was chosen for subsequent amplification as at this temperature three samples were amplified with the least non-specific bands. In order to remove the non-specific bands the template concentration for faecal samples was lowered from 130 ng to 120 ng. The optimised PCR amplification products are shown in Figure 13B.

TotalLab (Newcastle, UK) was used to determine the size of the bands present in Figure 13B. The average size of the bands was 810 bp (SD±8.32); this is substantially larger than *in silico* predicted size of 471 bp. These samples for the black rhinoceros were sequenced by Inqaba (Biotech, SA). A multiple alignment, which included a reference sequence for the black (FJ905814.1) rhinoceros, was performed with the sequences obtained for the HV1 collected samples to determine if any genetic differences occurred between individuals (Figure 14).
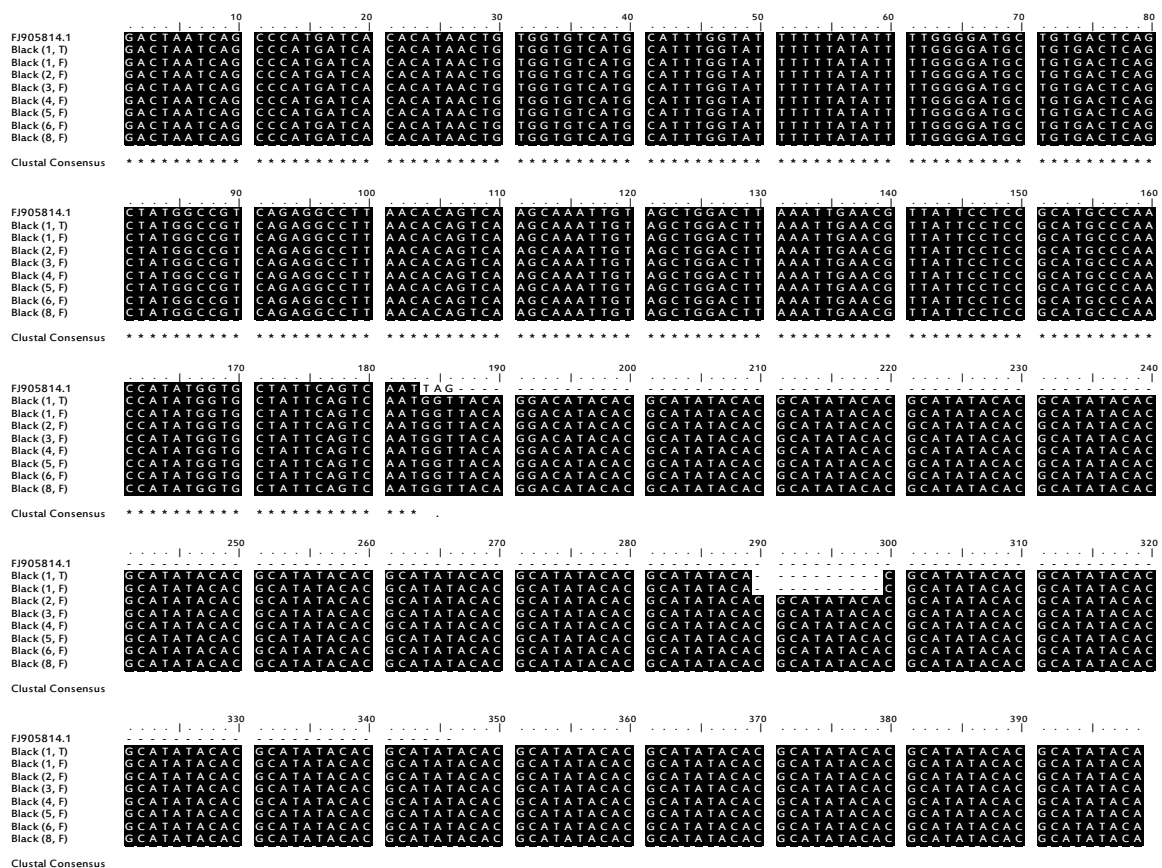
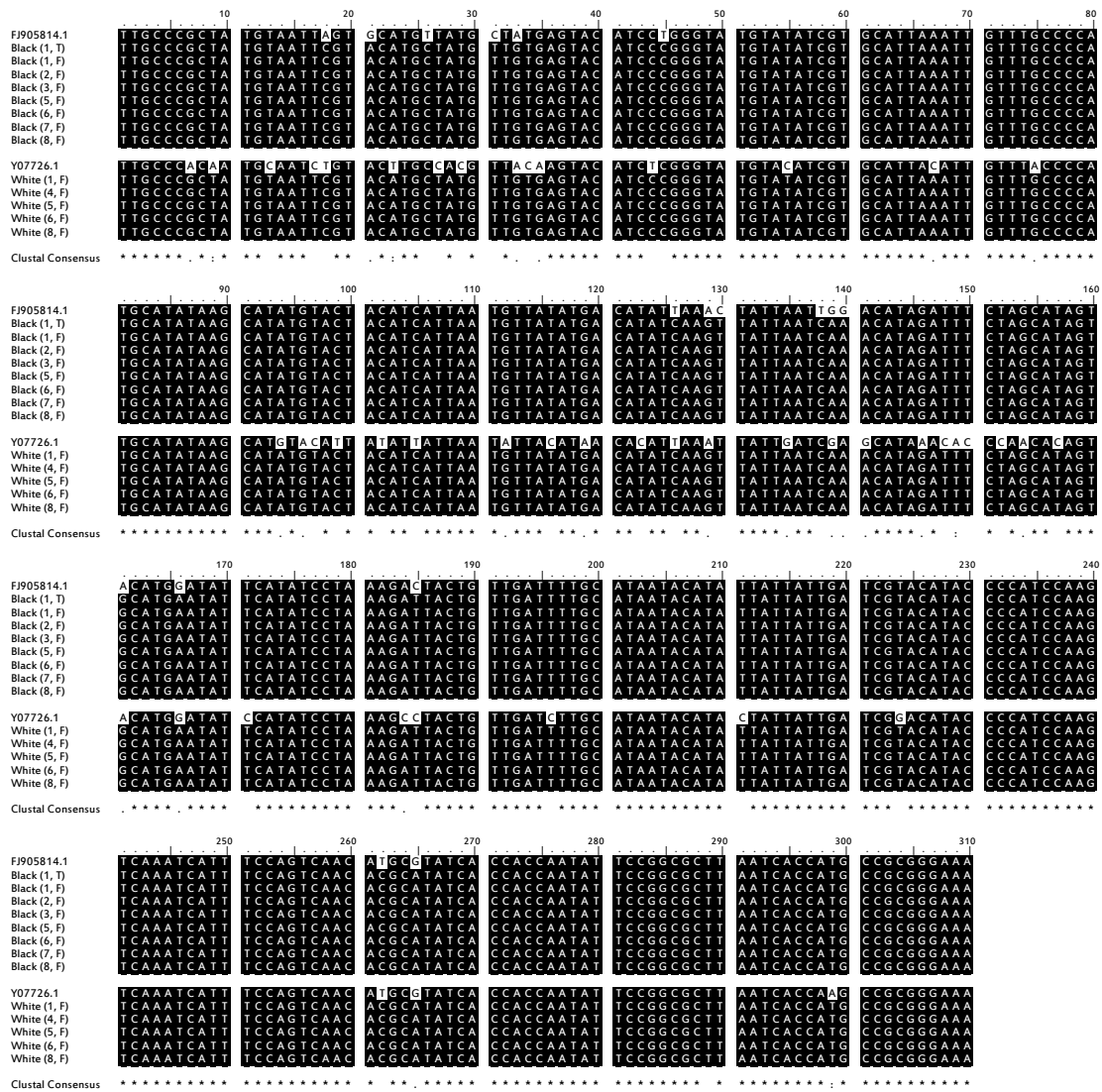**Figure 14: A multiple sequence alignment for the black rhinoceros HV1 sequences.** The sequence alignment included sequences obtained from one tissue and seven faecal samples as well as a reference sequence (FJ905814.1). The primers designed to amplify this region were not compatible for the white rhinoceros. The sequence identity matrix can be seen in Table AVII for the black rhinoceros.

The black rhinoceros sequences resulted in a BLAST identity of 100% to the *Diceros bicornis* complete mitochondrial DNA sequence (FJ905814.1) with a maximum score of 368, a query cover of 39% and an E-value of $4e^{-98}$. The isolated DNA sample sequences, except for Black (1, T) and Black (1, F), had a sequence similarity of 1.000 as no variation amongst these samples was observed in Figure 14. Isolated DNA sample sequences for Black (1, T) and Black (1, F) had a sequence similarity of 0.974 (sequence difference of 2.6%) as there was a deletion of 10 bp from position 289 bp. The reference sequence and isolated DNA sample sequences had a sequence similarity of 0.458 (with the exception of Black [1, T] and Black [1, F] which has a sequence identity of 0.470) as a deletion of 216 bp was observed (position 184-400 bp) in Figure 14. The elongated sequence supports the size of the bands present in Figure 13B. The sequence from 184-400 bp was examined by BLAST; however this resulted in no significant similarity matches being found. A total of 171 bp were lost as a result of poor sequence quality at the 5' and 3' ends.

Primers were designed to amplify a 580 bp region of the HV2. The forward and reverse primers had melting temperatures of 52.0℃ and of 60.0℃ respectively. A gradient PCR, with the annealing temperatures of 53.0℃, 55.0℃ and 57.0℃, was performed to determine the optimal annealing temperature for amplification (Figure 15A). Faecal sample PCR reactions included 35 cycles and tissue samples had a cycle number of 25.
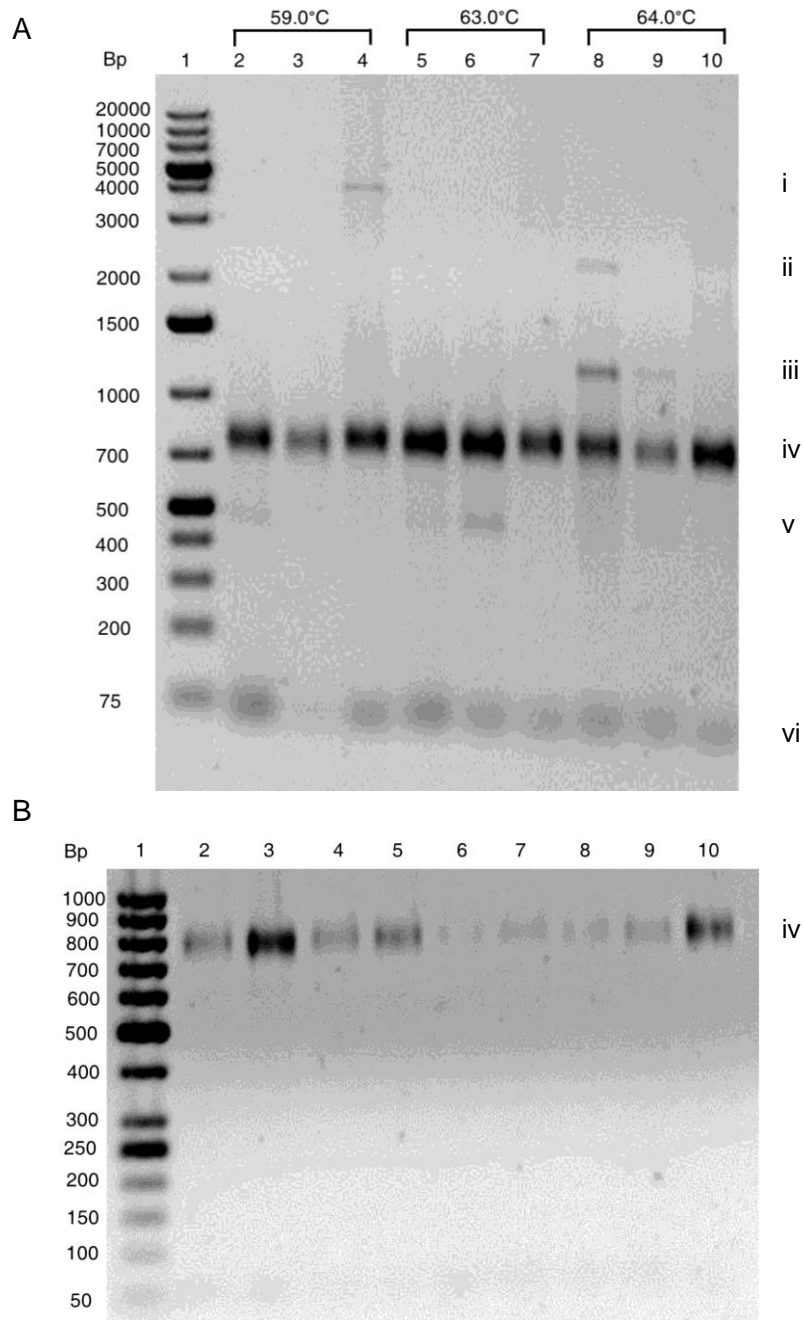
Figure 15: Amplified HV2 region for the black rhinoceros. (A) A gradient PCR for the HV2 region to determine the optimal annealing temperature for amplification. Lane 1: O'Gene Ruler 1 kb Plus DNA Ladder, Lane 2: black (1, F), Lane 3: black (2, F), Lane 4: black (1, F), Lane 5: black (2, F), Lane 6: black (1, F), Lane 7: black (2, F). (B) Optimized PCR products for the black rhinoceros HV2 region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, F), Lane 3: black (2, F), Lane 4: black (3, F), Lane 5: black (4, F), Lane 6: black (5, F), Lane 7: black (6, F), Lane 8: black (7, F), Lane 9: black (8, F), Lane 10: black (1, T). Agarose (1% w/v) gel electrophoresis was used. Target bands to determine optimal annealing temperature (i), non-specific bands (ii, iii, iv, v, vi, vii, viii, ix, x and xi) and optimised PCR product bands for the HV2 region (xii).

Non-specific bands were present in lanes 3 and 4 in Figure 15A. The annealing temperature of 53.0°C was chosen for subsequent amplification as both samples were amplified. The final PCR products for the black rhinoceros HV2 region were shown in Figure 15B.

TotalLab (Newcastle, UK) was used to determine the size of the target bands (ix) present in Figure 15B. The average size of the bands was 605 bp (SD±7.13), which correlates with the *in silico* size 580 bp. An unknown band (ix) was present for all faecal samples in Figure 15B. This band was absent from the positive control (tissue sample). The final PCR protocol for the HV2 region was performed on white rhinoceros samples. The final PCR product for the HV2 region was then sequenced for both the black and white rhinoceros.

