UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Scatology as non-invasive tool for conservation of cheetah (*Acinonyx jubatus* Von Schreber, 1775) in South Africa

by

Kiersten Sarah Herring

Thesis

Submitted in partial fulfilment of the requirements for the degree

Master of Science

in the

Faculty of Natural and Agricultural Science

University of Pretoria

September 2008

Supervisor: Dr E. van Marle-Köster

Co-supervisor: Prof P. Bloomer

"The moral progress of a nation can be judged by the way it treats its animals"

Gandhi

Table of Contents

Acknowledgements

This thesis would not exist without the encouragement and support of numerous people and organisations. I would like to thank Dr Este van Marle-Köster and Prof Paulette Bloomer not only for being my supervisors but for also being mentors and providing me with invaluable guidance. My sincere thanks to Prof Dan York and his team who put in many hours in assisting me with lab work, thank you for your enthusiasm and willingness to help. This research was conducted for De Wildt Cheetah and Wildlife Centre, thank you to Ann van Dyk, Vanessa Bezuidenhout and the team for your help and support, I hope that this research will enable you to break new boundaries! A special thanks to Kelly Marnewick for helping me with the Thabazimbi sample collections and to the farmers for allowing us access to their properties. Thank you to the MEEP team for answering my endless questions and for your advice when it came to lab analyses and stats. I'd also like to thank the following people and organisations for their input into this project: Beth Reiter, Jacqui Sommerville, Taryn Bigwood, Charles Hefer, the NZG and Jubatus.

Thank you to all of my friends and family for your encouragement.

Michael, thank you for your love and support. You encouraged me every step of the way and I am so grateful to you.

Most of all, I would like to thank my parents, Stephen and Judy Herring, for always believing in me and for constantly motivating and encouraging me. You taught me to always finish what I have started and to keep positive.

List of tables

List of figures

Abbreviations

%	Percent
2DE	Two dimensional gel electrophoresis
ºC	Degrees Celsius
ADO	Rate of Allelic dropout
AFLP	Amplified Fragment Length Polymorphism
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BSA	Bovine Serum Albumin
CITES	Convention on International Trade in Endangered Species of Wild Flora and Fauna
CTAB	Hexadecyltrimethyl-ammonium bromide
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotide triphosphate (A, G, C, T)
EDTA	Ethylene diamine tetracetic acid
FA	Rate of False Allele occurrence
FABI	The Forestry and Agricultural Biotechnology Institute
GPS	Global Positioning System
H_E	Expected Heterozygosity
H_O	Observed Heterozygosity
ISAG	International Society of Animal Genetics
IUCN	International Union for the Conservation of Nature and Natural resources
kb	Kilo-base pairs
km	Kilometres
km^2	Square kilometres
km/h	Kilometres per hour
m^2	Square meters
MHC	Major Histocompatibility Complex
ml	Millilitre
mM	Millimolar
mtDNA	Mitochondrial DNA
NCBI	National Centre for Biotechnology Information
ng	Nanograms
ng/µl	Nanograms per micro litre
nm	Nanometers
NZG	National Zoological Gardens of South Africa
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information Content
P_{ID}	Probability of identity
ρmol	Pico mole
qPCR	Quantitative PCR
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SD	Standard Deviation
SNP	Single Nucleotide Polymorphism

SS	Species-specific
SSR	Simple Sequence Repeats
Tm	Melting temperature
µl	Micro litre
µM	Micro molar
USA	United States of America
UV	Ultra violet
wBRC	Wildlife Biological Resource Centre

Abstract

Cheetah are extremely difficult to census due to their elusive nature and as a result there is only a limited amount of information regarding the status of cheetah populations across Africa. The ideal for genetic management of cheetah would be to obtain DNA from as many cheetah as possible for continued genetic profiling. This study was conducted to test the applicability of scatology, a non-invasive sampling approach, in population genetic studies of cheetah in South Africa. For the pilot study, DNA was extracted from cheetah faeces and the corresponding animals' blood samples. A species-specific mitochondrial DNA primer set was validated for captive cheetah for use in in-situ studies. PCR success rates of the species-specific primer set and 12 cat microsatellite markers were examined. The PCR amplification success of the species-specific primer set was high: all fresh faecal samples had a positive result (100%), four-day-old samples were 83% positive and two-week-old samples were 94% positive. Amplification success of the 12 microsatellites was 59% for fresh faecal samples, 20% for four-day-old samples and 2.4% for two-week-old samples. Genotyping errors were identified by comparing the genotypes obtained from microsatellite analyses of faecal DNA and the matching blood sample of each animal in the pilot study. Overall, rates of 16% and 7% were found for allelic dropout and false alleles respectively. Faecal samples were obtained from the Thabazimbi area to demonstrate the field application of molecular scatology for South African cheetah. These samples were analysed and the alleles obtained were within the same range as those from the pilot study. Data of 45 random De Wildt cheetah blood samples obtained from the NZG were analysed and compared with the results of the pilot study; as expected all alleles in the De Wildt pilot study animals occurred in the NZG samples. The results of this study indicate that, in addition to the methods currently employed in cheetah conservation, such as camera trapping and telemetry, scatology can provide a better understanding of the spatial distribution and levels and structure of genetic variation among wild cheetah in South Africa.

