An individual-based approach to genetic management in the game industry, with specific reference to parentage determination in free-ranging populations

by

Karen Ehlers

THESIS
Submitted in fulfilment of the requirements for the degree
of

Philosophiae Doctor

in

Zoology

in the

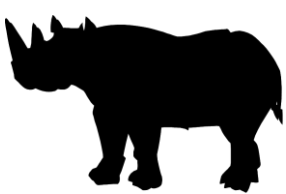
FACULTY OF SCIENCE & AGRICULTURE
(School of Molecular & Life Sciences)

at the

UNIVERSITY OF LIMPOPO

Promoter: Prof. J.P. Grobler (UL)
Co-promoter: Prof. A. Kotze (NZG)
Co-promoter: Prof. J.J. Spies (UFS)
Co-promoter: Dr. W. Luus-Powell (UL)

2012



Declaration

I declare that the thesis hereby submitted to the University of Limpopo, for the degree of *Philosophiae Doctor* in Zoology has not previously been submitted by me for a degree at this or any other university; that it is my work in design and in execution, and that all material contained herein has been duly acknowledged.

20/04/2012

K. Ehlers (Mrs) **Date**

[redacted]

Table of Contents

Chapter 2 – DNA profiling

Chapter 3 – Population structure

Chapter 4 – Individual identification

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to thank Prof. Paul Grobler, my promoter and co-worker, for his patience, time, guidance, expertise and support throughout my study.

Prof. Antoinette Kotze, my co-promoter for initiating this project and her input throughout this study.

Prof. Johan Spies, my co-promoter and boss, for his input and assistance.

Dr. Wilmien Luus-Powell, my support system at the University of Limpopo, with all her help regarding the registration and administration of my study.

My deepest gratitude to the following people and institutions who contributed either directly or indirectly to the completion of my study:

Mrs. Annelize Steyn
Mr. Brad Fick
Dorcas Lekgethisho
Agricultural Research Council
National Zoological Gardens of South Africa
University of the Free State
University of Limpopo

1.1 Introduction

Genetic management should form a vital part of overall strategies for the conservation of wildlife populations. This is especially true in South Africa, where the development of game farms and small reserves has led to small and fragmented populations. In such groups, both population structure and gene flow is largely dependent on artificial management (Caballero *et al*. 2010). To ensure the long term survival of these fragments, populations should thus be managed using sound scientific strategies.

The use of DNA markers and associated approaches for statistical analysis and management is well established in the field of conservation genetics. However, published studies in conservation genetics almost invariably follow a population approach (Caballero *et al*. 2010). Individual-based approaches similar to the intensive management strategies applied to domesticated species have been implemented for some endangered species and to manage populations in zoological gardens, but have not been developed for most wildlife species. Individual-based management strategies enable the farmer or conservation manager to apply scientifically founded programs to ensure that programs are sustainable and to reduce the risk of problems associated with inbreeding. The need for individual-based strategies to supplement population-based approaches for South African wildlife was first proposed by Kotze *et al.* (2004). In the current study, the feasibility of using DNA markers as a tool for individual-based wildlife management will therefore be further investigated. Compared to domesticated animals, a significant constraint during the genetic management of free-ranging wildlife is that individuals are in most instances not identifiable and breeding events are not observed. Information obtained from DNA profiles will assist managers in the decision making process by providing currently unavailable data on parentage and other relationships within populations. Furthermore, the establishment of genetic databases for wildlife will contribute to solving investigations related to cases of poaching (which is currently a very serious problem, particularly so for rhino species). In cases of poaching where biological evidence is collected from suspects by law-enforcement agencies, the DNA profile obtained from the evidence can be compared to the reference sample of the poached animal in the database and the identity of the animal can be proven in a court of law.

This thesis will focus on the use of DNA markers for the genetic management of wild black- and white rhino in four different reserves in South Africa. Short tandem repeat (STR) markers - that are the most suitable markers for answering specific questions relevant to population genetic management of rhino - will be investigated. Specifically, the suitability of the markers for individual identification, parentage verification and for determining of kinship between individuals will be tested. The application of the results obtained in future management decisions regarding translocation and breeding programs are then discussed.

In this chapter, the origin and size of the wildlife industry in South Africa and the potential contribution of genetic management to the conservation of wildlife species are first reviewed. Specific emphasis is placed on aspects that are relevant to the management of rhino species. Background on the history of rhino, and genetic studies related to rhino species by other authors are addressed. Information regarding molecular markers that can be used as a tool in conservation and the challenges it entails is also reviewed. In Chapter 2, the sampling and methods used to obtain DNA profiles of all the individuals included in this study is described. This chapter also include the DNA profiling analysis process and the detection of genotyping errors. The population structure of the rhino populations is addressed in Chapter 3, followed by a discussion of the individual identification of rhino in Chapter 4. The process of parentage verification and verifying siblings is examined in Chapter 5, and in Chapter 6 a summative discussion, conclusions and recommendations on the management application of all the information obtained in the preceding chapters are presented. Chapter 7 is a synoptic summary of this thesis, with Chapter 8 listing the literature cited.

1.2 The South African Game Industry

The South African game industry is a major enterprise in South Africa. Wildlife on private lands has shown a considerable growth in recent years due to various political and legal developments, for example, the decline in agricultural subsidies and the Game Theft Act of 1991 (Cousins *et al.* 2010).

1.2.1 Private ownership of wildlife

South Africa was one of the first countries in the world to allow the private ownership of wildlife species. Privatization of wildlife has been shown to advance conservation objectives (McGranahan 2008) and private landowners in South Africa have initiated recovery actions that lead to the saving of species, with the conservation of black wildebeest (*Connochaetes gnou*) being a prime example (Cousins *et al*. 2010). Private ownership also extends to endangered species such as white rhino (*Ceratotherium simum*). Prior to 1960, wildlife was the property of the state, with the state responsible for regulating and handling problem-animal control as well as the issuing of hunting licenses. There was no official incentive for farmers to encourage wildlife "ranching" (Emslie and Brooks 1999; Muir-Leresche and Nelson 2000), although a few individual private landowners did initiate significant conservation programs (McGranahan 2008).

Current South African wildlife legislation gives greater control of wildlife populations to private owners (Cousins *et al*. 2010). Provided that specific conditions are met, ownership of game animals is possible. Furthermore, if regulations for enclosure are met and approval is granted, oversight of hunting activities is transferred to the land owner and no hunting permits are then required for hunting (in most provinces). This led to a rapid growth of wildlife farming and expansion of the number of wildlife populations. Since the economics of wildlife farming made it more attractive than ordinary farming, land owners often formed conservancies which involved a group of land owners joining together to manage their land as a joint wildlife unit (Muir-Leresche and Nelson 2000; Cousins *et al.* 2010).

There are still arguments that private conservation is based on markets and that the volatile nature of markets means that long-term conservation cannot be guaranteed. However, experience thus far showed that markets have consistently supported the private supply of wildlife and habitat over a period of more than four decades (Krug 2001).

1.2.2 The impact of privatization on wildlife species

There are over 10 000 wildlife farms and 4 000 mixed wildlife/lifestock enterprises in South Africa (Cousins *et al*. 2010). Unfortunately, the isolation of many properties with game proof-fences, and extra-limital translocation, led to the formation of many isolated fragmented wildlife populations, often with low numbers of individuals in each population. These smaller populations have created a new set of problems in wildlife management.

Fragmentation and habitat destruction are most likely the major cause of the increase in the rate of species extinction in recent decades. Many species, due to the increased confinement and smaller population sizes, are placed at risk of inbreeding and loss of genetic variation because gene flow between such areas is often limited or completely absent. A population or species deprived of genetic variation may be more susceptible to diseases and may have lost the ability to survive climatic extremes, pollutants, diseases, pests and parasites (Frankham 1995; Hedrick and Kalinowski 2000). Inbred populations may also show reduced juvenile survival compared to outbred populations which is a growing concern for many endangered species conservation programmes (Frankham 1995; Spong *et al*. 2000).

There are a number of unresolved problems with population fragmentation, for example, there is insufficient data available on the effects of management actions on fragmented populations. Research results need to be translated into tools and guidelines for wildlife owners, in such a way that it is suitable for management decisions and can be practically implemented (Henly *et al*. 2004).

1.3 Conservation genetics and genetic management

Wallis (1994) stated that the central aim of conservation is the maintenance and eventual restoration of natural biodiversity. An endangered species will not recover if all the different entities such as genetics, ecology, behaviour, systematic and biogeography are not taken into consideration in the management plan of the species. For example, if only the habitat problem of the species is rectified, it is not to say the species will recover on its own, since inbreeding depression can have an effect if the population size is too low. If only genetic management is considered, the

numbers may be increased for a short time, but the increased population size will not be able to survive if their habitat is not restored as well. Habitat protection and enhancement, and species management, usually go hand in hand, though with varying degrees of emphasis on each. Conservation genetics is concerned with populations and individuals, and it thus forms a part of the overall conservation assessment process (Wallis 1994).

Conservation genetics deals with genetic factors that affect extinction risk. Genetic management is required to minimize these risks. Genetic tools, such as DNA markers and appropriate statistical approaches, can aid conservation efforts by minimizing inbreeding and loss of genetic diversity; identifying populations of concern; resolving population structures; resolving taxonomic uncertainties; defining management units within species; detecting hybridization; defining sites for reintroduction; choosing the best populations for reintroduction; helping to understand species biology; and be used in forensics (Frankham *et al*. 2002).

Comparative genetic analysis presents an objective means of assessing diversity at various levels. It can form a baseline for informed management decision in conjunction with other biological, economic and social considerations. In short, conservation geneticists try to maximize outbreeding to avoid inbreeding depression, maximize the population size to avoid loss of genetic variation by drift and to maximize subspecies and species diversity (Wallis 1994). Knowledge of a populations' history and genetic structure, will be useful to develop sound conservation plans for endangered species (Zhang *et al.* 2002). According to Hedrick *et al.* (2001) for the effective conservation of endangered species, it is important to firstly identify which taxa within closely related groups are different species. The next step is to identify evolutionarily significant units (ESUs) within species, in other words, populations that are evolving independently of other populations. Conservation actions may then be implemented in evolutionary important management units (MUs) within ESUs (Hedrick *et al.* 2001).

It is necessary to assist farmers and conservation authorities with genetic management strategies to maintain acceptable levels of genetic diversity in commercially used populations and to monitor representative populations periodically using molecular methods (Kotze *et al.* 2002). The characterization of the general population genetic structure is an important step in evaluating a populations' viability.

The relationships among some individuals in the population are, however, often not known. In many cases, this may not be a problem, but in small or captive populations, where each individual's contribution to past and future genetic diversity is critical, knowledge of relatedness is essential. Identification of relatedness among captive-, translocated-, reintroduced- or founder populations will thus help to guide future population management decisions. In some cases, identification of parentage alone will provide sufficient understanding of population structure (Haig 1998).

In February 2003, a workshop was held at the Agricultural Research Council (ARC) Animal Improvement Institute, Irene, South Africa, where the needs of the wildlife industry were addressed. Conservation authorities as well as game farmers supported the need to have an institution responsible for determining the degree of inbreeding, preserve genetic diversity, identify individuals who should have breeding priority in a herd and create pedigrees for various wildlife species (Kotze *et al.* 2003). According to Bothma (2002) a significant amount of research still needed to be done on the genetic management of wildlife. Molecular tools were available but needed to be adapted for the various wildlife species. To be able to render such a service to the industry, many research projects have been initiated in the past decade.

Diceros bicornis minor (Linnaeus 1758) (black rhino) and *Ceratotherium simum simum* (Burchell 1817) (white rhino) are two of the species in South Africa that can benefit significantly from genetic management, since these species have low numbers and populations are fragmented, and genetic diversity is thus potentially very limited. The illegal poaching of both black and white rhino has also re-appeared as a serious threat to the survival of these species in recent times. For these reasons, the black and white rhino were chosen as the species of interest for this study.

1.4 Current status of the black and white rhino

The black rhino consists of four subspecies or ecotypes, the Western (*Diceros bicornis longipes*); Eastern (*D.b. michaeli*); South-western (*D.b. bicornis*) and the South-central (*D.b. minor*) (Emslie and Brooks 1999).

The Western black rhino were the rarest and most endangered of the subspecies in the last decade, with only a few animals remaining in northern Cameroon (Emslie and Brooks 1999). During a survey in 2008, no evidence of the

Western black rhino could however be found and this subspecies may now be extinct (IUCN 2010). The Eastern black rhino is currently found in Rwanda and Tanzania. The South-western black rhino is currently found mainly in Namibia. Small populations have also been re-established in the south-western parts of South Africa (Emslie and Brooks 1999).

Diceros b. minor is the most abundant of the black rhino subspecies which used to occur from western and southern Tanzania through Zambia, Zimbabwe and Mozambique to the northern and eastern parts of South Africa. Today, most of these rhino occur in South Africa with smaller populations in Zimbabwe, Swaziland, southern Tanzania and Mozambique (Emslie and Brooks 1999).

During the 19th century, black rhino were hunted relentlessly. In southern Africa only two breeding populations had survived by 1933. The numbers continued to decline and between 1970 and 1992 the numbers had been reduced by 96% due to poaching and hunting. The estimated number of black rhino in South Africa in 1987 was 577. In the 10 years that followed the numbers increased to 1043 due to investments made in conservation strategies which included monitoring and law enforcement (Emslie and Brooks 1999). According to the IUCN red list of threatened species, South Africa now hosts an estimated black rhino population of 1 470 animals. This species is however still listed as critically endangered (IUCN 2010).

The white rhino is divided into two subspecies, the northern white rhino (*Ceratotherium simum cottoni*) and the southern white rhino (*C.s. simum*). The northern white rhino is nearly extinct and only a few animals may be left in the Democratic Republic of the Congo (Emslie and Brooks 1999). During a survey in 2003, the sighting of only four rhino was confirmed (IUCN 2010).

The southern white rhino is widely distributed through the bushveld areas of South Africa into Namibia, Botswana and Zimbabwe. In the 19th century, the species was hunted to the brink of extinction, except for a population in KwaZulu-Natal's Umfolozi region where there were only about 20 - 28 individuals left. In later years, white rhino were successfully re-established in the Pilanesberg-, Kruger National Park and other smaller nature reserves. There were about 7 095 white rhino in South Africa in 1997 (Mills and Hes 1997; Cunningham *et al*. 1999; Emslie and Brook 1999). According to the IUCN the estimated number of white rhino in the wild in 2007 was 17 480. By 2010, South Africa conserved 93% (16 255 individuals) of the global

population (IUCN 2010). An increase in numbers is however not necessarily an indication that the white rhino is no longer threatened.

1.4.1 Current management strategies for rhino

1.4.1.1 Ownership

The majority of rhino are still conserved in state-run protected areas although there are an increasing number of rhino that are managed by the private sector. Black rhino in private reserves are under custodianship of the state whereas white rhino can be privately owned (Emslie and Brooks 1999). In the case of custodianship, the ownership of the offspring is negotiable. The offspring can either be shared by the custodian and the state or the custodian can keep the rhino after 20 years of custodianship (Du Toit 2006). In South Africa, more than 20% of white rhino are privately owned. The move to private ownership increased the land available to rhino which contributed to the conservation of these animals (Emslie and Brooks 1999).

In South Africa and Namibia, owners of white rhino are allowed to sell animals and even have limited sport hunting opportunities, depending on the availability of surplus bulls. In 2005, an annual hunting quota of five black rhino bulls was approved for South Africa and Namibia, because of the "surplus male problem" in breeding populations. The level of offtake was thus less than 0.5%, which is sustainable (Du Toit 2006). In contrast, in countries such as Kenya, only non-consumptive use of rhino, such as tourism, is allowed (Emslie and Brooks 1999).

1.4.1.2 Biological management

Some rhino populations are managed using a meta-populations approach, with regular translocations, resulting in artificially induced gene flow. Managers try to achieve at least a 5% growth per annum. Rhino areas that are stocked below the carrying capacity can have an average population growth rate of up to 10–15%, if the sex-ratio is female-biased and the mortality rates are low (Du Toit 2006).

Biological management is however not restricted to the management of rhino numbers. The removal of animals from donor populations should also be carefully

managed since it could have a negative impact on these populations. Care should be taken that not only young females are removed since such a bias could skew the age and sex structure of the population in the long term. The introduction of competing browsers or grazers can also have a negative impact on the rhino population's performance (Du Toit 2006).

1.4.1.3 Translocation

To improve biological management, it is recommended that rhino numbers are kept below 75% of the ecological carrying capacity of an area. Alternatively, after 50% of the ecological carrying capacity has been reached, between 5–8% of the population should be translocated annually (Du Toit 2006).

In the case of black rhino, removals should be spread throughout the populations in an area, whereas in the case of white rhino, removals should be concentrated in one section of the locality. This type of removal will create an area of low density which will lead to the migration of surplus rhino to this area. The translocation of surplus animals has the advantage that it reduces population densities of donor populations and that new populations can be created or existing populations can be augmentated (Du Toit 2006).

Depending on the skills of the captive team, translocation mortality rates can vary between 2–5%. This is still acceptable since the reduction in population density will lead to an increase in the population growth rate, which is higher than the mortality rate (Du Toit 2006).

The most important question is, however, which individuals should be translocated. In most situations, the females should be unrelated to maximise the genetic diversity. The cows should not have a dependent calf, since the mortality rate for the translocation of calves is too high. They should be capable of breeding and in fair physical condition. The females must also preferably be about seven years old (Du Toit 2006; Emslie *et al.* 2009).

For bulls, the situation is more complex. Genetically it is better to remove dominant bulls since they have already contributed genetic material to the next generation. From a behavioural and demographic perspective, the removal of

dominant bulls will stimulate intra-species fighting for dominance which can lead to a higher mortality rate and lower population growth rate (Du Toit 2006).

1.4.1.4 Monitoring of rhino

The management of rhino entails the monitoring of rhino in reserves to ensure that none of these animals are missing due to illegal offtakes or other demographic impacts. Poaching can thus immediately be detected. Since adaptive management is required to maximize meta-population growth rate, there must be accurate data available. Data on population size estimates, measures of demographic performance, mortality patterns, behaviour and translocations are needed to ensure a successful management plan (Du Toit 2006).

1.4.1.5 Individual identification of rhino

To be able to monitor rhino successfully, monitoring staff should be able to identify every rhino in the population on a regular basis. Conservation authorities currently use three methods to identify rhino, namely ear tags, ear notches and microchips.

Brightly coloured ear tags are temporary and are mainly used for auction animals in bomas (i.e. enclosures at auction centres). Ear notches are for permanent identification and can be either natural of artificial. Triangular notches of about 2.5 cm x 2.5 cm x 1.5 cm are cut into the ear. The pieces of ear tissue provide an incidental source of DNA and are thus stored for future analysis. For identification, a system is used where the left ear markings are used for ones and the right ear markings for tens (See figure 1.1). For example, a rhino with identification number 18 will have marking number 10 on the right ear and markings five and three on the left ear (Du Toit 1998).

Ear notch numbering systems may vary, thus when translocation occurs, the ear notches have to be checked to ensure that the identification number is still unique to that specific rhino. Males and females are numbered in different series and ID numbers should not be re-used after an animal has died to avoid future confusion (Du Toit 2006).

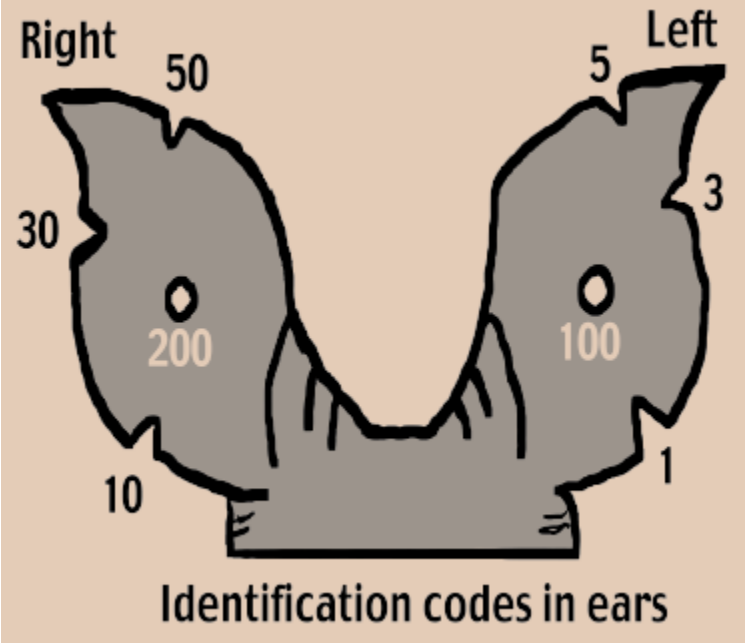


Figure 1.1 Identification codes in ears (Du Toit 1998)

As an alternative or complementary form of identification, microchips can be used. Microchips are inserted in the horn and right-hand side of the tail. The microchip has a unique code that can be read from a distance of 20-30 cm. This type of identification is mainly used when ownership is in dispute or in poaching cases where the poached horn can thus be identified (Du Toit 1998).

1.4.1.6 Reintroductions

To limit the loss of genetic variation and to ensure that new populations are viable, the population size of a new founder population should consist of at least 20 or more effective individuals (Du Toit 2006). An effective founder population requires that the individuals are unrelated and capable of breeding. Ideally the carrying capacity should not be less than 100 rhino. This is very seldom possible, and to overcome this problem, meta-population management can be applied. In other words, smaller populations will exist but on a regular basis, translocations between these populations will have to be made manually to simulate genetic connectivity. It is important that the gene pool of the founder populations is large and diverse enough for long-term sustainability. It is thus also desirable that more than one source population should be used when a new founder population is re-introduced into an

area. The animals should however be from the same subspecies (Du Toit 2006; Emslie *et al.* 2009).

Behavioural aspects of rhino should also be taken into consideration when re-introduction into an area is done. It is unavoidable that newly introduced males will fight for their territory, but if the death rates are too high, it should not be accepted as "normal". It could also be attributed to management problems. For this reason it is better to introduce the recommended 20 animals at the same time or at least within a year and not over an extended period. If this is not possible, the new rhino bulls should be released in areas of the reserve away from the dominant bulls (Du Toit 2006).

1.4.2 Previous genetic studies on rhino

Previous genetic studies have been done on both white and black rhino. O' Ryan and Harley (1993) did a comparative study between the mitochondrial DNA (mtDNA) of black (*D. b. minor*) and white rhino (*C. s. simum*). These authors used 24 black- (16 from the Hluhluwe Game Reserve, Natal, South Africa; six from Mkuzi, Natal, South Africa and two from Zimbabwe) and four white (from the Hluhluwe Game Reserve) rhino samples. Very little intraspecific variation was found in the *Sac II* and *Hpa I* sites for both species. Xu and Arnason (1997) compared mtDNA sequences between white rhino and Indian rhino (*Rhinoceros unicornis*) and found that the divergence between these two species occurred approximately 27 million years ago.

Swart *et al.* (1994) determined the genetic variation in four isolated populations of black rhino by using 30 protein-coding loci. A total of six *D. b. bicornis* was sampled at Etosha National Park, Namibia. *Diceros b. minor* was sampled at Mkuzi Game Reserve, Natal, South Africa (34); Hluhluwe-Umfolozi National Park, Natal, South Africa (25) and Zambesi Valley, Zimbabwe (90). These authors found that the genetic diversity of these populations were similar to those found in outbred populations that are not genetically depauperated.

Brown and Houlden (1999) isolated and characterized 11 microsatellite markers for black rhino. Between two and seven alleles were detected at each locus. These markers were tested on two samples of *D. b. michaeli* and a group of *D. b.*

minor (the exact sample size is not mentioned). Cunningham *et al.* (1999) isolated and characterized five black rhino microsatellites of which only three were polymorphic. These markers were also tested on white rhino, with only one marker found to be polymorphic. Garnier *et al.* (2001) extracted DNA from faecal samples of 35 wild black rhino in the Save Valley Conservancy, Zimbabwe. Ten microsatellite markers were used to verify the parentage of 19 calves successfully. There were no white rhino samples included in the study. Florescu *et al.* (2003) optimized five polymorphic microsatellite loci for white rhino, which were the first species-specific microsatellite markers for white rhino. The markers were tested on 30 white rhino samples from the Hluhluwe-Umfolozi National Park. Harley *et al.* (2005) compared the genetic diversity of the three subspecies of black rhino (*D. b. michaeli, D. b. minor, D. b. bicornis*) by using nine microsatellite markers. These authors found that *D. b. michaeli* had more genetic diversity compared to the other two subspecies. Only one individual from the fourth subspecies, *D. b. longipes* was profiled. This individual had the least genetic variation compared to all the other 121 rhino tested. Nielsen *et al.* (2008) tested 21 microsatellite markers on black and white rhino from the Hluhluwe-Umfolozi Reserve and found that 17 of these markers amplified in both species. Sixteen of the 17 markers were polymorphic in the white rhino and only 12 of these were polymorphic in the black rhino. These markers however were only tested on 22 white and 6 black rhino samples which could partly explain why more genetic diversity was found in the white rhino than in the black rhino. These authors also concluded that the quality of DNA extracted from faecal samples for both black and white rhino was not suitable for DNA profiling in contrast to the results published by Garnier *et al.* (2001). Scott (2008) reported on microsatellite variability in the four rhinoceros species. This author tested 24 microsatellite markers and found similar results to those found by Harley *et al.* (2005). *Diceros b. michaeli* had the highest level of genetic diversity, while the southern white rhino and the Indian rhino had the least genetic variability.

1.5 Aspects of genetic management that is applicable to rhino

1.5.1 Reintroduction and translocation

Currently in South Africa, game-proof fences prevent the free exchange of genetic material between animals of neighbouring game farms, which may lead to the genetic impoverishment of populations. A management strategy for populations with low heterozygosity values and suffering from inbreeding depression was thought to be the deliberate introduction of individuals from elsewhere into the populations (Frankham 1995).

The deliberate augmentation of an endangered population using a genetically distinct lineage from elsewhere, however raises two potential problems. Firstly, genetic augmentation may result in a short-term increase in fitness, but it may be followed by a subsequent decline, because there is the potential problem of outbreeding depression. Natural selection will tend to create populations that are adapted to the local conditions. Augmentation introduces genetic material that is possibly adapted to a different set of conditions. The newly introduced genetic resources can thus cause the population to be less suitable for specific conditions. The second problem encountered is to verify that the two populations are of the same species or subspecies and that hybridization is not induced artificially (Amos and Balmford 2001; Hedrick 2005a).

Wild animals, for example white rhino and sable antelope, can be bought from conservation authorities, boma auctions, catalogue auctions, game farmers or from wild animal capture and marketing enterprises. To minimize the risks involved when buying and translocating wild animals, only wild animals that are ecologically adapted to a particular region and that are known to have occurred naturally in that region before, should be bought and translocated. This will also ensure that competition between ecologically equivalent animals is eliminated, for example between sable antelope and gemsbok. Furthermore, in the case of antelopes, complete breeding herds should be re-established because it can take some time for new social groupings to form (Bothma 2002). Another major problem with adding new stock to existing herds to prevent inbreeding is that there may be no record keeping of the prior history of wild animals. In other words, there is still a chance that the new

animals will not genetically contribute to the herd, since they may come from the same origin, and the effect of inbreeding depression will continue.

It would be advantageous for owners of neighbouring game farms to swap or trade healthy breeding animals from time to time to induce gene flow and prevent inbreeding depression in their own herds. Indeed, it is a sound economic and genetic guideline to rather obtain animals from the available local stock than to import them at great expense from distant locations. Apart from the genetic implications, losses in animals that are kept in captivity for a long time or are transported over long distances may be unacceptably high (Bothma 2002).

Reintroduction and translocation of animals are a hazardous and expensive undertaking. The success of reintroduction is increased by selecting sites within the historical range of the species. Populations for re-stocking species should be those with the highest reproductive fitness in the wild habitat and the highest genetic diversity to optimize chances of survival (Frankham *et al.* 2002).

As mentioned before, the second problem encountered when translocating animals, is to verify that the two populations are of the same species and that hybridization is not induced (Amos and Balmford 2001; Hedrick 2005a). Vernesi *et al.* (2003) genetically studied the reintroduction of wild boar (*Sus scrofa*) in Central Europe after World War II. Nine microsatellite loci were used to study wild boar populations in Italy and Hungary. Results indicated that none of the populations showed any genetic evidence of demographic decline and that the three parental populations from Italy and Hungary were genetically distinct from each other. Most of the individuals hunted near Florence were genetically intermediate between the parental groups, suggesting that hybridization had occurred in this area. Vernisi *et al.* (2003) concluded that hybridization was a threat to the native groups and that the development of a conservation and management strategy for wild boars in Europe was essential.

In the case of rhino, individuals from different populations are chosen and translocated to a new area to ensure that the optimal genetic diversity for the new population can be obtained. Caution should be taken when rhino are translocated, because the donor population might be destabilized if the population has not reached a stage of definite genetic and demographic viability yet. Starting up new re-introduction projects should be avoided if existing ones are not adequately

consolidated. If the donor population was established from at least 50 animals and a few generations had passed, then translocations will not have a negative impact on that population. It is however suggested that sub-adult males should be translocated since this is the class that is most likely to disperse under natural conditions, even though from a genetic point of view, the dominant bulls would be better candidates for translocations (Du Toit 2006).

1.5.2 Genetic drift and gene flow

Fragmented populations are subjected to genetic drift, creating the possibility that the same deleterious allele may be absent in one population and present at high frequency in another (Frankham *et al.* 2002). Genetic drift can be described as the random variation in allele frequencies from one generation to the next. The outcome of genetic drift is unpredictable and the magnitude of the drift depends upon the population size. In the long-term, genetic drift reduces the genetic variation in populations and can cause populations to diverge from each other (Halliburton 2004).

By measuring rates of gene flow, or the lack thereof, populations that are most genetically depauperate, most fragmented, in terms of each other, can be identified. The assessment of populations can identify populations in need of augmentation, populations that could safely donate individuals to more vulnerable populations, or populations in need of further demographic or environmental consideration (Haig 1998). As stated by Morjan and Riesenberg (2004), when gene flow among populations exceeds four migrants per generation, neutral alleles are homogenized among populations, effectively producing a panmictic species. Conversely, species cohesion will break down when gene flow is reduced among population to less than one migrant per generation, allowing differentiation to occur through the fixation of alternative alleles through genetic drift. Populations, rather than species are thus the units of evolution (Morjan and Rieseberg 2004). It is therefore important to monitor gene flow between populations as can be seen in the following examples.

Norton and Ashley (2004) used six polymorphic microsatellites to determine whether there was a difference in the genetic variability between captive Baird Tapirs (*Tapirus bairdii*) in North and Central American zoos and wild populations. Microsatellite results indicated that the Central American zoos had a risk of losing

genetic variability due to inbreeding. North American zoos had an effective management plan to avoid inbreeding, thus, the populations had sufficient genetic variation. Significant levels of population differentiation were however, detected between the North American zoos and wild populations. The divergence between the North American zoos and the wild populations could probably be attributed to a founder effect or isolation. Norton and Ashley (2004) concluded that an exchange program between the North- and Central American zoos would be beneficial to both.

Jones *et al*. (2004) investigated the genetic diversity and population structure of the Tasmanian devils (*Sarcophilus ianiarius*) on the island of Tasmania. A set of 11 polymorphic microsatellites were used in the study and results indicated that the heterozygosity and allelic diversity values were low in all subpopulations. Substantial genetic structuring was observed for comparisons spanning unsuitable habitat, implying limited dispersal of Tasmanian devils between the well-connected eastern populations and a smaller northwestern population. The analyses of the Tasmanian devil populations across Tasmania are consistent with the hypothesis that habitat-related impedance to dispersal affects gene flow. The strongest genetic differentiation appeared between the eastern half of the island, which was more or less panmictic, and the northwestern corners, which was genetically distinct. Jones *et al.* (2004) concluded that translocation between these two regions should be restricted.