A multiple alignment, which included a reference sequence for the black (FJ905814.1) and white (Y07726.1) rhinoceros, was performed on the sequences obtained for the target amplicon (Figure 15B, xii) for the HV2 region (Figure 16). A sequence alignment for the unknown band (Figure 15B, ix) was shown Appendix 3 (page 85). No significant similarity was found on the NCBI database when the unknown band was BLAST.
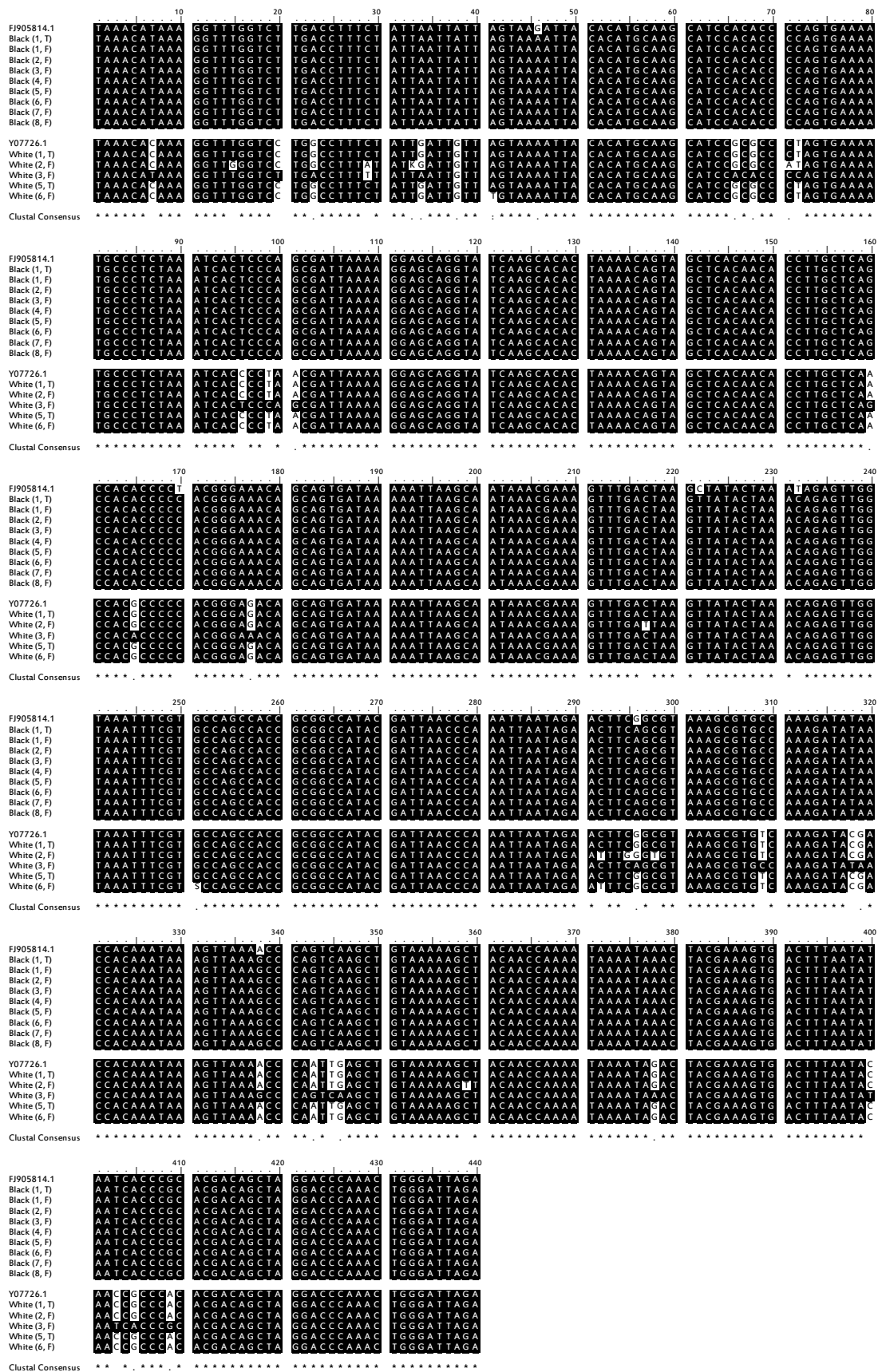
Figure 16: A multiple sequence alignment for the black and white rhinoceros HV2 sequences. The black rhinoceros sequence alignment included sequences obtained from one tissue and eight faecal samples as well as a reference sequence (FJ905814.1). The white rhinoceros sequence alignment included sequences obtained from two tissue and three faecal samples as well as a reference sequence (Y07726.1). The sequence identity matrix can be seen in Table AIX for the black rhinoceros and Table AX for the white rhinoceros.

A total of 140 bp were lost due to sequence quality being lost at the 5' and 3' ends of the isolated DNA sample sequences. The black rhinoceros HV2 sequences resulted in a BLAST identity of 99% to the *Diceros bicornis* complete mitochondrial DNA sequence (FJ905814.1) with a maximum score of 952, a query cover of 100% and an E-value of 0.0. The reference sequence and isolated DNA sample sequences had a sequence identity of 0.986 as six variations were observed in Figure 16 (variation at positions 46, 170, 222, 232, 296 and 338). This is a sequence difference of 1.4%. The isolated DNA samples had a sequence identity of 1.000 as no intraspecies variation was observed. The white rhinoceros sequences resulted in a BLAST identity of 99% to the *Ceratotherium simum* complete mitochondrial DNA sequence (Y07726.1) with a maximum score of 941, a query cover of 100% and an E-value of 0.0. Variation was observed between the reference sequence and isolated DNA sample sequence. Samples white (1, T) and white (5, T) had no sequence difference to the reference sequence while white (2, F) had a difference of 2.1%, white (3, F) had a difference of 6.2% and white (6, F) had a difference of 0.7% (see Table AX in the appendix for detail on sequence identity for this region). Sequence variation was observed for samples white (2, F), (3, F) and (6, F). The sample for white (2, F) had SNPs at the positions 15, 71, 217, 294, 297 and 359. The sample for white (2, F) and white (3, F) both contain a SNP at the position 29, however the SNPs were different (white [3, F] contained an A; white [3, F] contained a T). Sample for white (2, F) and white (6, F) contained a shared SNP at position 292. Sample white (3, F) contained several SNPs which were conserved in the samples obtained for the black rhinoceros. These were observed at the positions of 7, 20, 23, 34, 65, 67, 72, 95, 98, 101, 161, 165, 176, 296, 309, 318, 319, 338, 342, 344, 345, 378, 400, 403, 405 and 409. The samples white (1, T) and (5, T) did not contain variation.

4.4.4 Optimisation of PCR amplification for a region of the cytb gene

Primers were designed to amplify a 545 bp region of the *cytb* in order to determine if any genetic variation occurred between individuals. The forward and reverse primers had melting temperatures of 62.0℃ and of 68.0℃ respectively. A gradient PCR, with the annealing temperatures of 60.0℃, 62.0℃ and 64.0℃, was performed to determine the optimal annealing temperature for amplification (Figure 17A). Faecal sample PCR reactions included 35 cycles and tissue samples had a cycle number of 25.
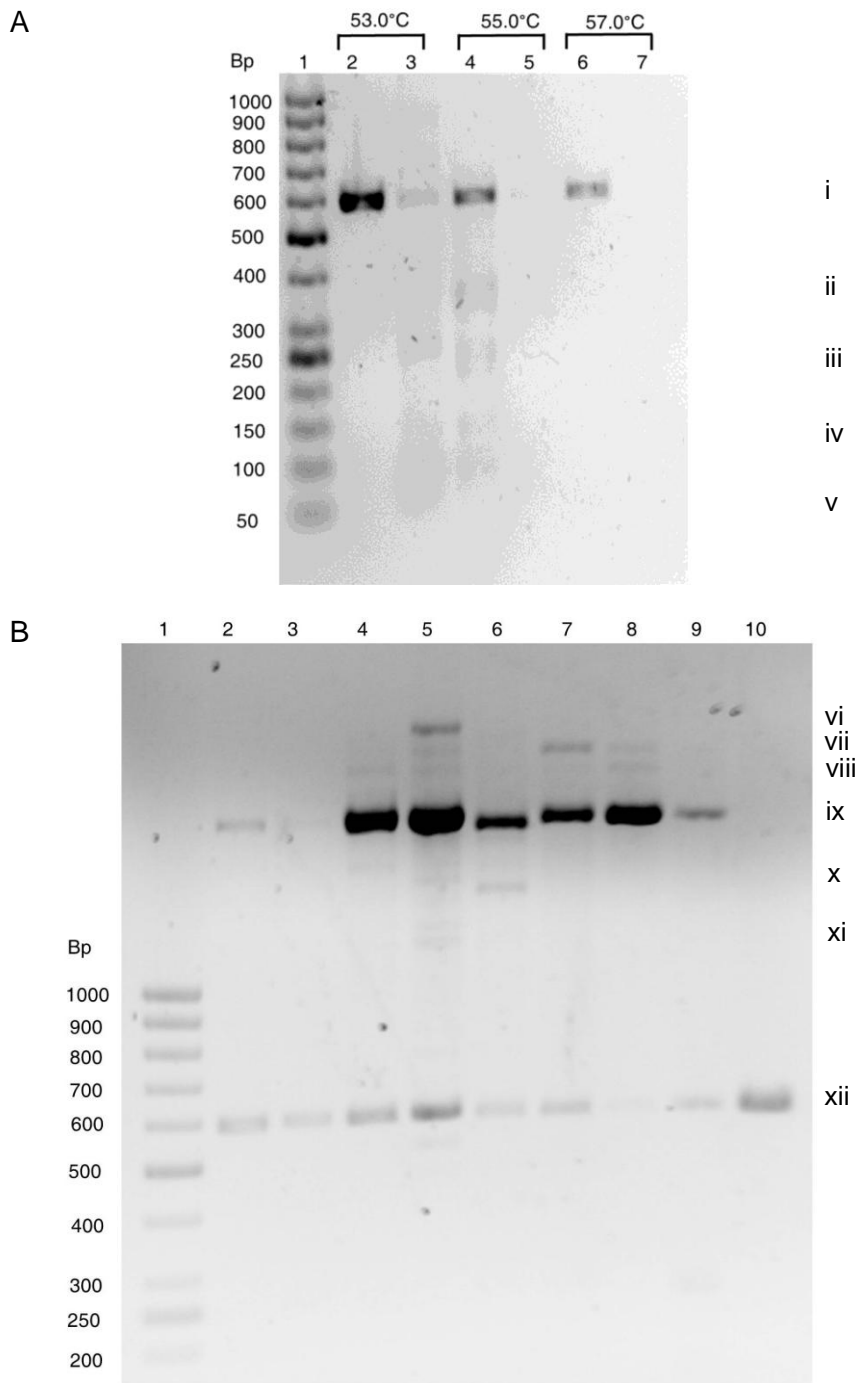
Figure 17: Amplified *cytb* region for the black rhinoceros. (A) A gradient PCR for the *cytb* region to determine the optimal annealing temperature for amplification. Lane 1: O'Gene Ruler 1 kb Plus DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (1, T), Lane 7: black (1, F), Lane 8: black (2, T), Lane 9: black (3, F), Lane 10: black (1, T), Lane 11: black (1, F), Lane 12: black (2, F), Lane 13: black (3, F). (B) Optimized PCR products for the black rhinoceros *cytb*. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (4, F), Lane 7: black (5, F), Lane 8: black (6, F), Lane 9: black (7, F), Lane 10: black (8, F). For both A and B: Agarose (1% w/v) gel electrophoresis was used. Non-specific bands (i, ii, iii, v, vii and ix), target bands to determine optimal annealing temperature (iv), primer dimers (vi and x) and optimised PCR product bands for the *cytb* region (viii).

Non-specific bands were observed, in lane 5 and 9 (i, ii, iii, v, vii and ix) (Figure 17A). Primer dimers were observed (vi). The annealing temperature of 64.0°C was chosen for subsequent amplification due to the primer dimer bands being the least dense at this temperature. The optimised PCR protocol was performed on all samples. The optimised *cytb*

54

PCR products for the black rhinoceros (viii) are shown in Figure 17B. Primer dimers (x) were observed in Figure 17B for the amplified faecal products. These were removed by performing gel extractions on all *cytb* products. TotalLab (Newcastle, UK) was used to determine the size of the bands present in Figure 17B. The average size of the bands was 580 bp (SD±6.78), which is in range with the *in silico* size of 545 bp. The optimised PCR protocol for the *cytb* region was performed on white rhinoceros samples (data not shown). The *cytb* PCR products for the black and white rhinoceros were sequenced.