Chapter 1

Introduction and Literature Review

Cheetah

Duma (KiSwahili), Ingulule (Zulu) or Kisakasaka (Kasanga) are just some of the names by which the world's fastest land animal is known – the cheetah. The southern African cheetah (*Acinonyx jubatus jubatus),* a subspecies of cheetah, is a member of the family Felidae in the order Carnivora. The family Felidae consists of three subfamilies namely the Acinonychinae which is solely represented by cheetah (*Acinonyx jubatus)*, Felinae which consists of 13 genera and 28 species, and the Pantherinae consisting of four genera and seven species (Skinner & Chimimba, 2005). Other subspecies of cheetah, based on geographical range, include the Asiatic cheetah (*A. j. venaticus)*, West African cheetah (*A. j. hecki*), central African cheetah (*A. j. soemmeringii)* and the East African cheetah (*A. j. raineyii)*. The word cheetah is thought to be from the Sanskrit word *chitraka* which gives rise to the Hindi word *chita* meaning speckled or spotted one (Hunter & Hamman, 2003).

Cheetah are phenotypically characterized by their pale yellow coats that are covered in small round black spots and white underbellies. Their black 'tear-lines' are also distinctive. The 'King cheetah' is a variation of this coat pattern, that results in a blotched pattern with stripes running down the back. King cheetah were originally thought to be a leopard-cheetah hybrid (Cooper, 1926), later they were recognised as cheetah but were thought to be a new species (*A. rex*) (Pocock, 1927). In 1981 the first King cheetah cubs were born in captivity at the De Wildt Cheetah Centre where it was confirmed that the King cheetah coat pattern is due to a mutation inherited as a single autosomal recessive allele, therefore simply a colour variant of *A. jubatus* (van Dyk, 1991). Other coat colour and pattern variations include cheetah with very pale coats or coats with few spots (reported in the desert regions) as well as albino specimens (van Aarde & van Dyk, 1986, Bottriell, 1987, van Dyk, 1991, Hunter & Hamman, 2003).

The cheetah is recognised as being the fastest land animal that can achieve speeds of approximately 115km/h. Certain adaptations such as the semi-retractable claws, enlarged heart muscles, enlarged adrenals and elongated legs enable it to reach top speed (O'Brien, 1994). Another adaptive trait studied by O'Regan (2002) is the cheetah's light skull. The skull weighs less due the fact that they have smaller teeth when compared to the tooth size of other big cats such as leopards. Cheetah teeth are on average 0.3mm smaller and the canine teeth are also shorter (O'Regan, 2002).

Female cheetah are often solitary and males usually form coalitions of four or five individuals. Cheetah have large overlapping home ranges. Urine and faecal markers are used to maintain territories by males while female urine has no territorial significance but attracts males when the female is in oestrus (Skinner & Chimimba, 2005). Prey usually consists of small to medium bovids as well as the young of larger bovids. Males can take larger prey than females due to their larger size and hunting by the coalition (Skinner & Chimimba, 2005). Cheetah are adapted to grassland, savannah, Nama-Karoo and Succulent Karoo biomes. They are often thought to frequent open plains but savannah woodland also forms part of their habitat. They do not occur in forest or in woodland with thick bush unless they seek shelter (Yalden *et al*., 1980, Skinner & Chimimba, 2005).