Bouzat *et al*. (1998) used six polymorphic microsatellites to measure the genetic variation of the greater prairie chicken (*Tympanuchus cupido*) before and after a bottleneck that was caused due to habitat destruction. Pre-bottleneck samples were collected from museums and post-bottleneck samples were sampled from natural populations. Results indicated that there were alleles lost due to the bottleneck. Johnson *et al.* (2003) used the same microsatellites markers as Bouzat *et al.* (1998), as well as mtDNA markers and concluded that the historic populations of prairie-chickens were once interconnected by gene flow but current populations were now isolated. Johnson *et al.* (2003) thus suggested that maintaining gene flow may be important for the long-term persistence of the prairie-chicken populations.

From the preceding, the need to monitor the gene flow and heterozygosity values between and within rhino populations is clear, since genetic diversity is essential for the long term survival of these species. To ensure that there is no further loss of genetic diversity, gene flow between different populations should be

maintained by suitable management actions, including the translocations of animals on a regular basis, if possible.

1.5.3 Inbreeding and inbreeding depression

One of the most important aspects of genetic management is the avoidance of inbreeding. Inbreeding is the mating of individuals related by ancestry. It is unavoidable in small populations, as all individuals become related by decent over time. Inbreeding may reduce reproductive fitness since it increases the homozygosity levels and exposes rare deleterious alleles (Frankham *et al*. 2002). Allozyme and microsatellite data have verified that 77% of 170 threatened taxa had lower heterozygosity and genetic diversity values than their related non-threatened taxa. A reduced genetic diversity value is an indication that the populations' reproductive fitness is already compromised and that their subsequent extinction risk is elevated in a changing environment (Spielman *et al*. 2004). Inbreeding is measured by using the inbreeding coefficient (*F*). The inbreeding coefficient can be defined as the probability that an individual has two alleles at a locus that are identical by descent (Frankham *et al*. 2002). The inbreeding coefficient of a taxon has a positive correlation with the populations' reproduction fitness. It is not possible to answer the question on when the genetic effects of lowered diversity are of sufficient magnitude that they must be directly managed, because the answer will depend on the inbreeding coefficient and thus on the effective population size and number of generations (Spielman *et al*. 2004).

Inbreeding depression can occur when the homozygosity value of a population is increased, usually through matings between related individuals, which leads to a loss of fitness (Amos and Balmford 2001). Inbreeding depression is most prominent for features associated with reproductive fitness and has an effect on all aspects of reproductive fitness. It includes offspring number, juvenile survival, longevity, interbirth interval, mating ability, sperm quantity and quality, competitive ability and developmental time in animals (Frankham *et al*. 2002; Spielman *et al*. 2004). Most genes in diploid organisms can operate satisfactorily with only a single copy, and loss of function mutations are generally recessive, showing their effect only in the homozygous state. The more homozygotic the population becomes, the greater the

chance that these deleterious alleles will be expressed. Deleterious alleles are usually removed by natural selection. There is thus a differential mortality among individuals carrying greater or lesser mutation loads. If a population starts to decline, the load will build up and the average fitness of the population will fall, also known as a mutational meltdown (Amos and Balmford 2001).

Populations with low genetic diversity tend to be more sensitive to environmental changes. Due to the lack of genetic variation, populations are sometimes unable to adapt which may lead to a decline in population size and consequently an increase in inbreeding (Frankham 1995).

A population bottleneck occurs when a population is suddenly reduced to a very small size. The consequences are a reduction in heterozygosity, rare alleles that are lost and a change in the allele frequencies. If a population increases very rapidly in numbers after going through a bottleneck, the reduction in the heterozygosity may be small, but it will take a very long time to restore the initial level of heterozygosity (Halliburton 2004) and specific alleles may be permanently lost.

Low genetic variation in a species might be indicative of a recent population bottleneck, and such a bottleneck may potentially indicate vulnerability to extinction. The more recent a bottleneck has been, the higher the expectation that the bottleneck will influence the future of the species (Zhang *et al.* 2002). The loss in mean heterozygosity can be minimal if population size increases rapidly following a single bottleneck of short duration. Some taxa are highly susceptible, but others relatively immune to fitness depression effects from consanguineous mating (Avise 2004).

A low heterozygosity value in a population is not necessarily the result of bottlenecking. Heterozygosity is only one measure of genetic diversity. Loss of allelic diversity tends to be a more sensitive indicator of historical bottlenecks than reductions in heterozygosity (Amos and Harwood 1998; Amos and Balmford 2001). It is important to be cautious when interpreting low molecular heterozygosities, as most of the reductions in genetic variation presumably are outcomes, rather than causes of population bottlenecks (Avise 2004).

Despite the importance of genetic variability, there are many exceptions to the rule that high levels of genetic variability equate evolutionary health. Some species have low variability but are exponentially increasing in population size, whereas others appear to be declining despite great variability (Amos and Harwood 1998). The

main consequence of reduced variability is however the lowering of a population's ability to react to novel challenges (Amos and Balmford 2001). Molecular analyses can help to identify natural or captive populations that display severe genetic impoverishment from past population bottlenecks or inbreeding. Less clear is the extent to which molecular heterozygosity is a reliable gauge of a population's short-term survival and long-term adaptive potential. Thus, managing captive or natural populations for genetic heterozygosity *per se* should not come at the expense of neglecting important behavioural, ecological, or environmental factors (Avise 2004).

1.5.4 Individual identification as a forensic application

The incidence of wildlife related crimes is increasing since the probability of apprehension and conviction are less compared to conventional burglary, and the financial gains are lucrative (Brown 2000). Poaching and illegal harvesting are threats to a wide variety of endangered species, especially large cats, elephant, bears, rhino, parrots, whales and some plants (Frankham *et al.* 2002). In 2006, between 25 000 and 29 000 kg of ivory were seized. Wasser *et al.* (2009) estimate that about 38 000 African elephants were slaughtered in 2006. Major crime syndicates are involved because of the growing demand and the low risk of prosecution (Wasser *et al.* 2009). Most countries have laws to protect endangered species, but it is difficult to obtain evidence to convict individuals suspected of illegally taking of, or trading in, threatened species (Frankham *et al.* 2002). The value of illegal trade in wildlife products is estimated at tens of billions of US dollars annually. Examples of large busts made by authorities include the confiscating of 55 000 reptile skins in India, 23 metric tons of pangolins in Asia and 2 000 Indian star tortoises in India (Wasser *et al.* 2009).

Molecular techniques have assumed an important and growing role in the detection of illegal hunting and trading of wildlife (Frankham *et al.* 2002). Examples include Kotze *et al.* (2008) who used microsatellite markers to determine the origin of cheetah which were confiscated in South Africa on suspicion of illegal import. The genotype assignment testing however indicated that no foul-play had occurred. A second example is a study by Marshall *et al*. (1999). These authors developed a set of microsatellites for the Arabian oryx (*Oryx leucoryx*) consisting of cross-species

markers. The microsatellite markers had enough variation for identification purposes, and had been used successfully in a forensic case (Marshall *et al*. 1999).

There are also reported murder cases where DNA profiling of animals have helped to solve the case, for example the murder case of Shirley Duguay in Canada. In 1994, the body of Shirlely Duguay, who was a resident from Prince Edward Island, was discovered by the Royal Canadian Mounted Police. A leather jacket was found at the crime scene covered with blood from Duguay and two hair samples from a cat. With the help of STR markers, the DNA profile from the cat fur found on the jacket could be traced to a cat named Snowball, whose owners were the parents of the victim's ex-husband, Douglas Beamish. Beamish was convicted of second-degree murder and sentenced to 15 years in jail (Day 2001).

Animal DNA forensic includes individual identification as well as parentage verification in cases where the female has been exposed to multiple males, or where switches of offspring took place. Since the demand for animal DNA forensic work has increased, it has become important to have a set of minimum guidelines for service laboratories (Budowle *et al.* 2005). In South Africa, the Animal Genetics laboratory at the ARC (Agricultural Research Council) in Irene, Pretoria, is forensically accredited to analyze DNA profiles for poaching cases that include domesticated animals. The different species that occur in wildlife forensic cases do however provide a problem regarding the use of DNA profiling. There are in most cases no microsatellite marker sets available to provide these services of identification and parentage verification. Even if the markers are available, genetic DNA databases are not available to statistically determine the profile probability that gives credibility to the evidence samples.

Despite the close monitoring and management of rhino, poaching is still a very serious threat to the survival of this species. According to Martin (2009) there has been an increase in rhino poaching cases in South Africa. The number of confirmed illegal rhino killings, were over 90 animals in 2008. More than 122 rhino were killed in 2009 (SAPA 2010a). On 23 January 2010 it was reported that 14 rhino were already killed for that year (SAPA 2010b). On 25 February 2010, the number had increased to 28 rhino killed and dehorned (SAPA 2010c). In July 2010, it was reported that the KwaZulu-Natal province had experienced 11 rhino poaching cases during which some of the rhino were killed with an AK47 assault rifle (SAPA 2010d).

There are even reported cases of rhino being de-horned while still alive and a month old calve being captured and loaded onto a helicopter (SAPA 2010e).

Rhino owners are getting desperate because poachers are members of well-resourced syndicates that are causing great harm to the rhino populations. The Kruger National Park is being patrolled by the military, while some private owners have contemplated the poisoning of rhino' horns. The costs for private security are around R30 000 per month (SAPA 2010a; SAPA 2010b). The Free State Nature Conservation authority is working together with an elite force of the South African Police Service to prevent further poaching of rhino in the Free State (SAPA 2010f).

The main reason for these illegal killings are for rhino horns, which are mainly sold in Vietnam and China (Milliken *et al.* 2009). The horns are used as a fundamental ingredient in traditional medicines. These medicines are used for the treatment of impotency, fever, pains, arthritis and for presumed spiritual purposes (demon possession). Research has however found that rhino horn contains no medicinal properties whatsoever (SAPA 2010a). One horn can be worth as much as (South African Rand) SAR 1.7 million (SAPA 2010g). A Vietnamese man was caught at OR Tambo International Airport in Johannesburg, South Africa with 16.1 kg of rhino horn. The estimated value of the rhino horn would have been SAR 2.2 million in Vietnam (SAPA 2010h).

As previously mentioned, most rhino have ear-notches to help identify them in the field. All information regarding a rhino is kept in a studbook, which includes the ear notch number (identification number), microchip numbers and other relevant information such as the estimated date of birth, gender, etc. Ear-notching is however not sufficient enough in poaching cases for identification, since the ears can be removed from the animal, as can the microchips. Currently, the pieces of ear-tissue that are removed during ear-notching, are stored for DNA profiling which will serve as a reference DNA sample for that animal. If the animal is thus pouched, the reference DNA profile will help to identify the animal if any biological evidence is recovered.

1.5.5 Parentage verification

To implement a successful conservation management plan for black- and white rhino, data is required on the paternity and genetic variability of rhino. Pedigree information will help with management decisions, such as decisions on the translocation of animals between reserves. This type of information can help managers to decide which specific individuals in a captive population of known pedigree should have breeding priority when the goal is to maximize population genetic variation (Avise 2004).

Genetic studies of parentage have played a major role in the study of evolution and behavioural ecology and have become one of the central themes in the field of molecular ecology (Jones and Ardren 2003). Quantitative genetics parameters such as heritability are best estimated with pedigrees, but this is rarely possible for wildlife species. Genetic markers can be used to study these aspects as well as the parentage, relatedness and/or fitness of wildlife populations (DeWoody 2005).

The use of pedigree analyses and genetic data is extremely important in helping to direct efforts to breed endangered species in captivity. Only in the past 20 years have managers of zoos and aquariums incorporated genetics-based tools to the current species-survival programmes. Conservation breeding programmes are also relevant to wild populations. Breeding programmes in this context entail assessing the genetic health and integrity of an endangered population or species (DeSalle and Amato 2004).

Pedigrees are used extensively in genetic management of captive populations, so it is important to verify the accuracy of such records. Pedigree information can also be used to calculate the inbreeding coefficient of an individual (Frankham *et al*. 2002). The information on parentage is essential to study the impact of inbreeding, to verify pedigrees used in genetic management of threatened species, and to determine the effective size of populations. Often, parentage cannot be determined from direct behavioural observation. Molecular genetic markers, including microsattelites, can however be used to resolve these uncertainties (Ely *et al*. 1999).

With the latest DNA technology available, such as STRs, breeding success in wildlife species can be determined with great accuracy and detail. Not only can the parentage of an individual be verified, but the reproductive success of each adult can

also be estimated. Also, estimations regarding the number of available partners of each gender can be made and the effective number of breeders in a population can be determined. These techniques provide tools that managers can use to conduct intensive breeding assessments for many wildlife species (DeWoody 2005).

In many wildlife species, the social structure and relatedness of individuals in a population may not be as clear as expected from observed field data. Molecular markers can thus be used to help clarify the relatedness between individuals. The power of estimating relatedness between individuals depends upon the genetic variability of the markers and the number of markers being used. Microsatellites are currently the best markers that can be used to verify parentage and estimate relatedness among individuals in a population in wildlife species (DeWoody 2005).

Individual identification and parentage verification are usually based on the 'exclusion principle' (Weller *et al.* 2006). This technique (based on Mendelian rules of inheritance) uses incompatibilities between parents and offspring to reject particular parent-offspring hypotheses. Exclusion is an appealing approach, because exclusion of all but one parent pair from a complete sample of all possible parents for each offspring in a population could be considered the paragon of parentage analysis. However, few studies have achieved this ideal. Using strict paternity exclusion criteria (a single mismatch excludes a putative parent) may result in false exclusions because of mutation, genotyping errors, or null alleles. Ironically, these problems become more acute as more data are generated to solve the problem, because the assay of additional loci (or additional individuals) increase the likelihood that a dataset will contain errors or mutations (Jones and Ardren 2003; Hedrick 2005b; Weller *et al*. 2006). These shortcomings will be discussed in Chapter 5.

Most systems in nature do not permit a perfect parentage analysis through complete exclusion. In the cases where complete exclusion is not possible, in other words when multiple males are not excluded as possible fathers of an offspring, a statistically based method to assign paternity is needed. Techniques have been developed that assign progeny to non-excluded parents based on a likelihood score derived from the genotypes of putative parents. The likelihood approach has been shown to be an efficient way to evaluate human relationships as well as paternity assignment in polygamous natural populations. When multiple loci are used, the likelihood ratios may be multiplied to obtain an overall paternity index. The natural

logarithm of the combined likelihood ratios can be used to obtain the likelihood of odds (LOD) score. In such analysis, a LOD score of zero indicates that the alleged father is as likely to be the father of the offspring as a random male from the population, and a positive LOD score indicates that the alleged male is more likely to be the father than a random male. Marshall *et al*. (1998) suggested a statistical approach to discriminate between non-excluded males in which Δ is the difference in LOD scores between the most likely male and the next most likely male, and thus developed the CERVUS software to calculate the Δ values and the statistical significance of Δ values. Determining a sibship relationship is more problematic than parentage testing since there are no obligatory alleles between siblings that make it possible to absolutely exclude a biological relationship. A lack of shared alleles at any locus does not exclude two individuals from being related. Reid *et al.* (2004) performed a study to determine the sensitivity of the ABI Identifiler kit to determine whether it would be possible to distinguish between full and non-siblings. Likelihood ratios in the form of combined sibship indices (CSI) were calculated by using DNA-View software v25.05 (Reid *et al.* 2004). A CSI value greater than one implies that the individuals tested are related. None of the non-siblings pairs tested had a CSI value high enough to be classified as siblings, but all of the siblings tested had a CSI value greater than ten. In 80% of the cases, the CSI values were greater than 1000. Statistical analysis can thus be used in many cases to determine if two individuals are related (Reid *et al.* 2004).

1.6 Molecular markers

Molecular markers that can be used for individual identification and parentage verification include allozymes, short tandem repeats (STRs) also known as microsatellites, single nucleotide polymorphisms (SNPs), mitochondrial DNA (mtDNA) and sex-chromosome markers. These markers (except for allozymes) are usually neutral markers with respect to natural selection and can thus be used to determine relatedness (DeWoody 2005).

1.6.1 Allozymes

Since 1966, genetic diversity was mainly measured indirectly by using protein electrophoresis until the late 1980's when microsatellites were discovered. Tissue samples are homogenized and placed in special buffers to extract the enzymes. Samples are applied to a gel (starch or polyacrylamide) and an electrical potential gradient is applied to the gel that causes the enzymes to migrate through the gel. Enzymes are separated by size, shape, and/or charge. Proteins from specific loci are usually detected by their unique enzymatic activity using a histochemical stain (Parker *et al.* 1998; Frankham *et al.* 2002).

There are two main disadvantages when using allozyme markers. Firstly the samples that can be used are mainly blood, or tissue samples. In the case of domestic animals, blood samples can be easily taken, but in the case of wildlife species, an animal must be killed or sedated by a veterinarian before a sample can be taken. There is also a risk involved, because wild animals are susceptible to stress and the animal may suffer or die which will be a tragic loss when working with endangered species. Secondly, the fact that a product of a functional gene is used means that the marker is not neutral in terms of natural selection. The results obtained may also result an underestimation of true genetic variation, because there are 64 codon combinations but only 20 amino-acids. The advantages of these markers are however that it is very easy and inexpensive to develop (DeWoody 2005).

1.6.2. Short Tandem Repeats (STRs)

STR loci were first discovered in the late 1980's. These markers are tandem repeats of short DNA segments, usually di-, tri- or tetranucleotides and are highly polymorphic. They are found scattered abundantly throughout the nuclear genomes of most plants, animals and humans (Avise 2004). STR markers are found mainly in the non-coding regions of the genome (Tóth *et al*. 2000). For this reason, STRs are neutral relative to natural selection (assuming that the marker is not closely linked to a functional gene). This is a great advantage for markers used for parentage verification because greater allelic diversity will be expressed and the markers will thus be more informative (DeWoody 2005). If the markers were not neutral and selection occurred,

the alleles of the marker will end up being similar for the whole population and no or little genetic variation will be detected.

STR markers are co-dominant and inherited in a Mendelian fashion. It also has the advantage of accuracy and reproducibility. For these reasons, this technology seems to be the current method of choice for individual identification, parentage analysis and the study of population structure (Chambers and MacAvoy 1999; Avise 2004; Fan and Chu 2007).

One of the most important advantages of using STR markers is that this technique is PCR-based, in other words only a very small sample is needed to generate a DNA profile. Thus various types of samples can be used, including blood, tissue, faecals, feathers and hair. The development of techniques such as biopsy-darting and creative non-invasive sampling methods, make the use of STR markers a very practical tool for population genetic studies on wildlife (Karesh *et al.* 1987).

One of the disadvantages of using STR markers is that the development costs are very high and time consuming. Firstly, a genomic library must be constructed for the target species, and then screened for clones that contain microsatellite repeats. The inserts of the positive clones must be sequenced and the information from the unique sequences flanking each repeat region is then used to synthesize PCR primers (Avise 2004). STR markers do however not necessarily have to be species-specific to be used for parentage verification studies.

Ely *et al.* (1999) compared results obtained from biochemical markers with results obtained from human STR markers that were used for parentage verification in a colony of rhesus monkeys. They concluded that biochemical markers should be abandoned in favour of the more informative STR markers for routine paternity testing (Ely *et al.* 1999). Bonnet *et al*. (2002) developed a set of 11 cross-species microsatellites for four tropical deer species: Eld's (*Cervus eldii*), Swamp (*C. duvaucelii*), Rusa (*C. timorensis russa*) and Vietnamese Sika (*C. nippon pseudaxis*) deer. Parentage exclusion probability and probability of identity were computed for the Swamp deer population. The probability of exclusion was found to be lower than in cattle breeds, but still sufficient enough to verify a parentage (Bonnet *et al.* 2002). The use of cross-species markers could save a lot of time and money in many research projects (Parker *et al.* 1998; Clifford *et al.* 1999; Bonnet *et al.* 2002; Lillandt *et al.* 2002).

The highly polymorphic nature of STRs, which is a result of high mutation rates, is invaluable for genetic studies, but present specific challenges. Many of the mutations that occur in STRs result in alleles of the adjacent size classes, so that the mutational process tends to be imperfectly stepwise or ladder-like. This characteristic of STRs can often cause serious interpretive difficulties, especially in studies of geographic population structure, because alleles that are identical in state are not necessarily identical by descent (Avise 2004; Morin *et al.* 2004). Because of their high mutability, microsatellites are thought to play a significant role in genome evolution by creating and maintaining quantitative genetic variation (Tóth *et al*. 2000).

Another disadvantage of STR makers is the occurrence of null alleles. Null alleles have been 'discovered' by the observation of different typing results when utilizing independent STR primer sets. Null alleles occur due to mutations (variants) at or near the 3'-end of a primer and results in little or no extension during PCR (Butler 2005). Null alleles have been observed in many studies done, for example, the study done by Paetkau and Strobeck (1995). These authors detected null alleles at one locus, in both Asiatic black bears (*Ursus thibetanus*) and in North American black bears (*U. americanus*) while conducting pedigree analysis. Results of a study done by Pemberton *et al.* (1995) indicated that three out of 16 microsatellites for red deer had non-amplifying alleles.

Null alleles are an important consideration for parentage analysis, because this phenomenon can cause false exclusions when heterozygotes are scored incorrectly as homozygotes. Null alleles can usually be detected as a significant departure from expected Hardy-Weinberg equilibrium. In studies in which a known parent is sampled with groups of offspring, null alleles are easy to detect because they result in incompatibilities between the known parent and offspring that invariably involve homozygous genotypes. Null alleles are present in species-specific STR markers and cross-species STR markers (Pemberton *et al*. 1995; Jones and Ardren 2003). If a null allele is detected at a STR locus, there are several possible solutions. First, the problem PCR primer could be redesigned and moved away from the problematic site. This solution is, however, undesirable because it is time consuming and labour intensive. A second solution is to drop the STR locus from the study rather than attempting to redesign the PCR primers. A third and more favourable solution is to

add a 'degenerate' primer that contains the known sequence polymorphism. This extra primer will then amplify alleles containing the problematic primer binding site sequence variant. A fourth possible solution is to re-amplify the sample with a lower annealing temperature and thereby reduce the stringency of the primer annealing. If the primer is only slightly destabilized, as detected by a peak height imbalance with a heterozygous sample, then it may be possible to correct the peak height imbalance by lowering the annealing temperature during PCR (Butler 2005).

Genotyping process cannot be completely automated due to presence of null alleles, dropped alleles and other artefacts, which is a disadvantage for service laboratories. The genotyping error rates are found to be between 1 to 5% for STR markers (Baruch and Weller 2008). The testing for genotyping errors will be discussed in Chapter 2.

1.6.3 Single Nucleotide Polymorphisms (SNPs)

SNPs are currently still relative new molecular markers that can be used for parentage verification. With SNPs, the differences that occur in nucleotides that occupy a specific site in the genome is analysed and used to determine relatedness between individuals. These markers can be either neutral to selection or not, depending on whether the sites are derived from a coding gene or not (DeWoody 2005).

Weller *et al*. (2006) suggested that microsatellites markers should be replaced by single nucleotide polymorphism (SNP) as the markers of choice for parentage verification and individual identification. SNP markers have a lower genotyping error rate, but since most of the time these markers are bi-allelic, more SNP markers are needed to yield the same power of exclusion as with microsatellite markers. Weller *et al.* (2006) carried out a study to determine the number of markers necessary to obtain a 99% probability that none of the samples would be erroneously matched, as a function of the number of individuals, accounting for genotyping mistakes. The results indicated that to give the same exclusion probability, for every four microsatellite markers used, eight SNP markers should be used, assuming that all the microsatellites had five alleles. The study also concluded that only four microsatellite

markers were necessary to identify an individual with a probability of 99%, assuming that there was no inbreeding in the population.

1.6.4 Mitochondrial DNA (mtDNA)

Mitochondrial DNA is found in the mitochondria, which is located in the cytoplasm of eukaryotic cells. The mitochondria are maternally inherited and markers are thus mainly used to study population structure and phylogenetic relationships, because it reveals the evolutionary footprint of differential dispersal between the sexes (DeWoody 2005). It can only be used to trace the female line of ancestry.

1.7 Advantages of STRs compared to SNP markers for parentage analysis

Coates *et al.* (2009) performed a comparative performance study of SNPs and microsatellite markers for population genetic analysis. These authors concluded that the data obtained from both marker sets resulted in similar conclusions with respect to population structure. SNPs provided a higher estimate of the inbreeding coefficient (F_{ST}) compared to microsatellite markers.

SNPs have the advantage of lower genotyping error rates (between 0.1% - 0.8%) and genotyping can be fully automated, which makes this test more cost effective. A disadvantage of these markers is the low Polymorphic Information Content (PIC) values (Baruch and Weller 2008).

Another advantage of SNP markers over microsatellite markers is that genotype comparisons can be done between laboratories without any adjustment of raw data. In the case of microsatellite markers, because the migration rate of fragments during electrophoresis can differ, collaborating laboratories have to standardize their results, before comparisons can take place which can be problematic (Coates *et al*. 2009).

Quality assurance and quality control are a requisite for service laboratories. These measures are well established in human forensic and paternity testing laboratories that make use of microsatellite markers. These quality control recommendations should also be applied for animal forensic laboratories, which will make these results also acceptable in a court of law (Budowle *et al.* 2005). To make

this a reality, sets of microsatellites must first be compiled and tested on the various wildlife species of South Africa, before it can be used in wildlife forensic cases.

According to Baruch and Weller (2008), SNPs are the future markers of choice for parentage analysis. For this study however, it was decided to rather use STR markers, since the forensic application of DNA profiles is one of the focus areas of this thesis. Results can still be obtained from mixture samples with microsatellite markers, which is often the case with forensic samples whereas SNP markers are not suitable for mixture samples. There were also previous published data available for STR markers for rhino which saved time and costs for this project. No information regarding SNPs for rhino were available.

1.8 Aims of this study

The overall aim of this study is to compile a set of species-specific and cross-species microsatellite markers for wildlife with specific reference to the rhino species in South Africa that can be used for individual identification and parentage verification. As seen with previously discussed studies, the information gained by these DNA profiles, can contribute to conservation management decisions. These sets of markers are thus a genetic molecular tool that will enable the establishment of a National Genetic Database for black and white rhino species in South Africa. Not only will the information from this database contribute to management decisions, but it can contribute to any future research project involving this species. The population structure can be determined, since parentage can be verified and thus the genetic diversity and inbreeding of populations can be monitored. With the establishment of a National Genetic Database for wildlife, translocation decisions involving isolated and fragmented populations can be made without the problem of outbreeding depression. The DNA database will also assist in DNA forensic cases where the statistical interpretation of results is crucial.

The specific aim of this study is to use the rhino species as an example of how an individual-based approached can be applied in the genetic management of wildlife with the following objectives:

- Determine the population structure of three white rhino populations in the Mpumalanga Province of South Africa, and one black rhino population in the Eastern Cape Province.

- Screening the DNA database for genotyping errors, null alleles and other possible artefacts.

- Determine the most effective STR markers to use and the number of STR markers necessary for individual identification.

- Determine the effectiveness of the selected STR markers for parentage analysis by testing the selected marker sets in a large scale parentage analysis for both black and white rhino populations.

2.1 Introduction

It is not always feasible to collect blood or tissue samples from free-ranging wild animals, especially in the case of rhino. An alternative to the ordinary sampling methods can therefore include non-invasive methods such as biopsy-darting or the collection of faecal samples. These techniques have been used successfully on some species in the past but, unfortunately, in the case of rhino these methods are not very feasible. For example, Cunningham *et al*. (2001) extracted DNA from black rhino dung but concluded that although the extracted DNA could be amplified the results were not reproducible, which is not acceptable when the DNA profiles are to be used for individual identification and parentage verification purposes. Garnier *et al.* (2001) also made use of faecal samples from black rhino to verify parentages in black rhino. The samples had to be extracted three times. The best extract was selected and amplified with STR markers six times. The implication is that the cost involved when using faecal samples is six times or more than the cost involved when a blood or tissue sample is used. There is also a higher risk of allelic dropout when working with low quality template DNA. According to a study done by Nielsen *et al.* (2008), the quality of DNA extracted from both black and white rhino faecal samples were not suitable for DNA profiling, since there were too many background peaks to reliably score size alleles. For this study it was decided to make use of blood and tissue samples.

2.2 Sample collection

The Agricultural Research Council and the Mpumalanga Parks Board initiated a project to verify the parentage of white rhino in three nature reserves in this province - Mthethomusha, Loskopdam and Songimvelo. For this study, blood and tissue samples were collected opportunistically, when black and white rhino in these national parks were ear-notched and microchipped as part of the identification system for rhino. Since 1997, ear notches, blood and hair samples were collected from the rhino in these nature reserves. From Songimvelo Nature Reserve (population 1) 8 adult males, 14 adult females and 31 sub-adults and calves have been sampled. In Loskopdam Nature Reserve (population 2), 14 adult males, 11 adult females and 8 sub-adults and

calves have been sampled and from Mthethomusha (population 3), 8 adult males, 15 adult females and 4 calves have been sampled (See Appendix A1-A3 for locations of Nature Reserves in South Africa).

A total of 72 black rhino samples were also collected from the Sam Knott nature reserve in the Eastern Cape, of which 23 were adult males and 32 adult females. The maternal parents of seventeen of these rhino are known, but the paternal parents are unknown. The ear notch samples were received from Mr. B. Fick in 2007 and 2008. See Appendix B1 – B4 for a listing of all the rhino samples received.

All animals were sedated with M99, an etorphine marketed by Norvartis. The dose for black rhino is usually higher per weight than that used for white rhino, since the black rhino can tolerate the negative physiological aspects better. Calves younger than 6 months should not be darted (Hofmeyr 2006).

After a few minutes the animal will show signs of the drug taking effect allowing the veterinarian and the rest of the crew to come closer (See Figure 2.1 and 2.2). The first priority is to cover the eyes and the ears of the rhino to help keep the animal calm. The heart rate is monitored at all times, the microchips are inserted into the horn and blood samples are taken from a vein in the ear for DNA analysis as shown in Figures 2.3 to 2.7. As soon as the rhino has been microchipped, ear-notched and samples taken, the rhino is injected with nalorphine and M50-50 which will immediately improve the respiration of the rhino and allow it to stand up (Hofmeyr 2006).

Figure 2.1: Locating white rhino in the reserve

The first step is to locate the rhino in the reserve either by helicopter or a 4X4 vehicle and dart it with M99. After the animal is darted, it will take approximately 10 minutes for the drug to take effect. Great care must be taken not to lose sight of the rhino otherwise, it may have serious implications. The rhino can fall down at an unsuitable location, for example near a cliff and get seriously injured. If track of the animal is lost, it may be drugged for too long and the rhino may die.

As the rhino begins to show signs of the drug taking effect, the crew can start moving closer to the rhino. Signs include the rhino standing still on one spot, unable to move and swaying as if intoxicated.

Figure 2.2: The rhino is showing signs that the M99 is taking effect

As soon as the rhino is down, the eyes have to be covered with a blanket and earplugs inserted into the ears. This will help to keep the rhino calm. Always be alert for other rhino in the area.

Figure 2.3: The eyes and ears of the rhino are covered to keep it calm

The general health of the animal is checked and all visible wounds are treated. The heart rate and breathing is monitored at all times to ensure that the animal does not succumb to the effect of the M99 drug.

Figure 2.4: All visible wounds are checked and treated

A cordless electric drill is used to drill a hole in the horn so that a microchip can be inserted and sealed with silicone. The microchip number is recorded together with the ear notch identification number for studbook record keeping.