A multiple alignment, which included a reference sequence for the black rhinoceros (FJ905814.1) and a reference sequence for the white rhinoceros (Y07726.1), was performed on the *cytb* sequences to determine if any genetic differences occurred between individuals (Figure 18).
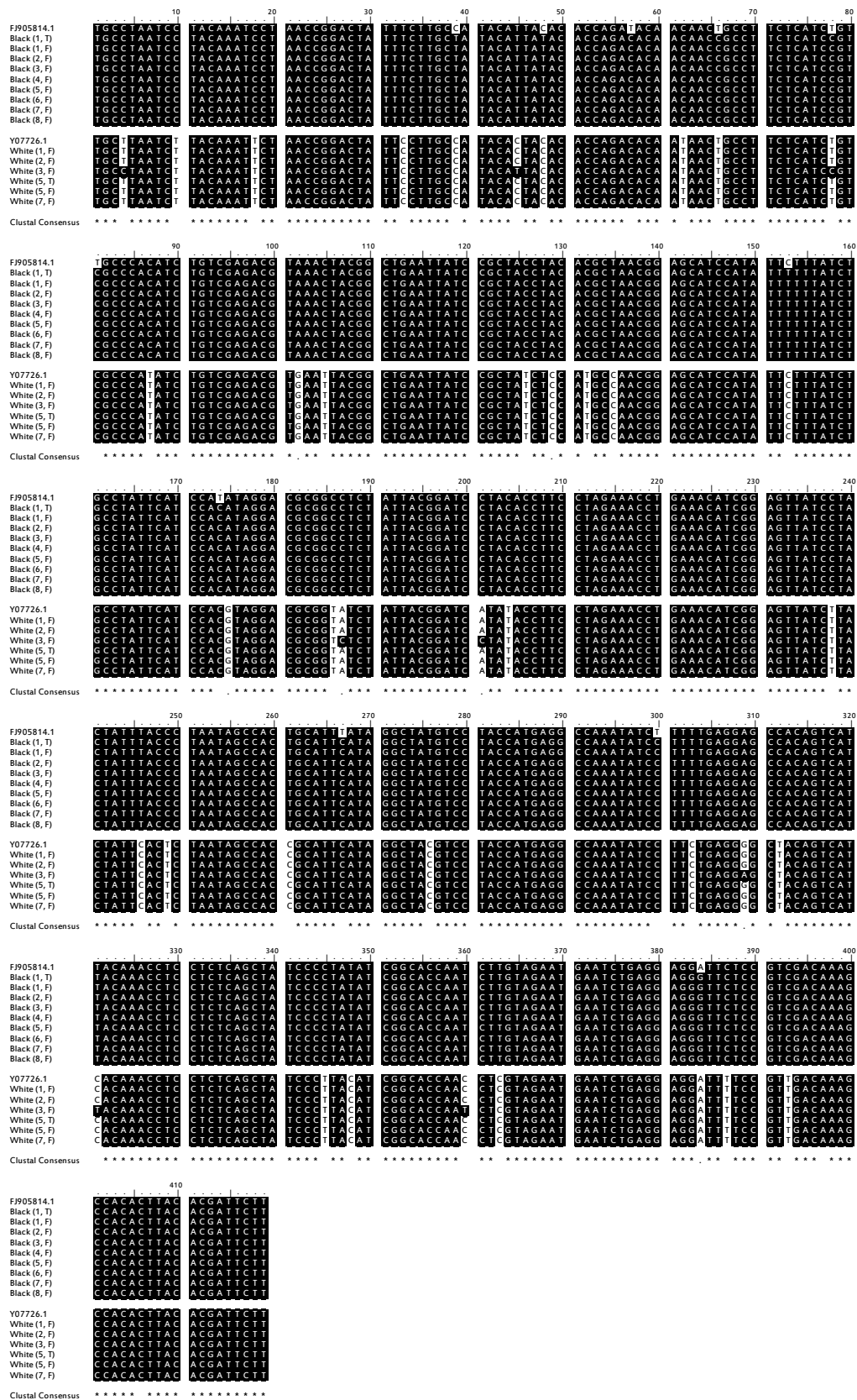
Figure 18: A multiple sequence alignment for the obtained black and white rhinoceros *cytb* sequences. The sequence alignment for the black rhinoceros included a reference sequence (FJ905814.1) and sequences obtained from one tissue and eight faecal samples. A reference sequence (Y07726.1) as well as sequences obtained from one tissue and five faecal samples. The sequence identity matrix can be seen in Table AXI for the black rhinoceros and Table AXII for the white rhinoceros.

56

The black rhinoceros sequences resulted in a BLAST identity of 100% to the *Diceros bicornis* isolate CT *cytb* gene (EU107377.1), with a maximum score of 921, a query cover of 97% and an E-value of 0.0. Sequences for the black rhinoceros collected samples were 100% conserved. The reference sequence and isolated DNA sample sequences had a sequence identity of 0.973 (variations in the reference sequence observed at positions 39, 48, 57, 66, 78, 81, 153, 174, 267, 300 and 384). The isolated DNA sample sequences had a sequence identity of 1.000 as no variation was observed in Figure 18. The white rhinoceros sequences resulted in a BLAST identity of 100% to the *Ceratotherium simum cytb* gene (JF718874.1), with a maximum score of 904, a query cover of 100% and an E-value of 0.0. The reference sequence and isolated DNA sample sequences had a sequence similarity of 1.000 with the exception of white (3, F) (variations at positions 4, 45, 78, 187, 201, 309, 320 and 360) which had a sequence identity of 0.980 (sequence difference of 2%). Variation was observed in Figure 18 between the black and white rhinoceros sequences at positions 4, 10, 18, 33, 39, 45, 48, 62, 66, 78, 87, 102, 105, 126, 129, 132, 135, 153, 175, 186, 187, 201, 204, 238, 246, 249, 261, 275, 303, 309, 311, 321, 345, 348, 360, 363, 384, 387 and 393. A total of 125 bp were lost due to a loss of sequence quality at the 5' and 3' ends of the isolated DNA sample sequecnes.

4.4.5 Optimisation of PCR amplification for a region of the 16s rRNA gene

A primer set was designed to amplify a 531 bp region of the *16s rRNA* gene. The forward and reverse primers had melting temperatures of 66.0℃ and of 62.0℃ respectively. A gradient PCR, with the annealing temperatures of 62.0℃, 64.0℃ and 66.0℃, was performed to determine the optimal annealing temperature for amplification (Figure 19A). Faecal sample PCR reactions included 35 cycles and tissue samples had a cycle number of 25.
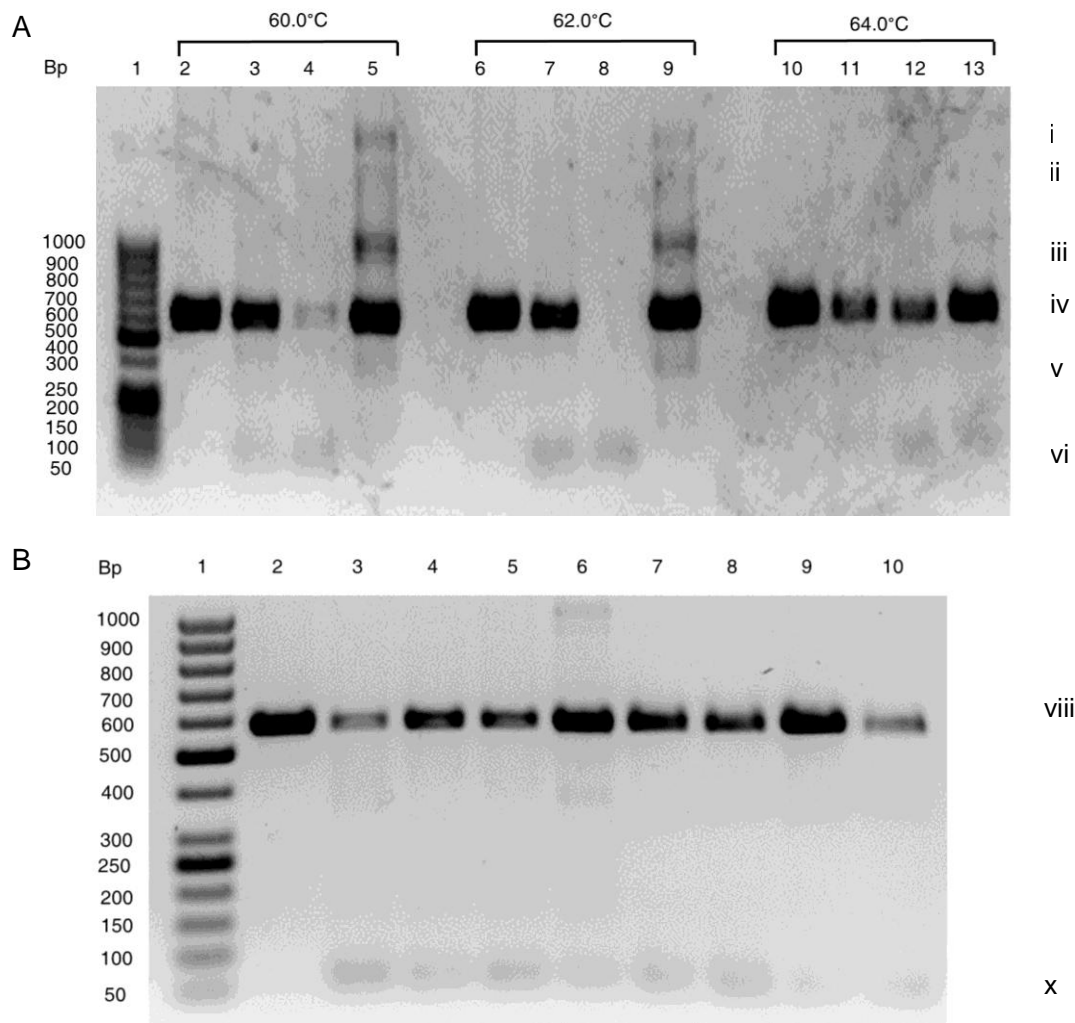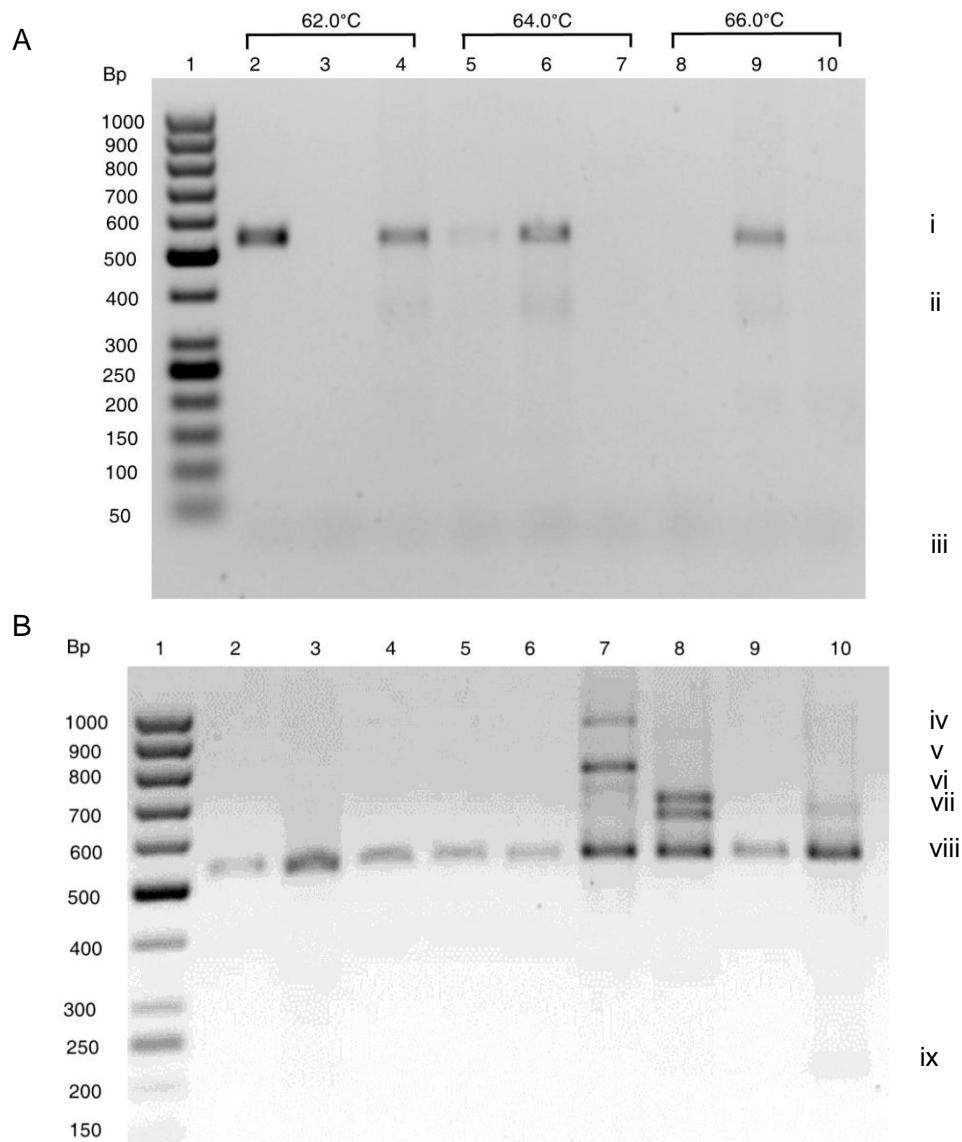
Figure 19: Amplified *16s rRNA* region for the black rhinoceros. (A) A gradient PCR for the *16s rRNA* region to determine the optimal annealing temperature for amplification. Lane 1: O'Gene Ruler 1 kb Plus DNA Ladder, Lane 2: black (1, F), Lane 3: black (2, F), Lane 4: black (3, F), Lane 5: black (1, F), Lane 6: black (2, F), Lane 8: black (3, F), Lane 8: black (1, F), Lane 9: black (2, F), Lane 10: black (3, T). Optimized PCR products for the black rhinoceros *16s rRNA* region. Lane 1: GeneRuler™ 50bp DNA Ladder, Lane 2: black (1, T), Lane 3: black (1, F), Lane 4: black (2, F), Lane 5: black (3, F), Lane 6: black (4, F), Lane 7: black (5, F), Lane 8: black (6, F), Lane 9: black (7, F), Lane 10: black (8, F). For both A and B: Agarose (1% w/v) gel electrophoresis was used. Target bands to determine optimal annealing temperature (i), non-specific bands (ii, iv, v, vi, vii and ix), primer dimers (iii) and optimised PCR product bands for the *16s rRNA* region (viii).

In Figure 19A, faint primer dimers were seen for all PCR products (iii) and non-specific bands were observed, in lane 4, 6 and 9 (ii). The annealing temperature of 62.0°C was chosen for subsequent amplification due this temperature having the most successful amplification of sample. The optimised *16s rRNA* PCR products for the black rhinoceros are shown in Figure 19B.

Primer dimers were absent from the optimised PCR products. Non-specific bands were observed in Lanes 7, 8 and 10 (iv, v, vi, vii and ix). TotalLab (Newcastle, UK) was used to determine the size of the target bands present in Figure 19B (viii). The average size of the bands was 555 bp (SD±4.32), which is similar in size compared to the *in silico* size of 531 bp. The black rhinoceros samples were for sequenced. Prior to sequencing the non-specific bands were removed by performing gel purifications, and therefore they did not interfere with sequencing.