Cheetah once ranged throughout Africa and the Middle East, across to Tajikistan in the north and India in the southeast (Massicot, 2005). Currently, they are almost exclusively found in sub-Saharan Africa following either complete extermination or near extinction from the remainder of their historical range (Nowell & Jackson, 1996). A small remnant population still exists in northern Iran and isolated populations consisting of no more than 50 mature individuals are found in North Africa (Cat Specialist Group, 2002). Today the global population is estimated to be less than 10000 breeding individuals with no subpopulation consisting of more than 1000 breeding individuals (Cat Specialist Group, 2002). Due to their limited numbers, cheetah have been listed as vulnerable on the IUCN Red List of Threatened Species and are under the protection status of CITES Appendix 1 (Convention on International Trade in Endangered Species of Wild Flora and Fauna, 1973). An Appendix 1 quota system was established under CITES in 1992 for live animals and trophies, with annual quotas for cheetah allocated as follows: 150 (Namibia), 50 (Zimbabwe), 5 (Botswana). Hunting of cheetah is prohibited in South Africa and a number of other African countries (Convention on International Trade in Endangered Species of Wild Flora and Fauna, 1973).

The South African cheetah population consists of less than 1000 animals with approximately 250 in captivity, 350 on reserves and 200-400 roaming free (Bouwer, 2003). The largest population of cheetah in South Africa thus occurs outside conservation areas (Figure 1.1). This may be due to the absence of lion and hyena and the high density of prey due to water supply and supplementary feeding on game farms (Marnewick & Cilliers, 2006). Lion and hyena predation is known to be partly responsible for over 90% of cheetah cubs dying before reaching independence. Adult cheetah are also known to be killed or forced to lose their prey to these predators (Durant, 1998, 2000).

Figure 1.1 Distribution of cheetah in South Africa including protected areas, previously protected areas and farming areas (Marnewick *et al.*, 2007)

In the past, the removal of live cheetah from the wild into captivity, sport hunting, and the fur market contributed to a decline in the species (Massicot, 2005). Potential threats to cheetah today include hunting, a change in land management practices (May, 1995), fragmentation of the population, shooting and trapping by farmers, competition with other predators (Nowell & Jackson, 1996) and illegal trade (Cilliers, 2005, personal communication, deon@dewildt.org.za).

This is not the first time in history that cheetah numbers have been low. Cheetah are thought to have passed through a population bottleneck 10 000 – 12 000 years ago (O'Brien *et al.*, 1983, 1985, 1987, Menotti-Raymond & O'Brien, 1993, Driscoll *et al.*, 2002). A bottleneck occurs when a population experiences a substantial reduction in numbers as a result of disease, natural disaster or over-killing, leaving only a few individuals to increase the population. A second, more recent, bottleneck is thought to have occurred due to inbreeding causing the very low genetic variability in southern African cheetah *(*O'Brien *et al.*, 1985, 1987, Nowell & Jackson, 1996*)*. It has been suggested that cheetah have 10 to 100 times less genetic variation than any other felid (Driscoll *et al.*, 2002).

This bottleneck hypothesis was based on allozyme data and on the separation of abundant soluble proteins in two dimensions on polyacrylamide gels (2DE) to estimate genetic diversity. None of the 18 allozymes tested were polymorphic in the cheetah samples and the average heterozygosity for all loci using 2DE was found to be low relative to other felids and other carnivore species (O'Brien *et al.*, 1983). In a later study skin grafts were attempted between 12 unrelated cheetah and two siblings (four from De Wildt Cheetah Breeding and Research Centre, two at the Johannesburg Zoo, South Africa and eight from Wildlife Safari in Oregon); 11 of these skin grafts were successful, also suggesting that there is little genetic variation between individuals (O'Brien *et al.*, 1985). Other studies measuring the variation in cheetah were based on morphological variation of cranial

characteristics of cheetah. These studies also indicated that the variation in cheetah was low due to limited variation in cranial morphology (Wayne *et al.*, 1986).

More recent studies suggest that cheetah have accumulated genetic variation since the population bottlenecks (Menotti-Raymond & O'Brien, 1993, 1995). In the study by Menotti-Raymond and O'Brien (1995), ten microsatellite markers from the domestic cat were used to demonstrate abundant polymorphism in puma, lion, cheetah and the domestic cat. The cheetah samples exhibited the lowest level of polymorphism for these markers. They were estimated to have approximately half (51%) of the heterozygosity observed in the domestic cat. This heterozygosity is higher than the results found in previous studies because coding loci (allozymes, MHC Class 1 genes), used in previous studies, evolve at a slower rate compared to the more rapidly evolving microsatellites.

Hedrick (1996) suggests an alternative potential explanation for the genetic structure of cheetah. Due to the fact that prey, predators (competitors) and habitat structure determine the range and structure of cheetah subpopulations, it may be possible that there were frequent extinction and recolonization events in some cheetah subpopulations as opposed to a population-wide bottleneck. The study concluded that if subpopulation changes are responsible for the genetic structure of cheetah then it can be assumed that there were numerous founder events and that certain subpopulations that became fixed for detrimental alleles most probably would have gone extinct. This means that if bottlenecks occurred then there was still a reasonable amount of genetic variation among the remaining individuals because of different detrimental alleles in the various subpopulations so that the overall metapopulation could be viable.