Figure 2.5: Microchips are inserted into the horn and the unique number is recorded for the studbook

Blood samples are taken for DNA profiling. The rest of the samples are stored in the biobank of the National Zoological Gardens for safekeeping and future research programs. The profiles will be used to verify the parentage of the calves and the results will be included into the studbook. This information will help management make informed decisions regarding the translocations of rhino and current breeding programmes.

Figure 2.6: Collecting a blood sample for DNA profiling

Measurements are taken from all animals. An estimation of the calves' age is made.

Figure 2.7: Measurements taken are included into the studbook

2.3 DNA extraction

White rhino samples were kept frozen until DNA was extracted using the commercial High Pure PCR Template Kit (Roche©). The prescribed protocol for mammalian tissue was followed. After extraction, the blood and tissue samples were aliquoted and the white rhino samples were stored at the BioStore at the Agricultural Research Council (ARC), Irene.

The DNA of the black rhino samples was extracted using the Qiagen DNA extraction kit using the tissue sample spin column protocol. The Qiagen kit protocol proved to be faster and easier to use than the High Pure PCR Template Kit. After extraction, the black rhino samples were aliqoted and stored at the biobank of the National Zoological Gardens of South Africa. All extracted DNA was stored at -5ºC.

2.4 Microsatellite markers

Species-specific as well as cross-species microsatellite markers were screened on the black and white rhino samples. Table 2.1 indicates the list of primers that were available. The use of cross-species microsatellite loci saves time and effort in the laboratory, as well as on developing costs (Bonnet *et al*. 2002). A total of 85 microsatellite markers were randomly selected and screened, of which fifteen were nominally species-specific for black rhino (Brown and Houlden 1999; Cunningham *et al.* 1999; Scott 2008) and only five were species-specific for white rhino (Florescu *et al.* 2003). Previously discussed genetic studies by Garnier *et al.* (2001); Harley *et al.* (2005) and Nielsen *et al.* (2008), used combinations of these published markers.

None of the equine, camelides, bovine, ovine or caprine markers amplified, but one of the porcine microsatellite markers (SW035) did amplify and was included in the set of markers used. The final set of markers chosen for the DNA profiling process is listed in Table 2.2. The markers were chosen depending on levels of polymorphism and reproducibility. There were nine polymorphic microsatellite markers selected for the black rhino and 11 polymorphic microsatellite markers for the white rhino.

Table 2.1: List of microsatellite loci tested on black and white rhino samples

White Rhino	Black Rhino	Porcine	Equine	Camelides	Bovine	Ovine	Caprine
7C	BR4	SW035	VHL20	VOLP10	BM1824	HSC	INRA231
32A	BR6	SO035	HTG4	VOLP67	BM2113	INRA23	SRCRSP5
32F	BR17	SW1041	AHT4	YWLL44	ETH10	OARFCB20	SRCRSP8
12F	DB1	SW21	HMS7	VOLP32	ETH225	CSSM36	INRA63
7B	DB5	SW2404	HTG6	YWLL08	ETH3	MGTG4B	SRCRSP24
	DB14	SW749	HMS6	YWLL38	INRA23	MCM185	BM1329
	DB23	SO006	AHT5	LCA66	SPS115	CMC527	BM1818
	DB30		ASB2	VOLP03	TGLA122	AGLA293	ILSTS87
	DB42		HMS2	LCA63	TGLA126	TGLA57	SRCRSP23
	DB44		HMS3		TGLA227	ILSTS002	OARFCB11
	DB49		HTG10		TGLA53		RM4
	DB52		HTG7				INRA005
	DB66						INRA172
	BIRH2B						INRA006
	BIRH37D						ILSTS19
							MAF65

Table 2.2 List of microsatellite loci used for DNA profiling for black and white rhino, including GenBank number, primer sequences (forward and reverse) and polymorphic (P) and monomorphic (M) indications

Locus	GenBank	F-primer sequence	R-primer sequence	Black rhino	White rhino
***7C	AY138543	TGA ACT CTG ATG GAA AG	AAA CAG GTC TTG ATT AGT GC	-	P
***32A	AY138541	CAG TCC TGC TGC ATA AAT CTC	GCA GTA CAG CTA GAA TCA CC	P	P
***7B	AY138544	CCT CTG TGA TTA AGC AAG GC	ATG AAC AGG AAG GAA GAC GC	-	P
*BR4	-	CCC CTA AAT TCT AGG AAC AC	CCA AAG ACC ACC AGT AAT TC	P	M
*BR6	-	TCA TTT CTT TGT TCC CCA TAG CAC	AGC AAT ATC CAC GAT ATG TGA AGG	P	P
*BR17	-	ACT AGC CCT CCT TTC ATC AG	GCA TAT TGT AAG TGC CCC AG	P	M
** DB1	AF129724	AGA TAA TAA TAG GAC CCT GCT CCC	GAG GGT TTA TTG TGA ATG AGG C	P	P
**DB23	-	CCT CAG CAA TAA GGG GAG GAT TAG C	GTT GAT TCT CTG CCC CTG AGT TTG GG	P	M
** DB42	AF129730	CCT GTT AGT GTA ACT TCT ATG CTC CC	CAT GGA TGT TAG CTC AGG GCT GAT C	-	P
**DB44	-	GGT GGA ATG TCA AGT AGC GG	CTT GTT GCC CCA TCC CTG	M	P
** DB49	AF129732	GTC AGG CAT TGG CAG GAA G	CAG GGT AAG TGG GGG TGC	-	P
** DB52		CAT GTG AAA TGG ACC GTC AGG	ATT TCT GGG AAG GGG CAG G	-	P
** DB66	AF129734	CCA GGT GAA GGG TCT TAT TAT TAG C	GGA TTG GCA TGG ATG TTA CC	-	P
¥SW035	–	TCA AGT TGG AGA GTC TGA GGC	AAG ACT GCC CAC CAA TGA G	P	P
****B1RH2B	AY606080	CCC TTT TCT CCC TTT ATC TAG	ATA CTG TGA AAT CCT GTT CC	P	-
****B1RH37D	AY606083	CCA CTC AGA ATG AGA AAT GG	TCT CCC TAC TTA ATC CCA CC	P	-

¥ Rohrer *et al.* 1994
*Cunningham *et al.* 1999
**Brown and Houlden 1999
***Florescu *et al.* 2003
**** Scott 2008

2.5 Optimization and multiplexing of STR primers

The approach used to minimize cost during fragment analyses was to multiplex reactions. A multiplex corresponds to the amplification of several markers in the same PCR (Bonnet *et al*. 2002). When multiplexing is applied, it is important that the primers that are used are 18–24 bp in length and have a GC content of 35-60%. The annealing temperature of the PCR will thus be 55-58˚C or higher. The concentration of primers that amplify well must be diluted and the concentration of the weaker primers must be increased to balance the amplifying products (Henegariu *et al.* 1997). Other complications regarding multiplexing may include the formation of primer dimers, poor sensitivity and specificity and the preferential amplification of certain specific targets (Markoulatos *et al.* 2002).

A series of dilutions were made to optimize the various primers tested. The optimum concentration of each primer needed to be established to ensure that the peak heights (fluorescent signals) were between 150 and 4000 RFU (relative fluorescence units). This is because the ABI PRISM® 377 instrument used can only convert a specific and limited range of fluorescent signals into digital values. Too low peaks make differentiation between samples peaks and background fluctuations problematic. If the peaks are too high, the instrument cannot measure the true value of the signal. The ABI instrument also cannot compensate for the spectral overlap among the dyes and as a result artefact peaks, called pull-up peaks, can appear in other colours (Applied Biosystems 1998; Butler 2005). Optimizing the concentration of the primers will ensure that genotyping errors are kept at a minimum.

Considering that cross-species application of primers was employed, the optimization of the primers was done using the DNA of species for which each was developed. The series of dilutions consisted of a 50 pmol, 25 pmol, 12,5 pmol, 6,25 pmol, 3,125 pmol, 1,262 pmol, 0.781 pmol, 0.391 pmol and 0.195 pmol primer concentration (forward and reverse primers in equal amounts). It is important that the primer-to-template ratio is optimal, to prevent the formation of primer dimers that will have a negative impact on the DNA profile analysis process (Markoulatos *et al.* 2002).

To make the profiling more economical, the microsatellites that amplified successfully were divided into three separate fluorescent labelled multiplex PCRs for

the white rhino. Multiplex combinations took into account the capacity of co-amplification with the same PCR conditions, the absence of artefacts and the absence of overlapping allelic size ranges of markers. The forward primers of the microsatellite markers were labelled with different fluorochromes (either Fam, Joe or Tamra). For the black rhino samples, the loci were amplified using separate PCR reactions for better amplification, but co-loaded on to the ABI instrument to save on time and running costs.

Multiplex PCR reactions were carried out with 10-50 ng of genomic DNA, 2U AmpliTaq Gold™, 1x Buffer (with 1.5 mM $MgCl_2$), 0.1-0.5 µM of each primer 250 µM per dNTP, with ddH_2O added to 7.5 µl, and using the following PCR conditions: 10 min denaturation step at 95°C; 35 cycles of 1 min at 94°C, 1 min at 60 - 63°C (depending on the multiplex); 1 min at 72°C, and a 60 min elongation step in a Perkin Elmer thermocycler (9600). The constituent of each multiplex is shown in Table 2.3.

The PCR products were diluted 5X with sterile water. One microliter of the diluted amplicon was mixed with 2 µl loading mixture which consisted of deionized formamide, loading buffer and Rox350 as an internal size standard. Samples were denaturated for 3 min at 95°C and loaded on a Long Range™ gel (Applied Biosystem) and separated on an ABI PRISM 377 genetic analyzer.

Table 2.3 Multiplexes used for white rhino PCR reactions

Plex 1	Plex 2	Plex 3
DB42	DB66	SW35
32A	7C	DB44
BR6	DB1	DB49
DB52		7B

2.6 DNA profile analysis

Genescan™ Analysis 2.1 and Genotyper™ 2.5 software were used to analyse the data after electrophoresis. GeneScan is used to size and quantify the DNA fragments and to adjust the spectral overlap of the dyes used. This makes it possible to scrutinize one dye color at a time.

The DNA fragments were sized relative to an internal size standard that was mixed with the DNA samples. The sizing and quantifying of DNA fragments were done automatically which allowed faster and more accurate analysis than other detection methods. Depending on the running conditions, sufficient resolution was achieved to differentiate between fragments that had apparent sizes of up to 5000 base pairs. The results of the gels were displayed as electropherograms and tabular data. Each electropherogram represented a single gel lane. The tabular data provided precise sizing and quantitative information (Applied Biosystems 1997; Butler 2005).

After automatic gel analyses, the gel tracking was checked. If a problem occurred, the misaligned lanes were manually retraced. The analysis parameters were set so that the analysis range was between 800 and 6000 datapoints. The peak detection thresholds were set for 100 RFU for the blue- (FAM), green- (JOE) and yellow peaks (TAMRA) and 50 RFU for the red peaks (ROX). The minimum peak half width was set as 3 points.

After the parameters were set, the size standard was defined. GeneScan ROX 350 was used as an internal size standard with the following peak sizes: 35; 50; 75; 100; 139; 150; 160; 200; 250; 300; 340; 350. These size standard fragments were used to calculate the sizes of the unknown DNA fragments by using the Southern method (Butler 2005). For the local Southern method to work there must be at least two size standard fragments larger than the largest unknown fragment and two smaller than the smallest unknown fragment. After the data was analyzed, the peak assignments for the size standard were verified in all the sample files by using the results control method. A new size standard was defined for those samples with incorrect peak assignments. The data was exported to the Genotyper® software (Applied Biosystems 1997; Butler 2005).

2.7 Scoring and binning of microsatellite alleles

Genotyper software was used to convert GeneScan fragment data to called alleles defined by the user, in other words it was used to determine each sample's genotype. The software labels the fragments with identifying labels. The alleles were thus displayed as labelled peaks in a plot display. Histograms and bins were compiled to characterize the peak data. Tables were created that correlated the peak data and

marker information (Applied Biosystems 1996). For alleles that did not fit in a particular bin, the allele name was given as unknown and allele size calling was done manually. See appendix C for DNA profiles.

2.8 Genotyping errors

A genotype error can be defined as an observed genotype of an individual that does not correspond to the true genotype of that individual. A percentage of genotyping errors is unavoidable during fragment analysis. Even a genotyping error rate of 1% (which is considered as a good value for most studies) has a significant impact on large sample sizes (Selkoe and Toonen 2006). An example of a study done on chimpanzee (*Pan troglodytes*), where genotyping errors led to false conclusions was discussed by Pompanon *et al.* (2005). It was found that males in the group were excluded as possible fathers for half of the offspring. The conclusion was thus made that half of the offspring were sired by males from outside the group. It was later found that the exclusion of males was made because of genotyping errors and that these males were indeed the fathers of the offspring.

There are three stages of the genotyping process which are crucial for the successful and accurate genotyping of samples. These include the matrix file, the internal size standard and the external size standard. The matrix file is used for proper colour separation in an electropherogram. If the colour separation is not accurate, pull-up peaks will be observed that can cause faulty genotyping. The internal size standard (ROX350 in this case) is necessary for proper sizing of DNA fragments. If some of the peaks of the internal size standard are under the detection values, the sample fragments will be incorrectly sized which will lead to an incorrect genotype. The external size standard is used to verify the genotyping process. The external size standard has known fragment sizes. The unknown samples are compared to the "known" sample to ensure that the genotyping process was done accurately (Butler 2005).

The main causes of genotyping errors include: a) low quantity and quality of DNA template which causes samples not to amplify very well which can lead to allele dropout; b) the results are misinterpreted because of biochemical artefacts such as pull-up peaks, split peaks and stutter peaks; c) contamination could have occurred; d)

human error such as mislabelling of samples or data that are incorrectly entered into the system; e) a mutation in the primer-binding area which can cause null alleles and thus incorrect genotypes (Pompanon *et al.* 2005; Selkoe and Toonen 2006).

2.8.1 Low quantity or quality DNA

Great care should be taken when genotypes are analysed from profiles that were obtained from low quantity or quality DNA, since there is an increase in allelic dropout and stutter peaks (Butler 2005). Allelic dropout is a chance event where one allele at a particular marker is preferentially amplified to the other. The result is that a homozygote is observed in the electropherogram, but the true genotype is actually a heterozygote (Butler 2005; Goodwin *et al.* 2007).

2.8.2 Biochemical artifacts

Several biochemical artifacts may be present in the profile when a genotype is analyzed. These include stutter, split and pull-up peaks, and the effect of overloading.

2.8.2.1 Stutter peaks

Stutter peaks are formed by strand slippage during the extension step of the PCR process and are one repeat unit smaller or larger than the true allele. The stutter peaks are smaller than the true allele in most cases. Different microsatellite markers have different tendencies to stutter depending on the core repeat. Di-nucleotide repeats are more prone to stutter than tetra-nucleotide repeats. Stutter peaks are given threshold values which simplifies the interpretation of microsatellites. Difficulty in analyzing a DNA sample which is heterozygous for adjacent alleles can however present a problem (Goodwin *et al.* 2007).

2.8.2.2 Split peaks

The *Taq* polymerase used in PCR reactions for adding nucleotides to the newly synthesized DNA molecule has an activity known as terminal transferase. A

nucleotide is added to the end of the amplified molecule which is non-template related. In other words an extra nucleotide (usually an adenine residue) is added to the fragment. If a split peak is observed in the electropherogram, it is an indication that half of the amplified DNA product has an addition of an adenine where as the rest of the amplified DNA does not. This can complicate the analysis procedure. To overcome this problem, an extra incubation step of between 45 – 60 minutes at 72°C was added to the PCR program to ensure that all the amplified DNA fragments had an extra adenine (Goodwin *et al.* 2007).

2.8.2.3 Pull-up peaks

Pull-up peaks occur when the matrix file, which is responsible for colour separation, is not of good quality. Consequently, the correction is not perfect and the peaks are composed of more than one colour. Pull-up peaks can also occur if over-amplification took place. Pull-up peaks are easy to recognize since they are a smaller sized product that are exactly the same size as the real microsatellite allele, only in a different colour (Goodwin *et al.* 2007).

.

2.8.2.4 Overloaded profiles

Overloading can also lead to difficulty in interpreting DNA profiles. The CCD camera of the instrument is saturated and the peak height will no longer be a good indicator of the amount of product that was amplified (Goodwin *et al.* 2007). This problem can be overcome by diluting the amplicon product and re-running it on a gel.

2.8.3 Contamination

The ability of PCR to amplify low quantities of DNA can present a problem if proper care is not taken. Contamination could originate from various sources and can cause allele drop-in to occur, which is the appearance of an allele that is not from the analysed sample (Budowle *et al.* 2009). Validated laboratory procedures should thus be followed at all times. There are three potential sources of contamination during the PCR process: 1) genomic DNA from the environment; 2) contamination between

samples during preparation; and 3) contamination with amplified product from a previous PCR reaction. Negative and positive controls should thus be included in all PCR reactions (Goodwin *et al.* 2007).

2.8.4 Human error

A study done by Hoffman and Amos (2005) found that 93% of genotyping errors occurred because of human error. Eighty percent of these mistakes were due to faulty scoring of alleles, 10.7% were errors that were made during the data input and only 6.7% were due to allelic dropout. The remaining 2.7% were due to pipetting errors, sample mix-ups and contamination. Allelic calling has thus been identified as the most important source of genotyping error. These authors found that the most common form of incorrect scoring occurred with adjacent allele heterozygotes where the one allele was mistakenly taken for a stutter peak.

2.8.5 Null alleles and allelic drop-out

Mutations in the flanking region sequences of a microsatellite marker may prevent the primer to anneal to the template DNA resulting in a failure to amplify. A heterozygote will thus be interpreted as a homozygote. Other possible causes of null alleles include the preferential amplification of short alleles or slippage during PCR amplification (Chapuis and Estoup 2007).

According to Chapuis and Estoup (2007) there are several studies that have suggested that the sequences of the flanking regions of microsatellites are less stable than those in other genomic regions. To date, no correlation has been found between the frequency of null alleles and the allele frequency or type of repeat unit. Null alleles are however likely to be found in populations with high levels of diversity in flanking sequences and in populations with large effective sizes. The presence of null alleles may lead to the overestimation of both F_{ST} and genetic distance values in cases of significant population differentiation.

Allelic dropout was initially thought to occur only with low quantity and quality template DNA samples. However, a study done by Soulsbury *et al.* (2007) suggested that allelic drop-out can also occur at specific loci and at specific pair-wise

combinations of alleles even though high quality DNA was used. This is an indication that other mechanisms are also responsible for allelic dropout. These authors found no significant relationship between DNA age and frequency of allelic drop-outs, nor were allelic drop-outs and amplification success related to median allele size. It was thus concluded that allelic drop-out is not a random process.

2.9 Limiting genotyping errors

Genotyping errors can have serious implications for individual identification studies as well as parentage verification studies, especially in this study where the DNA profiles will be used in the future as reference DNA profiles. In studies where non-invasive samples are used, the chances of genotyping errors are increased, for example the rhino study done by Nielsen *et al.* (2008) as mentioned in Chapter 1, where the authors made use of faecal samples and concluded that DNA profiles obtained were not reliable. Creel *et al.* (2003) carried out a study on the wolves in Yellowstone National Park. The aim of their study was to use non-invasive samples to obtain DNA profiles which would then be used to estimate the population size of the wolves in Yellowstone. Due to genotyping errors however, the population size was 5.5-fold overestimated. These results could have had serious conservation implications if it was used for management decisions.

Dakin and Avise (2004) did a study on the impact of genotyping errors and null alleles on molecular parentage analyses. They surveyed 233 articles and concluded that the most common approach was to report the detection of null alleles but with no further corrective action taking place. In parentage verification studies, the real parent can thus be excluded as a parent and this will have a serious impact on the accuracy of studbook keeping.

The impact of genotyping errors will be less on population studies where analyses are frequently based on allelic frequencies than studies that are based on individual identification. Overall, there is still a great need for studies on the effect of genotyping errors in interpretation of results (Pompanon *et al*. 2005).

To avoid genotyping errors, it is advised that good quality samples are obtained for studies. Pompanon *et al.* (2005) suggested that the process must be kept as automated as possible and that human manipulation should be avoided, since this is

the area where most of the genotyping errors occur. Semi-automated scoring followed by visually inspection is found to be the most reliable scoring procedure.

The addition of positive and negative controls with all samples is a good approach. To determine the error rate, 5–10% of samples can also be blindly repeated, but this will have a cost implication for the study project. Most genotyping errors are however detected when pedigrees are known and samples are checked for Mendelian inheritance and linkage analyses (Pompanon *et al.* 2005). The more polymorphic a marker is, the more caution should be applied when analyzing the profile. These markers will have a higher mutation rate which will thus include null alleles. It is also more convenient to use markers with smaller fragment sizes, because even poor quality DNA will amplify successfully (Hoffman and Amos 2005). Genotyping errors can be spotted by either duplicating results or by verifying results with known pedigree information (Pompanon *et al.* 2005).

The use of mother-offspring mismatches is useful for detecting genotyping errors, but should be used in conjunction with other approaches, because this approach relies critically on how many offspring the mother had and whether the relatedness is genuine (Hoffman and Amos 2005).

2.10 Testing profiles for genotyping errors and null alleles

Before continuing with the statistical analysis of profiles, it is thus very important to screen a dataset for genotyping errors and null alleles, since this could have a major impact on the interpretation of results. Micro-checker (Van Oosterhout *et al.* 2004), is a software program that aids in the identification of genotyping errors of microsatellites from diploid populations. This program helps with the identification of null alleles, drop-out alleles, the scoring of stutter peaks and it also detects typographic errors. Null alleles have its' own specific deficiency and excesses of particular genotypes as opposed to allelic drop-outs which is assumed to be largely independent of allele sizes (Van Oosterhout *et al.* 2004).

The DNA profiles of the black and white rhino populations were separately analysed with Micro-checker to scan for genotyping errors. The data was firstly screened for values that were out of range or zero values and secondly for values that

had inconsistent base-pair repeat units. The confidence interval for the Monte Carlo simulations was set at 95%.

There were a total of 36 alleles found in the white rhino population with a zero value, which were microsatellite loci that did not amplify. There were also six loci found in the black rhino population with a zero value. The analysis was continued with the option of omitting the faulty values from the analysis.

There were no microsatellite markers found with null alleles present in the black rhino population. There were however three markers in the white rhino populations that may have null alleles present namely DB42; DB66 and 7C. The excess of homozygotes at these three loci can be attributed to either stuttering that resulted in scoring errors or null alleles. Care should thus be taken with these three loci when verifying parentage. There is however no indication that allelic drop-out took place.

To continue with further statistical analysis, the software program *Create*, (Coombs *et al.* 2009), was used. This program can be used to generate input files for the different software programs that are usually used for multi-purpose analysis, specialized applications, quantitative genetic programs, sib-ship reconstruction and parentage reconstruction programs. The results will be discussed in the following chapters.

Information regarding the population structure and patterns of gene flow between wildlife populations are crucial to determine which populations are in need of human-assisted exchange of individuals to prevent inbreeding and loss of genetic diversity (Frankham 2002).

3.1 Statistical analysis of microsatellite data

As described in Chapter 1, microsatellites are either di-, tri- or tetranucleotide repeats units. These markers are highly polymorphic and have an unusually high mutation rate that renders them particularly useful for estimating levels of genomic diversity. The mutation rates are in the order of 10^{-3} per locus per generation (Väli *et al.* 2008).

There are important assumptions that are made when using molecular markers that should be kept in mind when results are interpreted. Firstly, it is assumed that the observed polymorphisms are neutral and secondly that the loci used are indeed a good predictor of the overall genomic diversity of the populations studied (Hanotte and Jianlin 2005). Väli *et al.* (2008) found that the microsatellite marker heterozygosity found in grey wolves (*Canis lupus*), coyotes (*Canis latrans*), Eurasian lynx (*Lynx lynx*) and wolverine (*Gulo gulo*) underestimated the significant difference in nucleotide diversity among these carnivore populations and the authors suggested that sequencing-based measures of genetic diversity, such as SNPs should be included in these types of studies. Ljungqvist *et al.* (2010) however stated that previous studies did show that heterozygosity at more variable markers would provide a stronger prediction of genome-wide genetic diversity than markers such as SNPs with little variation. SNP studies will also require a substantial effort in term of number of loci genotyped.

Microsatellite markers can be used to estimate population diversity, differentiation, and connectivity. Genetic characterization provides new information to guide and prioritize conservation decisions. The mean number of alleles (MNA) and observed (Ho) and expected (He) heterozygosity are the most commonly calculated population genetic parameters for assessing within population diversity. The classic parameters for assessing drift between populations using genetic markers are the genetic differentiation fixation indices (F_{ST} and related R_{ST}-statistics). The F_{ST}

measures the degree of genetic variation between subpopulations by calculating the standardised variances in allele frequencies amongst populations (Hanotte and Jianlin 2005). R_{ST} is an analogue of F_{ST}, with the difference that a stepwise mutation model is assumed. R_{ST} can thus be calculated from the variances of allele sizes (Balloux and Lugon-Moulin 2002).

3.2 Intra-and inter-population genetic variation in rhino populations

3.2.1 Genetic diversity

Genetic diversity is a key factor in the long-term survival of a species. It enables the species to adapt in a changing environment. Therefore, in conservation genetics, the evaluation of genetic diversity is crucial. The use of neutral markers provides an overall measure of the relative level of genomic variability of a population (Väli *et al.* 2008).

Gene diversity or expected heterozygosity is an important parameter in studies on the genetic structure of populations. Since microsatellite loci are co-dominant, traditional population genetic methods can be used for calculating allelic and genotype frequencies. Standard population genetic models can thus be applied (Gutiérrez-Espeleta *et al*. 2000). The mean number of alleles per locus (MNA), the observed heterozygosity (Ho) and the unbiased expected heterozygosity (He) under Hardy-Weinberg equilibrium were computed by using the software program MICROSATELLITE ANALYSER (MSA). MSA is a program that was specifically developed for large microsatellite data sets and inbred lines. This program uses the expected number of alleles under the infinite allele model as well as under the stepwise mutation model (Dieringer and Schlötterer 2003).

Only the markers (loci) that yielded results in both black and white rhino populations were used in the statistical analysis. The loci included in the statistical analysis are listed in Table 3.1. The black rhino had only one monomorphic marker (DB44) and the number of alleles for the other loci ranged from 2 to 6 alleles. The white rhino populations however, had three monomorphic markers (BR4, BR17 and DB23) and the number of alleles for the other loci ranged from 2 to 6 alleles as well.

The mean number of alleles (MNA) and the average of the expected- and observed heterozygosities from both black and white rhino populations are listed in Table 3.2. The mean number of alleles (MNA) and the observed heterozygosities is higher in the black rhino population than in the combined white rhino populations. The white rhino populations thus have less genetic diversity than the black rhino population which could be attributed to the fact that there were only about 28 white rhino left during the 1900s.

Table 3.1 Microsatellite loci used for statistical analysis of population structure for black and white rhino populations, the number of alleles and the size range in base pairs

Population	Loci	Number of alleles	Size range (bp)
Black rhino	BR4	4	124 – 130
	BR6	3	134 – 154
	BR17	6	122 – 140
	DB1	2	127 – 131
	DB23	2	180 – 182
	DB44	1	176
	32A	2	242 – 244
	SW35	4	123 - 131
White rhino	BR4	1	109
	BR6	6	133 – 153
	BR17	1	118
	DB1	4	127 – 141
	DB23	1	170
	DB44	5	171 – 181
	32A	2	234 – 248
	SW35	2	127 - 133

Table 3.2 The samples size, mean number of alleles (MNA), the average expected heterozygosity and the average observed heterozygosity of the black and white rhino populations studied

Population	Sample size	MNA	Average He	Average Ho
Black – Sam Knott	72	3.125	0.428	0.428
White – Mthethomusha	27	2.500	0.330	0.313
White – Loskopdam	33	2.125	0.242	0.269
White - Songimvelo	53	2.250	0.254	0.250

A fundamental point of departure in population genetics is to test the loci for conformation to Hardy-Weinberg equilibrium (HWE), in which the observed genotype frequencies are compared to the expected frequencies for an ideal population (Selkoe and Toonen 2006).

The Hardy-Weinberg principle states that after one generation of random mating in the absence of factors such as selection, genetic drift, gene flow and mutation, the genotype proportions will not change over time (Hedrick 2005b). In other words in a large random mating population in the absence of mutation, migration or selection the allele and genotype frequencies will attain an equilibrium, referred to as the Hardy-Weinberg equilibrium (Frankham *et al.* 2002). A heterozygote deficit (or homozygote excess) occurs when the data set contains more homozyotes than expected under HWE. This can be due to biological violations of the set criteria of an ideal population, for example inbreeding or selection could be occurring in the population. A phenomenon known as the Wahlund effect can also cause a heterozygote deficit. This occurs when two genetically distinct groups are lumped into a single sampling unit. It can also be caused by the spatial scale chosen for sampling, a site larger than the true scale of a population. A third factor that may result in a heterozygote deficit is the presence of null alleles, as was discussed in Chapter 2 (Selkoe and Toonen 2006).

Detecting significant deviations from HWE requires large sample sizes and strong disturbing forces. Lack of significance cannot be interpreted to mean that no evolutionary processes are at work in populations. Different processes may be acting in a way that are not detectable with a goodness of fit test, or they may be too weak to be detectable with the given sample size (Halliburton 2004).

Exact tests for deviations from Hardy-Weinberg equilibrium (HWE) were performed by using MSA. Significance levels were calculated per locus, per population (Table 3.3). As seen in Table 3.2, the observed heterozygosity and the expected heterozygosity for all the populations studied were very similar. When the loci were tested separately, DB1 and BR17 in the black rhino population and SW35 in two of the white rhino populations however deviated significantly from HWE (Table 3.3).

Table 3.3 Black and white rhino populations tested for Hardy-Weinberg equilibrium

Locus	Black rhino	White rhino
32A	NS	NS
BR6	NS	NS
DB1	P = 0.00033	NS
DB44	M	NS
SW35	NS	P = 0.01644 (Songimvelo), P = 0.01990 (Mthethomusha)
BR17	P = 0.00235	M
DB23	NS	M
BR4	NS	M

NS = not significant

P < 0.05 significant

M = Monomorphic

When two loci are situated very close together on a chromosome, they may not assort independently and will thus be transmitted to offspring as a linked pair. The loci can however also be functionally related to each other. This will also cause the loci to be inherited together, even though they are not situated on the same chromosome. This phenomenon is known as gametic disequilibrium or linkage disequilibrium (Selkoe and Toonen 2006).

Microsatellite genotypes were tested for linkage disequilibrium (LD) by using the software program ARLEQUIN ver. 3.1 (Excoffier *et al.* 2006). There was no LD observed in the white rhino populations, but there was significant LD between BR17 and DB23 and between DB23 and BR4 observed in the black rhino population.

The sample sizes in this study varied. Consequently it was necessary to test for differences in allelic richness between the different populations. MSA software was thus used (Table 3.4). The overall allelic richness was higher in the black rhino population than for any of the white rhino populations. This is once again an indication that the white rhino have less genetic diversity than the black rhino, even though the white rhino have larger numbers.

Table 3.4 Allelic richness of microsatellite loci used for black and white rhino populations

Population	32A	BR6	DB1	DB44	SW35	BR17	DB23	BR4	Average
Black – Sam Knott	1.2381	1.4375	1.5026	1	1.6025	1.4778	1.5034	1.6937	1.4349
White – Mthethomusha	1.5091	1.6184	1.4879	1.4827	1.5400	1	1	1	1.3248
White – Loskopdam	1.4611	1.5543	1.1678	1.2942	1.4611	1	1	1	1.242
White – Songimvelo	1.4528	1.5468	1.1878	1.3310	1.5118	1	1	1	1.2547

3.2.2 Genetic differentiation among rhino populations

Genetic drift can be defined as changes in the genetic composition of a population due to random sampling in small sub-populations. Chance events can thus have a major impact in small populations; for example it may lead to random changes in the allele frequencies from one generation to the next. It may also lead to a loss in genetic diversity and fixation of alleles within a population. Differentiation among fragmented populations from the same original source can also occur due to genetic drift (Frankham *et al.* 2002).