A multiple alignment, which included a reference sequence for the black rhinoceros (FJ905814.1), was performed on the sequences obtained for the *16s rRNA* collected samples to determine if any genetic differences occurred between individuals (Figure 20).
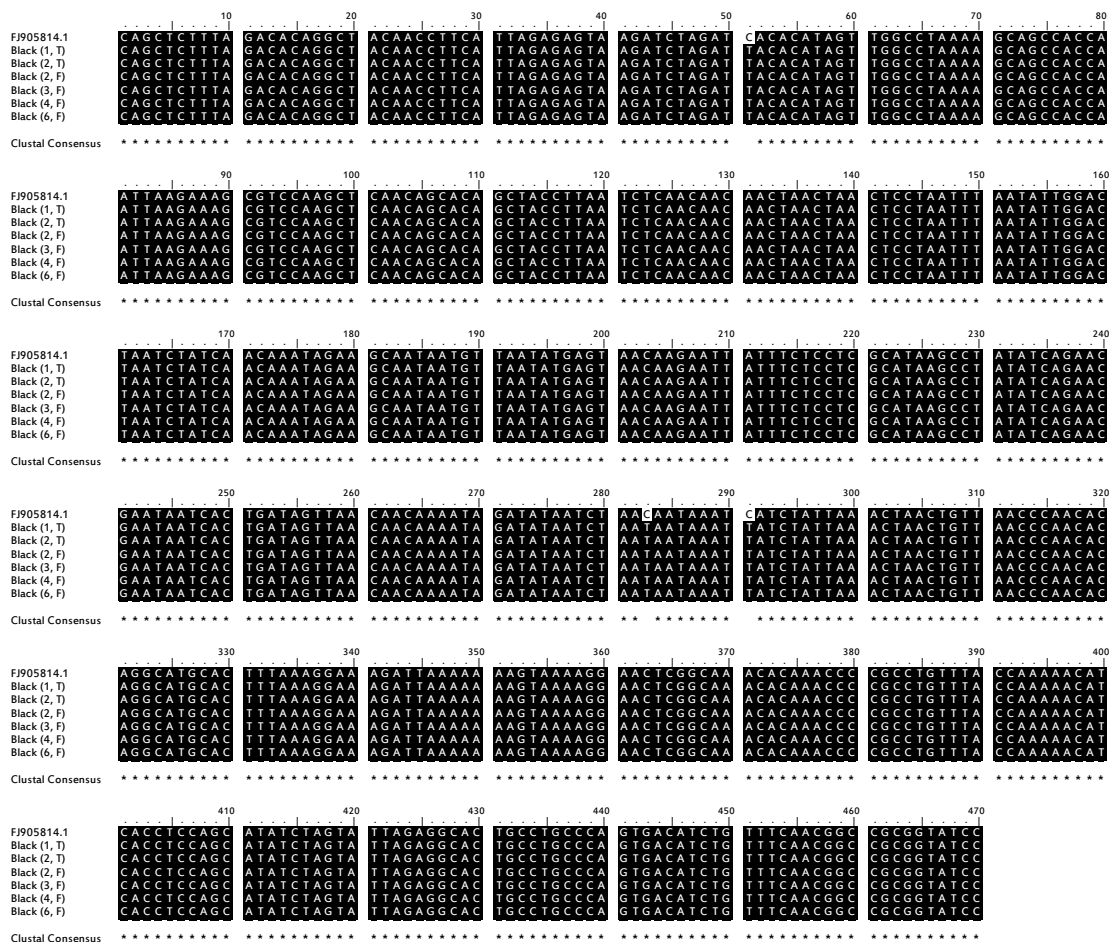
Figure 20: A multiple sequence alignment for the obtained black rhinoceros *16s rRNA* sequences. This sequence alignment includes two tissue and four faecal samples. It also includes a reference sequence (FJ905814.1). The primers designed for this region were not compatible with the white rhinoceros mtDNA genome. The sequence identity matrix can be seen in Table AVIII.

The obtained black rhinoceros sequences resulted in a BLAST identity of 99% to the *Diceros bicornis* complete mitochondrial DNA sequence (FJ905814.1) with a maximum score of 865, a query cover of 98% and an E-value of 0.0. The reference sequence and the isolated DNA sample sequences had a sequence identity of 0.993 as variations were observed at positions 51, 283 and 291 (sequence difference of 0.7%). The isolated DNA sample sequences had a sequence identity of 1.000 as no variation was observed for these samples in Figure 20. A total of 61 bp were lost as a result of poor sequence quality at the 5' and 3' ends.

CHAPTER FIVE

DISCUSSION

5.1 Faecal sample as a non-invasive source of DNA

For this study, faecal matter was investigated to determine if it could serve as a future source of non-invasive DNA. The interest in non-invasive sources of DNA amongst conservation biologists is increasing as the animals do not need to be observed or captured (Taberlet & Luikart, 1999). Gerloff *et al.* (1995) extracted nuclear DNA from free-living bonobos (*Pan paniscus*) faeces in order to amplify hypervariable regions which contained simple repeat sequences. The study performed DNA extractions from 33 bonobos and approximately 66% of the samples were successfully amplified on the first attempt. The authors conducted a consistency test which indicated that sometimes only one out of two alleles where amplified. The authors thought that this may be due to the limited amount of DNA present in the faecal samples, and therefore validation was required for results that indicated homozygous individuals (Gerloff *et al.,* 1995). It should be noted that the study conducted by Gerloff *et al.* (1995) focused on nuclear DNA extracted from faeces and not mtDNA. This study recommended that caution be taken when working with DNA isolated from faecal matter. In my current investigation, DNA isolated from faecal matter was 100% conserved with the DNA isolated from tissue samples. The mtDNA copy number may have played a positive role in this finding as it has a much higher (25 fold) copy number than nuclear DNA (Moritz & Cicero, 2004).

The results from this study showed that DNA isolated from faecal samples had concentrations that were comparable to that of the DNA extracted from matched blood samples (Table II). Tissue was chosen to serve as the positive control for experimentation due to higher concentrations of DNA being obtained for this source of DNA. The extraction products for the faecal samples for both species of rhinoceros did not migrate into the gel (Figure 2). This may be due to residual wash buffer not being removed from the extraction column (Zymo Research, 2014). DNA extractions performed on faecal sample may contain other DNA, for example DNA may be extracted from an animal's prey, or in the case of the rhinoceros DNA may be extracted from digested plant material (Marrero *et al.,* 2009).

Molecular studies which have utilized faecal DNA have been conducted on wildlife other than the rhinoceros. Yamashiro *et al.* (2010) extracted mtDNA from faecal samples to

differentiate between the sika deer (*Cervus nippon*) and Japanese serow (*Capricornis crispus*). The *cytb* gene was investigated. The authors experienced an 86% success rate for the amplification of the *cytb* gene from faecal samples. PCR-RFLP was performed on the DNA extracted from faecal samples, which discriminated between the sika deer and Japanese serow (Yamashiro *et al.,* 2010). Faecal sample was used by Hapke *et al.* (2001) to examine the D-loop in hamadryas baboons (*Papio hamadryas hamadryas*). The study aimed at investigating the transferring of female individuals between neighbouring baboon troops on a genetic level. This would in turn allow for a gene-flow to be examined. DNA was obtained from faecal matter, A total of 74 faecal samples were collected from 10 different geographical areas. The authors compared nucleotide sequences by utilizing an analysis of molecular variance and phylogenetic tree reconstructions. The authors were unsuccessful in finding a relationship between geographic patterns and population genetic structures, however female gene-flow could be detected (for populations that surpassed neighbouring subpopulations) (Hapke *et al.,* 2001), however DNA obtained from the faecal samples produced reproducible results. Jensen-Seaman & Kidd (2001) examined the D-loop region in four East African populations of gorillas (*Gorilla beringei*). DNA was extracted from non-invasive samples (faeces and hair). The aim for their study was to investigate genetic diversity between four population groups in order to infer interpopulation relationships and intrapopulation genetic structures. The authors reported that genetic diversity between the populations was low and that evidence of bottlenecking was present (Jensen-Seaman & Kidd, 2001). Brown & Houlden (2000) investigated genetic variation between two subspecies of the black rhinoceros (*D. b. michaeli* and *D. b. minor*) by examining the D-loop. Genetic information was obtained from faecal, blood or biopsy material. A nucleotide diversity of 0.43% and a haplotype diversity of 0.86 was found for the *D. b. minor*. The D-loop displayed a 2.6% differentiation between the two subspecies of the black rhinoceros. The authors concluded that genetic variation for the black rhinoceros had been conserved (Brown & Houlden, 2000).

The results obtained from our study support the results obtained by Yamashiro *et al.* (2010), Hapke *et al.* (2001), Jensen-Seaman & Kidd (2001) and Brown & Houlden (2000) as reliable faecal DNA was extracted from rhinoceros faeces. These extractions allowed for the seven mtDNA target regions to be amplified for the black rhinoceros and for the five regions that were compatible with the white rhinoceros to be amplified. A higher initial template concentration and cycle number was required for faecal sample (150 ng and 35 cycles)

amplification in comparison to that of tissue samples (60 ng and 20 cycles). This may be explained by the fact that the DNA extracted from faecal matter may have contained DNA from other source obtained through diet. This may have been a factor as to why a higher concentration of faecal DNA was needed for amplification (Wehausen *et al*., 2004).

Idaghdour *et al.* (2003) established a reliable method (silica-based) to extract DNA from faecal samples. The authors used this DNA to investigate a 542 bp region of the D-loop for the Great Bustards (*Otis tarda*) that inhabited Morocco. The authors found that their amplification from faecal samples success rate was directly linked to the composition and freshness of the sample. Dry faecal sample, which accounted for 30% of the study, could not be amplified, however fresh samples had a success rate of 100% (Idaghdour *et al.*, 2003). Fresh faecal sample was produced by 25 individuals. The D-loop for these individuals was investigated. The authors performed a validation study for the method using five other bird species. This resulted in an 80% success rate for amplifying DNA from faecal samples (Idaghdour *et al.*, 2003). The faecal samples used for this investigation were fresh as the sample was taken directly from the animals rectum, however in future studies this may not be the case as faecal matter is to serve as a non-invasive option for DNA. It would be advised that when using isolated DNA from faecal samples that one observes the animal for defecation as this would minimise the degradation level. In private game reserves, rhinoceros can be identified by unique ear notch patterns therefore in some cases this method of collection would also allow for the collector to validate which rhinoceros the sample belonged to. However this may not always be the case, therefore molecular techniques designed to utilize faecal matter needs to be reliable and accurate.

It was determined that faecal samples could be used as source of DNA for this study. The designed primers were able to amplify the seven mtDNA target regions for the black rhinoceros and five regions for the white rhinoceros. As expected, sequence quality was observed to be lost at the 5' and 3' ends of the isolated DNA faecal sample sequences, however the sequence quality observed in the middle of the sequences was viable (indicated by distinct bands). The positive control and its matched faecal sample was always 100% conserved (sequence identity of 1.00), indicating that DNA extracted from faecal sample was adequate in order to conduct biological investigations.

5.2 A molecular technique to differentiate between the black and white rhinoceros using the *COI* gene

Identifying species using the *COI* gene is becoming standard practice for wildlife biologists because dogma in the field suggests that all animals can be identified using this gene (Hebert *et al.,* 2003). However it should be noted that Mallet & Willmott (2003) expressed concern about using this gene for species identification as closely related species might not contain sufficient genetic variation to allow for discrimination. A sequence identity of 0.941 and 0.942 was observed between the black and white rhinoceros for the *COIi* and *COIii* regions, respectively. This represents a 5.9% and 5.8% sequence difference between the two species for the respective regions of the *COI* gene. Anderson-Lederer *et al.* (2012) has reported that large mammals exhibit genetic differences of less than 2%, therefore the sequence difference obtained for the white and black rhinoceros could be said to be substantial. The similar sequence differences between the black and white rhinoceros for the two regions of the *COI* gene is interesting as the *COIi* region investigated was 154 bp and the *COIii* region investigated was 470 bp in length. Therefore, attention should be drawn towards these findings as even though a substantial difference in size occurred between the two *COI* regions investigated, the sequence difference between the two species was similar.

A novel PCR-RFLP assay was developed in this study which allowed for species identification for the black and white rhinoceros. Figure 7 showed that the experimental assay could be conducted routinely in a laboratory that had the necessary equipment. These results were supported by the alignment in Figure 8. To the best of our knowledge, this developed technique is the first which allows for reliable identification of the black and white rhinoceros.

5.3 A biological investigation into the variation observed in the mitochondrial DNA target region sequences for the black and white rhinoceros

Black rhinoceros from southern Africa were used in a study conducted by O'Ryan *et al.* (1994) to investigate mtDNA genetic variation. A total of 33 individuals (Southern, South-central, Eastern and Chobe black rhinoceros) were used to establish cell cultures. The individuals represented three different geographic regions. A total of sixteen restriction enzymes were used to construct restriction maps (O'Ryan *et al.,* 1994). The restriction maps indicated that there were two sites that were different for individuals that represented two out of the three different geographical regions. Restriction maps for individuals that originated from the same geographic regions had no variation. The authors suggested that in order for

inbreeding depression to be avoided, black rhinoceros from wild (and captive) populations should be bred with individuals from other populations due to the genetic diversity being low (O'Ryan *et al.,* 1994). Several of the investigated target regions displayed genetic variation; these results support the findings by O'Ryan *et al.* (1994).

Sequence variation was observed for the black rhinoceros between the isolated DNA sample sequences and reference sequence for all seven examined target regions. Sequence differences between the isolated DNA sample sequences and the reference sequence were found to be 0.7%, 5.5%, 1.4%, 2.7 % and 0.7% for the *COIii*, D-loop, HV2, *cytb* and *16s rRNA* regions, respectively. A sequence difference for the HV1 region between the isolated DNA sample sequences and the reference sequence was found to be 54.2% with the expectation of one individual (represented by Black [1, T] and [1, F]) which had a sequence difference of 53.0%. Intraspecies variation was observed for *COIi* (one variation in individual Black [5, F]) and HV1 regions. The HV1 region showed variation for one individual (represented by Black [1, T] and [1, F]) as a result of a bp deletion. This individual had a sequence difference of 2.6% when compared to the other isolated DNA sample sequences (Figure 14).

No sequence differences were observed for the white rhinoceros between the isolated DNA sample sequences and the reference sequence for the *COIi*, *COIii* and *cytb* regions. A sequence difference of 15.5% was observed for the D-loop isolated DNA sample sequences and the reference sequence. The HV2 region had no difference in sequence between individuals White (1, T) and (5, T) however differences of 2.1%, 6.2% and 0.7% were observed for individuals White (2, F), (3, F) and (6, F) and the reference sequence. Intraspecies variation was observed in the HV2 and *cytb* regions for the white rhinoceros. The individual White (3, F) had a sequence difference of 2% for the *cytb* region when compared to the other isolate DNA sample sequences. Three out of the five isolated DNA sample sequences resulted in variation for the white rhinoceros with regards to the HV2 region.