In conservation genetics, genetic variation in a species is considered beneficial and most conservation biologists assume that increased genetic variation will increase a species' chance for survival. Low levels of genetic variation are thought to threaten a species' ability to respond to threats and changes such as disease, parasites and environmental change (Amos & Harwood, 1998). With an ever decreasing cheetah population, the number of viable breeding individuals is declining which increases the risk of inbreeding and inbreeding depression. Inbreeding is the mating of closely related individuals and results in physical and reproductive abnormalities such as a rise in infertility, a lower birth rate and an increase in abnormal sperm. Cheetah have been found to have sperm counts that are ten times lower than other related felid species (such as the domestic cat) and 70 percent of their sperm is morphologically abnormal (O'Brien, 1983). Inbreeding can cause the fixation of deleterious alleles in small populations which can also result in certain populations having a lower fitness (inbreeding depression) when compared with other populations, reducing survival at various life stages (Hedrick & Kalinowski, 2000).

The degree to which inbreeding depression affects a population largely depends on the genetic constitution of the species, the number and location of the genes, the distribution of their effects and dominance, their interaction and the effect of the environment on lowered fitness in inbred individuals. One example of the impact of the environment on fitness is that the viability of juveniles in captivity is lower than in the wild (Hedrick & Kalinowski, 2000). Another example is a study by Jiminez *et al.* (1994) where inbred and non inbred white-footed mice (*Peromyscus leucopus noveboracensis*) were released into a natural study site and the rate of survival was determined by capture-recapture methods. The results showed that inbred mice had a lower survival than non inbred mice indicating decreased fitness in inbred mice (Hedrick & Kalinowski, 2000).

In order to genetically restore populations with reduced fitness, further genetic deterioration should be avoided. It has been proposed that maintaining heterozygosity within a population could be one method to ensure the probability of a species' survival in terms of its variability over time. This is usually achieved by controlled breeding programs that avoid intense inbreeding by maintaining a closed population of a certain minimum size or by exchanging breeding individuals. In a less intensive situation an adequate breeding population should be maintained to reduce the effects of inbreeding (Avise, 1994). This can be done in a number of ways, firstly by the introduction of unrelated individuals from other populations (Hedrick & Kalinowski, 2000) or by creating corridors between genetically isolated populations. In some populations, inbreeding is difficult to avoid especially if the population consists of a very limited number of individuals but through carefully controlled breeding it may still be possible to purge inbreeding depression. An example of this is the study by Templeton and Read (1983) on Speke's gazelle; they successfully managed to produce progeny from only three females and one male using breeding strategies. This shows that it is possible to increase the fitness and reduce the effect of inbreeding of a population by proper management and breeding strategies either by manipulating breeding or by translocating individuals. Another method to determine inbreeding depression and then genetically restore a population is to use a molecular-marker based approach (Glémin *et al.*, 2006). In a study by Glémin *et al.* (2006), on *Brassica insularis,* microsatellites were used to determine inbreeding depression and to test for the occurrence of epistasis.

DNA markers in conservation genetics

There are many different DNA markers at our disposal today with their own advantages, disadvantages and areas of application. The different markers include restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), minisatellites, microsatellites and single nucleotide polymorphisms (SNP). There are different categories and classifications of molecular markers based on the techniques used and the characteristics of the markers, these categories are described by O'Brien (1991) and

Karp and Edwards (1998). A brief summary including the main attributes of some of these markers is given in the review by Sunnucks (2000).

There are a number of aspects to be taken into consideration when selecting the appropriate DNA marker, these include the sensitivity of the marker, whether there is a need for a single-locus or multi-locus marker (multilocus markers can visualise many genes simultaneously as opposed to single locus markers), the ability of a marker to yield gene genealogies and frequencies, the type of DNA being used (e.g. mitochondrial DNA vs nuclear DNA) and the efficiency of development and use of the marker in the specific field of study. For example, molecular markers are useful for determining identity and sex. Markers for the *SRY* gene on the Y chromosome can be used to distinguish male from female DNA (Kohn & Wayne, 1997) and other methods include using the amelogenin gene which exists on both X and Y chromosomes and the ZFX and ZFY genes (Kurose *et al.*, 2005). The amelogenin maker was originally used for sex determination in cattle and humans and then later in many other species including sheep, red deer and Bengal tigers (Sullivan *et al*., 1993, Ennis & Gallagher, 1994, Pfeiffer & Brenig, 2005, Bagavatula & Singh, 2006)