It is common practice to analyze the population structure to test hypotheses concerning gene flow and isolation within species in population genetics. The most frequently used method is by estimating F_{ST}, which is a measure of population differentiation (Meirmans 2006). The estimation and comparison of both F- and R-statistics are especially relevant for critical comparison and careful interpretation of data and may give the most valuable information about the genetic structure of a population (Oliveira *et al.* 2006). The values of both of these estimates are dependent on the amount of within-population genetic variation.

Wright (1965) used inbreeding coefficients to describe the distribution of genetic diversity within and among population fragments. F-statistics is the measure of total inbreeding in a population (F_{IT}), partitioned into 1) that due to inbreeding within sub-populations (F_{IS}) and 2) that due to differentiation among sub-populations (F_{ST}). F_{IS} is the probability that two alleles in an individual are identical by descent and is calculated as the inbreeding coefficient F, averaged across all individuals from all population fragments. F_{ST}, the fixation index (sometimes referred to as G_{ST}), is the effect of the population sub-division on inbreeding. It is the probability that two

alleles drawn randomly from a population fragment (either from different individuals, or from the same individual) are identical by descent. With high rates of gene flow among fragments, this probability is low. With low rates of gene flow among fragments, populations diverge and possibly become inbred, and F_{ST} increases (Frankham *et al.* 2002). High levels of genetic variation therefore generally lead to lower F_{ST} estimates (Meirmans 2006).

Fixation index (F_{ST}) measurements of genetic differentiation are based on the infinite allele (IAM) or *k* allele models (KAM) of mutation. In these models, mutation generates new alleles which were not present in the population before the mutation event. These models are thus not 100% appropriate, but for microsatellites where most of the mutations involve the addition or subtraction of a repeat unit, the models can still be used. The alleles change according to a stepwise model where the size of new mutant alleles depends upon its progenitor (Goodman 1997).

Slatkin (1995) stated that F_{ST} will tend to underestimate the true level of genetic differentiation when microsatellite markers are used. He proposed a model which will take into account some of the features involved with microsatellite evolution. The R_{ST} coefficient, developed by Slatkin (1995) is an analogue measure to $F_{ST,}$ but is more appropriate for estimating population structure at microsatellite loci with many alleles (Halliburton 2004; Avise 2005). R_{ST} differs from F_{ST} in that it considers allele sizes, which, according to the stepwise mutation model, contain information about the relationships among alleles. Goodman (1997) described a computer program, RST CALC, which calculates R_{ST}. It adjusts for unequal sample sizes and unequal variances in different population (Halliburton 2004).

Gene flow is fundamental for metapopulation management because it allows genetic diversity to be maintained by acting directly on the population structure and against random genetic drift (Oliveira *et al.* 2006). Consequently gene flow, rather than migration rate, needs to be estimated (Frankham *et al.* 2002). The product *Nm*, a demographic parameter describing the effective number of migrants per population and generation (gene flow), can be inferred from F_{ST}. One migrant per generation, $Nm = 1$, is considered enough to prevent the effects of genetic drift among populations (Hedrick 2005b). Gaggiotti *et al.* (1999) suggested that for most typical sample sizes (moderate populations size of N = 50 with 10 loci studied) and genetic parameters encountered in experimental studies, F_{ST} should be preferred over R_{ST} to estimate

gene flow parameters with microsatellites because it generally gave a lower mean square error of *Nm* estimates. Balloux and Goudet (2002) showed that F_{ST} is more efficient in the case of high levels of gene flow whereas R_{ST} better reflects population differentiation under low gene flow (Hardy *et al*. 2003). Oliveira *et al.* (2006) found that with sample sizes larger than 50 individuals and 20 or more microsatellite markers used, R_{ST} performed better than F_{ST} for estimating *Nm*.

To investigate the population genetic structure of rhino, both F_{ST} (Weir and Cockerham 1984) and R_{ST} (Slatkin 1995) statistics were thus calculated as implemented in ARLEQUIN version 3.1 and RST-CALC (Goodman 1997), respectively. The results, as well as the P-values and Bonferroni correction for multiple tests of the P-values, are listed in Table 3.4 and Table 3.5.

Additional hierarchical analyses of variation among locations were conducted using analysis of molecular variance (AMOVA) as described by Excoffier *et al*. (1992) (See Table 3.6). AMOVA was initially introduced as an extension of the analysis of gene frequencies to estimate population genetic structure from molecular haplotype frequencies in haploid organisms, using an analysis of variance framework. It can however also be applied to microsatellite data to obtain an analogue of the R_{ST} statistic in diploid organisms (Michalakis and Excoffier, 1996).

The low F_{ST} and R_{ST} values listed in Table 3.5 and Table 3.6 for the white rhino populations is an indication that the three populations could be considered as a single population. This may reflect the recent management history of these reserves, since regular translocations potentially led to some artificial maintenance of gene flow among fragments There is thus little differentiation between these three reserves which is very important since the genetic diversity of white rhino are already compromised and any further loss of diversity in any of these populations is to be avoided.

Results from AMOVA conform to expectations. The high percentage of variation (64.85% - Table 3.7) between the black and white rhino is expected, since these are different species. The percentage of variation attributed to variation among individual white rhino populations was only 1%, which confirms the conclusion that the white rhino populations can be considered as a single population. The percentage of variation between individuals within the white rhino populations is 96.9% (Table 3.8).

Table 3.5 F_{ST} values (below the diagonal) among black and white rhino populations and the corresponding P-values (above the diagonal)

	Black – Sam Knott	White - Mthethomusha	White - Loskopdam	White - Songimvelo
Black – Sam Knott		0.001	0.001	0.001
White – Mthethomusha	0.597		0.001	0.004
White – Loskopdam	0.638	0.043		0.0001
White - Songimvelo	0.646	0.036	0.040	

Table 3.6 R_{ST} values (below the diagonal) and the *Nm*-values (above the diagonal) among black and white rhino populations

	Black – Sam Knott	White - Mthethomusha	White - Loskopdam	White - Songimvelo
Black – Sam Knott		0.091	0.089	0.0925
White – Mthethomusha	*0.734		18.81	121.751
White – Loskopdam	*0.737	0.013		3.937
White - Songimvelo	*0.730	0.002	0.060	

*P<0.05 significant

Table 3.7 Hierarchical distribution of total genetic diversity among the black rhino population and the three white rhino populations

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	453.397	2.530	64.85
Among populations within groups	2	8.289	0.0.95	1.01
Within populations	366	487.382	1.332	34.13
Total	369	949.068	3.901	

Table 3.8 Hierarchical distribution of total genetic diversity among the three white rhino populations

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	8.289	0.043	3.88
Among populations within groups	109	115.854	-0.009	-0.78
Within populations	112	121.000	1.080	96.90
Total	223	245.143	1.115	

3.2.3 Individual-based analysis

Several studies have shown that microsatellite data can be used to identify the population of origin of an individual (Maudet *et al*. 2002; Kotze *et al*. 2008).

Population assignment tests are very useful for assigning an individual to its population of origin, which is very effective specifically in poaching cases. Poaching and illegal trafficking of wildlife and their products are among the most serious threats to the survival of many wildlife species, including the rhino species as discussed in Chapter 1. There are three possible software packages that can be used to conduct assignments - GENECLASS, WHICHRUN and STRUCTURE (Maudet *et al.* 2002). For this study it was decided to use STRUCTURE for further analysis since this programme is currently used by most population geneticists.

STRUCTURE was developed by Pritchard *et al.* (2000). This program is used to assess the genotypes of individuals in a clustering method (fully Bayesian approach). It involves the placing of individuals into K (number) of populations of origin and it simultaneously assigns individuals to the population with explicit estimates of their 90% confidence intervals. The results provide a probability value that can be interpreted directly as a probability of origin for each individual tested in each population. K is chosen in advance and can be varied across independent runs of the algorithm. Ten individual repetitions of K = (1-4) were run in order to check for consistency in the results. The standard deviation for K = 1-4 was also calculated and plotted in Figure 3.1. The results showed a clear plateau in –Ln(Pr) values at K=2 (Figure 3.1), with this also being the highest probability value overall. Since there is a clear plateau of probability estimates from K=2 onwards, followed by little real difference among probabilities, K=2 was regarded as the true K and it was not considered necessary to use other methods to elucidate true K, such as the ΔK approach of Evanno *et al.* (2005). The rhino populations studied were thus divided into two real genetic populations corresponding to species boundaries, i.e. a black rhino population and a white rhino population. Outcomes for K=2-4 are shown in Figure 3.2a–c. As can be deducted from the results, the three white rhino population should thus be considered as a single, genetically interchangeable population. All the black rhino individuals were clustered together, but in the case of white rhino only nine individuals from the Mthethomusha reserve (population 2) had a similar genotype profile.

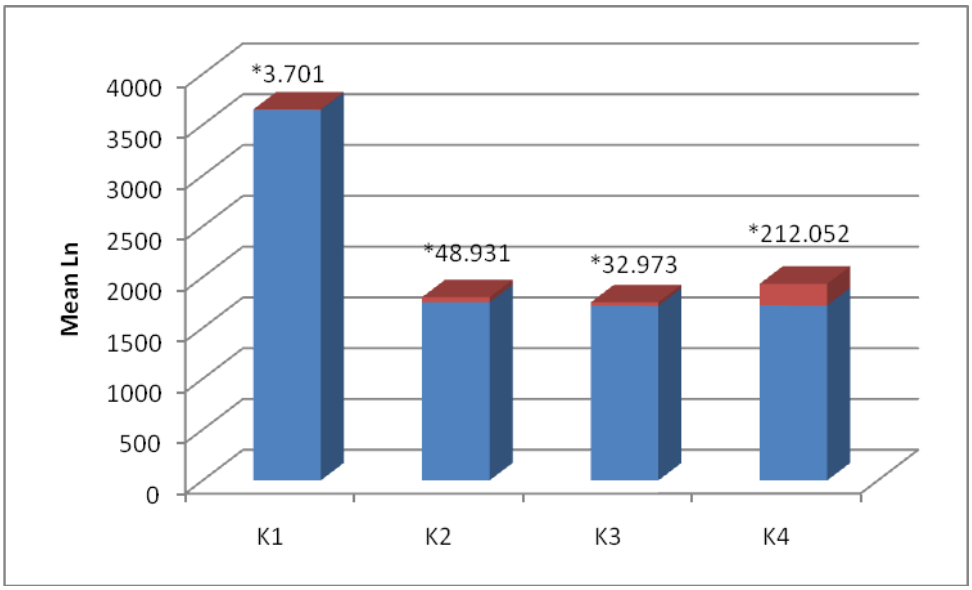
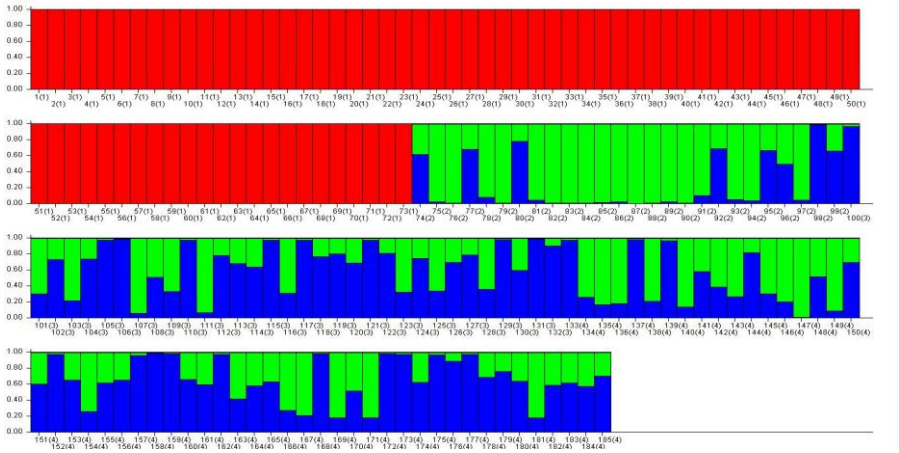


Figure 3.1 Mean –Ln(Pr) (in blue) and *standard deviation (red) for K = 1 – 4. SD values are multiplied by a factor of 10 for clarity of interpretation

a) K = 2

b) K = 3



c) K = 4

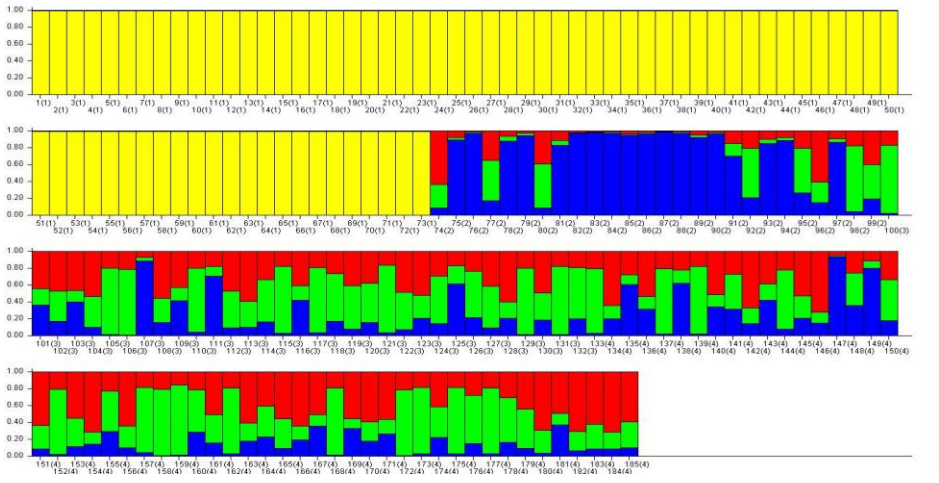


Figure 3.2 a-c Summary plots of STRUCTURE results. Each individual is represented by a single vertical line broken into coloured segments, with lengths proportional to each of the K = 1 - 4 inferred clusters

3.3 Discussion

The eight markers that were genotyped for both black and white rhino's were used during the population structure analysis. Seven of the eight markers were polymorphic for the black rhino whereas only five were polymorphic for the white rhino populations (see Table 3.1). Locus DB44 was monomorphic (176 bp) for the Sam Knott black rhino population, even though previous studies done on black rhino found this marker to be polymorphic. Brown and Houlden (1999) observed six alleles (170 – 184 bp) and Garnier *et al.* (2001) observed three alleles (172 – 176 bp) for marker DB44. Since the black rhino population of the Sam Knott nature reserve was founded from only a few rhino in the same area (personal communication Mr. B. Fick) it can be concluded that because of a founder effect, perpetuated by drift, allele 176 from marker DB44 is fixed for this black rhino population.

There were three monomorphic markers (BR4, BR17 and DB23) identified in the three white rhino populations. Marker 32A, which was also used in studies done by Florescue *et al.* (2003) and Nielsen *et al.* (2008), had only two alleles in this study. Florescue *et al.* (2003) found 3 alleles and Nielsen *et al.* (2008) 5 alleles for the same marker which can be an indication that the three white rhino populations have less genetic diversity than previously studied white rhino populations.

The average observed heterozygosity (Ho = 0.42808) for the black rhino population was found to be lower than for any of the previous studies done on black rhino except for the study of Nielsen *et al.* (2008) that included six black rhino. The observed heterozygosity values for previous genetic studies done on black rhino included: Brown and Houlden, (1999) Ho = 0.686; Cunningham *et al.* (1999), Ho = 0.587; Garnier *et al.* (2001), Ho = 0.72; Harley *et al.* (2005), Ho = 0.436; Nielsen *et al.* (2008), Ho = 0.322 and Scott (2008), Ho = 0.477.

The genetic variation found in the three white rhino population studied, were even less than the genetic variation found in the black rhino population. The observed heterozygosity ranged from 0.25 to 0.32. There were only two previous genetic variation studies done on white rhino, both from the Hluhluwe Umfolozi Game reserve (Florescue *et al.* 2003; Nielsen *et al.* 2008). Florescu *et al.* (2003) found the average observed heterozygosity to be 0.596 for a population of 30 white rhino, whereas Nielsen *et al.* (2008) tested 22 white rhino and found a Ho = 0.463. Both of

these studies indicated that the Hluhluwe Umfolozi Game reserves' white rhino population had more genetic diversity than the three white rhino populations tested in the nature reserves in Mpumalanga.

The average allelic richness of the Sam Knott black rhino population was found to be 1.43. Scott (2008) estimated the allelic richness for *Dicornis bicornis minor* to be 3.1, which is much higher than the value found for the black rhino population of the Sam Knott nature reserve. It can thus be concluded that the lack of genetic variation in the black rhino population of the Sam Knott nature reserve needs to be addressed with management decisions regarding the translocation of animals.

The allelic richness for the white rhino were found to vary between 1.24 – 1.32 which is lower than the estimate for black rhino populations. Scott (2008) found an allelic richness value of 2.2 in 59 white rhino. However, the latter rhino came from three different source populations, Hluhluwe Umfolozi (South Africa), Waterberg National Park (Namibia) and Metro Toronto zoo (Canada) which can most likely explain the difference in allelic richness. White rhino in all case studies tends to have less genetic diversity than black rhino. This is due to the fact that the white rhino population numbers had decreased to about 28 rhino whereas the black rhino were at approximately 400 when conservation efforts were implemented.

The average F_{ST} value between the black rhino population and the three white rhino populations was 0.627, which clearly indicated that these are two different species. The F_{ST} values among the three white rhino populations was however very low. The lowest F_{ST} value (0.036) was found between the rhino from Mthethomusha and Songimvelo. There have been a higher number of rhino translocations taking place between the three reserves which is clearly supported by the low F_{ST} values. R_{ST} and *Nm* values supported the trend that most gene flow has occurred between Mthethomusha and Songimvelo. The summary plot of the STRUCTURE results (Figure 3.2a-c) also clearly indicates that there are no significant differences between the three white rhino populations. The AMOVA performed on the data also only gave 1.01% of variation among the three white rhino populations, which is an indication that the three populations can be considered as a single population.

The F_{ST} and R_{ST} values, as well as the results obtain from STRUCTURE and the AMOVA testing all indicate that the three white rhino populations can be considered as a single, genetically interchangeable population. Even though meta-

population management has only been implemented for a short period of time for the white rhino populations of Mthethomusha, Loskopdam and Songimvelo, the results given is an indication that the rhino population is being successfully managed as a meta-population from a genetic point of view. It should however also be taken into account that these reserves have existed for only a few decades, providing little time (in evolutionary terms) for significant drift. The within population variation was found to be 34.13% which was expected since all the other previously mentioned white rhino studies clearly indicated a lack of genetic variation in white rhino populations. The loss of genetic diversity is thus a serious threat to the survival of this species.

In conclusion, the black rhino population has less genetic variation than previously studied populations of the same species. This problem can be overcome with a few well thought out translocations and with the genetic information obtained in this study regarding parentage and kinship (see Chapter 5), more effective management decisions can be made. The white rhino populations have low genetic diversity, but are being successfully managed as a meta-population. Translocation of white rhino from outside into these three reserves should however be considered to improve the genetic diversity of these populations.

4.1 Introduction

In recent times, convicted rhino poachers and smugglers of rhino products have been sentenced to imprisonment of up to 10 years in South Africa (SAPA 2010h; SAPA 2010i). One of the problems encountered in convicting rhino poachers is to prove the identity of the stolen biological products that were found on the poacher or smuggler in a court of law. In other words, that the animal skin, meat or horns found on the suspect, belonged to the animal poached. DNA profiling can be used to solve this problem. If the poached rhino has been previously genotyped, the DNA profile can be used as a reference sample. If the profiles match, the identity of the rhino has been confirmed. The genetic identification however relies on accurate reference samples and data. Validation of reference data for population and individual identification is thus of fundamental importance.

The practical application and difficulty of keeping reference material and DNA profiles of wild animals on a database or studbook has limited the exchange of data and development of DNA profiling techniques in wildlife forensics (Ogden *et al.* 2009). One of the difficulties include that in species such as springbuck and kudu, most individuals look alike and are thus not visually distinguishable from one another. It is thus difficult to collect reference samples that can be traced back to the original animal. Owners of wildlife prefer not to make use of ear tags for identification purposes, because ear tags create the impression of domestication. Since these animals are sold mainly for hunting purposes, the impression of domestication should be avoided. If the visual identification of an animal is not possible, it will not be possible to compare evidence material with a reference sample, and identification and parentage verification will thus be impossible. The assignment of an individual to a specific population of origin based on their DNA profile is however still possible.

Individual identification by using microsatellite markers is becoming very important in wildlife forensic cases, since this technique can provide key evidence to wildlife crime investigations (Waits *et al.* 2001). Not only is DNA profiling currently being used in cases where animals were poached for their meat, horns, tusks, bones or skin, but it is also being used in cases that involve the illegal trading of animals. For example, the profiles can be used to prove that an animal was bred in captivity and

not illegally taken from the wild. The ability to verify familial relatedness is an important application in wildlife forensics (Ogden *et al.* 2009).

There are two additional factors that can cause problems with individual identification. The first problem is the use of a low number of loci, since there is a greater possibility that two individuals will have the same DNA profile when a low number of loci are used. Secondly, genotyping errors, such as those that were discussed in Chapter 2, can cause a negative exclusion. Both of these problems can be addressed by using the appropriate number of microsatellite markers (Kalinowski *et al.* 2007). In this chapter, a set of markers will be evaluated with respect to their ability to differentiate between individuals as well as the minimum number of loci necessary for identification purposes in both black and white rhino.

4.2 Individual identification and probability of identity (P_{ID})

Individual identification by using a group of co-dominant nuclear DNA microsatellite loci is well established in human populations, but the estimation of match probability has been controversial. One reason for controversy is the potential biases in data sets due to violations of Hardy-Weinberg equilibrium and linkage disequilibrium. In these circumstances the assumption of the product rule that the loci are independent, are violated and the estimation of the probability of a match will be lower than the "true" probability of a match. Theoretical evaluation of match probability has also been found to be underestimated, in cases where the individuals are related or in cases where the populations have sub-structure. In other words, the chance that two individuals will have the same DNA profile is higher in cases where the individuals are related to one another (Waits *et al.* 2001).

If two samples are excluded as being from the same source, in other words, the two DNA profiles that were obtained did not match, no further statistical analysis is necessary. In the case where two DNA profiles do match, it is however necessary to calculate the probability that the two individuals have the same DNA profile by chance. The probability is influenced by various factors such as the number and variability of markers used to obtain the DNA profiles, how common the alleles are in the species (allele frequency) and how related the individuals are in the population where the samples were taken (Ogden *et al.* 2009).

A theoretical estimator that is increasingly used to access the statistical confidence for individual identification is the probability of identity (P_{ID}). P_{ID} is the probability that two individuals drawn at random from a population will have the same genotype at multiple loci (Waits *et al.* 2001). The probability of identity (P_{ID}) is estimated from allele frequencies in a population using established formulae. This probability is a standard statistic used in forensic science to evaluate how well a set of molecular markers discriminates between individuals (Kalinowski *et al.* 2007).

The average probability of identity (PI_{ave}) is used to determine the number of loci that is required to identify an individual with a reasonable level of confidence. Small values of PI_{ave} are desirable as identical DNA profiles for different individual are then highly unlikely. A balance between the confidence level of identification and the number of microsatellites used is important because with an increase in number of loci used, there is also an increase in laboratory costs as well as potential errors to consider (Ayres and Overall 2004).

GIMLET (Valière 2002) is a program that can be used to construct consensus genotypes and to estimate error rates for a set of genotypes. This is particularly useful for genotypes that have been obtained from non-invasive samples, where allelic drop-out and genotyping errors can cause analysis problems. Even though tissue and blood samples were used in this study, some of the samples were stored for a long period of time and the quality of the DNA obtained from these samples was potentially not optimal. For this reason as well for the fact that the program can be used to estimate the probability of identity, GIMLET was used to analyse the data (Valière 2002). The probability that two individuals in the populations share the same genotype (PI) is computed using the equations of biased PI, unbiased PI (with sample size correction) and PI for sibs. The equation used for populations where individuals randomly mate is given in Paetkau and Strobeck (1994) and the formula used to correct for small samples of individuals are given in Kendall and Stewart (1977). The equation used for populations that consists of sibs, are given in Evett and Weir (1998) and Taberlet and Luikart (1999).

4.3 Probability of identity results for rhino

The data files for the black and white rhino populations were analyzed separately with GIMLET software. In the black rhino population, the monomorphic marker DB44 was removed and two additional polymorphic markers, B1RH37D and B1RH2B were added to the marker set. In the case of the white rhino, the three monomorphic markers, BR4, BR17 and DB23 were removed and replaced with six polymorphic markers, DB42, DB49, DB52, DB66, 7B and 7C. Since the marker sets differs from the set used in Chapter 3, the number of alleles, the expected heterozygosity (He) and observed heterozygosity (Ho) had to be re-calculated by using the MICROSATELLITE ANALYSER (MSA) program. All the markers were scanned for possible genotyping errors and null alleles (See Chapter 2).

The number of alleles, expected heterozygosity (He) and observed heterozygosity (Ho) values per locus are listed in Table 4.1 and Table 4.2 for black and white rhino, respectively. Expected heterozygosity values in the black rhino population ranged between 0.24 and 0.69 and for the white rhino populations between 0.26 and 0.65. The total number of alleles per locus varied between 2 and 6 alleles. As previously mentioned in Chapter 2, the three loci DB42, DB66 and 7C had an excess of homozygotes (as calculated by Micro-checker) which was also confirmed with results from GIMLET. The loci DB66 and 7C also deviated significantly from the HWE, which could be an indication of the presence of null alleles at these two loci.

Table 4.1 The expected (He) and observed (Ho) heterozygosities, and the total number of alleles in black rhino (# alleles) tested with nine polymorphic loci

Locus name	He	Ho	# alleles
B1RH37D	0.421	0.370	2
BR4	0.698	0.741	4
32A	0.240	0.273	2
BR6	0.444	0.480	3
DB1	0.472	0.466	2
B1RH2B	0.502	0.527	2
DB23	0.507	0.476	2
SW35	0.607	0.548	4
BR17	0.471	0.452	6

Table 4.2 The expected (He) and observed (Ho) heterozygosities, and the total number of alleles in white rhino (# alleles) calculated from 11 polymorphic loci

Locus name	He	Ho	# alleles
DB42	0.601	0.694	4
32A	0.468	0.562	2
BR6	0.556	0.596	6
DB52	0.663	0.751	4
***DB66**	0.510	0.162	3
***7C**	0.578	0.229	5
DB1	0.174	0.196	4
DB44	0.305	0.280	5
DB49	0.672	0.721	4
7B	0.408	0.345	4
SW35	0.479	0.533	2

*loci not in HWE

The new sets of markers were tested for linkage disequilibrium (LD) using ARLEQUIN version 3.1. The LD values are listed in Table 4.3 for black rhino and Table 4.4 for white rhino. A pair-wise interclass correlation test was performed for all possible two-locus combinations.

Table 4.3 Linkage disequilibrium for nine microsatellite loci used for black rhino

	B1RH37D	BR4	32A	BR6	DB1	B1RH2B	DB23	SW35	BR17
B1RH37D									
BR4	-								
32A	-	-							
BR6	-	-	-						
DB1	-	-	-	-					
B1RH2B	-	-	-	-	-				
DB23	-	-	-	-	-	-			
SW35	-	-	-	-	-	-	-		
BR17	+	-	-	-	-	-	+	-	

+ Significant linkage disequilibrium (P < 0.001 after the Bonferroni correction was applied)

Table 4.4 Linkage disequilibrium for 11 microsatellite loci used for white rhino

	DB42	32A	BR6	DB52	DB66	7C	DB1	DB44	DB49	7B	SW35
DB42											
32A	-										
BR6	-	-									
DB52	-	-	-								
DB66	-	-	-	-							
7C	-	-	-	-	-						
DB1	-	-	-	-	-	-					
DB44	-	-	-	-	-	-	-				
DB49	-	-	-	+	-	-	-	-			
7B	-	-	-	-	-	-	-	-	-		
SW35	-	-	-	-	-	-	-	-	-	-	

+ Significant linkage disequilibrium ($P < 0.0009$ after the Bonferroni correction was applied)

For the black rhino population, six out of 36 pairwise loci combinations showed LD. After the Bonferroni correction was applied, only two out of 36 comparisons showed a significant deviation. Significant linkage was thus found only between marker BR17 and B1RH37D and between BR17 and DB23. Locus BR17 was therefore excluded from further statistical analysis.

For the white rhino populations, nine out of 55 loci combinations showed possible LD. After the Bonferroni correction was applied, only one pair of loci showed significant linkage. This linkage was found between markers DB49 and DB52. Since both loci had four alleles, DB49 was chosen to be excluded from further statistical analysis because DB52 amplified successfully even with poor quality DNA.

The Probability of identity (PI) was calculated for both black and white rhino populations by using only the polymorphic unlinked loci. GIMLET version 1.3.3 was used and the results are listed in Table 4.5 for the black rhino populations and in Table 4.6 for the white rhino populations.

Table 4.5 Probability of identity per locus for the black rhino population

Locus	PI_{biased}	$PI_{unbiased}$	PI_{sibs}
B1RH37D	0.428	0.421	0.649
BR4	0.153	0.146	0.444
32A	0.611	0.601	0.785
BR6	0.361	0.349	0.623
DB1	0.391	0.386	0.612
B1RH2B	0.377	0.374	0.597
DB23	0.375	0.372	0.593
SW35	0.238	0.231	0.510

Table 4.6 Probability of identity per locus for the white rhino populations

Locus	PI_{biased}	$PI_{unbiased}$	PI_{sibs}
DB42	0.189	0.185	0.472
32A	0.376	0.374	0.595
BR6	0.228	0.221	0.520
DB52	0.192	0.188	0.471
DB66	0.365	0.362	0.590
7C	0.255	0.250	0.524
DB1	0.563	0.555	0.759
DB44	0.439	0.431	0.679
7B	0.390	0.383	0.637
SW35	0.385	0.382	0.606

The PI_{biased} values does not include correction for sample size differences whereas the $PI_{unbiased}$ is calculated using a formula with sample size correction. The PI_{sibs} is based on an equation for co-dominant loci and provides an upper bound estimator. The PI_{biased} was used to form the lower bound and the PI_{sibs} the upper bound for identification estimation in the rhino populations studied.

Eichmann *et al.* (2005) investigated the number of microsatellite markers necessary to attain a PI_{sibs} value suitable for forensic applications in canine forensics. These authors concluded that at least ten microsatellite markers are needed to achieve a PI_{sibs} value of 0.0001.

For the accurate identification of individuals in natural animal populations, low PI_{sibs} values are desirable and these should be between 0.001 and 0.0001 as

suggested by Waits *et al.* (2001). The multiloci PI_{biased} and PI_{sibs} values were rearranged from the most informative locus to the least informative locus. The PI values for black rhino population are listed in Table 4.7 and for white rhino populations in Table 4.8. The value given for each locus is in combination with the previous loci in the table, for example, if only locus BR4 was used to identify a black rhino, about 1 out of 10 rhino will fit the profile. If you used loci BR4 and SW35 in combination to identify a rhino, the probability that a rhino will fit the profile is 3 out of 100 (0.153 x 0.238 = 0.036 or 3 out of 100).