Ashley *et al.* (1990) investigated mtDNA variation for the black rhinoceros. Two subspecies of black rhinoceros were used for the study, this included eleven Eastern black (Kenyan origin) and twelve South-central black rhinoceros (of which eleven originated from Zimbabwe and one originated from South Africa). Blood and tissue samples were used as the sources of DNA. Restriction enzymes were used to digest the mtDNA. Three haplotypes

were identified for the 23 individuals. The first haplotype identified individuals that originated from Kenya, the second identified individuals that originated from Zimbabwe and the last identified the individuals that originated from Zimbabwe and South Africa (Ashley *et al.,* 1990). The identified haplogroups were very similar with sequence variation being approximately 0.17%. The average sequence variation between the two subspecies was 0.29%. A close genetic relationship was suggested by the authors for the investigated individuals. Due to the fast evolving rate of mtDNA, the authors suggested that a common ancestral relationship was shared by the investigated individuals. The sequence difference obtained in our study showed values greater than that reported by Ashley *et al.* (1990), however not all of the examined individuals contained genetic variation. One black rhinoceros and three white rhinoceros showed variation in several regions. Interestingly, the sequences obtained from the black rhinoceros isolated DNA varied from the reference sequence for all of the target regions. This was also true for the white rhinoceros, but with the exception of the *COIii* and *cytb* region where no variation was shown. Wilson *et al.* (1985) suggests that the evolutionary rate of mtDNA is 2% per million years. Therefore, Ashley *et al.* (1990) suggested that their recent ancestry lineages could date back to 100 000 years.

5.3.1 Investigation of the COI

The *COIi* region displayed a single SNP for the black rhinoceros sequences at position 39 (Figure 8). This preliminary finding was indicative that the *COI* gene may have been a potential site for intraspecies variation. Li *et al.* (2014) conducted a study in which the main objective was to determine the amount of genetic variation found in five different geographical populations of cuttlefish (*Sepiella japonica*). The authors amplified a 681 bp region of the *COI* gene in 96 individuals collected off the coast of China (Li *et al.,* 2014). The authors identified 22 different haplotypes for the 96 samples. Of these 22 haplotypes, fifteen appeared in a single population, six appeared in two populations and one was found in all five populations. The authors concluded that the genetic variation levels for the cuttlefish investigated were low (Li *et al.,* 2014). As previously mentioned, the *COI* gene is commonly used for species identification, therefore intraspecies variation would not be expected. However as Li *et al.* (2014) observed variation in cuttlefish and variation was observed for the black rhinoceros in the *COIi*, the *COIii* region was investigated to determine if this gene housed intraspecies variation for the black and white rhinoceros. The *COIii* gene resulted in no sequence difference within the black or white rhinoceros isolated DNA sample sequences as each species had a sequence identity of 1.000. Therefore, the *COI* gene did not provided

substantial variation to allow for individualisation of the different isolated DNA sample sequences.

5.3.2 Investigation of the D-Loop

The D-loop has been used for the discrimination of human individuals (Greenberg & Aquardo, 1983; Wilson & Cann, 1983; Fridman *et al*., 2011; Bataille *et al.* & Andréasson *et al.,* 2002) however when used for animals such as the rhinoceros, gelada baboon and American alligators low levels of genetic variation have been observed (Anderson-Lederer *et al*., 2012; Ashley *et al*., 1990 & Glenn *et al.*, 2002).

Anderson-Lederer *et al*. (2012) investigated the level of variation found in several black rhinoceros D-loops. This group amplified a 406 bp region of the D-loop from 101 individuals. These 101 individuals comprised of 65 South-central black rhinoceros from South Africa, eleven South-central black rhinoceros from outside South Africa, 21 Eastern black rhinoceros and four South black rhinoceros. Ear notches and blood samples were used for experimentation. The authors found that the South-central black rhinoceros population from KwaZulu-Natal had a single mtDNA haplotype. This trend was also observed for the Sumatran rhinoceros (Morales *et al*., 1997). However, the Sumatran rhinoceros inhabited separate geographic regions for approximately 10 000 years, whereas the South-central black Kwa-Zulu Natal rhinoceros population has only become separate from other black rhinoceros subspecies in recent years. Anderson-Lederer *et al*. (2012) raised the question as to whether the loss of genetic variation was because of a recent bottleneck of the population, or if the lineage originated from related ancestors and the separate genetic lineage goes back for longer than had originally been assumed. The data obtained in this investigation suggests that Anderson-Lederer *et al*. (2012) may have been correct in raising concern with regards to low levels of genetic variation in the D-loop. In our study, the investigated D-loop region allowed for family determination as the black and white sequences were identified as rhinoceros. No intraspecies variation was observed for the black and white rhinoceros DNA sequences, and no sequence variation occurred between the two species.

The D-loop had been identified as a target region for intraspecies variation by investigations conducted of several species. A study conducted by Belay & Mori (2006) examined the D-loop in four populations of gelada baboon (*Theropithecus gelada*) to determine the presence of genetic variation. These populations originated from the Northern and Southern regions of the Rift Valley (Ethiopia). PCR-RFLP was used to investigate the D-loop. The Northern and

Southern populations were able to be separated due to the populations each having a different haplotype, which had a sequence divergence of 9.8% (Belay & Mori, 2006). The authors did not find any variation between or amongst local populations. The authors were able to conclude that gene flow ceased between the Northern and Southern populations due to a historical geographic distribution break for the species. The authors therefore believed that the Southern population had become a new subspecies of the gelada baboon (Belay & Mori, 2006). Our study supports the data found by Belay & Mori (2006) as no intaspecies variation was observed for the D-loop isolated sample sequences. Future work would have to be conducted in order to determine if this region would be able to identify separate population groups. However, the fact that no variation was observed between the black and white rhinoceros causes one to doubt if variation within the same species would yield variation.

5.3.3   *Investigation of the hypervariable regions*

Investigations have been conducted on several species in order to determine if the hypervariable regions contain intraspecies variation. Aranishi & Okimoto (2005) discovered a novel non-coding region of the mtDNA that could be used for individualisation amongst Pacific oyster (*Crassostrea gigas*). The investigated region was amplified with the valine *tRNA* and glycine *tRNA* that lay adjacent to the targeted region. A comparison between 29 Pacific oysters was conducted to identity polymorphic sites. A total of thirty polymorphic sites were identified, this led to 22 haplotypes being identified, with most of the haplotypes being specimen specific. The nucleotide alteration frequency was calculated to be 4.68% and the genetic diversity was calculated to be 0.988 for the 22 haplotypes (Aranishi & Okimoto, 2005). This analysis was deemed successful for intraspecies phylogenetic studies for the Pacific oyster. As the hypervariable regions showed intraspecies variation for other species, variation was expected for the rhinoceros HV1 and HV2 regions.

The HV1 region allowed for the rhinoceros family to be identified from the isolated DNA sample sequences. The isolated DNA sample sequences had a large addition of bp when aligned with the reference sequence (Figure 14). This unreported sequence region did not result in any BLAST matches. This potentially shows a sequencing error for the reference sequence that was created by Willerslev *et al.* (2009) as there was a large deletion in the reference sequence. Alternately this could show a novel region for the examined black rhinoceros individuals from South Africa.

Intraspecies variation was observed for the white rhinoceros individuals (2, F), (3, F) and (6, F) for the HV2 region. Interestingly, no variation was observed for the DNA isolated from tissue samples. No variation has been observed between the positive control and its matched faecal sample, therefor this may simply be coincidental. This variation may be indicative of familial relationships for the white rhinoceros however this variation is insufficient for individualisation.

5.3.4 Investigation of cytb

For closely related species, the *cytb* region should have sequence conservation as these species would have originated from the same ancestry line (Linacre *et al.,* 2009). A study conducted in 1998, showed that the *cytb* gene frequently had a sequence variation of more than 2% for closely related vertebrate species (Johns & Avise, 1998). Variation was observed between the two rhinoceros species in Figure 18. The sequence difference obtained when comparing the isolated DNA sample sequences from the black and white rhinoceros was 9.4%. Sequence variation between species is expected due to evolution of mtDNA and therefore the *cytb* gene can be used to show intraspecies variation (Linacre *et al.,* 2009). Linacre *et al.* (2009) stated that for the *cytb* region, two closely related species may have limited sequence variation. For example, four base differences may occur between the wolf (*Canis lupus*) and the dog (*Canis lupus familiaris*). Therefore, there is a need to develop tests which incorporate significant variations. A total of 39 variation sites were observed between the black and white rhinoceros. This is indicative that this region of the *cytb* gene incorporates sufficient variation sites in order to correctly identify the subspecies of rhinoceros.

Variation was observed in the *cytb* region for one of the white rhinoceros sample sequences. Individual white (3, F) contained 8 variation sites (2%) from the other isolated DNA samples sequences (Figure 18). No other white rhinoceros sample sequences displayed sequence variation.

Genetic investigations have been conducted on species other than the rhinoceros. Bradley & Baker (2001) investigated sequence variation in the *cytb* gene for seven genera of bats and four genera of rodents. The study looked at mtDNA variation within a species (intraspecific, intrasubspecific and intrapopulational variation) and variation that occurred within a family (intrageneric and sister-species variation). The intraspecific variation, the genetic variation within single species, was investigated for eighteen species of rodents and fourteen species of

bats. These results indicated an average genetic distance that ranged from 0.0–8.7 with the average being 2.49. This showed that the *cytb* gene was a possible site for genetic variation (Bradley & Baker, 2001).

Variation between closely related species can be seen using not only the *COI* gene but also the *cytb* gene (Linacre *et al.,* 2009). These two genes are most commonly used for species identification in wildlife forensics (Alacs *et al.,* 2009). In this study of the 545 bp region amplified for the *cytb* gene, 420 bp were used for a multiple sequence alignment. The *cytb* gene allowed for family determination, with the *cytb* sequences identified as rhinoceros and species was also determined from the collected samples. The black rhinoceros *cytb* resulted in a 100% sequence match to the *Diceros bicornis* (EU107377.1) and the white rhinoceros samples resulted in a 100% sequence match to the *Ceratotherium simum* (JF718874.1). This data, with regards to the rhinoceros, does not support the hypothesis put forward by Bradley & Baker (2001) as this site did not indicate viable intraspecies variation for the examined rhinoceros samples.

5.3.5   *Investigation of 16s rRNA*

An investigation was conducted on the *16s rRNA* gene to determine the level of genetic variation (Marín *et al.,* 2013). Although this study was not on rhinoceros, it showed genetic insights for the target gene. Marín *et al.* (2013) used a 530 bp region of the *16s rRNA* gene to investigate the population genetic structure of the Peruvian scallop (*Argopecten purpuratus*). For the study, a total of 69 individuals were used. Samples from three different geographic populations of scallops were collected along the coast of Peru. The authors identified nineteen polymorphic sites, which led to eighteen different haplotypes being identified (Marín *et al.,* 2013).

Marín *et al.* (2013) indicated that the *16s rRNA* gene had potential for showing intraspecies variation. Therefore, this region was investigated for the rhinoceros. The *16s rRNA* gene was investigated for the black rhinoceros. The isolated DNA sample sequences resulted in a sequence match of 99% to the *Diceros bicornis* (FJ905814.1). This is indicative that this region can be used for family identification; however this region would need to be investigated for the white rhinoceros. No intaspecies variation was observed for the collected sample sequences in Figure 20. However, the collected sequences varied from the reference sequence at three nucleotide positions.

5.4 Evaluating familial relationships and individualisation for the investigated black and white rhinoceros

Two rationales should be considered as to why the black rhinoceros isolated DNA sample sequences displayed variation from the reference sequence for all target regions investigated. The first is that the published mtDNA genome reference sequence is not reliable. The second is that these target regions may infer that the isolated DNA sample sequences and the reference sequence originated from different geographic regions.

The rhinoceros samples utilised for this study were collected from two regions of South Africa. The first region has had animals translocated from the second region. It is possible that not enough time has passed for the mtDNA to mutate, and therefore variation was not observed between the isolated DNA sample sequences. As previously mentioned, the DNA used to construct the reference sample was obtained from a specimen collected in a Zoo in Denmark. The origin of this specimen is not known. It is recommended that these regions be further investigated for the black rhinoceros. Samples from distinctly different geographic regions (where no translocation has occurred) should be taken for the black rhinoceros in order to determine if these regions can separate individuals based on where they originated from.

The D-loop and HV2 displayed variation between the white rhinoceros isolated DNA sample sequences and the reference sequence. These regions should therefore be investigated further in order to determining if they can infer geographic origin for white rhinoceros samples.

The translocation of the rhinoceros individuals, who produced samples for this investigation, between the different regions in South Africa may also provide insight as to why no investigated region could infer individualisation. Linacre & Tobe (2009) suggested that the D-loop and hypervariable regions allowed for human individualisation. The human population is significantly larger than the current rhinoceros populations. Rhinoceros mtDNA may therefore not have had sufficient time to allow for significant mutations to arise.

CHAPTER SIX

CONCLUSIONS AND FUTURE WORK

The aim of this project was to develop a forensic technique to study DNA from non-invasive samples in order to identity familial relations and for individualisation of the black and white rhinoceros. Primers were designed to amplify target regions of the black rhinoceros mtDNA genome. These target regions included regions of the *cytb*, D-loop, HV1, HV2, *16s rRNA,* and two regions of the *COI*. Five of these seven regions were compatible for the white rhinoceros (*COIi*, *COIii*, D-loop, HV2 and *cytb*). DNA was extracted from faecal, tissue and blood samples. Faecal samples served as the non-invasive source of DNA and tissue was chosen to serve as the positive control.

A novel biological assay was established for discriminating between the black and white rhinoceros. This was achieved by amplifying the *COIi* region and digesting it with *Hin*dIII. This resulted in the discrimination between the two species as the white rhinoceros contained the restriction enzyme and the black rhinoceros did not.

All investigated target regions allowed for the rhinoceros family to be identified. The *COIi*, *COIii*, HV2 and *cytb* regions allowed for the subspecies of rhinoceros to be identified. The *16s rRNA* and HV1 regions were only investigated for the black rhinoceros. These regions identified the subspecies of rhinoceros correctly however subsequent testing will need to be conducted on the white rhinoceros in order for this to be confirmed. The D-loop did not allow for subspecies determination as the white rhinoceros was incorrectly identified.