Microsatellites are commonly used molecular markers in conservation genetics because of their high levels of variability, ease and reliability of scoring, codominant inheritance and short lengths, these factors make them useful in studies involving DNA from sources such as bones, faeces and hair (Luikart & England, 1999). Microsatellites are composed of multiple copies of tandemly arranged simple sequence repeats (SSRs) that range in size from two to six base pairs but are generally less than 5bp long (Tautz, 1989, Liu & Cordes, 2004) consisting of di-, tri- or tetra-nucleotide repeats. The repeat sequences are classified into three families namely pure, compound and interrupted repeats (Jarne & Lagoda, 1996). Microsatellites are usually evenly distributed throughout a genome on all chromosomes and regions of the chromosome in eukaryotic species (Liu & Cordes, 2004). Microsatellite polymorphism is based on size differences due to varying numbers of repeat units contained by alleles at a given locus (Liu & Cordes, 2004) that are smaller than 1kb and vary over a narrow size range (Bruford & Wayne, 1993).

In cheetah, microsatellites were shown to be useful for determining genetic diversity and parentage. Driscoll *et al.* (2002) assessed the patterns of genomic diversity for ninety feline microsatellite markers in certain populations of cheetah, lion and puma. They found that it was possible to use microsatellites to interpret historic population bottlenecks. Their results were supported by previous conclusions regarding the cheetah's population bottleneck in the late Pleistocene age. Microsatellites can therefore be used to detect and to estimate historic population bottlenecks. Depending on polymorphism and the level of heterozygosity of microsatellite markers, 20 microsatellite markers are recommended to determine paternity and kinship and to distinguish relatives from non-relatives (Kohn & Wayne, 1997). The analysis of faecal DNA with the aid of

microsatellites has also been used in cheetah research by Gottelli *et al.* (2007). The aim of the study was to determine whether multiple paternity occurred in cheetah. In order to obtain DNA from the cubs, faecal sampling was done directly after the cheetah had moved away from the defecation site therefore all samples were considered fresh.

Scatology

Molecular scatology is the study of faeces using molecular approaches. In the past, a variety of sources of DNA such as hair, bones, feathers (in birds), saliva, skin and nails have been collected for non-invasive genetic analysis but these samples were often difficult to obtain (Kohn & Wayne, 1997). Molecular scatology is a non-invasive technique for obtaining information on wild mammal populations (Kohn & Wayne, 1997). This technique is of special interest when studying rare, nocturnal or elusive species because there is no need to capture, harm or cause the animal any unnecessary stress. Faeces contain gut mucosal cells from the intestinal lining therefore DNA can be isolated and analysed. Once the DNA is purified from the faeces, specific DNA sequences are amplified using PCR (Saiki *et al.,* 1985).

In conservation genetics there is a strong trend toward non-invasive genetic sampling methods. Due to the complexity of the method, various researchers recommend testing the sampling method to ensure that it is compatible with the scientific goal before extensive sampling is conducted (Taberlet & Luikart,1999). It is important to determine that there will be enough amplifiable DNA for the required genetic analyses. A pilot study is necessary to determine and avoid any errors that may be associated with the non-invasive sampling and individual identification.

A number of studies using scatology with molecular markers have been done on various species such as mountain lions (*Puma concolor*) (Ernest *et al*., 2000), wolves (*Canis lupus*) (Creel *et al.*, 2003, Lucchini *et al.*, 2002), jaguars (*Panthera onca*)(Farrell *et al.*, 2000; Farrell, 2001), brown bears (*Ursus arctos*) (Kohn & Wayne, 1997), otters (*Lutra lutra*) (Dallas *et al.*, 2003), American mink (*Mustela vison*), polecat (*Mustela putorius*) (Hansen & Jacobsen, 1999), Taiwanese Macaques (*Macaca cyclopis*) (Chu *et al.*, 2006), Eurasian badgers (*Meles meles*) (Frantz *et al.*, 2003), sympatric carnivores on the Tsushima Islands (Kurose *et al.*, 2005), red wolves (*Canis rufis*), coyotes (*Canis latrans*) (Adams *et al.*, 2003), Barbary Macaques (*Macaca sylvanus*) (Lathuilliere *et al.*, 2001) and seals (Reed *et al.,* 1997). These studies, on a variety of species, concluded that faecal DNA analysis as a non-invasive technique was an effective method for detecting and identifying a species of interest and for use in population genetics studies. This information will help to establish management strategies in order to preserve these animals in an ever-decreasing habitat due to the rapidly increasing human population. The mitochondrial DNA (mtDNA) extracted from the faecal samples can be used for species and individual identification with the use of molecular markers (Farrell, 2001). Population size can also be estimated using

molecular scatology. Faecal samples are collected, typed for a number of microsatellite markers and genotypes counted to estimate the minimum population size (Creel *et al.*, 2003, Frantz *et al.*, 2003).