The combination of all eight loci used in the black rhino population gave a PI_{sib} value of 0.0156. The value is above the optimum value for individual identification. When the linked locus BR17 was added a PI_{sibs} value of 0.009 was achieved. For further statistical analysis presented in chapter 5, all nine loci were thus included. In the white rhino populations, a PI_{sibs} value of 0.004 was achieved with the 10 loci. This value is also above the desired value of 0.001. When the linked locus DB49 was added, a PI_{sibs} value of 0.002 was achieved. All 11 loci were thus included for further statistical analysis.

Graphs comparing the PI_{sibs} and $PI_{unbiased}$ values against the number of loci (arranged from the most to the least informative) for both black and white rhino populations are presented in Figure 4.1 and Figure 4.2 respectively. As can be seen in both graphs, the lowest PI_{sibs} and $PI_{unbiased}$ values that were achieved are indicated by the green arrow, which is still above the desired PI_{sibs} and $PI_{unbiased}$ values.

Table 4.7 The range of multi locus PI values for the eight unlinked polymorphic microsatellite loci used for the black rhino population (rearranged from the most to the least informative locus)

	PIbiased	PIsibs
BR4	1.53E-01	4.44E-01
SW35	3.65E-02	2.27E-01
DB23	1.37E-02	1.35E-01
B1RH2B	5.16E-03	8.02E-02
DB1	2.02E-03	4.91E-02
BR6	7.28E-04	3.06E-02
B1RH37D	3.11E-04	1.99E-02
32A	1.90E-04	1.56E-02

Table 4.8 The range of multi locus PI values for the 10 unlinked polymorphic microsatellite loci used for the white rhino populations (rearranged from the most to the least informative locus)

	PIbiased	PIsibs
DB52	1.92E-01	4.71E-01
DB42	3.63E-02	2.22E-01
BR6	8.29E-03	1.16E-01
7C	2.11E-03	6.06E-02
DB66	7.71E-04	3.57E-02
32A	2.90E-04	2.13E-02
SW35	1.12E-04	1.29E-02
7B	4.35E-05	8.21E-03
DB44	1.91E-05	5.57E-03
DB1	1.08E-05	4.23E-03

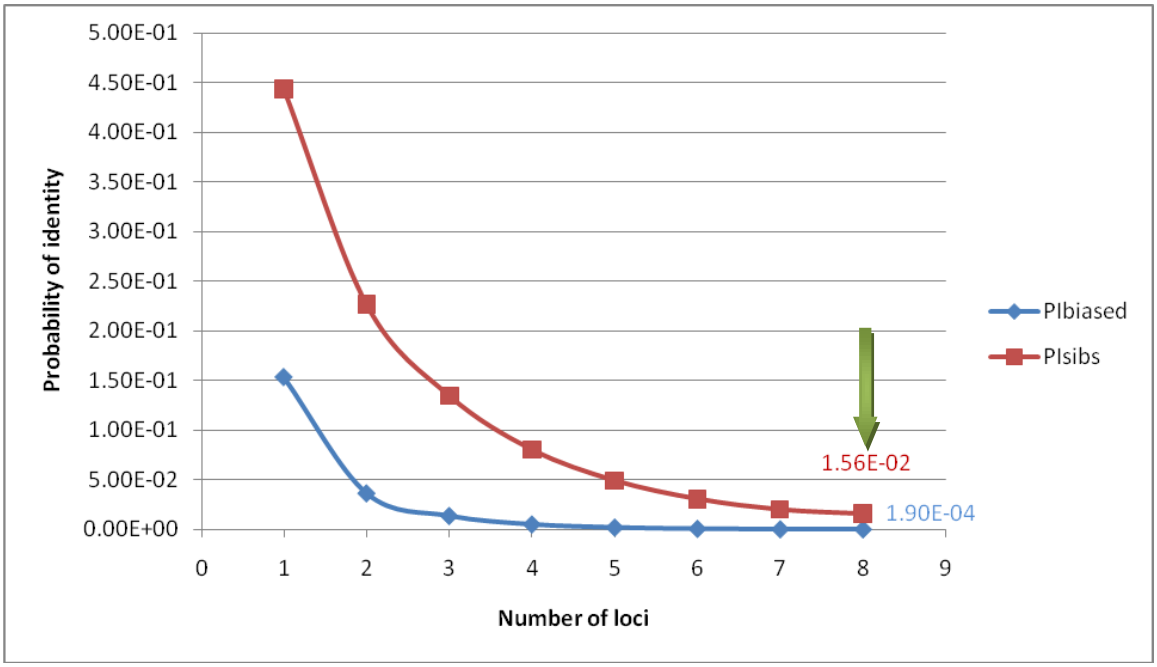


Figure 4.1 Multiloci PI_{biased} and PI_{sibs} in increasing order of single–locus values for black rhino population. Values are rearranged from the most informative locus to the least informative locus. The lowest PI_{sibs} and $PI_{unbiased}$ values that were achieved are indicated by the green arrow.

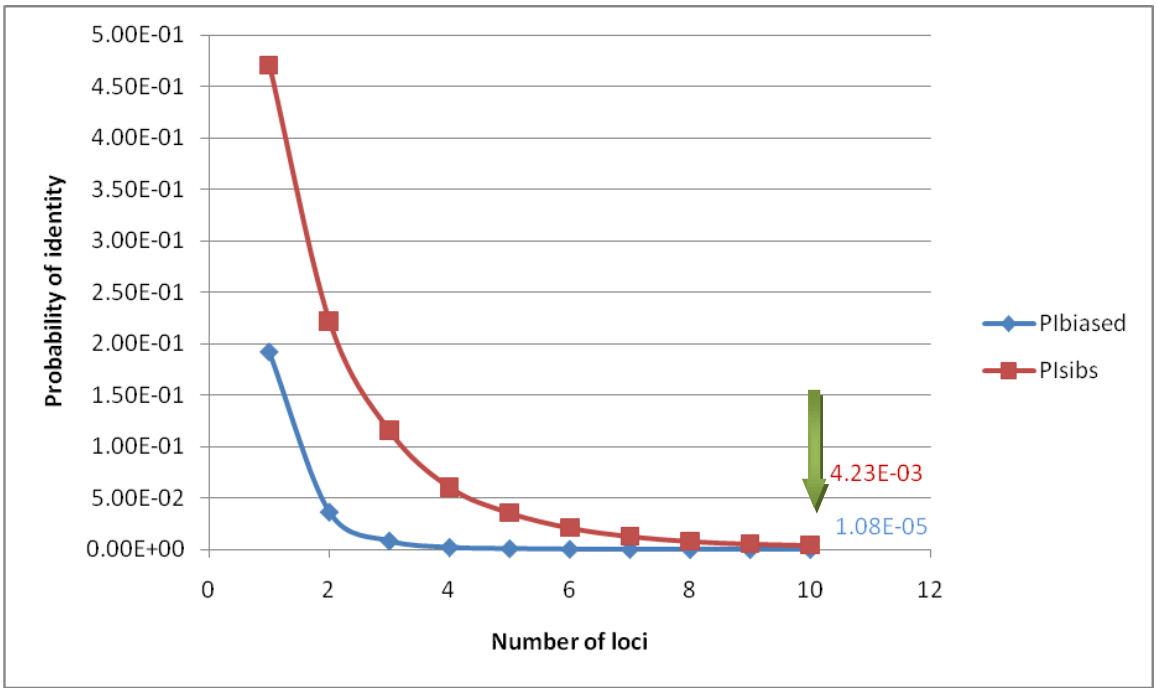


Figure 4.2 Multiloci PI_{biased} and PI_{sibs} in increasing order of single–locus values for the white rhino populations. Values are rearranged from the most informative locus to the least informative locus. The lowest PI_{sibs} ane $PI_{unbiased}$ values that were achieved are indicated by the green arrow.

4.4 Discussion

The panel of eight microsatellite markers tested in Chapter 3 was not enough for the individual identification and parentage verification of both black and white rhino populations (See Figure 4.1 and Figure 4.2). Extra polymorphic markers were thus added to both black and white rhino populations.

The probability of identity for unrelated black rhino was found to be 1 / 5 000 (1.90E-04) for unrelated individuals and 1 / 100 (1.56E-02) for related individuals. There is an estimate of 1 450 wild black rhino in South Africa, which means that the current set of loci may not be fully effective for discrimination. However, the black rhino population used in this study is isolated from other black rhino populations and the probability of identity could thus be underestimated, since these animals are related to each other.

The probability of identity for unrelated white rhino was 1 / 100 000 (1.08E-05) and for related individuals 1 / 250 (4.23E-03). There is an estimate of 17 480 wild

white rhino in South Africa. If the rhino are thus unrelated, the 11 loci available for white rhino would be sufficient for identification.

The optimal PI_{sibs} values were not achieved for both black and white rhino populations, but if more unrelated individuals are added to the dataset, the values should be close enough to ensure that the rhino will be identifiable for future analyses. To our knowledge this is the first study on rhino to include PI values and hence no comparisons could be made.

5.1 Introduction

The power of parentage studies was first recognised when existing paradigms in behavioural ecology were overturned during various studies on birds (Jones and Ardren 2003). Genetic parameters like heritability are best estimated with the aid of pedigrees, but this is rarely possible for wild (free-ranging) animals. Genetic markers can be used to study various aspects of organismal biology associated with parentage, relatedness and / or fitness. Parentage analyses have been used in studies concerning the identification of effective breeders in a population, dispersal and in hybridization studies (Selkoe and Toonen 2006).

With the availability of modern techniques and current software, breeding success can be assessed in detail. If all adults in a population are genotyped, estimates can be made regarding the reproductive success of each adult and the mean number of mates per male and female can be determined. The effective number of breeders can also be determined as field observations of breeding are not always accurate. Such disparities between social (observed) and genetic (actual) mating systems occur in most species. Data on the effective number of breeding adults versus the total number of adults has major implications for the maintenance of genetic diversity in a population (DeWoody 2005).

As an alternative to traditional frequency-based measures such as gene flow (from F_{ST}), dispersal can also be inferred directly from parentage analyses if sufficient data is available. An added advantage of parentage-based approaches is that real gene flow is considered as opposed to ecological or demographic dispersal. From a geneticist's point of view, dispersal is unimportant unless the disperser's gametes are transmitted to the next generation. By verifying parentage, the genetic contribution of the disperser to the next generation can be monitored (DeWoody 2005). There are various genetic approaches available to study and detect dispersal between different populations, for example assignment testing (see Chapter 3, STRUCTURE results page 58 - 60).

Parentage analysis can also assist in studies regarding hybridisation. Inter-specific hybridization threatens the integrity of many wildlife species and animals often hybridize because of human influences. Typical of this is the hybridisation between domestic cats and wildcats (Randi *et al.* 2001). Molecular analyses of

parentage and relatedness can quantify the incidence of such events (DeWoody 2005). For example, Robinson *et al.* (2005) performed a confirmation study on a F_1 hybrid between a black rhino bull and a white rhino cow. The confirmation was done by using a range of genetic markers (karyotype, FISH and microsatellites). The microsatellite markers used in their study included BR4, BR6, BR17, DB1, DB44, DB49, DB52 and DB66. The hybrid was heterozygous at all nine loci tested. The rhino displayed a heterozygosity value of one compared to the mean values of 0.52 for black rhino (n=117) and 0.32 for white rhino (n=7) according to the authors. The rhino calf also possessed alleles at these loci that were species-specific for both black and white rhino.

5.2 Parentage verification

The earliest and simplest technique used to verify parentage was based on the exclusion principle (see Chapter 1 page 24 - 25). Some of the weaknesses of this method include that it is not always possible to profile all individuals in a population and that it does not make any allowance for mutations, null alleles or genotyping errors which are unavoidable in the DNA profiling process (Jones and Ardren 2003).

The success of parentage verification largely depends upon the samples obtained. The less knowledge there is of the pedigree and the more incomplete the samples are, the more difficult the parentage verification process becomes in a population. The presence of related individuals in the pool of candidate parents can also influence the parentage assessment greatly. If the relatedness values are too high, the number of loci required for exclusion can increase. As more loci are used, the risk for genotyping errors however becomes even greater. When perfect parentage matches are not possible, statistical analyses are necessary to assess to what degree a researcher should feel confident about the reliability of parentage assignments (Jones and Ardren 2003).

Furthermore, the quality of parentage verification depends upon various factors which include the management of the laboratory, expertise of personnel and the quality assurance systems in place. The testing methods and the methods of validation are key points in paternity testing. In forensic cases that involve paternity testing, the Paternity Testing Committee (PTC) of the International Society for

Forensic Genetics (ISFG) recommends that the results should be presented in an objective way by means of commonly agreed statistical methods that are based on the likelihood ratio principle (Morling *et al.* 2002).

DNA assays of parentage and relatedness can empower wildlife biologists, conservationist authorities and private wildlife owners with data that will allow more informed management decisions. As a source of genotypic data, microsatellite markers are still the most informative molecular marker that can be used for the study of relatedness and parentage in natural populations (DeWoody 2005).

5.3 Genotyping errors caused by null alleles

Null alleles are one of the genotyping errors that have the most profound impact on parentage verification studies. Dakin and Avise (2004) defined a microsatellite null allele as any allele at a microsatellite locus that consistently fails to amplify to detectable levels via the polymerase chain reaction (PCR). Individuals that are heterozygous for a locus are thus observed as homozygous for that particular locus.

There are various factors that can cause null alleles. One potential cause is poor primer annealing due to a mutation event that took place within the primer binding area. A mutation event in the 3'-end of the priming site where extension begins, is particularly detrimental and will lead to null alleles arising. Null alleles can also appear because of the differential amplification of size-variant alleles. In other words, short length alleles often tend to amplify more efficiently than larger ones. The results can be that only the smaller of the two alleles are detected in a heterozygous individual. These alleles can be made visible by loading more of the amplified product during fragment analysis. A third source of null alleles involves PCR failure due to the poor quality of the DNA template. This problem is very difficult to analyze, because in some cases only a few loci fail to amplify where as the rest amplify with relative ease (Dakin and Avise 2004).

There is a variety of methods that can be used to detect null alleles. The most common approach is to screen for a significant deficit of observed heterozygotes relative to expected Hardy-Weinberg equilibrium. The rational for this is that loci with null alleles present would have been scored as homozygous loci (See Chapter 2

page 45 - 48). There are however population genetic factors that can equally cause an increase in homozygotes, for example inbreeding, bottlenecks, selection or even population substructure. Genetic parentage analysis can give more accurate information regarding the presence of null alleles (Dakin and Avise 2004). An example of the influence of null alleles on parentage studies is represented in Figure 5.1. If a rhino bull has for example three calves, and all three calves are excluded only at one particular locus and both the bull and the three calves are homozygotic for that locus, then it can be assumed that the locus has a null allele present.

	B1RH37D	**BR4**	**32A**	**BR6**	**DB1**	**B1RH2B**	**DB23**	**SW35**	**BR17**
Male	148 150	124 126	242 242	154 154	127 131	235 235	180 182	123 131	122 126
Female	148 148	124 126	242 242	134 134	127 131	229 229	180 180	129 131	132 132
Calf 1	148 148	124 124	242 242	134 134	131 131	229 235	180 182	123 131	126 132
	✓	✓	✓	✗	✓	✓	✓	✓	✓

	B1RH37D	**BR4**	**32A**	**BR6**	**DB1**	**B1RH2B**	**DB23**	**SW35**	**BR17**
Male	148 150	124 126	242 242	154 154	127 131	235 235	180 182	123 131	122 126
Female	148 148	124 126	242 242	134 134	127 131	229 229	180 180	129 131	132 132
Calf 2	148 150	126 126	242 242	134 134	127 131	229 235	180 182	129 131	122 132
	✓	✓	✓	✗	✓	✓	✓	✓	✓

	B1RH37D	**BR4**	**32A**	**BR6**	**DB1**	**B1RH2B**	**DB23**	**SW35**	**BR17**
Male	148 150	124 126	242 242	154 154	127 131	235 235	180 182	123 131	122 126
Female	148 148	124 126	242 242	134 134	127 131	229 229	180 180	129 131	132 132
Calf 3	148 148	124 126	242 242	134 134	127 127	229 235	180 180	123 131	122 132
	✓	✓	✓	✗	✓	✓	✓	✓	✓

Figure 5.1 The parentage of three rhino calves from the same male and female. The female was included as the possible mother at all loci tested, whereas the male was excluded at locus BR6 for all three calves. Since the father and all three calves are homozygotic for locus BR6, it can be assumed that a null allele is present and that all three calves inherited a null allele from the father.

The two methods discussed for detecting null alleles, (either by screening for an increase in homozygosity or by parentage testing), are however based on assumptions. The most accurate way to verify the presence of a null allele is to either

redesign the primer by removing it from the mutated area or to sequence the primer area (Dakin and Avise 2004). This method is however relatively expensive and time consuming and thus not used very often.

5.4 Parentage exclusions

Paternity exclusion needs to be carefully applied, particularly when there are multiple possible parents that may not be excluded, due to relatedness of individuals or inbreeding. Using strict paternity exclusion criteria (a single mismatch excludes a putative parent) may result in false exclusions because of mutation, genotyping errors, or null alleles. These problems may become an even greater concern when more loci are used because the probability of errors or mutations also increases. In a number of situations in natural populations, such as in many non-human vertebrates, paternity exclusion except for one putative father may not be possible (Hedrick 2005b).

Exclusion probabilities are calculated in several different ways, but the main aim is to describe the power of a locus to genetically exclude candidate individuals as parents. Dakin and Avise (2004) suggested that under realistic situations, microsatellite null alleles are uncommon and rare with a frequency of less than 0.2. However, if the frequency of the null allele is higher than 0.2, it could have an influence on the estimation of the average exclusion probability and the locus should rather be dropped from the parentage analysis. Sample size does not have a significant impact on the estimation of the mean exclusion probabilities. There is however the potential to falsely exclude a true parent when an offspring in question inherited a null allele.

An exclusion of one allele is allowed for parentage analysis for those markers that are not in Hardy-Weinberg equilibrium. As mentioned before, if a locus is not in HWE because of an excess of homozygous, it may be because of the presence of a null allele for that particular locus (Pemberton *et al.* 1995). Inbred populations tend to be more homozygous and subsequently cases do exist where complete exclusion is not possible. When multiple males are not excluded, a statistically based method is used to assign a most probable value which is derived from the genotypes (Jones and Ardren 2003).

5.5 Statistical parentage analysis of black and white rhino calves

In the current study, a categorical approach was used that is based on allocations that use likelihood-based approaches to select the most likely parent from a pool of non-excluded parents. A logarithm of the likelihood ratio (LOD score) was calculated by using CERVUS software (Marshall *et al.* 1998). The latest edition of this program (CERVUS 3.0) accommodates genotyping errors, which increases the success in paternity assignment (Kalinowki *et al.* 2007). The LOD is determined by determining the likelihood of an individual being the parent of a given offspring divided by the likelihood of these individuals being unrelated. The most likely parent is the one with the highest LOD score. The offspring can be unassigned if all parent-offspring relationships show zero likelihoods (Jones and Ardren 2003). The approaches underlying CERVUS 3.0 software do not find the candidate parents with the lowest number of mismatches, because null alleles and genotyping errors may lead to a false exclusion. Instead, the candidate parents are statistically analysed to find the most likely parents based on likelihood equations (Kalinowski *et al.* 2007).

The equations used are extremely sensitive to the estimate of the proportion of candidate adults sampled, which is rarely known. The requirement to assign both parents simultaneously makes the technique vulnerable to null alleles, linkage disequilibrium, mutations and genotyping errors. The presence of family members in the pool of candidate parents can also influence parentage assessment greatly. If the relatedness values are too high, more loci will have to be added and the number of loci required for exclusion could be prohibitive. A critical value needs to be determined that will produce a desired level of confidence in assignments of parentage, especially in data where errors exist (Jones and Ardren 2003).

Listed in Table 5.1 is the number of adult male and female rhino genotyped at the various nature reserves and the number of calves of which the parentage had to be verified.

Table 5.1 Number (#) of adult female, adult male rhino and number of calves genotyped at four nature reserves in South Africa

Nature reserve	# Adult females	# Adult males	# Sub-adults and calves	Total # rhino
Songimvelo	14	8	31	53
Loskopdam	11	14	8	33
Mthethomusha	15	8	4	27
Sam Knott	32	23	17	72

Table 5.2, 5.3 and 5.4 list the white rhino calves, and candidate mother and fathers with LOD scores and the confidence level as determined by CERVUS 3.0. Table 5.5 contains the list of the black rhino calves with their candidate parents. The relaxed confidence level (+) was set at 80% and the strict confidence level (*) was set at 95%. For the white rhino calves, the critical LOD score for strict confidence level was 8.33 and 12.38 for the black rhino calves. The relaxed confidence level for the white rhino calves were 1.31 and 6.82 for the black rhino calves.

5.5.1 Songimvelo Nature Reserve

The parentage of 25 of the 31 sub-adults and calves of the Songimvelo Nature Reserve could be resolved by using DNA profiling (see Table 5.2). The parentage of 23 calves was verified without any exclusions occurring, with four at a strict confidence level of 95%. The possible fathers of two calves (WR98/02 and WR100/02) were verified in the absence of the DNA profiles of the mothers. These two cows have not been genotyped to date. The trio confidence level could thus not be determined but the confidence of the pair combination between father and calves were determined. For calf WR98/02 a LOD score of 3.82E+00 was obtained, indicating that WR69/02 is probably the father of the calf with a relaxed confidence level of 80%.

The remaining six calves contained possible exclusions of parentage. Four of the calves had the same mother (WR83/02) but based on the presence and absence of specific alleles in the DNA profiles, this individual was excluded as the possible mother. Re-extraction and testing did not alter the profile of the mother. Human error could thus have occurred during sampling and it is suggested that this cow be re-

sampled and genotyped before the parentage of these calves can be determined. The DNA profile of the juvenile WR105/02 is similar to the DNA profile of WR83/02 which is also an indication that a sampling error had occurred.

Table 5.2 Verified parentage of 31 sub-adults and calves from Songimvelo Nature Reserve of which the mothers were known and the fathers were unknown. Analysis was done by using CERVUS 3.0.

Off/spring ID	Mother ID	Possible Father ID	Trio loci compared	Trio loci mismatched	Trio LOD score	Trio Confidence
WR88/02	WR77/02	WR69/02	11	0	9.88E+00	*
WR89/02	WR81/02	WR70/02	11	0	4.97E+00	+
WR90/02	WR82/02	WR70/02	11	0	8.65E+00	*
WR91/02	WR80/02	WR69/02	11	0	5.92E+00	+
WR92/02	WR81/02	WR70/02	11	0	4.13E+00	+
WR95/02	WR80/02	WR69/02	11	0	5.82E+00	+
WR97/02	WR86/02	WR70/02	11	0	5.05E+00	+
WR98/02	Not available	WR69/02	11	0	0.00E+00	
WR99/02	WR76/02	WR70/02	11	0	1.06E+01	*
WR100/02	Not available	WR70/02	11	0	0.00E+00	
WR101/02	WR87/02	WR70/02	11	0	3.81E+00	+
WR102/02	WR78/02	WR71/02	11	0	5.46E+00	+
WR103/02	WR86/02	WR70/02	11	0	5.24E+00	+
WR106/02	WR77/02	WR69/02	11	0	5.83E+00	+
WR107/02	WR82/02	WR70/02	11	0	4.88E+00	+
GA315/03	WR93/02	WR69/02	11	0	1.14E+01	*
GA316/03	WR98/02	WR69/02	11	0	6.32E+00	+
GA317/03	WR77/02	WR69/02	11	0	7.09E+00	+
GA318/03	WR94/02	WR70/02	11	1	4.35E+00	+
GA319/03	WR79/02	WR69/02	11	0	4.87E+00	+
GA320/03	WR80/02	WR69/02	11	0	5.96E+00	+
GA321/03	WR86/02	WR70/02	11	0	3.11E+00	+
GA322/03	WR78/02	WR70/02	11	0	6.70E+00	+
GA323/03	WR84/02	WR70/02	11	0	8.11E+00	+
GA326/03	GA328/03	WR75/02	11	1	3.35E+00	+

+ relaxed confidence level of 80%

*strict confidence level of 95%

In the case of the sub adult rhino WR96/02 the parentage could not be verified since two exclusions occurred. It could not be concluded whether the possible mother or possible father should be excluded as the parent (see Figure 5.2). Data obtained from monitoring these rhino should be taken into consideration in this cases. In future research, if SNP markers are developed for white rhino, it could be used in association with the STR markers to solve the parentage of these types of cases.

	DB42	**32A**	**BR6**	**DB52**	**DB66**	**7C**	**DB44**	**DB49**	**7B**	**SW35**	**DB1**
WR69/02	326 326	234 248	133 135	217 219	201 201	255 255	181 181	159 161	265 265	127 133	129 129
WR87/02	326 330	234 234	133 153	217 219	201 203	255 255	181 181	159 161	265 265	127 133	129 129
WR96/02	326 326	234 248	133 133	217 219	201 201	253 255	181 181	159 161	265 265	133 133	129 131
	✓	✓	✓	✓	✓		✓	✓	✓	✓	

Figure 5.2 DNA profiles of calf WR96/02 with candidate father WR69/02 and candidate mother WR87/02.

As were seen in Table 5.2, a single exclusion occurred in the parentage verification of both of calves GA318/03 and GA326/03. One exclusion is however acceptable for parentage analysis. These exclusions can probably not be contributed to null alleles but rather dropped alleles as indicated in Figure 5.3. Calf GA318/03 is a heterozygote for locus SW35 which can be an indication that either one of the parents could have a dropped allele. Calf GA326/03 could have a dropped allele at locus 7C, since the father is a heterozygote for that locus. Re-amplification at a lower annealing temperature did however not yield a different result. Field observations regarding die movement and behaviour of the rhino, for example, is the calf still drinking milk from the cow or is this the only bull in close vicinity of the cow, should be taken into consideration for the final decision whether the exclusion is a true exclusion or caused by either a dropped allele or a mutation. Furthermore, the DNA profiles will be used in future parentage analysis. If all the DNA profiles are re-confirmed at least three times in other parentage analysis, the exclusion should be considered as a true exclusion. If the parentage of these two calves are found to be incorrect, a correction should be made in the studbook keeping.

	DB42	32A	BR6	DB52	DB66	7C	DB44	DB49	7B	SW35	DB1
WR70/02	330 330	234 234	133 153	217 221	201 203	253 255	173 181	159 163	263 265	133 133	129 129
WR94/02	326 330	248 248	135 153	217 217	201 201	255 255	181 181	159 159	265 267	133 133	129 129
GA318/03	330 330	234 248	133 135	217 217	201 201	253 255	173 181	159 159	263 265	127 133	129 129
	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓

	DB42	32A	BR6	DB52	DB66	7C	DB44	DB49	7B	SW35	DB1
WR75/02	330 330	234 248	133 135	219 221	203 203	247 255	181 181	159 161	265 267	127 133	129 129
GA328/03	330 330	234 234	133 153	219 221	203 203	253 255	181 181	161 163	265 267	127 127	129 129
GA326/03	330 330	234 234	133 153	219 221	203 203	253 253	181 181	161 163	265 267	127 127	129 129
	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓

Figure 5.3 DNA profiles of calf GA318/03 and GA326/03 with candidate fathers (WR70/02 and WR75/02) and candidate mothers (WR94/02 and GA326/03). The LOD score for the trio combinations were at a relaxed confidence level of 4.35E+00 for calf GA318/03 and 3.35E+00 for calf GA326/03.

5.5.2 Loskopdam Nature Reserve

The parentage of five of the eight sub-adults rhino and calves from the Loskopdam Nature Reserve are listed in Table 5.3.

Table 5.3 Verified parentage of five sub-adults and calves from Loskopdam Nature Reserve of which the mothers were known and the fathers were unknown. Analysis was done by using CERVUS 3.0

Offspring ID	Mother ID	Possible Father ID	Trio loci compared	Trio loci mismatched	Trio LOD score	Trio Confidence
WR34/02	WR43/02	WR39/02	11	0	6.92E+00	+
WR44/02	WR38/02	WR42/02	11	0	4.79E+00	+
WR47/02	WR53/02	WR35/02	11	1	6.28E-01	-
WR56/02	WR48/02	WR59/02	11	0	8.89E+00	*
WR63/02	WR51/02	WR42/02	11	1	3.06E+00	+

+ relaxed confidence level of 80%

*strict confidence level of 95%

- most likely candidate parents is not assigned

Parentage was assigned to five (WR34/02; WR44/02; WR47/02; WR56/02 and WR63/02)) of the eight calves studied with a relaxed confidence level (1.31) for calves WR34/02, WR44/02 and WR63/02 and a strict confidence level (8.33) for calf WR56/02.

The mother of calf WR63/02 was excluded only at locus 7C (see Figure 5.4). This exclusion can however be attributed to a possible null allele since both the calf and mother were homozygous for locus 7C. In Chapter 2, when the raw molecular data set were screened for possible null alleles, marker 7C tested positive for the possibility of null alleles being present.

	DB42	**32A**	**BR6**	**DB52**	**DB66**	**7C**	**DB44**	**DB49**	**7B**	**SW35**	**DB1**
WR42/02	324 330	234 248	133 153	219 221	203 203	255 255	181 181	161 163	265 267	133 133	129 129
WR51/02	324 330	234 248	133 135	219 221	203 203	253 **?**	181 181	161 163	265 265	133 133	129 131
WR63/02	324 330	234 248	135 153	219 219	203 203	255 **?**	181 181	161 161	265 265	133 133	129 129
	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓

Figure 5.4 DNA profile of calf WR63/02 with candidate father WR42/02 and candidate mother WR51/02. Locus 7C contains a possible null allele. The LOD score for the trio combination was at a relaxed confidence level of 3.06E+00.

In the case of calf WR47/02 (Figure 5.5) the candidate father WR35/02 and mother WR53/02 combination gave a single exclusion which cannot be attributed to a possible null allele, since the calf is heterozygous for locus 7C. The LOD score also indicates that these are probably not the most likely candidates for parents for this calf. Separately tested, both parents could be assigned as the possible parent of the calf, but the LOD score of the mother and calf pair alone, gave a strict confidence level where as the candidate father only gave a relaxed LOD score of 2.30. This exclusion can be attributed to a possible genotyping error or the occurrence of a dropped allele at locus 7C for the mother WR53/02 or this may not be the father of the calf. The true father may rather be closely related to the candidate father WR35/02.

	DB42	**32A**	**BR6**	**DB52**	**DB66**	**7C**	**DB44**	**DB49**	**7B**	**SW35**	**DB1**
WR35/02	324 330	234 248	135 153	219 221	203 203	**253** 255	181 181	161 163	265 267	133 133	129 131
WR53/02	324 330	234 248	133 153	219 221	201 203	253 **253**	181 181	161 163	265 265	133 133	129 129
WR47/02	324 324	234 248	133 135	219 221	201 203	251 **253**	181 181	161 161	265 265	133 133	129 129
	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓

Figure 5.5 DNA profile of calf WR47/02 with candidate father WR35/02 and candidate mother WR53/02. Locus 7C excluded the candidate parents. According to the LOD score, the mother is the most likely parent of the calf.

In the other three unresolved parentage cases, the putative mothers were excluded as the candidate mothers and the exclusions could not be attributed to null alleles since the calves or mothers were heterozygous for these loci (see Figure 5.6). The candidate mother, WR61/02, was excluded as the mother for both her calves WR50/02 and WR62/02. Even though the samples were re-extracted and re-tested in duplicate, this result was consistent. It is thus suggested that the mother should be re-sampled to ensure that no human error occurred. In the case of calf WR49/02, the mother was excluded at two alleles and the father at none. The scenario that the two female samples (WR61/02 and WR54/02) could have been swopped when taken were tested by switching the DNA profiles, but the results for parentage was still negative.