The isolated DNA sample sequences for the black rhinoceros differed to the reference sequence for all investigated target regions, while only the HV2 and D-loop differed for the white rhinoceros. A single SNP was observed for a black rhinoceros individual for the *COIi* region, while the HV1 region had a bp deletion for all individuals which resulted in a sequence difference of 54.2% when compared to the reference sequence. Intraspecies variation was observed in the HV2 and *cytb* region for the white rhinoceros. Three out of five isolated DNA sample sequences displayed variation for the HV2 region. These sequences were different to the reference sequence by 2.1%, 6.2% and 0.7%. One isolated DNA sample sequence displayed a 2% sequence difference compared to the other obtained sequences for the *cytb* gene.

Strict laboratory protocols were followed to ensure no cross contamination occurred. Black and white rhinoceros sequences differed indicating that samples were not contaminated. A larger size of faecal samples, which originated from different distinct geographical areas, could be used to re-examine the target regions investigated in this study in order to determine geographic location of samples could be determined. The target regions could be re-examined using a different source of DNA. This source could be tissue or blood samples. This would allow for confirmation of the lack of genetic diversity for the black and white rhinoceros in South Africa. Future studies should also include blind testing in order further validate sample identification.

From this study, it can be concluded that the target mtDNA regions investigated are not suitable for identifying individual black and white rhinoceros. However, they may have potential for establishing geographic origin. Faecal sample sequences were consistent with those for the positive control. Therefore, faecal samples proved to be a reliable non-invasive source of DNA.

Loss of genetic variation can result in decreased animal fitness. This would contribute to further declines in the population (Harley *et al.,* 2005). This investigation indicated a lack of genetic diversity for the black and white rhinoceros in South Africa. Therefore a closely monitored breeding system is required in order to assess the level of genetic diversity and to avoid inbreeding depression

REFERENCES

Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E. & Carelli, V. 2004, "The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool", *American Journal of Human Genetics*, vol. 75, no. 5, pp. 910-918.

Alacs, E.A., Georges, A., FitzSimmons, N.N. & Robertson, J. 2009, "DNA detective: a review of molecular approaches to wildlife forensics", *Forensic Science, Medicine, and Pathology,* vol. 6, no. 3, pp. 180-94.

Alibhai, S., Jewell, Z. & Towindo, S. 2001, "Effects of immobilization on fertility in female black rhino (*Diceros bicornis*)", *Journal of Zoology,* vol. 253, no. 3, pp.333–345.

Amin, R., Bramer, M. & Emslie, R. 2003, "Intelligent data analysis for conservation: experiments with rhino horn fingerprint identification", *Knowledge-Based Systems,* vol. 16, no. 5, pp. 329-336.

Anderson, S., De Bruijn, M., Coulson, A., Eperon, I., Sanger, F. & Young, I. 1982, "Complete sequence of bovine mitochondrial DNA", *Journal of Molecular Biology*, vol. 156, pp. 683- 717.

Anderson-Lederer, R., Linklater, W.L. & Ritchie, P.A. 2012, "Limited mitochondrial DNA variation within South Africa's black rhino (*Diceros bicornis minor*) population and implications for management", *African Journal of Ecology,* vol. 50, no. 4, pp. 404-413.

Andréasson, H., Asp, A., Alderborn, A., Gyllensten, U. & Allen, M. 2002, "Mitochondrial sequence analysis for forensic identification using pyrosequencing technology", *BioTechniques,* vol. 32, no. 1, pp. 124-126.

Aranishi, F. & Okimoto, T. 2005, "Sequence polymorphism in a novel noncoding region of Pacific oyster mitochondrial DNA", *Journal of Applied Genetics,* vol. 46, no. 2, pp. 201-206.

Ashley, M.V., Melnick, D.J. & Western, D. 1990, "Conservation Genetics of the black rhinoceros (*Diceros bicornis*), I: Evidence from the DNA mitochondrial of three populations", *Conservation Biology,* vol. 4, no. 1, pp. 71-77.

Baker, C.S., Cooke, J.G., Lavery, S., Dalebout, M.L., Yu, M., Funashi, N. 2007, "Estimating the number of whales entering trade using DNA profiling and capture-recapture analysis of market products", *Molecular Ecology*, vol. 16, pp. 2617-2626.

Ball, M.C., Pither, R., Manseau, M., Clark, J., Petersen, S.D., Kingston, S., Morrill, N. & Wilson, P. 2007, "Characterization of target nuclear DNA from faeces reduces technical issues associated with the assumptions of low-quality and quantity template", *Conservation Genetics,* vol. 8, no. 3, pp. 577-586.

Bataille, M., Crainic, K., Leterreux, M., Durigon, M. & de Mazancourt, P. 1999, "Multiplex amplification of mitochondrial DNA for human and species identification in forensic evaluation", *Forensic Science International,* vol. 99, no. 3, pp. 165-70.

Belay, G. & Mori, A. 2006, "Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, *Theropithecus gelada*, in Ethiopia", *Biochemical Systematics and Ecology,* vol. 34, no. 7, pp. 554-561.

Bendall, K.E., Macaulay, V.A. & Sykes, B.C. 1997, "Variable levels of a heteroplasmic point mutation in individual hair roots", *The American Journal of Human Genetics,* vol. 61, no. 6, pp. 1303-1308.

Berger, L. & Cunningham, C. 1994, "Phenotypic alterations, evolutionarily significant and structures, conservation rhino", *Conservation Biology,* vol. 8, no. 3, pp. 833–840.

Bradley, B.J., Stiller, M., Doran-Sheehy, D.M., Harris, T., Chapman, C.A., Vigilant, L. & Poinar, H. 2007, "Plant DNA sequences from feces: potential means for assessing diets of wild primates", *American Journal of Primatology*, vol. 69, pp. 1-7.

Bradley, R.D. & Baker, R.J. 2001, "American society of mammalogists a test of the genetic species concept: *cytochrome b* ", *Journal of Mammology,* vol. 82, no. 4, pp. 960-973.

Brown, S.M. & Houlden, B.A. 2000, "Isolation and characterization of microsatellite markers in the black rhinoceros (*Diceros bicornis*)", *Molecular Ecology*, vol. 8, pp. 1559–1561.

Brown, S.M. & Houlden, B.A. 2001, "Conservation genetics of the black rhinoceros (*Diceros bicornis*)", *Conservation Genetics,* vol. 1, pp. 365-370.

Brown, W.M. 1980, "Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis", Proceedings *of the National Academy of Sciences,* vol. 77, pp. 3605-3609.

Budowle, B., Wilson, M.R., DiZinno, J.A., Stauffer, C., Fasano, M.A., Holland, M.M. & Monson, K.L. 1999, "Mitochondrial DNA regions HVI and HVII population data", *Forensic Science International,* vol. 103, no. 1, pp. 23-35.

Cardillo, M., Mace, G.M., Jones, K.E., Bielby, J., Bininda-Emonds, O.R., Sechrest, W., Orme, C.D. & Purvis, A. 2005, "Multiple causes of high extinction risk in large mammal species", *Science (New York, N.Y.),* vol. 309, no. 5738, pp. 1239-1241.

Carracedo, A., Bär, W., Lincoln, P., Mayr, W., Morling, N., Olaisen, B., Schneider, P., Budowle, B., Brinkmann, B. & Gill, P. 2000, "DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing", *Forensic Science International,* vol. 110, no. 2, pp. 79-85.

Chui, Y., Esaw, B. & Laviolette, B. 1994, "Investigation of the metabolism of azaperone in the horse", *Journal of Chromatography B: Biomedical Sciences and Applications,* vol. 652, no. 1, pp. 23-33.

Coulombe, P.A. & Omary, M.B. 2002, "Hard and soft principles defining the structure, function and regulation of keratin intermediate filaments", *Current Opinion in Cell Biology,* vol. 14, no. 1, pp. 110-122.

Dalton, D.L. & Kotze, A. 2011, "DNA barcoding as a tool for species identification in three forensic wildlife cases in South Africa", *Forensic Science International,* vol. 207, pp. 1-3.

Department of Environmental Affairs. 2014, "Update on the rhino poaching statistics", [Online]: https://www.environment.gov.za/mediarelease/rhinopoaching_wildlifetrade Accessed on the 10th of January 2015.

Dieffenbach, C.W., Lowe, T.M. & Dveksler, G.S.. 1993, "General concepts for PCR primer design", *Genome Research*, vol. 3, pp. 30-37.

Emslie, R. 2012, "Ceratotherium simum" *IUCN Red List of Threatened Species*. [Online]: www.iucnredlist.org Accessed on the 6th of November 2013.

Emslie, R. & Brooks, M. 1999, "African Rhino. Status survey and conservation action plan", IUCN/SSC African Rhino Specialist Group. IUCN, Gland, Switzerland and Cambridge, UK, pp. 22.

Ferguson, J. & Swart M. 1997, "Conservation implications of genetic differentiation in Southern African populations of black rhinoceros (*Diceros bicornis*)", *Conservation Biology,* vol. 11, no. 1, pp. 79-83.

Ferris, S. D., Brown, W. M., Davisdon, W. S. & Wilson, A.C. 1981, "Extensive polymorphism in the mitochondrial DNA of apes", *Proceedings of the National Academy of Sciences,* vol. 78, pp. 6319-6323.

Frankham, R., Briscoe, D.A. & Ballou, J.D. 2002, "Introduction to conservation genetics", *Cambridge University Press*.

Fridman, C., Cardena, M.M.S.G., Kanto, E.A., Godinho, M.B.C. & Gonçalves, F.T. 2011, "SNPs in mitochondrial DNA coding region used to discriminate common sequences in HV1–HV2–HV3 region", *Forensic Science International: Genetics Supplement Series,* vol. 3, no. 1, pp. 75-76.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., Defelice, M., Lochner, A. & Aggart, M. 2002, "The structure of haplotype blocks in the human genome", *Science*, vol. 296, pp. 2225-2229.

Garnier, J.N., Bruford, M.W. & Goossens, B. 2001, "Mating system and reproductive skew in the black rhinoceros", *Molecular Ecology,* vol. 10, pp. 2031-2041.

Gerloff, U., Schlotterer, C., Rassmann, K., Rambold, I., Hohmann, G., Fruth, B. & Tautz, D. 1995, "Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*)", *Molecular Ecology,* vol. 4, pp. 515–518.

Gilpin, M.E. & Soulé, M.E. 1986, "Minimum viable populations: processes of species extinction", *Conservation Biology*: *The Science of Scarcity and Diversity,* pp 19-34.

Glenn, T.C., Staton, J.L., Vu, A.T., Davis, L.M., Bremer, J.R.A., Rhodes, W.E., Brisbin, I.L. & Sawyer, R.H. 2002, "Low mitochondrial DNA variation among American alligators and a novel non-coding region in crocodilians", *The Journal of Experimental Zoology,* vol. 294, no. 4, pp. 312-324.

Greenberg, B.D. & Aquadro, C.F. 1983, "Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals", *Genetics Society of America,* vol. 103, pp. 287-312.

Groves, C. & Grubb, P. 2011, "Ungulate taxonomy", *JHU Press*, pp. 317.

Gyllensten, U., Wharton, D., Josefsson, A. & Wilson, A.C. 1991, "Paternal inheritance of mitochondrial DNA in mice", *Nature*, vol. 352, pp. 255-257.

Hall, T.A. 1999, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT", *Nucleic Acids Symposium Series,* vol. 41, pp. 95-98.

Hanner, R. 2009, "Data Standards for BARCODE Records in INSDC (BRIs)", *Database Working Group, Consortium for the Barcode of Life Report*, pp. 1-4.

Hapke, A., Zinner, D. & Zischler, H. 2001, "Mitochondrial DNA variation in Eritrean hamadryas baboons (*Papio hamadryas hamadryas*): life history influences population genetic structure", *Behavioral Ecology Sociobiology*, vol. 50, pp. 483-492.

Harley, E.H., Baumgarten, I., Cunningham, J. & O'Ryan, C. 2005, "Genetic variation and population structure in remnant populations of black rhinoceros, *Diceros bicornis*, in Africa", *Molecular Ecology,* vol. 14, no. 10, pp. 2981-2990.

Heard, D.J., Olsen, J.H. & Stover, J. 1992, "Cardiopulmonary changes associated with chemical immobilization and recumbency in a white rhinoceros (*Ceratotherium simum*)", *Journal of Zoo and Wildlife Medicine,* pp. 197-200.

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. 2003, "Biological identifications through DNA barcodes", *Proceedings Biological Sciences / The Royal Society,* vol. 270, no. 1512, pp. 313-321.

Hoelzel, A. 1993, "Evolution by DNA turnover in the control region of vertebrate mitochondrial DNA", *Current Opinion in Genetics & Development,* vol. 3, no. 6, pp. 891-895.

Hsieh, H., Huang, L., Tsai, L., Kuo, Y., Meng, H., Linacre, A. & Lee, J.C. 2003, "Species identification of rhinoceros horns using the *cytochrome b* gene", *Forensic Science International,* vol. 136, no. 1-3, pp. 1-11.

Idaghdour, Y., Broderick, D. & Korrida, A. 2003, "Faeces as a source of DNA for molecular studies in a threatened population of great bustards", *Conservation Genetics,* vol. 4, pp. 789-792.

International Union for Conservation of Nature and Natural Resources. 2013, "IUCN Red List". Available [online]: http://www.iucnredlist.org Accessed on the 6th of November 2013.

Johns, G. C. & Avise, J. C. 1998, "A comparative summary of genetic distances in the vertebrates from the mitochondrial *cytochrome b* gene", *Molecular Biology and Evolution*, vol. 15, pp. 1481–1490.

Khanuja, S.P., Shasany, A.K., Darokar, M.P. & Kumar, S. 1999, "Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils", *Plant Molecular Biology Reporter,* vol. 17, no. 1, pp. 74.

Kock, R. & Pearce, P. 1985, "Anaesthesia in zoo ungulates", *Veterinary Anaesthesia and Analgesia,* vol. 13, no. 1, pp. 59-88.

Kraytsberg, Y., Schwartz, M., Brown, T.A., Ebralidse, K., Kunz, W.S., Clayton, D.A., Vissing, J. & Khrapko, K. 2004, "Recombination of human mitochondrial DNA", *Science (New York, N.Y.),* vol. 304, no. 5673, pp. 981.