Dietary habits can be determined non-invasively through the analysis of scats. Although not all prey is always detected, the technique can certainly be used to give an indication of the preferred prey of a species or of a specific population (Höss *et al.*, 1992). A study by Farrell (2001) used scats to determine the preying habits of the puma (*Puma concolor*) and the jaguar (*Panthera onca*). Similar to the cheetah, these large felids were being persecuted by cattle ranchers in Venezuela for alleged predation on livestock. By determining the prey and managing it properly, the attacks on the livestock could be reduced, consequently reducing the persecution on the puma and the jaguar (Farrell *et al.*, 2000). Disease and parasite infestation can be studied by conventional analyses of faeces but the use of molecular techniques can increase the resolution of the analysis by allowing more precise identification of food or pathogen species (Kohn & Wayne, 1997).

Molecular scatology has several limitations that need to be taken into account when planning a study that uses this technique. Extracting DNA of satisfactory integrity is a major difficulty for this technique and there is a large possibility for PCRs to fail due to the presence of degraded DNA. The Silica extraction method is a popular method that has proven to be relatively successful (Lathuilliere *et al.*, 2001). The commercial Qiagen® QIAamp stool mini kit (Southern Cross Biotechnology (Pty) Ltd, South Africa) which is based on the silica method, was also designed for DNA extraction from faeces (Taberlet & Luikart, 1999). An example demonstrating the difficulty involved in extracting pure DNA of good quality and quantity was seen in initial studies on herbivore faeces. It was found that herbivore faeces contain, in addition to microorganisms, undigested food, digestive enzymes, mucus, bile salts and bilirubin and plant polysaccharides that inhibit the *Taq* polymerase during PCR. A solution to this problem has been found, Hexadecyltrimethyl-ammonium bromide (CTAB) is used to remove the plant polysaccharides for DNA extracts enabling the *Taq* polymerase to function normally during PCR (Kohn & Wayne, 1997).

During DNA extraction and PCR preparation, contamination is a potential problem. Separate facilities for DNA extractions and PCR set-up should be used. Blank extractions and PCR blanks should be used at all stages of the trial (Kohn & Wayne, 1997). Taberlet and Luikart (1999) also recommended avoiding having concentrated DNA extracts in the extraction room, using dedicated pipettes with aerosol-resistant pipette tips and continuously monitoring reagents for DNA contamination.

A low DNA quality and quantity can result in errors such as allelic dropout (one allele of a heterozygous individual is not amplified during a positive PCR) and false alleles (the presence of spurious bands) (Lathuilliere *et al.*, 2001, Frantz *et al.*, 2003). Genotyping errors such as these are especially associated with the use of hair or faeces (Taberlet & Luikart, 1999). By amplifying shorter DNA segments more consistent amplifications are obtained compared to the amplification of longer DNA segments. Multiple extractions and PCR amplifications from each faecal sample may also reduce incorrect genotyping (Kohn & Wayne, 1997).

According to Lathuilliere *et al.* (2001), allelic dropout is common in microsatellites and there are a number of factors that may cause it. These factors include DNA degradation, preferential amplification of the shorter allele when DNA polymerase is limited and fluctuation in the number of copies of each allele when there is a limited amount of DNA available. Another possible factor is the GC content because templates that have a GC content higher than 70% can be difficult to replicate during PCR. This could be related to the higher melting temperature of the DNA caused by the higher proportion of GC base pairs (Applied Biosystems, 2000). A possible reason for the presence of false alleles is the occurrence of slippage (strand mispairing). Microsatellites evolve over time and the mutation rate is affected by the length of the microsatellite. Longer microsatellites mutate more often than shorter ones and the occurrence of point mutations cause slippage. Several mutation models are referred to in the literature for estimation of microsatellite evolution (Sainudiin *et al.,* 2004).

Genotyping errors have the potential of creating several genotypes per individual which can cause an overestimation when trying to determine population size. It can affect the ability to differentiate between individuals, alter allele frequencies, overestimate inbreeding and affect parentage analysis (Soulsbury *et al*., 2007). It has been suggested that overestimation can be reduced by decreasing the number of loci included in the genotype thus reducing the opportunity for errors to create false genotypes (Miller *et al.*, 2002; Creel *et al.*, 2003) but when conducting paternity and kinship studies a larger number of loci are recommended (Kohn & Wayne, 1997).