(i)

	DB42	32A	BR6	DB52	DB66	7C	DB44	DB49	7B	SW35	DB1
WR59/02	324 330	248 248	133 153	219 221	203 203	253 253	181 181	161 163	265 265	127 133	129 129
WR54/02	324 326	234 248	133 133	217 219	201 201	251 253	179 181	159 159	265 267	127 133	129 129
WR49/02	324 324	248 248	133 153	219 221	203 203	251 253	181 181	161 163	265 267	127 133	129 129
	✓	✓	✓	✓		✓	✓		✓	✓	✓

(ii)

	DB42	32A	BR6	DB52	DB66	7C	DB44	DB49	7B	SW35	DB1
WR59/02	324 330	248 248	133 153	219 221	203 203	253 253	181 181	161 163	265 265	127 133	129 129
WR61/02	324 330	234 248	135 153	219 219	203 203	253 253	181 181	161 161	263 265	133 133	129 129
WR50/02	324 330	248 248	133 133	219 221	203 203	253 253	181 181	159 161	265 265	127 127	129 129
	✓	✓		✓	✓	✓	✓		✓		✓

(iii)

	DB42	32A	BR6	DB52	DB66	7C	DB44	DB49	7B	SW35	DB1
WR59/02	324 330	248 248	133 153	219 221	203 203	253 253	181 181	161 163	265 265	127 133	129 129
WR61/02	324 330	234 248	135 153	219 219	203 203	253 253	181 181	161 161	263 265	133 133	129 129
WR62/02	324 324	234 248	133 133	221 221	203 203	253 253	181 181	163 163	265 265	127 127	129 129
	✓	✓			✓	✓	✓		✓		✓

Figure 5.6 DNA profile of (i) calf WR49/02 with candidate father WR59/02 and candidate mother WR54/02 and (ii) calf WR50/02 with candidate father WR59/02 and candidate mother WR61/02, as well as (iii) calf WR62/02 with candidate father WR59/02 and candidate mother WR61/02.

5.5.3 Mthethomusha Nature Reserve

In Table 5.4 the candidate parents of four white rhino calves of the Mthethomusha Nature Reserve are listed. Parentage could be assigned to all four calves with calves WR23/02 and WR26/02 at strict confidence levels.

Table 5.4 Verified parentage of four calves from Mthethomusha Nature Reserve of which the mothers were known and the fathers unknown. Analysis was done by using CERVUS 3.0.

Offspring ID	Mother ID	Possible Father ID	Trio loci compared	Trio loci mismatched	Trio LOD score	Trio Confidence
WR20/02	WR21/02	GA395/03	11	0	2.490E+00	+
WR23/02	WR22/02	GA395/03	11	0	9.54E+00	*
WR25/02	GA108/04	GA390/03	11	1	1.85E+00	+
WR26/02	WR24/02	GA390/03	11	0	8.38E+00	*

Calf WR25/02 has a single exclusion at locus SW35 with the candidate father GA395/03 (Figure 5.7), which can be attributed to a possible null allele. The LOD score also indicated a relaxed confidence level (1.31) when all three profiles were taken into account.

	DB42	32A	BR6	DB52	DB66	7C	DB44	DB49	7B	SW35	DB1
GA395/03	324 330	234 248	133 141	217 219	203 203	245 253	175 181	159 161	265 267	133 **?**	129 129
GA108/04	326 326	234 248	133 133	217 217	201 203	253 255	181 181	159 159	265 267	127 133	129 129
WR25/02	324 326	248 248	133 141	217 217	203 203	253 253	181 181	159 159	265 265	127 **?**	129 129
	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓

Figure 5.7 DNA profile of calf WR25/02 with candidate father GA395/03 and candidate mother GA108/04. Locus SW35 contains a possible null allele.

5.5.4 Sam Knott Nature Reserve

In Table 5.5 the candidate parents of the seventeen black rhino calves of the Sam Knott nature reserve are listed. For the black rhino calves, the critical LOD score for strict confidence was 12.38 and 6.82 for relaxed confidence. Parentage assignments were problematic since the whole population have not yet been genotyped.

The parentages of five calves (BLR19/08; BLR24/08; BLR29/08, BLR38/08 and BLR39/08) were assigned, but with a low confidence level. BLR21/08 is a male born in 2005, who was a positive match for the parentage of two of the calves (BLR24/08 and BLR38/08) which were born in 2003 and 2004 respectively. It is thus impossible for that bull to be the father of those two calves. The trio LOD score also indicated that the parents assigned were not the most likely parents of these two calves. The father of the calves are most likely closely related to this male, either an older brother or father. The parentages of these calves can thus only be concluded when the whole population has been genotyped.

The mothers of the calves were verified according to the studbook information. There were however four calves of which the expected mothers were excluded as the possible mothers. These calves included BLR14/08; BLR17/08; BLR33/08 and BLR43/08. When the DNA profiles of calves BR17/08 and BLR14/08 were swopped, a positive match between mother and calves were found. A human error had thus occurred either in the laboratory or during sampling. DNA should firstly be re-extracted and profiled to verify the results and to ensure that the mistake did not occur in the laboratory. If the same results are obtained, these animals will have to be re-sampled and re-tested. Calf BLR33/08 was excluded at two loci from the mother. These exclusions cannot be attributed to null alleles since the mother is a heterozygous for one of the mismatch loci. In the case of calf BLR43/08, the possibility of a null allele cannot be excluded and since only one locus does not match, the cow can still be the possible mother of the calf. Four mothers were assigned to their calves with a strict confidence level of 95%, and eight mothers were assigned to their calves with a relaxed confidence level of 80%.

Table 5.5 Parentage analysis results for 17 black rhino calves from the Sam Knott Nature Reserve with the mothers known and the fathers unknown. Analysis was done by using CERVUS 3.0.

Offspring ID	Mother ID	Possible Father ID	Trio loci compared	Trio loci mismatched	Trio LOD score	Trio Confidence
BLR1/08	BLR21/07	BLR21/08	9	1	2.88E+00	-
BLR2/08	BLR16/08	BLR41/08	9	1	-6.91E-01	-
BLR10/08	BLR17/08	BLR25/08	9	2	-4.63E+00	-
BLR11/08	BLR2/08	BLR10/07	9	1	-4.99E-01	-
BLR13/08	BLR2/08	BLR22/07	9	3	-9.08E+00	-
BLR14/08	BLR16/08	BLR 10/07	9	3	-1.02E+01	-
BLR17/08	BLR32/08	BLR4/07	9	2	-6.97E+00	-
BLR19/08	BLR16/08	BLR10/07	9	0	2.38E-01	-
BLR23/08	BLR4/08	BLR9/07	9	2	-2.90E+00	-
BLR24/08	BLR21/07	BLR21/08	9	0	6.97E+00	-
BLR29/08	BLR7/08	BLR25/08	9	0	2.64E+00	-
BLR33/08	BLR32/07	BLR20/08	9	2	-6.73E+00	-
BLR37/08	BLR20/07	BLR10/07	9	1	1.46E+00	-
BLR38/08	BLR21/07	BLR21/08	9	0	4.43E+00	-
BLR39/08	BLR3/07	BLR18/08	9	0	5.29E+00	-
BLR40/08	BLR24/08	BLR8/08	9	0	0.00E+00	
BLR43/08	BLR8/07	BLR31/08	9	2	-4.49E+00	-

+ relaxed confidence level of 80%

*strict confidence level of 95%

- most likely candidate parents is not assigned

5.6 Summary of parentage results for the different populations

A schematic diagram was drawn of the verified parentages of each reserve. Figure 5.8 lists all the adult males of the Loskopdam Nature Reserve and the calves. The lines indicate the parentage as determined by using CERVUS 3.0. Similar results were obtained for the Mthethomusha Nature Reserve (Figure 5.9) and for the Songimvelo Nature Reserve (Figure 5.10). Calves GA318/03 and GA326/03 were not included, since one exclusion was observed in both cases (See Table 5.2). Figure 5.11 is a summary of the results for the Songimvelo Nature Reserve. From these diagrams, one can conclude which individuals are the most effective breeders in the population.

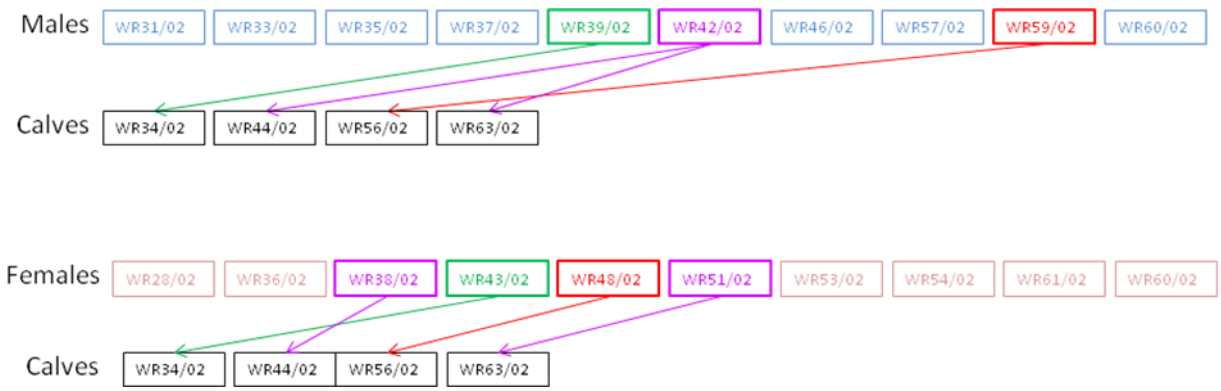


Figure 5.8 A schematic pedigree indicating the possible parents of each calf in the Loskopdam Nature Reserve white rhino population (parents of a calf are indicated in the same colour)

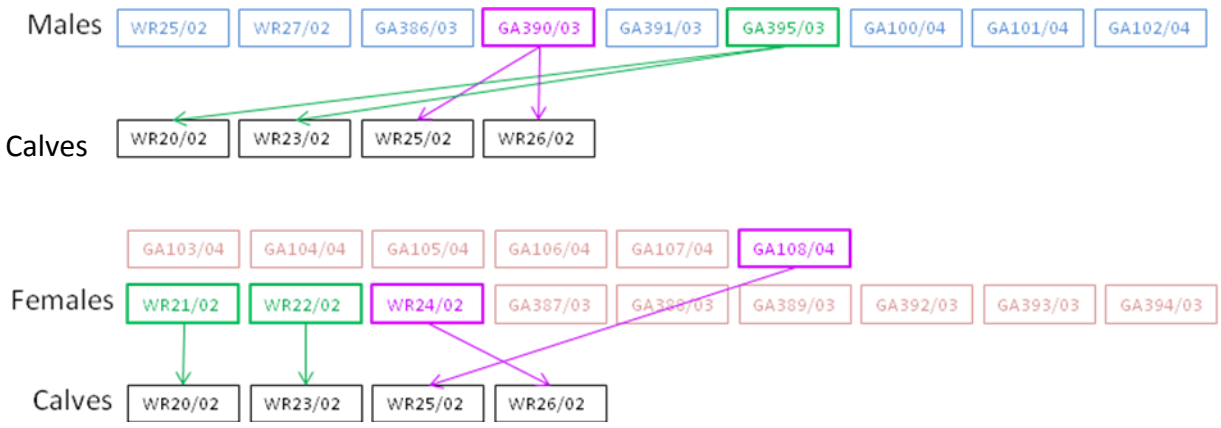


Figure 5.9 A schematic pedigree indicating the possible parents of each calf in the Mthethomusha Nature Reserve white rhino population (parents of a calf are indicated in the same colour)

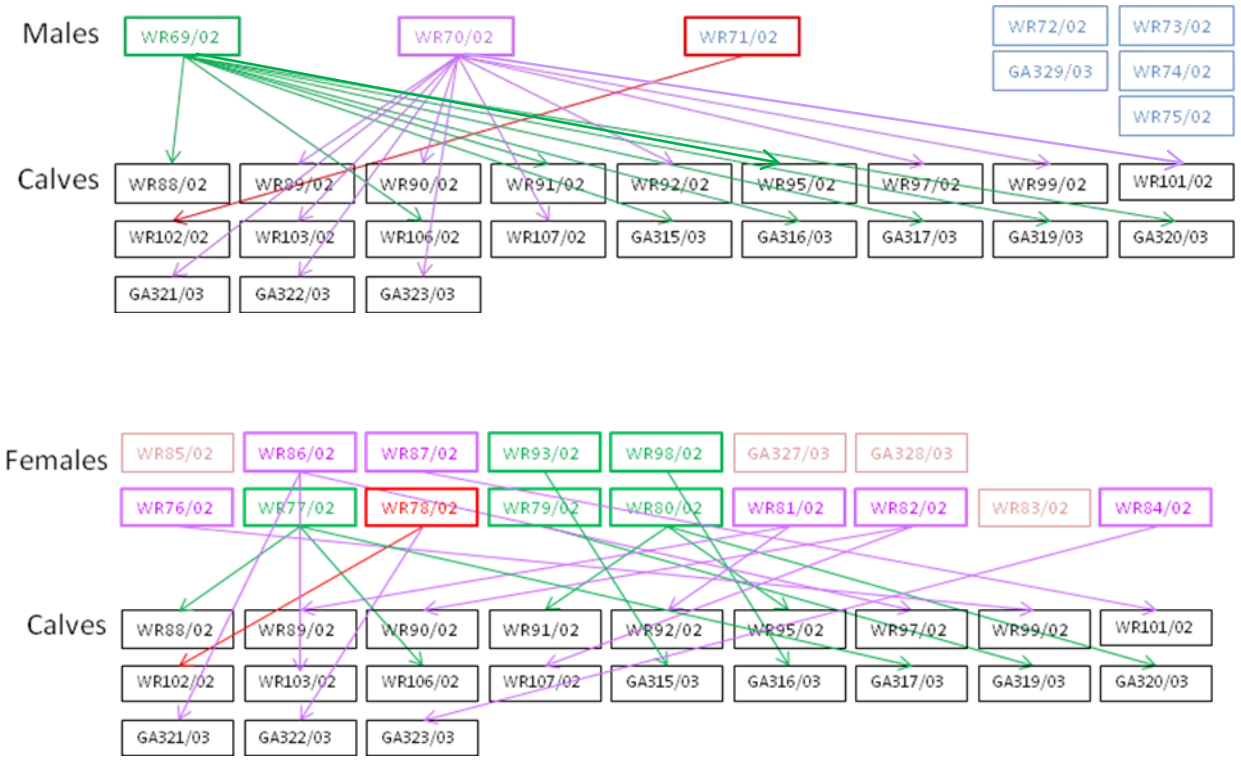


Figure 5.10 A schematic pedigree indicating the possible parents of each calf in the Songimvelo Nature Reserve white rhino population (parents of a calf are indicated in the same colour)

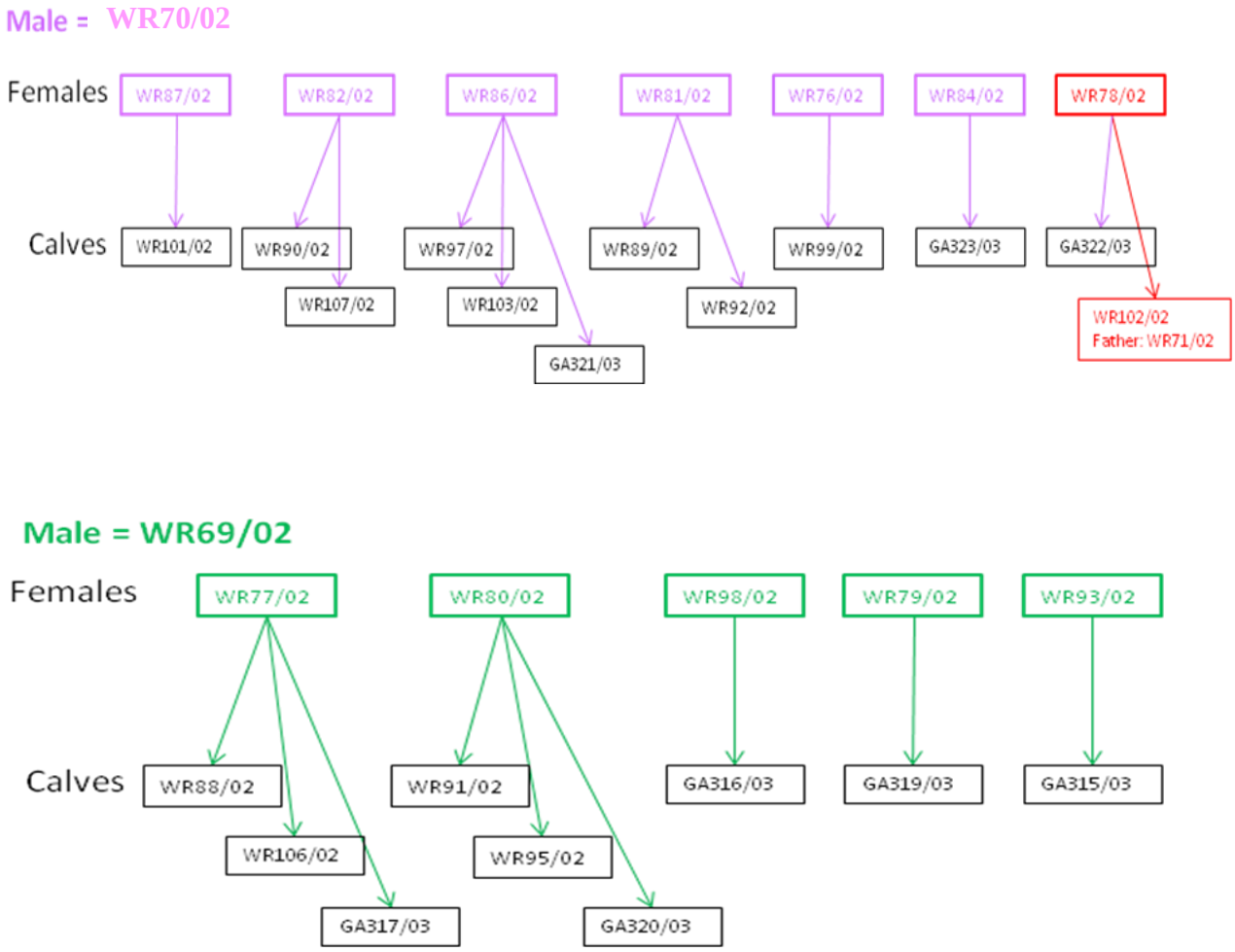


Figure 5.11 A schematic diagram of mothers and their calves in the Songimvelo Nature Reserve white rhino population

5.7 Discussion

The genetic and social structures of a population are inextricably linked. In many wildlife species, the social structure is known (at least superficially), but little is known regarding the relatedness of individuals within the social groups.

The power to estimate relatedness using DNA markers is positively correlated with the genetic variability of the markers used and with the number of markers employed. Microsatellite markers are one of the best available markers to use for the estimation of relatedness and verification of parentage since these markers are highly polymorphic and abundant in the genome (DeWoody 2005).

When the pedigree of all the individuals in a population are known and verified via DNA profiling, inbreeding can be avoided with proper management

decisions being made to avoid further inbreeding. In other words decisions based on translocations are now based on facts and not only on field observations.

In total, the parentage of 65% of the white rhino calves could be verified by using DNA profiling. In the Loskopdam Nature Reserve the calves were the offspring of three of the 10 potential bulls. There are thus only three bulls contributing to the genetic diversity of the next generation of white rhino. The parentage of four of the eight calves could be resolved by using the DNA profiles. To effectively solve the overall parentage pattern, the whole population needs to be sampled and profiled.

In the Mthethomusha Nature Reserve the parentage of all four the calves were solved. Two out of the nine potential bulls were the fathers of these calves. The bulls GA380/03 and GA395/03 both had thus two offspring from two different females (See Figure 5.9).

In the Songimvelo Nature Reserve, the parentage of 21 calves was verified for both parents. There were eight possible bulls available for parentage, but all the calves were the offspring of only two of these bulls. Calf WR102/02 was the only offspring of bull WR71/02. The rest of the calves were the offspring of either bull WR69/02 or WR70/02.

All three calves of cow WR77/02, as well as the three calves of cow WR80/02, were the offspring of bull WR69/02. The two calves of the cow WR82/02, the three calves of WR86/02 and the two calves of WR81/02 were the offspring of bull WR70/09. Only cow WR78/02 had two calves with different fathers, the rest of the females only had calves from the same bull.

The information gained from the parentage verification, can also provide information regarding kinship between calves. For example the results showed that calves WR88/02, WR106/02 and GA317/03 are full sibs whereas WR88/02 and WR91/02 are half sibs (sharing only the same father).

The information can be used to calculate the effective population size for each of the nature reserves studied (see figure 5.8; 5.9 and 5.10). Knowledge of the effective population size is important since the census size is invariably much higher than the effective population size. This can lead to an overestimation of the available genetic diversity for a population. For example in the case of the Songimvelo Nature Reserve, management may be under the impression that the reserve have eight rhino bulls contributing to the genetic material of the next generation whereas in reality

only two bulls are the genetic contributors of the next generation. The genetic diversity of the next generation will thus be lower than expected since the offspring will be related and the genetic material from the bulls not contributing to the next generation will be lost. Management should thus take pedigree information into consideration when decisions regarding the translocations of rhino are made. To ensure that inbreeding is kept at a minimum at these reserves, the dominant bulls (effective bulls) should thus be removed after two to three generations from the reserve to give other bulls the opportunity to contribute genetic material to the next generation. This will ensure that the current genetic diversity is maintained.

The parentage of the black rhino calves at the Sam Knott nature reserve could not be resolved successfully due to the fact that all the rhino were not genotyped. The possible fathers of the calves are probably not genotyped yet. The possible mother of calf BLR33/08 as listed in the studbook was however excluded as the mother of the calf. The other mother/calves exclusions could probably be explained either due to human error or null alleles. Depending on the accuracy of the field observation, this cow should be re-sampled and tested if possible otherwise the results should be accepted as a true exclusion.

As mentioned before, the male BLR21/08 fitted the profile of the father for two of the other calves, which was impossible, since he was younger than the calves. This is probably an indication that a closely related male to the calf (a father or full-sib) is the true parent of these calves. This can only be answered when all the rhino on the reserve have been genotyped. This scenario also highlights the importance of studbook keeping and field observations regarding the movement and behaviour of rhino of which the value should never be underestimated when verifying parentage.

Parentage analysis should thus always take studbook information into consideration when parentage is assigned. Even though the set of markers did prove sufficient for individual identification, it also showed limitations in solving parentage of the calves. The difficulty could be attributed to the fact that the whole population had not been genotyped yet and that due to inbreeding, the genetic diversity of the particular populations were very low. The suggestion is therefore that SNPs are developed for black rhino that could be used together with STR markers to solve the parentage of these calves.

6.1 Rhino poaching

The rhino populations of South Africa are facing a real danger of extinction due to an increase in the number of poaching cases. Syndicates equipped with helicopters, tranquilization-drugs and guns can kill several rhino in a single evening. Methods of poaching extend to poisoning, based on poisoned cabbage halves that are placed in the vicinity of rhino middens, since rhino have the habit of visiting their dung piles and eating whatever food has been planted there (SAPA 2010j). Owners of rhino are becoming very frustrated since they not only suffer significant financial losses, but the conservation of these two species is put at risk as well. Questions on the sustainability of rhino farming arise. This could put the long-term survival of the rhino populations in jeopardy. The cost involved in protecting these animals is high, placing an additional burden on game-owners. Some private owners hire private security firms, some de-horn their own rhino while some owners have gone as far as considering the arsenic poisoning of their rhino' horns. Managers of nature reserves have attempted to translocate their rhino to a single reserve to facilitate easier protection of populations and to save costs regarding security. This however had led to fighting among rhino bulls which led to even further losses in the case of white rhino.

On 21 September 2010, a syndicate of 11 suspects were arrested on suspicion of rhino poaching. These suspects included the owner of a game farm in Polokwane and his wife, two wildlife veterinarians and the wife of one of these veterinarians, a pilot as well as five professional hunters. This group faces charges ranging from assault, defeating the ends of justice, fraud, corruption, malicious injury to property, illegal possession of firearms and ammunition, contravention of the National Environmental Management Biodiversity Act, No 10 of 2010, contravention in terms of the Medicines and Related Substances Act, No 101 of 1965 as well as contravention of the Prevention of Organised Crime Act, No 121 of 1998. The case has been postponed until 11 April 2011 for further investigations (SAPA 2010k; SAPA 2010l; SAPA 2010m). It is hoped that successful prosecutions will serve as deterrent to all other syndicates involved with rhino poaching.

To help address the problem of rhino poaching, it has become crucial that apprehended perpetrators will be severely punished to serve as a deterrent. Gareth

Morgan, South African shadow environment minister is of the opinion that there needs to be a strong signal that killing this country's rhino will be met with severe consequences (SAPA 2010m). To prove guilt in a rhino poaching case can however be difficult. It is necessary to prove the identity of the poached animal, which can be problematic with wildlife species. As mentioned in Chapter 1, rhino are traditionally either identified via ear-notches or microchips. These are however easily changed or removed by those involved when a rhino is poached. DNA profiling is thus an excellent alternative since it cannot be tampered with or changed. A DNA profile can also be obtained from most biological material such as blood, tissue and hair. Biological evidence found at the poaching crime scene can thus be compared to biological evidence found on the suspect, for example blood stains found on the suspects clothes or vehicle. If the DNA profiles match, the identity of the rhino is established and real evidence for prosecution of the suspect(s) is at hand.

6.2 Forensic application of STR markers

Microsatellite markers have been used since the 1990s for DNA casework. This technique has been proven and is acceptable in a court of law. For this reason, and the fact that STR markers can be used in mixture samples as well as degraded samples, STRs were chosen as the marker of choice for the identification of rhino during this part of the study, rather than SNP or other available markers.

The quality control that is applied in human forensic cases should also be applied in wildlife forensic cases. Forensic quality control measures include that the evidence samples should be collected by an officer of the SAPS. Samples should be individually wrapped and packaged for the forensic laboratory to ensure that contamination does not occur. The evidence samples should also be clearly labelled and in such a way that no tampering with evidence can take place. The package should be sealed by the officer before it is dispatched to the laboratory.

To ensure that the chain of custody is maintained, laboratory personnel should make sure that no tampering took place with the evidence samples. All samples must be processed separately to avoid contamination in the laboratory. DNA extraction, amplification and electrophoresis laboratories should be separated to avoid contamination between evidence and reference samples. Profile analysis should be

done independently by two analysts and their results should be compared. A DNA forensic report should be compiled and sent to either the investigating officer or the prosecutor of the case.

6.3 DNA database

In the DNA forensic report that follows from investigation, the statistical probability that another rhino may have the same DNA profile as the poached rhino, should also be addressed. To prove the identity of a rhino, a DNA genetic database similar to the National DNA Database of South Africa (NDDSA) should exist for rhino. This database can be used to calculate the profile probability. The database will consist of reference samples and evidence samples. All rhino with DNA profiles have to be added onto the database. This will facilitate identification of rhino when evidence samples are screened against the reference database, (if the rhino was previously genotyped). The database can also be used to determine kinship if the poached rhino was not previously genotyped, by using various approaches related to the verification of siblings or parentage.

One of the aims of this study was to select a panel of STR markers to establish a DNA genetic database for white and black rhino that is suitable for forensic purposes. To achieve this aim, 85 STR markers were screened for both species of which 16 amplified successfully (Table 2.2).

6.4 Screening of STR markers in rhino

The conservation history of rhino in South Africa presents an obstacle during rhino forensic cases. Since both black and white rhino were hunted to the brink of extinction, there is very little genetic diversity remaining in current rhino populations. Genetically it is thus a significant challenge to distinguish genetically between two individuals, especially when related animals are involved.

Most of the available STR markers for rhino were thus screened for both black and white rhino during this study. Since allelic diversity at most of the rhino STR markers tested were restricted to two alleles per locus in most cases, with only two markers with six alleles present, the rhino were also screened with STR markers

developed for domesticated species such as bovine, equine, ovine, caprine and porcine markers. From the 65 cross-species markers screened, only one porcine marker, SW035, amplified successfully and showed polymorphism in both black and white rhino.

6.5 Population structure

Eight STR markers were used to screen a total of 72 black rhino and 113 white rhino. Statistical analysis to determine the population structure of the black rhino population of the Sam Knott Nature Reserve and white rhino populations of Songimvelo, Loskopdam and Mtethamusha Nature Reserves were done. Markers that amplified successfully for both black and white rhino were selected for this analysis.

Results obtained from the black rhino population indicated that this particular population had less genetic diversity than any previously studied black rhino populations. The notion that white rhino have less genetic diversity than black rhino is also supported by the results of this study. There were no significant genetic differences found between the three white rhino populations studied. This could be explained due to the fact that the populations are managed as a meta-population and that translocations between these reserves are made on a regular basis. The relatively low overall levels of genetic diversity in white rhino may also be a contributing factor.

The fact that the genetic diversity was found to be very low in both species can cause a problem with identity testing and parentage verification. There were only a small number of polymorphic loci available. The profiles of related individuals will thus be even more similar which can complicate parentage analysis even further.

6.6 Identification

The DNA profiles of all the rhino used in this study was screened for any possible genotyping errors and null alleles present. The accuracy of DNA profiles is crucial for identification and parentage verification studies. The genotyping process is complicated by factors such as the formation of artefacts, dropped alleles and null alleles. An incorrect DNA profile can have major implications. For example, if a

reference sample has been genotyped incorrectly and evidence samples are compared to the reference sample, the results will be an exclusion and a guilty suspect may be exonerated in a court of law. With parentage analysis (which is complicated by the fact that Mendelian inheritance has to be taken into consideration), a male could be excluded as the father even though he is the father of the offspring. For these reasons it is important that a dataset is screened for genotyping errors before any further analysis is attempted.

No genotyping errors were detected during the current study, but there were, however, three markers in the white rhino populations that may have null alleles present namely DB42; DB66 and 7C. There were no null alleles present in any of the markers used in the black rhino population.

To optimize the STR marker sets for identification and parentage verification purposes, the monomorphic marker DB44 was removed and two additional polymorphic markers, B1RH37D and B1RH2B were added to the black rhino marker set. For the white rhino STR marker set, the three monomorphic markers, BR4, BR17 and DB23 were removed and replaced with six polymorphic markers, DB42, DB49, DB52, DB66, 7B and 7C. The number of alleles per locus still ranged between two to six alleles, but the overall average number of alleles per locus improved.

For the rhino populations studied, it was found that the minimum number of loci necessary for individual identification was nine loci for the black rhino and 11 loci for the white rhino. The optimal PI_{sibs} values for identification were not achieved, but the values were close enough to ensure that the rhino were identifiable. The black rhino studied came from a single isolated population and the white rhino came from three populations that are managed as a single meta-population. A proportion of the animals screened in each species will thus likely be related to one another. It follows that if it is possible to distinguish these individuals from one another reasonably effectively despite the close relatedness, it will without doubt be possible to distinguish unrelated individuals from different populations. It is thus most likely that the optimal PI_{sibs} values will be achieved if more rhino populations are included in this study.

While non-invasive samples have proved valuable in many studies in the field of conservation genetics, such samples should not be considered as a potential source of reference material. In general, even with high quality tissue and blood samples,

genotyping errors cannot be avoided. Reference samples are used specifically for identification and parentage verifications and if the sample sources cannot be trusted the DNA profile would even be less trustworthy since the incidence of artefacts and dropped alleles will increase.

The original biological sample collected should always be stored in duplicate and preferably this should be stored at different facilities. This will ensure that samples do not get lost, for example malfunctioning of a freezer can cause the samples to degrade. Availability of the original biological sample means that the DNA profile can always be re-tested if needed.

6.7 Parentage verification

Discussions with game farmers and conservation managers revealed that studbooks for rhino populations are kept. The mothers of calves are identified via field observations and the possible fathers are listed based on the location of dominating bulls in that area. The assumed parentage can then be verified by using DNA profiles.