Kwok, P.W. 2002, "Single nucleotide polymorphisms: methods and protocols", *AACC Press*, San Franscisco, CA, USA.

Leader-Williams, N. 1992, "World trade in rhino horn: a review"*, Traffic*, pp. 40.

Lee-Thorp, J., van der Merwe, N. & Armstrong, R. 1992, "Final project report: ZA 309 Source area determination of rhino horn by isotopic analysis", *Unpublished report*.

Li, J., Ye, Y., Wu, C., Guo, B. & Gul, Y. 2014, "Genetic diversity and population structure of *Sepiella japonica* (Mollusca: *Cephalopoda: Decapoda*) inferred by mitochondrial DNA (*COI*) variations", *Biochemical Systematics and Ecology,* vol. 56, pp. 8-15.

Linacre, AS. & Tobe, S. 2009: "Forensic Science in wildlife investigations international forensic science and investigation series", Edited by: Linacre A. Boca Ratan: CRC Press.

Lindeque, M. 1990, "The case for dehorning the black rhinoceros in Namibia"*, South African Journal of Science,* vol. 86, pp. 226-227.

Liu, P., Seo, TS., Beyor, N., Shin, K., Scherer, JR. & Mathies, RA. 2007 "Integrated portable PCR-capillary electrophoresis microsystem for rapid forensicshort tandem repeat typing", *Analytical Chemistry*, vol 79, no. 5, pp. 1881-1889.

Mallet, J & Willmott, K. 2003, "Shortcuts in systematics? A commentary on DNA-based taxonomy", *Trends in Ecology & Evolution,* vol. 18, no. 2, pp. 63-65.

Marín, A., Fujimoto, T. & Arai, K. 2013, "Genetic structure of the Peruvian scallop *Argopecten purpuratus* inferred from mitochondrial and nuclear DNA variation", *Marine Genomics,* vol. 9, pp. 1-8.

Martin, E.B. 1984, "North Yemen and the rhino horn trade today", *Swara 7*, pp 28–33.

Martin, E.B. 1990, "Medicines from Chinese treasures", *Pachyderm13,*pp. 12-13.

Martin, E.B. & Martin, C.B. 1982, "Run rhino run", *Chatto and Windus,* pp.136.

Marrero, P., Fregel, R., Cabrera, VM. & Nogales, M. 2009. "Extraction of high-quality host DNA from faeces and regurgitated seeds: a Useful Tool for Vertebrate Ecological Studies", *Biological Research,* vol. 42, pp. 147–151.

Meyer, S., Weiss, G. & Haeseler, A.V. 1999, "Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA", *Genetics Society of America,* vol. 152, pp. 1103-1110.

Miller, C.R. & Waits, L.P. 2003, "The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conservation", *Proceedings of the National Academy of Sciences*, vol. 100, pp. 4334–4339.

Milliken, T. & Shaw, J. 2012, "The South Africa–Vietnam rhino horn trade nexus", Traffic.

Mills, L.S. 2012, "Conservation of wildlife populations: demography, genetics, and management", *John Wiley & Sons*.

Morales, J.C., Andau, P.M., Supriatna, J., Zainuddin, Z. & Melnick, D.J. 1997, "Mitochondrial DNA variability and conservation genetics of the Sumatran rhinoceros", *Conservation Biology,* vol. 11, no. 2, pp. 539-543.

Morin, P.A., Luikart, G. & Wayne, R.K. 2004, "SNPs in ecology, evolution and conservation", *TRENDS in Ecology and Evolution*, vol. 19, pp. 208-216.

Moritz, C. & Cicero, C. 2004, "DNA barcoding: promise and pitfalls." *PLOS Biology,* vol. 2, no. 10, pp.1529–1531.

Nabholz, B., Gle, S. & Galtier, N. 2008, "Strong variations of mitochondrial mutation rate across mammals: the longevity hypothesis", *Molecular Biology and Evolution,* vol. 25, no. 1, pp. 120-130.

Nielsen, L. 2007, "Genetic diversity of black and white rhinoceros using RAPD and microsatellite genetic markers: can faecal samples serve as a non-invasive tool?", North Grafton, Cummings School of Veterinary Medicine at Tufts University (Thesis), pp. 1-10.

Nielsen, L., Meehan-Meola, D., Kilbourn, A. & Alcivar-Warren, A. 2008, "Characterization of microsatellite loci in the black rhinoceros (*Diceros bicornis*) and white rhinoceros

(*Ceratotherium simum*): their use for cross-species amplification and differentiation between the two species", *Conservation Genetics,* vol. 9, no. 1, pp. 239-242.

Nowell, K., Chyi, W. & Pei, C. 1992, "The Horns of a Dilemma: The Market for Rhino Horn in Taiwan: Final Report for WWF Project II3637. 03", *Traffic International*, pp. 44.

O'Ryan, C.O., Flamand, J.R.B. & Harley, E.H. 1994, "DNA variation mitochondrial in black rhinoceros (*Diceros bicornis*): conservation management implications", *Conservation Biology,* vol. 8, no. 2, pp. 495-500.

Ogden, R., Dawnay, N. & McEwing, R. 2009, "Wildlife DNA forensics – bridging the gap between conservation genetics and law enforcement", *Endangered Species Research,* vol. 9, pp. 179–195.

Parsons, T.J. & Coble, M.D. 2001, "Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome", *Croatian Medical Journal*, vol. 42, pp. 304–309.

Peppin, L., McEwing, R., Ogden, R., Hermes, R.,, Harper, C., Guthrie, A. & Carvalho, R. 2010, "Molecular sexing of African rhinoceros", *Conservation Genetics,* vol. 11, no. 3, pp. 1181–1184.

Pienaar, D., Hall-Martin, A. & Hitchins, P. 1991, "Horn growth rates of free-ranging white and black rhinoceros", *Koedoe,* vol. 34, no. 2, pp. 97-105.

Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. 2005, "Genotyping errors: causes consequences, and solutions", *Nature Reviews Genetics*, vol. 6, pp 847-859.

Potgieter, D.W. 1995, "Contraband: South Africa and the international trade in ivory and rhino horn", *Queillerie, Cape Town, South Africa,* pp.195.

Roussel, Y.E. & Patenaude, R. 1975, "Some physiological effects of M99 etorphine on immobilized free-ranging moose", *The Journal of Wildlife Management,* vol. 39, no. 3, pp. 634-636.

Roux, K. 1995. "Optimization and troubleshooting in PCR." *Cold Spring Harbor Laboratory Press*: pp. 185–194.

Saccheri, I.J., Wilson, I.J., Nichols, R.A., Bruford, M.W. & Brakefield, P.M. 1999, "Inbreeding of bottlenecked butterfly populations. Estimation using the likelihood of changes in marker allele frequencies", *Genetics,* vol. 151, no. 3, pp. 1053-1063.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. & Arnheim, N. 1985, "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia", *Science (New York, N.Y.),* vol. 230, no. 4732, pp. 1350-1354.

Save the Rhino International. 2015. Rhino population figures. Available [Online]: http://www.savetherhino.org/rhino_info/rhino_population_figures Accessed on the 10th of January 2015.

Stuart, C. & Stuart T. 2006, "Field guide to the lager mammals of Africa", *Struik Publishers*, Cape Town.

Taberlet, P. & Luikart, G. 1999, "Non-invasive genetic sampling and individual identification", *Biological Journal of the Linnean Society,* vol. 68, pp. 41-55.

Van Coeverden de Groot, P.J., Putnam, A.S., Erb, P., Scott C., Melnic, D., O'Ryan, C. & Boag, P.T. 2011, "Conservation genetics of the black rhinoceros *Diceros bicornis bicornis* in Namibia" *Conservation Genetics*, no. 12, pp. 783–792.

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A.C. 1991, "African populations and the evolution of human mitochondrial DNA", *Science (New York, N.Y.),* vol. 253, no. 5027, pp. 1503-1507.

Wasser, S., Houston, C., Koehler, G., Cadd, G. & Fain, S. 1997, "Techniques for application of faecal DNA methods to field studies of ursids", *Molecular Ecology,* vol. 6, no. 11, pp.1091–1097.

Wehausen, J., Ramey, R. & Epps, C. 2004, "Experiments in DNA extraction and PCR amplification from bighorn sheep faeces: the importance of DNA extraction method", *The Journal of Heredity,* vol. 95, no. 6, pp. 503–509.

Wenger, S., Boardman, W., Buss, P., Govender, D. & Foggin, C. 2007, "The cardiopulmonary effects of etorphine, azaperone, detomidine, and butorphanol in field-anesthetized white rhinoceroses (*Ceratotherium simum*)", *Journal of Zoo and Wildlife Medicine,* vol. 38, no. 3, pp. 380-387.

Western, D. 1982, "Dehorn or not dehorn", *Swara,* vol. 5, pp. 22-23.

Willerslev, E., Gilbert, M.T.P., Binladen, J., Ho, S.Y.W., Campo, P.F., Ratan, A., Tomsho, L.P., da Fonseca, R.R., Sher, A., Kuznetsova, T.V., Nowak-Kemp, M., Roth, T.L., Miller, W. & Schuster, S.C. 2009, **"**Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution**",** *BMC Evolutionary Biology,* vol. 9, no 95, pp. 1471-2148.

Wilson, A.C. & Cann, R.L. 1983, "Length mutations in human mitochondrial DNA", *Genetics Society of America,* vol. 165, no. 104, pp. 699-711.

Wilson, A. C., Cann, R. L., Carr, A. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. & Stoneking, M. 1985, "Mitochondrial DNA and two perspectives on evolutionary genetics", *Biological Journal of the Linnear Society*, vol. 26, pp. 375-400.

World Wildlife Fund, 2013a. Black rhinoceros. Available [Online]: http://wwf.panda.org/what_we_do/endangered_species/rhinoceros/african_rhinos/black_rhin oceros/ Accessed on the 6th of November 2013.

World Wildlife Fund, 2013b. Northern white rhinoceros. Available [Online]: http://wwf.panda.org/what_we_do/endangered_species/rhinoceros/african_rhinos/white_rhin oceros/northern_white_rhinoceros/ Accessed on the 6th of November 2013.

Yamashiro, A., Yamashiro, T., Baba, M., Endo, A. & Kamada, M. 2010, "Species identification based on the faecal DNA samples of the Japanese serow (*Capricornis crispus*)", *Conservation Genetics Resources,* vol. 2, pp. 409-414.

Xu, X. & Arnason, U. 1997, "The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum* and comparison with the mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis*", *Moecular Phylogenetic Evolution,* vol.7, pp. 189-194.

Zymo Research. 2014, "ZR fecal DNA MiniPrep™ protocol". [Online]: http://www.zymoresearch.com/dna/microbial-environmental-dna-isolation-1/soil-fecal-plant-dna/zr-fecal-dna-miniprep Accessed on the 21st of October 2014.