A multiple-tubes approach was proposed by Taberlet *et al.* (1996) in order to overcome the problem of obtaining erroneous genotypes. Although the approach is successful, it is expensive and time consuming (Taberlet *et al.*, 1996). Another approach is the use of a maximum-likelihood method (Miller *et al*., 2002) to assess genotype reliability and to strategically minimize the number of PCR replicates used. This model is based on the assumptions that: a) both alleles at a heterozygous locus are equally likely to drop out, b) allelic dropout rates are even across loci, and c) all false alleles can be detected and eliminated from the data set. It is important that these assumptions are met in the data set under investigation for the model to be accurate (Frantz *et al.*, 2003). An alternative quantitative polymerase chain reaction (qPCR) method can be used

because it modifies the number of replicates of target DNA during PCR in response to information about template concentration and dropout rates (Morin *et al.*, 2001, Buchan *et al.*, 2005).

Another potential limitation of scatology is the difficulty involved in collecting the samples. Firstly, it can be very difficult to identify the faeces of a specific species in areas that have sympatric carnivores with similar body sizes and feeding habits (Paxinos, 1997) and secondly, some environments do not allow for the easy detection of the faeces. A solution to this potential problem is to use species-specific primer sets that are designed to solely amplify DNA of a target species (Kohn & Wayne, 1997). Another option to correctly identify faeces of a target species is to use dogs that are trained to find scats of a specific species. In a study by Smith *et al.* (2003) detection and accuracy rates of dogs trained to find scats of San Joaquin kit foxes (*Vulpes macrotis mutica*) were determined. The results showed that the accuracy of the dogs was 100% for detecting kit fox scats. Most of the dogs used were also highly accurate in choosing kit fox scat when red fox scats were also present. The dogs were also successful in finding up to 5.37 assumed kit fox scats per kilometre. This is a significantly higher detection success rate than humans. A Staffordshire terrier is currently being trained for the detection of cheetah scats in the Thabazimbi area of the Limpopo Province (South Africa) where the De Wildt Wild Cheetah Project is conducting research. Already the dog is able to recognize cheetah scat in the presence of other carnivore faeces (Marnewick, 2007, personal communication, kelly@dewildt.org.za). Although the dogs are known to be rather accurate, DNA confirmation of the target species with the use of species-specific primer sets is recommended to eliminate the possibility of error.

Cheetah conservation in South Africa

There are a number of non-governmental centres in South Africa dedicated to cheetah conservation; such centres include the De Wildt Cheetah and Wildlife Centre, Cheetah Outreach and Hoedspruit Cheetah Centre. The De Wildt Cheetah and Wildlife Centre is dedicated to ensuring the breeding and long term survival of cheetah. They have implemented many projects to ensure the success of their mission including an education project. The De Wildt education team aims to reach children and families that might come into contact with cheetah on the farms where they live or work (Bouwer, 2005, personal communication, cheetah@dewildt.org.za).

The Wild Cheetah Project is also one of De Wildt Cheetah Centre's conservation efforts. This team travels to farms in the Limpopo Province teaching the farmers in the area how to manage cheetah on their farms. The team is involved in farmers' conferences and makes use of any other occasions to teach the farmers and farm workers more about the cheetah and cheetah conservation. They are also involved in the capture of certain cheetah which may be in danger or are 'problem cheetah', that are subsequently relocated to a suitable protected area. Many farmers are under the impression that cheetah are responsible for most of their stock losses. Research to

date has concluded that although cheetah sometimes prey on farmers' livestock, stock losses are often due to predation by caracal and jackal. Contrary to popular belief, cheetah can in some situations curb stock losses by caracal and jackal (Cilliers, 2005, personal communication, deon@dewildt.org.za). This was shown at the Rietvlei Nature Reserve near Pretoria, where it was observed that every year the impala numbers were decreasing due to predation on the lambs and older animals by the jackal and caracal. After the release of two cheetah from the De Wildt Cheetah centre the impala numbers slowly started increasing. The reason for this increase is that cheetah prey on adult animals, but only eat certain parts of the carcass. Cheetah eat the buttocks, thighs, heart, liver and lungs – the soft meats. The remains of the carcass are then eaten by the scavengers who would rather conserve energy by scavenging, than hunting for themselves (Cilliers, 2005, personal communication, deon@dewildt.org.za).