The parentage of 34 white rhino calves were verified successfully. The verification of the parentage of the black rhino calves were less successful due to the fact that all the rhino in the population were not on the DNA database yet. The close relatedness of all individuals in the black rhino populations also led to males not being excluded as the possible father even though the studbook records indicated that it is impossible for that male to be the father of those calves. The mothers of the calves were successfully verified by the DNA profiles except for one calf where the possible mother was excluded. The marker set for black rhino should however be more successful for other black rhino populations since the population used in this study had less genetic variation than previously studied populations. This is the result of a lack of diversity since no augmentations from other populations have occurred since the founding of this particular population.

Parentage testing potentially provides a very effective method to verify the accuracy of DNA profiling. Multiple successful applications of a profile parentage verification leads to increased confidence regarding the accuracy of the DNA profile. Tracking a number of offspring from the same individual can be used to make

conclusions regarding artefacts and null alleles. For example, if a DNA profile of a rhino bull has been used to successfully, without any exclusions, verify the parentage of 10 calves and one exclusion is found with calve 11, the chances are better that the exclusion is a true exclusion and not due to an artefact or null allele, since none of the other calves had that same type of exclusion.

The parentage verification results also indicated that even though some of the reserves, for example Songimvelo Nature Reserve, had eight adults bulls, only three bulls contributed genetic material to the next generation. These results should be taken into consideration when decisions are made regarding the translocation of animals.

6.8 Conclusion and recommendations

The selected STR sets for both black and white rhino were successfully applied to determine the population structure for all the populations involved in this study. The marker sets also proved to be suitable for identification and parentage verification purposes and all aims of this study could be achieved. Since genetic diversity is a great concern in rhino and since this does affect the accuracy of parentage verification results, the development of SNP markers in addition to the STR markers used should be investigated.

A drawback from using DNA profiling in forensic cases, is that rhino horns cannot be used to identify an individual. During a study by Hsieh *et al.* (2003), mtDNA was extracted from rhino horn and by using the cytochrome B gene, the species could be identified. No successful attempts to extract genomic DNA from rhino horn have however been reported at the time this study was concluded. Since the horn is the most valuable part of the rhino and indeed the motivation for poaching, it is also the most likely piece of biological evidence that will be recovered from the suspect. Methods for extracting DNA from rhino horn should therefore be further investigated. Avenues of research in other species and scenarios could guide some research: for example, there are currently new methods available to extract DNA from fingernail clippings in humans (Yoshida-Yamamoto *et al.* 2010). Since rhino horns are biologically very similar to fingernails, this technique should be further investigated in rhino.

A core genetic database for both black and white rhino has been established as an outflow of this study, and this has now also been validated for forensic purposes. Blind spot checks of database information should be performed to ensure the accuracy and reliability of the database. After using a profile three-times successfully in a parentage verification scenario, a DNA profile should be accepted as an accurate profile with no known artefacts or null alleles present. If a single exclusion is encountered, it should be accepted as it is and the analyst should not have undue concerns about genotyping errors that might have occurred for that particular DNA profile.

7.1 Summary

The rhino populations in South Africa are facing a great danger of extinction due to an increase in the number of poaching cases. Proving the identity of a rhino in a court of law can be difficult. DNA profiling has been used successfully in the past two decades for humans, and thus the overall aim of this study was to select a panel of microsatellite markers that can be used for both black and white rhino to prove identity and parentage and to establish a DNA database.

A total of 109 white rhino were sampled at three nature reserves in Mpumalanga; Songimvelo, Loskopdam and Mthethomusha. Seventy two black rhino were sampled at the Sam Knott Nature Reserve in the Eastern Cape. A total of 85 STR markers were screened of which eight were selected to determine the population structure of the rhino populations based on the fact that these markers amplified for both black and white rhino DNA. Regarding identity and parentage testing, the monomorphic markers were replaced with polymorphic markers and a total of nine STR markers were selected for the black rhino population and 11 for the white rhino population. All the markers were screened for possible genotyping errors, null alleles and linkage disequilibrium.

The average observed heterozygosity (Ho = 0.42808) for the black rhino population is found to be lower than for any of the previous genetic studies done on black rhino. The genetic variation found in the three white rhino population studied, are even less than the genetic variation found in the black rhino population. The observed heterozygosity ranged from 0.25 to 0.32. The F_{ST}-value between the three white rhino populations is very low. The lowest F_{ST}-value (0.036) was found to be between the rhino from Mthethomusha and Songimvelo. There have been a higher number of rhino translocations taking place between the three reserves which is clearly supported by the low F_{ST}-values. R_{ST} and *Nm* values indicated that the most gene flow has occurred between Mthethomusha and Songimvelo. The F_{ST} and R_{ST} values, as well as the results obtain from STRUCTURE and the AMOVA testing all indicate that the three white rhino populations can be considered as a single population.

The probability of identity for unrelated black rhino is found to be 1 / 5 000 for unrelated individuals and 1 / 100 for related individuals. There is an estimate of

1 450 wild black rhino in South Africa. The black rhino population used in this study was isolated from other black rhino populations and the probability of identity could thus be underestimated, since these animals are related to each other.

The probability of identity for unrelated white rhino was 1 / 100 000 and for related individuals 1 / 250. There is an estimate of 17 480 wild white rhino in South Africa. If the rhino are thus unrelated, the 11 loci for white rhino would be sufficient. The optimal PI_{sibs} values were not achieved for either black or white rhino populations, but if more unrelated individuals are added to the dataset, the values should be close enough to ensure that the rhino will be identifiable.

In total, the parentage of 65% of the white rhino calves could be verified by using DNA profiling. The parentage of the black rhino calves at the Sam Knott nature reserve could not be resolved successfully due to the fact that all the rhino were not genotyped. Parentage analysis should always take studbook information into consideration when parentage is assigned. Even though the set of markers did prove sufficient for individual identification, it did however show limitations in solving parentage of the calves. The difficulty could be attributed to the fact that all the individuals in the populations have not been genotyped yet and that due to inbreeding, the genetic diversity of these populations are very low. I suggest therefore that SNPs are developed for black rhino that could be use together with STR markers to solve the parentage of these calves.

Blind spot checks of database information should also be performed to ensure the accuracy and trustworthiness of the database. After using a profile three times successfully in a parentage verification scenario, a DNA profile should be accepted as an accurate profile with no known artefacts or null alleles present. If a single exclusion is encountered, it should be accepted as it is and the analyst should not have undue concerns about genotyping errors that might have occurred for that particular DNA profile.

Amos W. & Balmford A. 2001. When does conservation genetics matter? *Heredity* 87, 257-265

Amos W. & Harwood J. 1998. Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions of The Royal Society / Biological Sciences* 353, 177-186

Applied Biosystems. 1996. ABI Prism™ Genotyper® 2.0 User's Manual. The Perkin-Elmer Corporation Perkin Elmer applied biosystems

Applied Biosystems. 1997. GeneScan®Reference Guide. The Perkin-Elmer Corporation. The Perkin-Elmer Corporation Perkin Elmer Applied Biosystems

Applied Biosystems. 1998. ABI Prism®377 DNA Sequencer User's Manual. The Perkin-Elmer Corporation Perkin Elmer Applied Biosystems

Avise J.C. 2004. Molecular markers, natural history and evolution. Second edition. Sinauer Associates, 684

Ayres K.L. & Overall A.D.J. 2004. API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives. *Molecular Ecology Notes* 4, 315-318

Balloux F. & Goudet J. 2002. Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology* 11, 771-783

Balloux F. & Lugon-Moulin N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11, 155-165

Baruch E. & Weller J.I. 2008. Estimation of the number of SNP genetic markers required for parentage verification. *Animal Genetics* 39, 474-479

Bonnet A., Thévenon S., Maudet F. & Maillard J.C. 2002. Efficiency of semi-automated fluorescent multiplex PCR with 11 microsatellite markers for genetic studies of deer populations. *Animal Genetics* 33, 343-350

Bothma J. du P. 2002. Game ranch management. Fourth edition. Van Schaik publishers, Pretoria, 709 pages

Bouzat J.L., Lewin H.A. & Paige K.N. 1998. The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *The American Naturalist* 152, 1-6

Brown P. 2000. Clamp down on wildlife criminals. *The Guardian*. February 17, 2000. Guardian Newspapers Limited

Brown S.M & Houldens B.A. 1999. Isolation and characterization of microsatellite markers in the black rhinoceros (*Diceros bicornis*). *Molecular Ecology* 8, 1551-1561

Budowle B., Garofano P., Hellman A., Ketchum M., Kanthaswamy S., Parson W., Van Haeringen W., Fain S. & Broad T. 2005. Recommendations for animal DNA forensic and identity testing. *International Journal of Legal Medicine* 119, 295-302

Budowle B., Eisenberg A.J. & Van Daal A. 2009. Review: Validity of low copy number typing and applications to forensic science. *Croatian Medical Journal* 50, 207-217

Butler J.M. 2005. Forensic DNA typing. Biology and technology behind STR markers. Second edition. Academic press, Burlington, 660 pages

Caballero A., Rodriguez-Ramilo S.T., Avila V. & Fernández J. 2010. Management of genetic diversity of subdivided populations in conservation programmes. *Conservation Genetics* 11, 409-419

Chambers G.K. & MacAvoy E.S. 1999. Molecular genetic analysis of hybridization. *Science for Conservation* 105. Department of Conservation, Wellington, New Zealand

Chapuis M. & Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24, 621-631

Clifford S.L., Jeffrey K., Bruford M.W. & Wickings E.J. 1999. Identification of polymorphic microsatellite loci in the gorilla (*Gorilla gorilla gorilla*) using human primers: application to noninvasively collected hair samples. *Molecular Ecology* 8, 1551-1561

Coates B.S., Sumerford D.V., Miller N.J., Kim K.S., Sappington T.W., Siegfried B.D. & Lewis L.C. 2009. Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity* 100, 556-564

Coombs J.A., B.H. Letcher, & K.H. Nislow. 2008. CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. *Molecular Ecology Resources* 8, 578-580

Cousins J.A., Sadler J.P. & Evans J. 2010. The challenge of regulating private wildlife ranches for conservation in South Africa. *Ecology and Society* online www.ecologyandsociety.org/volXX/issYY/artZZ/, 2010/03/25

Creel S., Spong G., Sands J.L., Rotella J., Zeigle J., Joe L., Murphy K.M. & Smith D. 2003. Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. *Molecular Ecology* 12, 2003-2009

Cunningham J., Harley E.H. & O'Ryan C. 1999. Isolation and characterization of microsatellite loci in black rhinoceros (*Diceros bicornis*). *Electrophoresis* 20, 1778-1780

Cunningham J., Morgan-Davies A.M. & O'Ryan C. 2001. Counting rhinos from dung: estimating the number of animals in a reserve using microsatellite DNA. *South African Journal of Science* 97, 293-294

Dakin E.E. & Avise J.C. 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93, 504-509

Day A. 2001. Nonhuman DNA testing increases DNA's power to identify and convict criminals. *Silent witness* 6, 1-2

DeSalle R. & Amato G. 2004. The expansion of conservation genetics. *Nature Reviews Genetics.* 5, 702-712

DeWoody J.A. 2005. Molecular approaches to the study of parentage, relatedness and fitness: practical applications for wild animals. *Journal of Wildlife Management* 69, 1400-1418

Dieringer D. & Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3, 167-169

Du Toit J.G. 1998. Rhino Ranching. Management manual for owners of white rhino. South African Veterinary Foundation. Africa Publishers, Pretoria, 64 pages

Du Toit R. (ed). 2006. Guidelines for implementing SADC rhino conservation strategies. SADC Regional Programme for Rhino Conservation. Harare (Zimbabwe), 95 pages

Eichmann C., Berger B., Steinlechner M. & Parson W. 2005. Estimating the probability of identity in a random dog population using 15 highly polymorphic canine STR markers. *Forensic Science International* 151, 37-44

Ely J.J., Aivaliotis M.J., Kalmin B., VandeBerg J.L. & Stone W.H. 1999. Comparison of biochemical polymorphisms and short tandem repeat (STR) DNA markers for paternity testing in Rhesus monkeys (*Macaca mulatta*). *Biochemical Genetics* 37, 323-334

Emslie R. H. & Brooks M. 1999. Status survey and conservation action plan. African Rhino. IUCN/SSC African rhino specialist group. IUCN. Information Press, Oxford, 92 pages

Emslie R.H., Amin R. & Kock R. 2009. Guidelines for the *in situ* re-introduction and translocation of African and Asian Rhinoceros. *IUCN Species Survival Commission,* Gland, 115 pages

Evanno G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology* 14, 2611-2620

Evett I. & Weir, B. 1998. Interpreting DNA evidence: Statistical Genetics for Forensic Scientists. Sinauer Associates, Sunderland, 278 pages

Excoffier L. & Heckel G. 2006. Computer programs for population genetics data analysis: a survival guide. *Nature Reviews Genetics* 7, 745-758

Excoffier L., Laval G. & Schneider S. 2006. Arlequin ver 3.1. An integrated software package for population genetics data analysis. *User manual*, 145 pages

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

Fan H., & Chu J.Y. 2007. A brief review of Short Tandem Repeat mutation. *Genomics, Proteomics & Bioinformatics* 5, 7-14

Florescu A., Davila J.A., Scott D., Fernando P., Kellner K., Morales C., Melnick D., Boag P.T. & Coeverden de Groot P. 2003. Polymorphic microsatellites in white rhinoceros. *Molecular Ecology Notes* 3, 334-345

Frankham R. 1995. Conservation Genetics. *Annual Review of Genetics* 29, 305-327

Frankham R., Ballou J.D. & Briscoe D.A. 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge, 640 pages

Gaggiotti O.E., Lange O., Rassmann K. & Gliddon C. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* 8, 1513-1520

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

Gutiérrez-Espeleta G.A., Kalinowski S.T., Boyce W.M. & Hedrick P.W. 2000. Genetic variation and population structure in desert bighorn sheep: implications for conservation. *Conservation Genetics* 1, 3-15

Goodman S.J. 1997. RST CALC: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology* 6, 881-885

Goodwin W., Linacre A. & Hadi S. 2007. *An introduction to Forensic Genetics.* John Wiley & Sons, Oxford, 162 pages

Haig S.M. 1998. Molecular contribution to conservation. *Ecology* 79, 413-425

Halliburton R. 2004. Introduction to population genetics. International edition. Pearson Education International, Upper Saddle River, 672 pages

Hanotte O. & Jianlin H. 2005. Genetic characterization of livestock populations and its use in conservation decision-making. The role of Biotechnology, Villa Gualino, Turin, Italy – 5-7 March, 2005, 131-136

Hardy O.J., Charbonnel N., Frévill H. & Heuertz M. 2003. Microsatellite allele sizes: A simple test to assess their significance on genetic differentiation. *Genetics* 163, 1467-1482

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* 14, 2981-2990

Hedrick P.W. & Kalinowski, S.T. 2000. Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics* 31, 139-162

Hedrick P.W., Parker K.M. & Lee R.N. 2001. Using microsatellite and MHC variation to identify species, ESUs and MUs in the endangered Sonoran topminnow. *Molecular Ecology* 10, 1399-1412

Hedrick P. 2005a. 'Genetic restoration:' a more comprehensive perspective than 'genetic rescue'. *Trends in Ecology and Evolution* 20, 109

Hedrick P.W. 2005b. Genetic of populations. Third edition. Jones and Bartlett Publishers, Sudbury, 737 pages

Henegariu O., Heerema N.S., Dlouhy S.R., Vance G.H. & Vogt P.H. 1997. Multiplex PCR: critical parameters and step-by-step protocol. *BioTechniques* 23, 504-511

Henly K., Lindenmayer D.B., Margules C.R., Saunders D.A. & Wissel C. 2004. Species survival in fragmented landscapes: where are we now? *Biodiversity and Conservation* 13, 1-8

Hoffman J.I. & Amos W. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* 14, 599-612

Hofmeyr M.S. 2006. Black and white rhino capture in the wild. *Proceedings of the North American Veterinary Conference.* Volume 20. January 7-11, 2006, Orlando, Florida

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* 136, 1-11

IUCN. 2010. IUCN Red list of threatened species. Version 2010.2. www.iucnredlist.org

Johnson J.A., Toepfer J.E. & Dunn P.O. 2003. Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Molecular Ecology* 12, 3335-3347

Jones A.G. & Ardren W.R. 2003. Methods of parentage analysis in natural populations. *Molecular Ecology* 12, 2511-2523

Jones M.E., Paetkau D., Geffen E. & Moritz C. 2004. Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore. *Molecular Ecology* 13, 2197-2209

Kalinowski S.T., Taper M.L. & Marshall T.C. 2007. Revising how the computer program CERVUS accomodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16, 1099-1106

Karesh W.B., Smith F. & Frazier-Taylor H. 1987. A remote method for obtaining skin biosy samples. *Conservation Biology* 1, 261-262

Kendall M. & Stewart A. 1977. The advanced theory of statistics, volume 1. Macmillan, New York, 676 pages

Kotze A., Grobler J.P. & Buduram P. 2002. The application of genetic markers in game farming in South Africa. Proceedings of the 7th World Congress on Genetics Applied to Livestock Production. Montpellier, France 19-23 August, Session 7

Kotze A., Harris L., Grobler P. & Ehlers K. 2003. Workshop on the establishment of a national database for game. ARC Irene, South Africa, 18 February 2003

Kotze A., Ehlers K., Cilliers D.C. & Grobler J.P. 2008. The power of resolution of microsatellite markers and assignment test to determine the geographic origin of cheetah (*Acinonyx jubatus*) in Southern Africa. *Mammalian Biology* 73, 457-462

Krug W. 2001. Private supply of protected land in Southern Africa: A review of markets, approaches, barriers and issues. Workshop on market creation for biodiversity products and services. Paris, 25 and 26 January (Centre for Social and Economic Research on the Global Environment)

Lillandt B., Bensch S., Hansson B., Wennerberg L. & Von Schantz T. 2002. Isolation and cross-species amplification of microsatellite loci in the Siberian jay (*Perisoreus infaustus*). *Hereditas* 137, 157-160

Ljungqvist M., Åkesson M. & Hansson B. 2010. Do microsatellites reflect genome-wide genetic diversity in natural populations? A comment on Väli *et al.* (2008). *Molecular Ecology* 19, 851-855

Markoulatos P., Siafakas N. & Moncany M. 2002. Multiplex polymerase chain reaction: A practical approach. *Journal of Clinical Laboratory Analysis* 16, 47-51

Marshall T.C., Slate J., Kruuk L.E.B. & Pemberton J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639-655

Marshall T.C., Sunnucks P., Spalton J.A., Greth A. & Pemberton J.M. 1999. Use of genetic data for conservation management: the case of the Arabian oryx. *Animal Conservation* 2, 269-278

Martin V. 2009. A dramatic rise in rhino poaching in South Africa. *Wild.* 17 January 2009, *www.wild.org/blog/a-dramatic-rise-in-rhino-poaching-in-south-africa*

Maudet C., Miller C., Bassano B., Breitenmoser-Würsten C., Gauthier D., Obexer-Ruff G., Michallet J., Taberlet P & Luikart G. 2002. Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex *Capra ibex (ibex). Molecular Ecology* 11, 421-436

McGranahan D.A. 2008. Managing private, commercial rangelands for agricultural production and wildlife diversity in Namibia and Zambia. *Biodiversity Conservation* 17, 1965-1977

Meirmans P.G. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60, 2399-2402

Michalakis Y. & Excoffier L. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142, 1061-1064

Milliken T., Emslie R.H. & Talukdar B. 2009. African and Asian Rhinoceroses – Status, conservation and trade. Report from IUCN/SSC, Asian Rhino Specialist Groups and TRAFFIC, 18 pages

Mills G. & Hes L. 1997. The complete book of Southern African mammals. Struik Publishers. Cape Town, 356 pages

Morin P.A., Luikart G., Wayne R.K. & the SNP workshop group. 2004. SNPs in ecology, evolution and conservation. Review. *TRENDS in Ecology and Evolution*. 19, 208-216

Morgan D.L. & Rieseberg L.H. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* 13, 1341-1356

Morling N., Allen R.W., Carracedo A., Geada H., Guidet F., Hallenberg C., Martin W., Mayr W.R., Olaisen B., Pascali V.L. & Schneider P.M. 2002. Paternity Testing Commission of the International Society of Forensic Genetics: recommendations on genetic investigations in paternity cases. *Forensic Science International* 129, 148-157

Muir-Leresche K. & Nelson R.H. 2000. Private property rights to wildlife: The Southern African Experiment. Competitive Enterprise Institute . International Centre for Economic Research. ICER Working Papers number 02-2000. http://ideas.repec.org/p/icr/wpicer/02-2000.html

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* 9, 239-242

Norton J.E. & Ashley M.V. 2004. Genetic variability and population differentiation in captive Baird's Tapirs (*Tapirus bairdii*). *Zoo Biology* 23, 521-531

O'Ryan C. & Harley E.H. 1993. Comparisons of mitochondrial DNA in black and white rhinoceroses. *Journal of Mammalogy* 74, 343-346

Ogden R., Dawnay N. & McEwing R. 2009. Wildlife DNA forensics – bridging the gap between conservation genetics and law enforcement. Endangered Species research preprint January 2, 1-17

Oliveira E.J., Pádua J.G., Zucchi M.I., Vencovsky R. & Vieira M.L.C. 2006. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology* 29, 294-307

Paetkau D. & Strobeck C. 1994. Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology* 3, 489-495

Paetkau D. & Strobeck C. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* 2, 519-520

Parker P.G., Snow A.A., Schug M.D., Booton G.C. & Fuerst P. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* 79, 361-382

Pemberton J.M., Slate J., Bancroft D.R. & Barrett J.A. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* 4, 249-252

Pompanon F., Bonin A., Bellermain E. & Taberlet P. 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics,* 1-13

Pritchard J.K., Stephens M. & Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945-959

Pritchard J.K., Wen X. & Falush D. 2009. Documentation for *Structure* software: Version 2.3. Software from www.pritch.bsd.uchicago.edu/structure.htm

Randi E., Pierpaoli M., Beaumont M., Ragni B. & Sforzi A. 2001. Genetic identification of Wild and Domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology and Evolution* 18, 1679-1693

Reid T.M., Wolf C.A., Kraemer C.M., Lee S.C. Baird M.L. & Lee R.F. 2004. Specificity of sibship determination using the ABI identifiler Multiplex system. *Journal of Forensic Sciences* 49, 1-3

Robinson T.J., Trifonov V., Espie I. & Harley E.H. 2005. Interspecific hybridization in rhinoceroses: Confirmation of a Black x White rhinoceros hybrid by karyotype, fluorescence *in situ* hybridization (FISH) and microsatellite analysis. *Conservation Genetics* 6, 141-145

SAPA. 2010a. Rhino owner at wits end. *News 24.com* 2010-07-23

SAPA. 2010b. 14 Rhino killed in 2010. *News 24.com* 2010-01-23

SAPA. 2010c. Skokkende statistieke oor renosters. *Volksblad: Briewe* 2010-02-25

SAPA. 2010d. Rhino poachers strike in KZN. *News 24.com* 2010-07-25

SAPA. 2010e. Poachers de-horn live rhino. *News 24.com* 2010-06-29

SAPA. 2010f. Dis oorlog na stroping van horings. *Volksblad* 2010-02-04

SAPA. 2010g. Cops nab 3 over rhino horns. *News 24.com* 2010-06-15

SAPA. 2010h. Rhino horn smuggler jailed. *News 24.com* 2010-07-01

SAPA. 2010i. Rhino poachers get 10 years. *News 24.com* 2010-03-25

SAPA 2010j. Rhinos are now being poisoned. *News 24.com* 2010-09-29

SAPA. 2010k. Renosterstropery: 'Baasbreine' gevang. *Volksblad* 2010-09-22

SAPA. 2010l. Callous rhino poaching – 11 get bail. *News 24.com* 2010-09-22

SAPA. 2010m. Fears poaching accused will flee. *News 24.com* 2010-09-22

Scott C.A. 2008. Microsatellite variability in four contemporary rhinoceros species; implications for conservation. *MSc. Thesis.* Queen's University Kongston, Ontario, Canada, 109 pages

Selkoe K.A. & Toonen R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9, 615-629

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457-462

Soulsbury C.D., Iossa G., Edwards K.J., Baker P.J. & Harris S. 2007. Allelic dropout from a high-quality DNA source. *Conservation Genetics* 8, 733-738

Spielman D., Brook B.W. & Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America* 101, 15261-15364

Spong G., Johansson M. & Björklund M. 2000. High genetic variation in leopards indicates large and long-term stable effective population size. *Molecular Ecology* 9, 1773-1782

Swart M.K.J., Ferguson J.W.H., du Toit R. & Flamand J.R.B. 1994. Substantial genetic variation in southern African black rhinoceros (*Diceros bicornis*). *Journal of Heredity* 85, 261-266

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

Tóth G., Gáspári Z. & Jurka J. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* 10, 967-981

Väli Ü., Einarsson A., Waits L. & Ellegren H. 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology* 17, 3808-3817

Valière N. 2002. GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes* 2, 377-379

Van Oosterhout C., Hutchinson W.F., Wills D.P.M. & Shipley P. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4, 535-538

Vernisi C., Crestanello B., Pecchioli E., Tartari D., Caramelli D., Hauffe H. & Bertorelle G. 2003. The genetic impact of demographic decline and reintroduction in the wild boar (*Sus scrofa*): a microsatellite analysis. *Molecular Ecology* 12, 585-595

Waits L.P., Luikart G. & Taberlet P. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* 10, 249-256

Wallis G.P. 1994. Population genetics and conservation in New Zealand: A hierarchical synthesis and recommendations for the 1990s. *Journal of The Royal Society of New Zealand* 24, 143-160

Wasser S.K., Clark B. & Laurie C. 2009. The ivory trail. *Scientific American*. July, 68-76

Weir B.S. & Cockerham C.C. 1984. Estimating F-Statistics for the analysis of population structure. *Evolution* 38, 1358-1370

Weller J.I., Seroussi E. & Ron M. 2006. Estimation of the number of genetic markers required for individual animal identification accounting for genotyping errors. *Animal Genetics* 37, 387-389

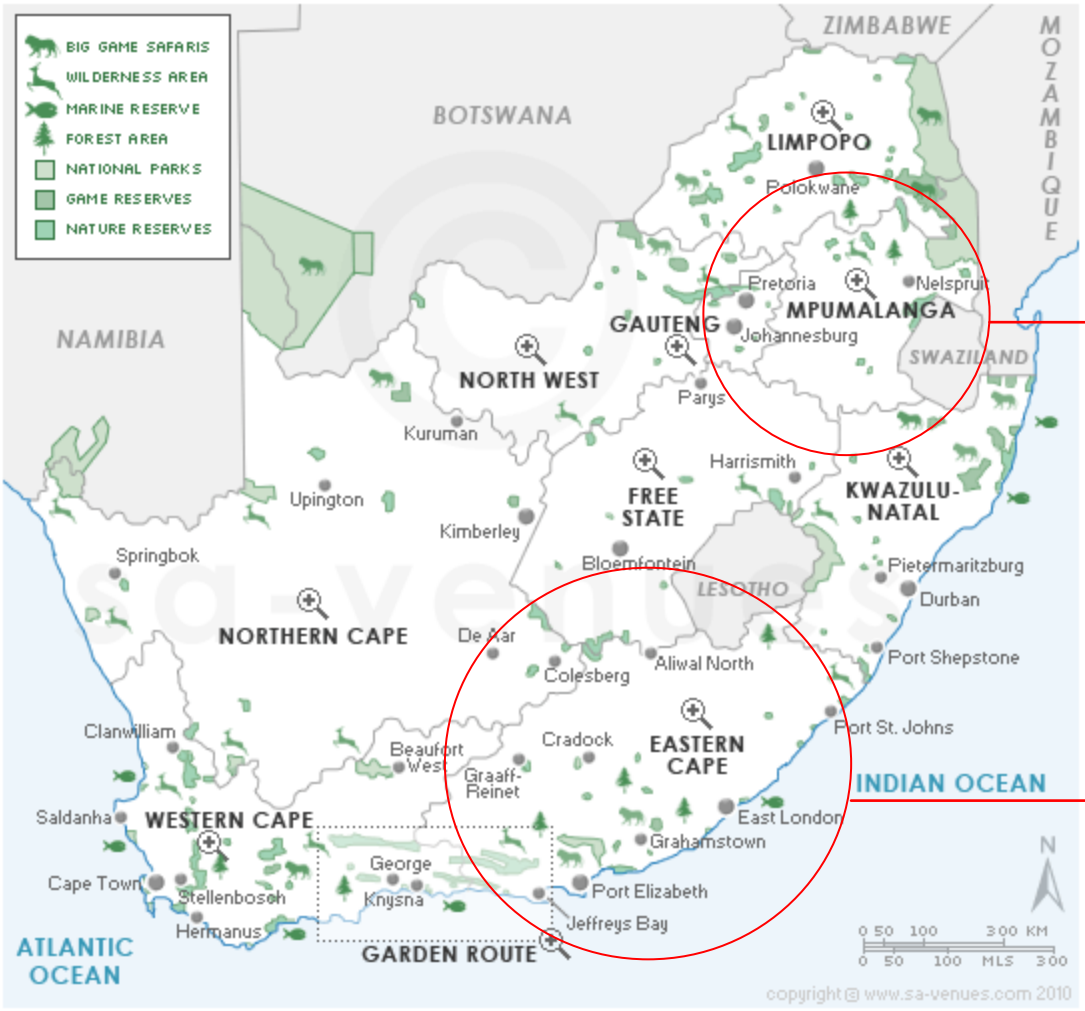
Wright S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395-420

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, *Rhinocerso unicornis*. *Molecular Phylogenetics and Evolution* 7, 189-194

Yoshida-Yamamoto S., Nishimura S., Okuno T., Rakuman M. & Takii Y. 2010. Efficient DNA extraction from nail clippings using the protease solution from *Cucumis melo*. *Molecular Biotechnology*. 46, 41-48

Zhang Y., Wang X., Ryder O.A., Li H., Zhang H., Yong Y. & Wang P. 2002. Genetic diversity and conservation of endangered animal species. *Pure and Applied Chemistry* 74, 575-584