APPENDIX
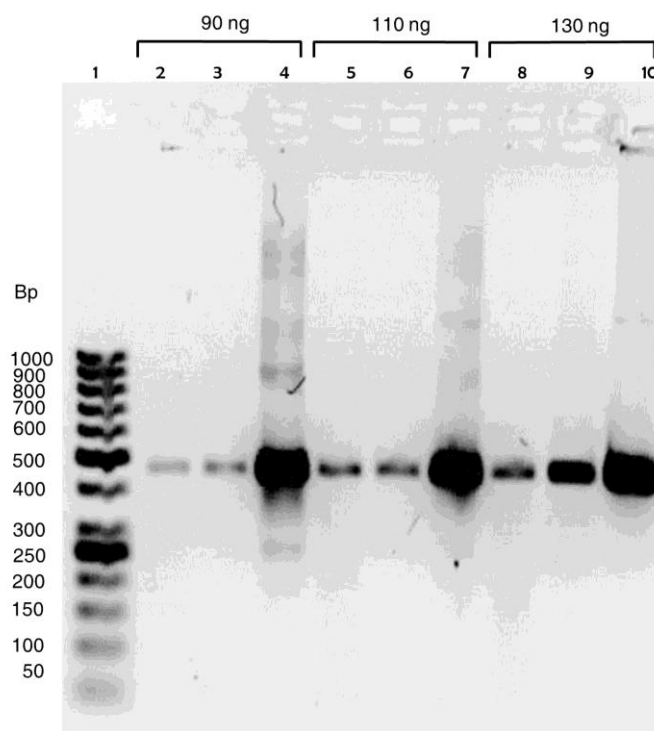
Appendix 1: Varying template concentration for the optimisation of the D-Loop regions amplification. Lane 1: O'Gene Ruler 1 kb Plus DNA Ladder, Lane 2: black (1, F), Lane 3: black (2, F), Lane 4: black (1, T), Lane 5: black (1, F), Lane 6: black (2, F), Lane 8: black (1, T), Lane 8: black (1, F), Lane 9: black (2, F), Lane 10: black (1, T). Agarose (1% w/v) gel electrophoresis was used.
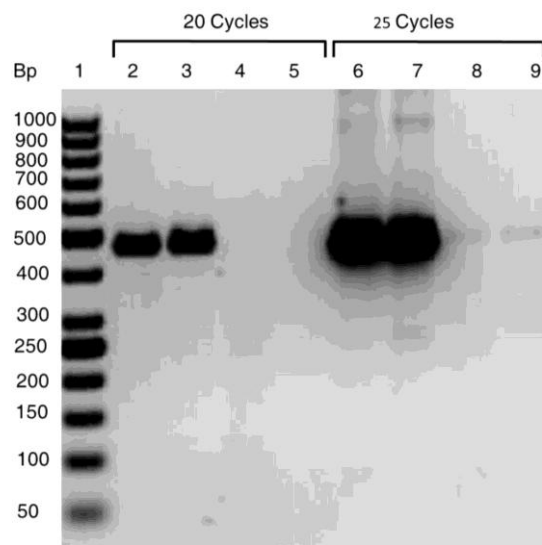
Appendix 2: Varying cycle numbers for the optimisation of the D-Loop regions amplification. Lane 1: GeneRuler™ 50bp DNA Ladder; Lane 2: black (1, T), Lane 3: black (2, T), Lane 4: black (1, F), Lane 5: black (2, F), Lane 6: black (1, T), Lane 7: black (2, T), Lane 8: black (1, F), Lane 9: black (2, F). Agarose (1% w/v) gel electrophoresis was used.
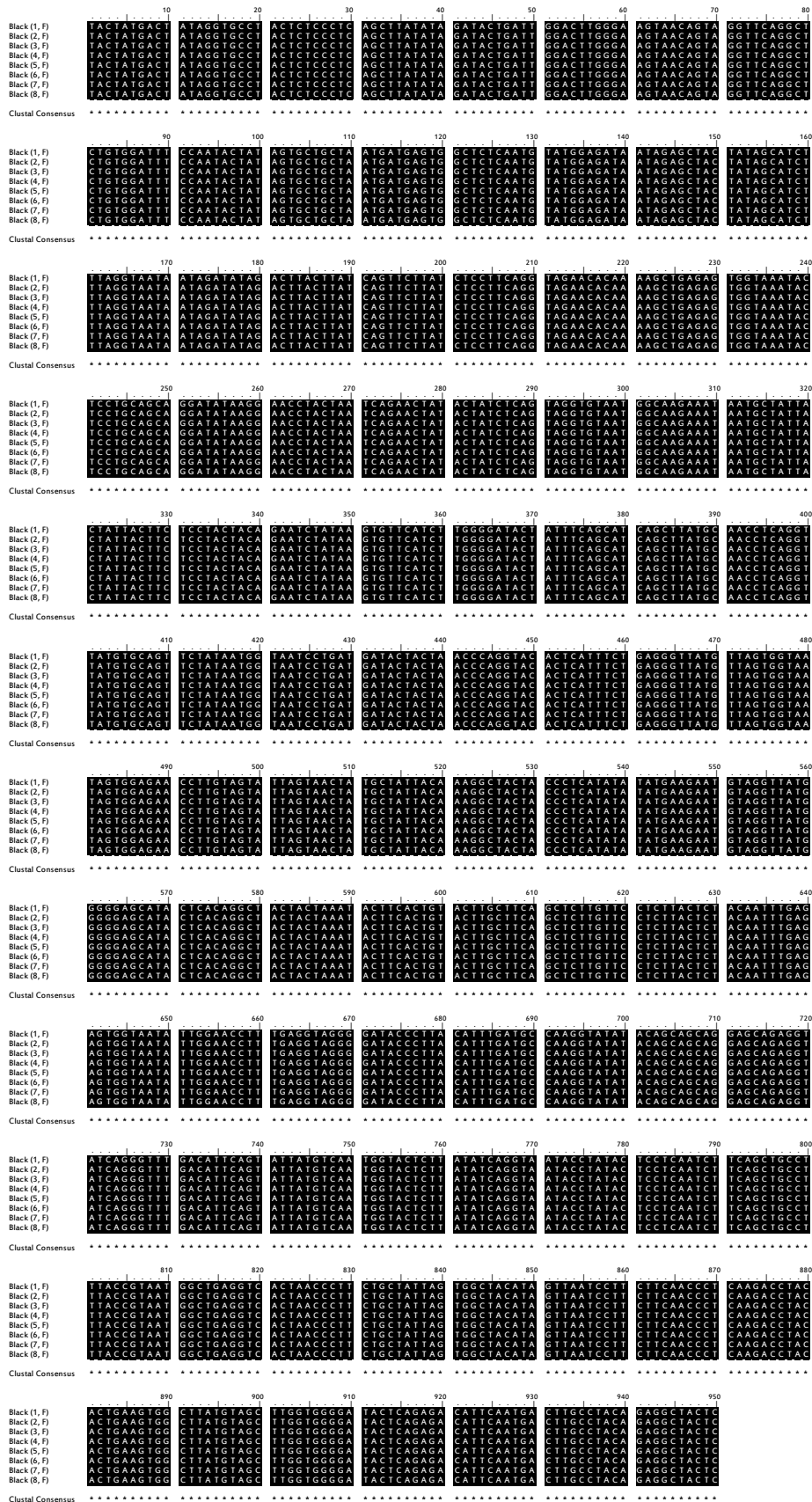
Appendix 3: A multiple sequence alignment for the unknown band present in the amplification of the HV2 region in black rhinoceros. This includes a black Rhinoceros reference sequence (FJ905814.1).

Table AI: The black rhinoceros sequence identity matrix for the *COIi* region.

	FJ905 814.1	Black (1, F)	Black (2, T)	Black (2, F)	Black (3, F)	Black (4, F)	Black (5, F)	Black (6, F)	Black (7, F)	Black (8, F)
FJ905 814.1	ID	0.987	0.844	0.987	0.987	0.987	0.98	0.987	0.987	ID
Black (1, F)	0.987	ID	0.844	0.987	0.987	0.987	0.987	0.987	0.987	0.987
Black (2, T)	0.844	0.844	ID	0.85	0.85	0.85	0.844	0.85	0.85	0.844
Black (2, F)	0.987	0.987	0.850	ID	1	1	0.993	1	1	0.987
Black (3, F)	0.987	0.987	0.850	1	ID	1	0.993	1	1	0.987
Black (4, F)	0.987	0.987	0.850	1	1	ID	0.993	1	1	0.987
Black (5, F)	0.98	0.987	0.844	0.993	0.993	0.993	ID	0.993	0.993	0.98
Black (6, F)	0.987	0.987	0.850	1	1	1	0.993	ID	1	0.987
Black (7, F)	0.987	0.987	0.850	1	1	1	0.993	1	ID	0.987
Black (8, F)	ID	0.987	0.844	0.987	0.987	0.987	0.980	0.987	0.987	ID

Table AII: The white rhinoceros sequence identity matrix for the *COIi* region.

	Y07 726.1	White (1, T)	White (1, F)	White (2, F)	White (3, F)	White (4, F)	White (5, T)	White (5, F)	White (6, F)	White (7, F)	White (8, F)
Y077 26.1	ID	1	1	1	1	1	1	1	1	1	1
White (1, T)	1	ID	1	1	1	1	1	1	1	1	1
White (1, F)	1	1	ID	1	1	1	1	1	1	1	1
White (2, F)	1	1	1	ID	1	1	1	1	1	1	1
White (3, F)	1	1	1	1	ID	1	1	1	1	1	1
White (4, F)	1	1	1	1	1	ID	1	1	1	1	1
White (5, T)	1	1	1	1	1	1	ID	1	1	1	1
White (5, F)	1	1	1	1	1	1	1	ID	1	1	1
White (6, F)	1	1	1	1	1	1	1	1	ID	1	1
White (7, F)	1	1	1	1	1	1	1	1	1	ID	1

Table AIII: The black rhinoceros sequence identity matrix for the *COIii* region.

	FJ905 814.1	Black (1, T)	Black (1, F)	Black (2, F)	Black (3, F)	Black (4, F)	Black (5, F)	Black (6, F)	Black (7, F)	Black (8, F)
FJ905 814.1	ID	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993
Black (1, T)	0.993	ID	1	1	1	1	1	1	1	1
Black (1, F)	0.993	1	ID	1	1	1	1	1	1	1
Black (2, F)	0.993	1	1	ID	1	1	1	1	1	1
Black (3, F)	0.993	1	1	1	ID	1	1	1	1	1
Black (4, F)	0.993	1	1	1	1	ID	1	1	1	1
Black (5, F)	0.993	1	1	1	1	1	ID	1	1	1
Black (6, F)	0.993	1	1	1	1	1	1	ID	1	1
Black (7, F)	0.993	1	1	1	1	1	1	1	ID	1
Black (8, F)	0.993	1	1	1	1	1	1	1	1	ID

Table AIV: The white rhinoceros sequence identity matrix for the *COIii* region.

	Y07726.1	White (1, T)	White (1, F)	White (3, F)	White (4, F)	White (5, F)	White (6, F)
Y07726.1	ID	1	1	1	1	1	1
White (1, T)	1	ID	1	1	1	1	1
White (1, F)	1	1	ID	1	1	1	1
White (3, F)	1	1	1	ID	1	1	1
White (4, F)	1	1	1	1	ID	1	1
White (5, F)	1	1	1	1	1	ID	1
White (6, F)	1	1	1	1	1	1	ID

Table AV: The black rhinoceros sequence identity matrix for the D-loop region.

	FJ9058 14.1	Black (1, T)	Black (1, F)	Black (2, F)	Black (3, F)	Black (5, F)	Black (6, F)	Black (7, F)	Black (8, F)
FJ9058 14.1	ID	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945
Black (1, T)	0.945	ID	1	1	1	1	1	1	1
Black (1, F)	0.945	1	ID	1	1	1	1	1	1
Black (2, F)	0.945	1	1	ID	1	1	1	1	1
Black (3, F)	0.945	1	1	1	ID	1	1	1	1
Black (5, F)	0.945	1	1	1	1	ID	1	1	1
Black (6, F)	0.945	1	1	1	1	1	ID	1	1
Black (7, F)	0.945	1	1	1	1	1	1	ID	1
Black (8, F)	0.945	1	1	1	1	1	1	1	ID

Table AVI: The white rhinoceros sequence identity matrix for the D-loop region.

	Y07726.1	White (1, F)	White (4, F)	White (5, F)	White (6, F)	White (8, F)
Y07726.1	ID	0.845	0.845	0.845	0.845	0.845
White (1, F)	0.845	ID	1	1	1	1
White (4, F)	0.845	1	ID	1	1	1
White (5, F)	0.845	1	1	ID	1	1
White (6, F)	0.845	1	1	1	ID	1
White (8, F)	0.845	1	1	1	1	ID

Table AVII: The black rhinoceros sequence identity matrix for the HV1 region.

	FJ9058 14.1	Black (1, T)	Black (1, F)	Black (2, F)	Black (3, F)	Black (4, F)	Black (5, F)	Black (6, F)	Black (8, F)
FJ9058 14.1	ID	0.47	0.47	0.458	0.458	0.458	0.458	0.458	0.458
Black (1, T)	0.47	ID	1	0.974	0.974	0.974	0.974	0.974	0.974
Black (1, F)	0.47	1	ID	0.974	0.974	0.974	0.974	0.974	0.974
Black (2, F)	0.458	0.974	0.974	ID	1	1	1	1	1
Black (3, F)	0.458	0.974	0.974	1	ID	1	1	1	1
Black (4, F)	0.458	0.974	0.974	1	1	ID	1	1	1
Black (5, F)	0.458	0.974	0.974	1	1	1	ID	1	1
Black (6, F)	0.458	0.974	0.974	1	1	1	1	ID	1
Black (8, F)	0.458	0.974	0.974	1	1	1	1	1	ID

Table AVIII: The black rhinoceros sequence identity matrix for the *16s rRNA* region.

	FJ905814.1	Black (1, T)	Black (2, T)	Black (2, F)	Black (3, F)	Black (4, F)	Black (6, F)
FJ905814.1	ID	0.993	0.993	0.993	0.993	0.993	0.993
Black (1, T)	0.993	ID	1	1	1	1	1
Black (2, T)	0.993	1	ID	1	1	1	1
Black (2, F)	0.993	1	1	ID	1	1	1
Black (3, F)	0.993	1	1	1	ID	1	1
Black (4, F)	0.993	1	1	1	1	ID	1
Black (6, F)	0.993	1	1	1	1	1	ID

Table AIX: The black rhinoceros sequence identity matrix for the HV2 region.

	FJ905 814.1	Black (1, T)	Black (1, F)	Black (2, F)	Black (3, F)	Black (4, F)	Black (5, F)	Black (6, F)	Black (7, F)	Black (8, F)
FJ905 814.1	ID	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.986	0.986
Black (1, T)	0.986	ID	1	1	1	1	1	1	1	1
Black (1, F)	0.986	1	ID	1	1	1	1	1	1	1
Black (2, F)	0.986	1	1	ID	1	1	1	1	1	1
Black (3, F)	0.986	1	1	1	ID	1	1	1	1	1
Black (4, F)	0.986	1	1	1	1	ID	1	1	1	1
Black (5, F)	0.986	1	1	1	1	1	ID	1	1	1
Black (6, F)	0.986	1	1	1	1	1	1	ID	1	1
Black (7, F)	0.986	1	1	1	1	1	1	1	ID	1
Black (8, F)	0.986	1	1	1	1	1	1	1	1	ID

Table AX: The white rhinoceros sequence identity matrix for the HV2 region.

	Y07726.1	White (1, T)	White (2, F)	White (3, F)	White (5, T)	White (6, F)
Y07726.1	ID	1	0.979	0.938	1	0.993
White (1, T)	1	ID	0.979	0.938	1	0.993
White (2, F)	0.979	0.979	ID	0.92	0.979	0.977
White (3, F)	0.938	0.938	0.92	ID	0.938	0.931
White (5, T)	1	1	0.979	0.938	ID	0.993
White (6, F)	0.993	0.993	0.977	0.931	0.993	ID

Table AXI: The black rhinoceros sequence identity matrix for the *cytb* region.

	FJ905 814.1	Black (1, T)	Black (1, F)	Black (2, F)	Black (3, F)	Black (4, F)	Black (5, F)	Black (6, F)	Black (7, F)	Black (8, F)
FJ905 814.1	ID	0.973	0.973	0.973	0.973	0.973	0.973	0.973	0.973	0.973
Black (1, T)	0.973	ID	1	1	1	1	1	1	1	1
Black (1, F)	0.973	1	ID	1	1	1	1	1	1	1
Black (2, F)	0.973	1	1	ID	1	1	1	1	1	1
Black (3, F)	0.973	1	1	1	ID	1	1	1	1	1
Black (4, F)	0.973	1	1	1	1	ID	1	1	1	1
Black (5, F)	0.973	1	1	1	1	1	ID	1	1	1
Black (6, F)	0.973	1	1	1	1	1	1	ID	1	1
Black (7, F)	0.973	1	1	1	1	1	1	1	ID	1
Black (8, F)	0.973	1	1	1	1	1	1	1	1	ID

Table AXII: The white rhinoceros sequence identity matrix for the *cytb* region.

	Y07726. 1	White (1, F)	White (2, F)	White (3, F)	White (5, T)	White (5, F)	White (7, F)
Y07726.1	ID	1	1	0.98	1	1	1
White (1, F)	1	ID	1	0.98	1	1	1
White (2, F)	1	1	ID	0.98	1	1	1
White (3, F)	0.98	0.98	0.98	ID	0.98	0.98	0.98
White (5, T)	1	1	1	0.98	ID	1	1
White (5, F)	1	1	1	0.98	1	ID	1
White (7, F)	1	1	1	0.98	1	1	ID