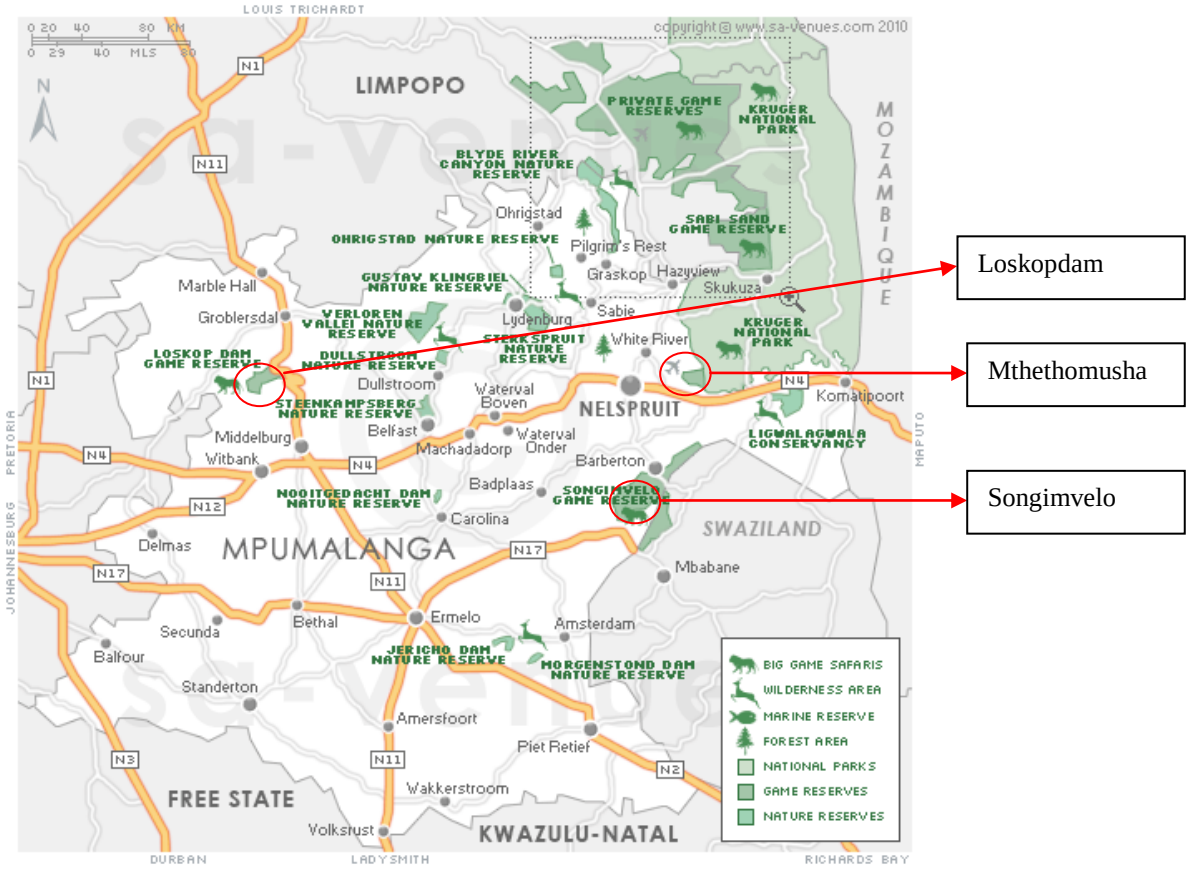
A1: Map of the major nature reserves in the nine provinces of South Africa (www.sa-venues.com)



See A2

See A3

A2: Map of Mpumalanga Province indication the location of the three nature reserves, Songimvelo, Loskopdam and Mthethomusha (www.sa-venues.com)



A3: Map of the Eastern Cape Province indicating the location of the Sam Knott Nature Reserve (www.sa-venues.com)

B1: List of the white rhino sampled from the Songimvelo Nature Reserve

DNA number	Male / Female
WR69/02	Male
WR70/02	Male
WR71/02	Male
WR72/02	Male
WR73/02	Male
WR74/02	Male
WR75/02	Male
WR76/02	Female
WR77/02	Female
WR78/02	Female
WR79/02	Female
WR80/02	Female
WR81/02	Female
WR82/02	Female
WR83/02	Female
WR84/02	Female
WR85/02	Female
WR86/02	Female
WR87/02	Female
WR88/02	Calf (Sub-adult male)
WR89/02	Calf (Sub-adult male)
WR90/02	Calf (Sub-adult female)
WR91/02	Calf (Sub-adult female)
WR92/02	Calf (Sub-adult female)
WR93/02	Calf (Sub-adult female)
WR94/02	Calf (Sub-adult female)
WR95/02	Calf (Sub-adult female)
WR96/02	Calf (Sub-adult male)
WR97/02	Calf (Sub-adult female)
WR98/02	Calf (Sub-adult female)
WR99/02	Calf (Sub-adult female)
WR100/02	Calf (Sub-adult female)
WR101/02	Calf (Sub-adult male)
WR102/02	Calf (Yearling male)
WR103/02	Calf (Yearling female)
WR104/02	Calf (Yearling female)
WR105/02	Calf (Juvenile)
WR106/02	Calf (Juvenile male)
WR107/02	Calf (Juvenile male)
GA314/03	Calf (female)
GA315/03	Calf (female)
GA316/03	Calf (male)
GA317/03	Calf (male)
GA318/03	Calf (female)
GA319/03	Calf (female)
GA320/03	Calf (male)
GA321/03	Calf (female)
GA322/03	Calf (male)
GA323/03	Calf (male)
GA326/03	Calf (male)
GA327/03	Female
GA328/03	Female
GA329/03	Male

B2: List of the white rhino sampled from the Loskopdam Nature Reserve

DNA number	Male / Female
WR28/02	Female
WR30/02	Male
WR31/02	Male
WR33/02	Male
WR34/02	Calf (male)
WR35/02	Male
WR36/02	Adult female
WR37/02	Male
WR38/02	Adult female
WR39/02	Male
WR40/02	Male
WR41/02	Male
WR42/02	Male
WR43/02	Adult female
WR44/02	Calf (Sub-adult female)
WR45/02	Yearling male
WR46/02	Male
WR47/02	Calf (Sub-adult female)
WR48/02	Adult female
WR49/02	Calf (Sub-adult female)
WR50/02	Male
WR51/02	Adult female
WR53/02	Adult female
WR54/02	Adult female
WR56/02	Calf (Sub-adult male)
WR57/02	Male
WR58/02	Female
WR59/02	Male
WR60/02	Male
WR61/02	Adult female
WR62/02	Calf (male)
WR63/02	Calf (male)
WR68/02	Female

B3: List of the white rhino sampled from the Mthethomusha Nature Reserve

DNA number	Male / Female
WR20/02	Calf (Sub-adult female)
WR21/02	Female
WR22/02	Female
WR23/02	Calf (Sub-adult female)
WR24/02	Female
WR25/02	Calf (Sub-adult male)
WR26/02	Calf (Yearling female)
WR27/02	Male
GA386/03	Male
GA387/03	Female
GA388/03	Female
GA389/03	Female
GA390/03	Male
GA391/03	Male
GA392/03	Female
GA393/03	Female
GA394/03	Female
GA395/03	Male
GA100/04	Male
GA101/04	Male
GA102/04	Male
GA103/04	Female
GA104/04	Female
GA105/04	Female
GA106/04	Female
GA107/04	Female
GA108/04	Female

B4: List of the black rhino sampled from the Sam Knott Nature Reserve

DNA number	Male / Female
BLR1/07	Female
BLR2/07	Male
BLR3/07	Male
BLR4/07	Male
BLR5/07	Male
BLR6/07	Female
BLR7/07	Female
BLR8/07	Female
BLR9/07	Male
BLR10/07	Male
BLR11/07	Female
BLR12/07	Female
BLR13/07	Female
BLR14/07	Female
BLR15/07	Female
BLR16/07	Female
BLR17/07	Female
BLR18/07	Female
BLR19/07	Female
BLR20/07	Female
BLR21/07	Female
BLR22/07	Male
BLR23/07	Male
BLR26/07	Female
BLR27/07	Female
BLR28/07	Male
BLR29/07	Female
BLR1/08	Calf (male)
BLR2/08	Calf (female)
BLR3/08	Female
BLR4/08	Female
BLR5/08	Male
BLR6/08	Male
BLR7/08	Female
BLR8/08	Male
BLR9/08	Male
BLR10/08	Calf (male)
BLR11/08	Calf (female)
BLR12/08	Female
BLR13/08	Calf (female)
BLR14/08	Calf (male)
BLR15/08	Female
BLR16/08	Female
BLR17/08	Calf (male)
BLR18/08	Male
BLR19/08	Calf (female)
BLR20/08	Male
BLR21/08	Male
BLR22/08	Female
BLR23/08	Calf (female)
BLR24/08	Calf (female)
BLR25/08	Male
BLR26/08	Male
BLR27/08	Female

BLR28/08	Female
BLR29/08	Calf (male)
BLR30/08	Male
BLR31/08	Male
BLR32/08	Female
BLR33/08	Calf (female)
BLR34/08	Female
BLR35/08	Female
BLR36/08	Female
BLR37/08	Calf (male)
BLR38/08	Calf (male)
BLR39/08	Calf (female)
BLR40/08	Calf (female)
BLR41/08	Male
BLR42/08	Male
BLR43/08	Calf (male)
BLR44/08	Male
BLR45/08	Male

C1: DNA profiles of the white rhino sampled from the Songimvelo Nature Reserve

DNA number	DB42		32A		BR6		DB52		DB66		7C		DB1		DB44		DB49		7B		SW35	
WR69/02	326	326	234	248	133	135	217	219	201	201	255	255	129	129	181	181	159	161	265	265	127	133
WR70/02	330	330	234	234	133	153	217	221	201	203	253	255	129	129	173	181	159	163	263	265	133	133
WR71/02	326	330	234	234	133	133	217	221	201	201	253	255	129	131	173	181	159	163	265	265	127	127
WR72/02	326	326	234	248	133	133	219	221	203	203	255	255	129	129	173	181	161	163	265	267	127	127
WR73/02	330	330	234	234	135	153	217	219	203	203	255	255	129	131	181	181	159	161	265	265	127	133
WR74/02	330	330	234	248	133	133	219	219	203	203	247	255	129	131	173	181	159	159	265	267	127	133
WR75/02	330	330	234	248	133	135	219	221	203	203	247	255	129	129	181	181	159	161	265	267	127	133
WR76/02	324	326	234	248	133	153	217	221	201	201	253	255	129	129	173	173	159	163	261	265	133	133
WR77/02	324	326	234	248	133	133	219	219	201	203	255	255	129	131	181	181	161	161	265	265	127	127
WR78/02	324	330	234	248	153	153	217	217	201	203	247	255	129	129	173	181	159	159	265	265	127	133
WR79/02	326	330	234	248	133	153	219	221	203	203	253	255	129	131	173	181	161	163	265	267	127	127
WR80/02	326	330	248	248	133	153	217	219	201	203	253	255	129	131	181	181	159	161	263	265	133	133
WR81/02	324	330	234	248	133	133	217	217	203	203	253	255	129	129	179	181	159	159	267	267	127	133
WR82/02	326	330	234	234	133	153	217	217	201	203	255	255	129	129	173	181	159	159	265	267	127	127
WR83/02	330	330	234	234	133	133	217	217	203	203	253	255	129	129	173	181	159	159	265	265	000	000
WR84/02	326	330	234	234	133	133	219	219	203	203	247	253	129	131	181	181	161	161	265	267	133	133
WR85/02	326	330	248	248	133	133	217	221	203	203	247	247	129	129	175	181	159	159	265	265	127	133
WR86/02	326	330	234	248	133	133	217	221	201	203	255	255	129	129	181	181	159	163	263	265	133	133
WR87/02	326	330	234	234	133	153	217	219	201	203	255	255	129	129	181	181	159	161	265	265	127	133
WR88/02	326	326	234	248	133	135	219	219	201	201	255	255	129	129	181	181	161	161	265	265	127	133
WR89/02	324	330	234	234	133	153	217	217	203	203	253	253	129	129	181	181	159	159	265	267	133	133
WR90/02	326	330	234	234	133	153	217	217	201	203	255	255	129	129	173	181	159	159	263	265	127	133
WR91/02	326	326	234	248	133	133	217	219	201	201	253	255	129	131	181	181	159	161	265	265	133	133
WR92/02	324	330	234	234	133	153	217	221	203	203	253	255	129	129	181	181	159	163	265	267	127	133
WR93/02	326	326	234	234	133	135	219	219	201	203	255	255	129	129	181	181	161	161	265	265	127	133
WR94/02	326	330	248	248	135	153	217	217	201	201	255	255	129	129	181	181	159	159	265	267	133	133
WR95/02	326	326	248	248	133	135	217	219	201	203	255	255	129	129	181	181	159	161	265	265	133	133
WR96/02	326	326	234	248	133	133	217	219	201	201	253	255	129	131	181	181	159	161	265	265	133	133
WR97/02	330	330	234	234	133	133	221	221	203	203	253	255	129	129	181	181	163	163	265	265	133	133

DNA number	DB42		32A		BR6		DB52		DB66		7C		DB1		DB44		DB49		7B		SW35	
WR98/02	326	330	234	234	133	135	217	219	201	201	255	255	129	129	181	181	159	161	265	267	127	133
WR99/02	326	330	234	248	153	153	221	221	201	201	253	255	129	129	173	181	163	163	261	265	133	133
WR100/02	330	330	234	248	133	153	219	221	201	203	253	255	129	129	171	181	159	163	265	265	133	133
WR101/02	330	330	234	234	133	153	217	219	201	201	255	255	129	129	181	181	159	161	265	265	133	133
WR102/02	324	330	234	234	133	153	217	217	201	201	255	255	129	129	173	181	159	163	265	265	127	133
WR103/02	330	330	234	234	133	133	217	221	201	203	255	255	129	129	173	181	159	163	265	265	133	133
WR104/02	326	330	248	248	133	135	217	217	201	201	255	255	129	129	181	181	159	159	265	267	127	133
WR105/02	330	330	234	234	133	133	217	217	203	203	253	255	129	129	173	181	159	159	265	265	127	133
WR106/02	324	326	234	234	133	133	217	219	201	203	255	255	129	129	181	181	159	161	265	265	127	127
WR107/02	330	330	234	234	133	133	217	221	201	201	253	255	129	129	173	181	159	163	265	265	127	133
GA314/03	324	324	248	248	135	153	217	219	201	201	255	255	129	129	181	181	159	161	265	267	127	133
GA315/03	326	326	234	234	133	135	219	219	201	201	255	255	129	129	181	181	161	161	265	265	133	133
GA316/03	326	326	234	248	133	133	217	217	201	201	255	255	129	129	181	181	159	159	265	267	127	133
GA317/03	324	326	248	248	133	135	217	219	201	203	255	255	129	129	181	181	159	161	265	265	127	127
GA318/03	330	330	234	248	133	135	217	217	201	201	253	255	129	129	173	181	159	159	263	265	127	133
GA319/03	326	326	248	248	133	135	217	221	201	203	253	255	129	131	181	181	159	163	265	265	127	133
GA320/03	326	326	234	248	133	153	217	219	201	201	253	255	129	131	181	181	159	161	265	265	127	133
GA321/03	330	330	234	248	133	153	217	217	201	201	255	255	129	129	181	181	159	159	265	265	133	133
GA322/03	324	330	234	234	153	153	217	217	203	203	247	253	129	129	181	181	159	159	265	265	127	133
GA323/03	330	330	234	234	133	133	219	221	203	203	247	253	129	129	173	181	161	163	265	265	133	133
GA326/03	330	330	234	234	133	153	219	221	203	203	253	253	129	129	181	181	161	163	265	267	127	127
GA327/03	324	330	234	234	133	153	217	219	203	203	253	255	129	129	181	181	159	159	265	265	127	133
GA328/03	330	330	234	234	133	153	219	221	203	203	253	255	129	129	181	181	161	163	265	267	127	127
GA329/03	330	330	234	248	133	153	217	219	201	201	247	255	129	129	181	181	159	159	265	265	127	127

C2: DNA profiles of the white rhino sampled from the Loskopdam Nature Reserve

DNA number	DB42		32A		BR6		DB52		DB66		7C		DB1		DB44		DB49		7B		SW35	
WR28/02	324	324	234	248	133	135	217	219	000	000	000	000	129	129	181	181	159	161	265	267	133	133
WR30/02	328	328	234	248	133	133	217	221	201	201	255	255	129	129	173	181	159	163	265	267	133	133
WR31/02	324	330	234	248	131	153	217	219	201	201	253	255	129	129	181	181	159	161	265	265	127	133
WR33/02	326	328	248	248	133	153	219	223	201	201	255	255	129	129	173	173	159	163	265	265	133	133
WR34/02	330	330	234	248	133	153	219	221	201	201	255	255	129	129	181	181	161	163	265	265	127	133
WR35/02	324	330	234	248	135	153	219	221	203	203	253	255	129	131	181	181	161	163	265	267	127	133
WR36/02	324	330	248	248	135	153	217	219	203	203	255	255	129	129	181	181	159	161	265	267	133	133
WR37/02	330	330	234	248	133	133	221	221	201	201	255	255	129	131	175	181	163	163	267	267	127	133
WR38/02	324	324	234	234	133	133	217	221	199	201	255	255	129	129	181	181	159	163	265	267	127	133
WR39/02	330	330	248	248	133	133	219	221	201	201	255	255	129	129	173	181	161	163	265	267	127	133
WR40/02	324	324	234	234	133	135	217	217	201	201	255	255	129	129	181	181	159	159	265	265	127	133
WR41/02	324	330	234	248	133	153	221	221	203	203	255	255	129	129	175	181	163	163	265	265	127	133
WR42/02	324	330	234	248	133	153	219	221	203	203	255	255	129	129	181	181	161	163	265	267	133	133
WR43/02	324	330	234	248	133	153	219	221	201	201	255	255	129	129	181	181	161	163	265	265	127	127
WR44/02	324	330	234	248	133	133	217	219	201	203	255	255	129	129	181	181	159	161	267	267	133	133
WR45/02	324	324	234	248	133	135	217	219	203	203	255	255	129	129	181	181	159	161	265	265	127	133
WR46/02	324	330	248	248	133	133	217	221	201	201	255	255	129	129	173	181	159	163	267	267	127	133
WR47/02	324	324	234	248	133	135	219	221	201	203	251	253	129	129	181	181	161	163	265	267	127	133
WR48/02	324	330	248	248	133	133	217	221	201	203	253	253	129	129	181	181	159	163	265	267	133	133
WR49/02	324	324	248	248	133	153	219	221	203	203	251	253	129	129	181	181	161	163	265	267	127	133
WR50/02	324	330	248	248	133	133	219	221	203	203	253	253	129	129	181	181	159	161	265	265	127	127
WR51/02	324	330	234	248	133	135	219	221	203	203	253	253	129	131	181	181	161	163	265	265	133	133
WR53/02	324	330	234	248	133	153	219	221	201	203	253	253	129	129	181	181	159	161	265	265	133	133
WR54/02	324	326	234	248	133	133	217	219	201	201	251	253	129	129	179	181	159	159	265	267	127	133
WR56/02	324	330	248	248	133	133	221	221	203	203	253	253	129	129	181	181	163	163	265	265	133	133
WR57/02	324	330	248	248	133	133	219	221	203	203	251	253	129	131	173	181	161	163	265	265	133	133
WR58/02	324	330	248	248	133	133	221	221	203	203	253	253	129	131	181	181	159	163	265	265	127	133
WR59/02	324	330	248	248	133	153	219	221	203	203	253	253	129	129	181	181	161	163	265	265	127	133
WR60/02	324	326	234	248	133	153	217	219	203	203	251	253	129	129	173	181	159	161	265	265	127	133

DNA number	DB42		32A		BR6		DB52		DB66		7C		DB1		DB44		DB49		7B		SW35	
WR61/02	324	330	234	248	135	153	219	219	203	203	253	253	129	129	181	181	161	161	263	265	133	133
WR62/02	324	324	234	248	133	133	221	221	203	203	253	253	129	129	181	181	163	163	265	265	127	127
WR63/02	324	330	234	248	135	153	219	219	203	203	255	255	129	129	181	181	161	161	265	265	133	133
WR68/02	326	330	248	248	133	135	217	219	201	201	255	255	131	129	173	181	159	159	265	265	127	133

C3: DNA profiles of the white rhino sampled from the Mthethomusha Nature Reserve

DNA number	DB42		32A		BR6		DB52		DB66		7C		DB1		DB44		DB49		7B		SW35	
WR20/02	330	324	234	234	133	153	217	219	203	203	253	255	129	129	181	181	159	161	265	265	127	133
WR21/02	330	330	234	234	133	153	219	221	203	203	253	255	129	129	181	181	161	163	265	267	127	127
WR22/02	324	330	234	248	133	141	217	219	203	203	253	253	129	129	173	181	159	161	265	267	127	127
WR23/02	324	330	234	234	141	141	217	219	203	203	253	255	129	131	173	181	159	161	265	267	127	133
WR24/02	324	330	234	248	133	153	217	221	203	203	253	253	129	131	181	181	159	163	261	265	127	133
WR25/02	324	326	248	248	133	141	217	217	203	203	253	253	129	129	181	181	159	159	265	265	127	127
WR26/02	324	330	234	248	141	153	217	217	203	203	253	253	129	131	175	181	159	159	265	267	127	133
WR27/02	330	330	248	248	133	153	221	221	203	203	253	255	129	129	181	181	163	163	261	267	127	133
GA386/03	330	330	234	248	133	133	217	221	203	203	253	253	129	129	175	181	159	163	265	267	127	127
GA387/03	324	326	234	248	133	141	217	219	201	201	253	253	129	129	181	181	159	161	265	265	000	000
GA388/03	324	330	234	234	141	141	217	219	203	203	251	253	127	129	173	181	159	161	265	267	127	133
GA389/03	324	324	234	248	133	151	217	221	203	203	251	253	127	127	181	181	159	163	267	267	133	133
GA390/03	324	330	234	248	133	141	217	219	203	203	245	253	129	129	175	181	159	161	265	267	133	133
GA391/03	324	324	234	248	133	133	217	217	201	201	251	253	127	129	175	181	159	159	265	267	133	133
GA392/03	330	330	234	248	133	151	219	221	201	201	251	253	127	127	173	173	161	163	265	265	133	133
GA393/03	330	330	234	248	133	141	217	219	203	203	245	251	127	129	173	181	159	159	265	267	127	133
GA394/03	330	330	234	248	133	133	217	219	203	203	245	251	129	129	173	175	159	161	265	265	133	133
GA395/03	324	328	234	234	133	141	217	219	203	203	253	255	129	131	175	181	159	161	265	267	133	133
GA100/04	326	330	234	248	153	153	217	219	201	201	255	255	129	131	175	181	159	161	265	265	133	133
GA101/04	324	326	248	248	133	133	217	221	201	201	253	255	129	131	181	181	159	163	265	265	127	133
GA102/04	326	326	234	248	133	141	217	217	201	201	253	253	129	129	181	181	159	159	265	265	127	127
GA103/04	326	330	234	234	133	133	217	217	203	203	253	255	129	131	175	181	159	159	267	267	133	133
GA104/04	330	330	248	248	133	133	219	219	203	203	253	255	129	131	181	181	157	157	265	265	127	127
GA105/04	324	324	234	234	133	133	217	221	201	201	255	255	129	129	181	181	159	163	261	265	127	127
GA106/04	330	330	234	248	141	153	217	221	203	203	255	255	129	129	173	181	159	163	265	265	127	133
GA107/04	324	330	248	248	133	135	219	221	199	203	253	255	129	141	181	181	161	163	267	267	133	133
GA108/04	326	326	234	248	133	133	217	217	201	203	253	255	129	129	181	181	159	159	265	267	127	133

C4: DNA profiles of the black rhino sampled at the Sam Knott Nature Reserve

DNA number	B1RH37D		BR4		32A		BR6		DB1		B1RH2B		DB44		DB14		DB23		SW35		BR17	
BLR01/07	148	148	126	126	242	242	134	134	127	131	229	229	176	176	282	282	182	182	123	131	124	140
BLR02/07	150	150	124	126	242	242	134	134	127	131	229	235	176	176	282	282	180	182	123	131	124	124
BLR03/07	150	150	124	126	242	242	134	154	127	131	229	235	176	176	282	282	180	180	123	123	124	124
BLR04/07	148	148	128	130	242	244	134	134	127	127	229	235	176	176	282	282	180	182	129	131	132	132
BLR05/07	148	148	124	128	242	242	140	140	127	127	235	235	176	176	282	282	180	182	131	131	132	132
BLR06/07	148	150	126	128	242	242	134	140	131	131	235	235	176	176	282	282	180	180	123	131	122	132
BLR07/07	148	148	124	128	242	242	134	140	131	131	229	235	176	176	282	282	182	182	123	131	132	132
BLR08/07	148	148	124	130	242	242	134	134	127	127	229	235	176	176	282	282	182	182	129	131	132	132
BLR09/07	148	150	126	126	242	242	134	154	127	127	229	235	176	176	282	282	180	180	131	131	132	132
BLR10/07	148	150	124	126	242	242	134	140	127	131	235	235	176	176	282	282	180	182	123	131	122	126
BLR11/07	148	148	126	130	242	242	134	154	127	131	229	235	176	176	282	282	180	182	123	131	126	132
BLR12/07	148	148	124	128	242	242	140	154	127	131	235	235	176	176	282	282	180	182	131	131	132	132
BLR13/07	148	148	124	128	242	242	154	154	127	131	229	235	176	176	282	282	180	182	123	125	122	132
BLR14/07	148	148	124	128	242	244	134	154	127	127	229	235	176	176	282	282	180	182	123	125	132	132
BLR15/07	148	150	126	126	242	242	134	134	127	131	229	235	176	176	282	282	182	182	123	131	132	132
BLR16/07	148	148	124	126	242	242	134	140	127	131	229	229	176	176	282	282	182	182	131	131	132	132
BLR17/07	148	148	124	126	242	242	134	134	127	131	229	229	176	176	282	282	180	182	131	131	132	132
BLR18/07	148	150	124	126	242	242	134	154	127	127	229	229	176	176	282	282	180	180	123	131	126	132
BLR19/07	148	148	126	128	242	242	134	140	127	131	235	235	176	176	282	282	182	182	123	131	132	132
BLR20/07	148	148	124	126	242	242	134	134	131	131	229	229	176	176	282	282	182	182	131	131	132	132
BLR21/07	150	150	126	130	242	244	134	134	127	131	229	235	176	176	282	282	180	180	123	131	122	132
BLR22/07	148	150	124	124	242	244	134	134	127	127	235	235	176	176	282	282	180	182	123	131	126	132
BLR23/07	148	148	126	126	242	244	134	140	127	131	229	235	176	176	282	282	180	182	131	131	132	132
BLR26/07	148	148	128	128	242	244	134	140	127	127	229	235	176	176	282	282	182	182	131	131	132	132

DNA number	B1RH37D		BR4		32A		BR6		DB1		B1RH2B		DB44		DB14		DB23		SW35		BR17	
BLR27/07	148	150	126	126	242	242	134	134	127	131	229	229	176	176	282	282	180	182	131	131	122	132
BLR28/07	148	148	128	128	242	244	134	140	127	127	229	235	176	176	282	282	182	182	131	131	132	132
BLR29/07	148	150	124	126	242	242	134	134	127	131	229	235	176	176	282	282	180	182	123	129	132	132
BLR01/08	148	150	128	130	242	244	134	134	127	127	229	235	176	176	282	282	180	180	131	131	122	126
BLR02/08	148	148	124	126	242	242	134	134	127	131	229	229	176	176	282	282	180	180	129	131	132	132
BLR03/08	150	150	124	126	242	242	134	154	127	131	235	235	176	176	282	282	180	182	123	123	122	132
BLR04/08	148	148	124	126	242	242	134	134	127	127	229	229	176	176	282	282	180	180	131	131	132	132
BLR05/08	148	150	124	126	242	242	134	154	000	000	235	235	176	176	282	282	182	182	131	131	126	132
BLR06/08	148	148	124	124	242	242	134	134	127	127	229	229	176	176	282	282	180	180	123	123	132	132
BLR07/08	148	150	124	126	242	244	134	140	127	131	229	229	176	176	282	282	180	182	123	125	126	132
BLR08/08	148	148	126	130	242	242	134	134	127	127	229	235	176	176	282	282	182	182	123	125	132	132
BLR09/08	148	150	124	126	242	244	134	134	127	127	229	235	176	176	282	282	180	182	131	131	132	132
BLR10/08	148	150	126	130	242	244	134	134	131	131	229	229	176	176	282	282	180	182	123	123	122	132
BLR11/08	148	148	124	124	242	242	134	154	131	131	229	235	176	176	282	282	180	182	123	131	126	132
BLR12/08	148	148	124	128	242	242	134	134	131	131	235	235	176	176	282	282	180	182	123	131	132	132
BLR13/08	148	150	126	130	242	244	134	134	131	131	229	229	176	176	282	282	180	182	123	131	126	132
BLR14/08	148	150	124	130	242	242	134	154	131	131	229	235	176	176	282	282	180	182	123	131	122	126
BLR15/08	148	148	124	126	242	242	134	154	127	131	235	235	176	176	282	282	180	182	123	123	122	132
BLR16/08	148	148	126	128	242	242	134	154	127	127	229	235	176	176	282	282	180	180	123	123	132	132
BLR17/08	148	148	126	126	242	242	134	140	127	131	229	229	176	176	282	282	180	182	125	131	132	132
BLR18/08	148	150	126	130	242	242	134	140	127	127	229	235	176	176	282	282	180	182	129	131	122	132
BLR19/08	148	150	124	126	242	242	134	140	127	127	229	235	176	176	282	282	180	180	123	125	126	132
BLR20/08	148	150	126	128	242	242	134	140	127	131	229	229	176	176	282	282	182	182	123	131	132	132
BLR21/08	148	150	124	128	242	242	134	134	127	127	235	235	176	176	282	282	180	180	131	131	122	132
BLR22/08	150	150	124	126	242	244	134	134	127	131	229	235	176	176	282	282	180	182	123	131	132	132
BLR23/08	148	150	126	126	242	242	134	154	127	127	229	235	176	176	282	282	180	182	123	123	132	132

DNA number	B1RH37D		BR4		32A		BR6		DB1		B1RH2B		DB44		DB14		DB23		SW35		BR17	
BLR24/08	150	150	128	130	242	242	134	134	127	127	229	235	176	176	282	282	180	180	123	131	122	132
BLR25/08	148	150	124	126	242	244	134	134	127	127	229	229	176	176	282	282	180	180	131	131	122	132
BLR26/08	148	148	126	128	242	242	134	140	127	127	229	235	176	176	282	282	182	182	123	131	130	132
BLR27/08	148	150	126	126	242	244	134	140	127	127	235	235	176	176	282	282	180	182	131	131	132	132
BLR28/08	148	148	124	126	242	242	134	134	127	127	229	235	176	176	282	282	180	180	123	125	132	132
BLR29/08	148	150	124	126	242	242	134	134	000	000	229	229	176	176	282	282	180	182	131	131	122	132
BLR30/08	148	150	128	128	242	242	134	140	127	127	235	235	176	176	282	282	182	182	123	125	130	132
BLR31/08	148	148	126	128	242	244	134	134	127	131	235	235	176	176	282	282	182	182	123	131	132	132
BLR32/08	150	150	124	128	242	242	134	154	127	131	229	235	176	176	282	282	180	182	123	131	122	132
BLR33/08	148	148	124	126	242	244	134	140	127	127	229	229	176	176	282	282	182	182	123	123	132	132
BLR34/08	148	148	124	126	242	242	134	134	127	131	229	235	176	176	282	282	180	182	129	131	130	132
BLR35/08	148	148	124	128	242	244	134	140	127	131	229	235	176	176	282	282	182	182	131	131	130	132
BLR36/08	148	148	126	126	242	242	134	134	127	127	229	229	176	176	282	282	180	182	131	131	132	132
BLR37/08	148	150	126	126	242	242	134	154	131	131	229	235	176	176	282	282	182	182	131	131	126	132
BLR38/08	150	150	124	130	242	242	134	134	127	131	229	235	176	176	282	282	180	180	123	131	132	132
BLR39/08	148	150	124	126	242	242	134	134	131	131	229	235	176	176	282	282	180	180	123	129	122	122
BLR40/08	148	148	126	130	242	242	134	134	127	131	229	229	176	176	282	282	182	182	123	123	126	132
BLR41/08	148	148	124	124	242	244	134	134	127	131	229	229	176	176	282	282	180	180	123	131	132	132
BLR42/08	148	148	126	124	242	242	134	134	127	131	229	229	176	176	282	282	180	182	125	125	132	132
BLR43/08	148	148	126	126	242	242	134	134	127	131	229	235	176	176	282	282	182	182	123	123	132	132
BLR44/08	148	148	128	130	242	244	134	134	127	131	229	235	176	176	282	282	180	182	131	131	126	132
BLR45/08	148	150	124	126	242	242	134	140	127	131	229	235	176	176	282	282	180	182	123	125	126	132