The De Wildt Cheetah and Wildlife Centre is currently implementing a national 'metapopulation' management plan for the cheetah in order to ensure that there is gene flow between the subpopulations that exist in protected areas. Genetic management and census of cheetah subpopulations is of great importance as many of the game reserves in South Africa that have cheetah are relatively small and can therefore sustain only a limited number of these animals.

The De Wildt Wild Cheetah Project is involved in the relocation of 'problem' cheetah from farm lands to conservation areas. To date, they have received 137 'problem' cheetah from farming areas, 53 of which were female and 84 male. Sixteen of these animals had to be placed in captivity due to serious injuries such as broken jaws or an amputated limb, making them un-releasable. These injuries are often caused during the capturing of the cheetah. People have been known to set their hunting dogs on the cheetah, catching them in snares or harming them intentionally in other ways. Although certain animals cannot be released back into the wild, to date 107 have been released into conservation areas and 64 have survived and adapted very well. The relocations have been successful with relocated cheetah producing 51 cubs (Marnewick *et al.*, 2007). The Metapopulation Management Project aims to develop a successful plan to manage these populations effectively to ensure that the genetic integrity of the sub-populations is not compromised.

A study by Marnewick and Cilliers (2006), members of the Wild Cheetah Project team, was conducted to determine the range use of male cheetah coalitions in the Thabazimbi district of the Limpopo Province in South Africa. Two coalitions, consisting of two and three individuals respectively, were caught, collared (satellite/cellular/radio telemetry collars) and monitored. The study concluded that the estimated range of the coalition of two cheetah was 310km^2 and 190km^2 for the coalition of three cheetah. This differs greatly from the ranges of up to 1608.4 km^2 that

were reported in Namibia (Marker, 2002). Habitat and prey have a seemingly large influence on the range use of cheetah.

A cheetah census is being conducted in the cheetah friendly areas (Limpopo Province) by using remote-triggered cameras that have been placed on certain farms to obtain a more accurate estimate as to how many cheetah are in fact still free-roaming (Marnewick & Cilliers, 2006). Cheetah are known to use specific scratching trees – the cameras are set up at these trees and record footage of the animal, providing valuable information such as the sex, number of animals in the coalition (group of males) and estimated age of the animals in the area (Figure 1.2) (Marnewick & Cilliers, 2006).

Cheetah are extremely difficult to census due to their elusive nature and as a result there is only a limited amount of information regarding the status of cheetah populations across Africa. With limited data regarding the number, distribution and genetic structure of cheetah, it is not possible to identify and address potential threats and to plan effective conservation strategies. Rarity makes the direct observation of the animal difficult but the need for genetic, demographic and life-history information for species facing extinction is more vital than ever (Kohn & Wayne, 1997).

Figure 1.2 A free-roaming cheetah photographed at a camera trap site that was set up by the De Wildt Wild Cheetah Project on a farm in the Thabazimbi area (Marnewick, De Wildt Cheetah Centre)

Aim of this study

The ideal for genetic management of cheetah would be to obtain DNA from as many cheetah as possible. Several factors complicate the sampling of cheetah whether they are captive, relocated or wild. Blood samples can be obtained from captive animals with some effort, but obtaining samples for genetic analyses from wild animals is extremely difficult. In the case of blood

sampling, not only is this impractical, it is also very stressful for the animal. A non-invasive method is therefore vital for the study of species, including cheetah, that are elusive, wide ranging or highly endangered because it can be implemented without having to capture or harm the animal or place it under unnecessary stress (Creel *et al.*, 2003). Comprehensive sampling can allow estimates of home range, reproductive patterns, kinship structure and population size (Kohn & Wayne, 1997).

The De Wildt Cheetah and Wildlife Centre requested that a project be implemented to investigate the use of scatology as a conservation tool for cheetah in South Africa. The aim of the present study was thus to determine if scatology could be used as a non-invasive method for genetic analysis of cheetah in South Africa. The feasibility and accuracy of the technique was tested by genotyping individuals at certain microsatellite markers and determining the degree of error associated with the technique using a control group consisting of blood samples from the same individuals. A preliminary field application of the technique, given the likely error obtained in the control group, was also tested by analysing scats from the Thabazimbi area.

Chapter 2

Material and Methods

2.1 Origin of samples

The cheetah individuals that formed part of this pilot study are from the De Wildt Cheetah Centre in the foothills of the Magaliesberg near Hartebeespoort dam in South Africa's North West Province (GPS coordinates: S 25 40 25.1 E 27 55 25.4). Twenty cheetah were selected that were kept in separate enclosures, as it was essential to collect faeces and blood of the same animal. In three situations, two cheetah were kept in the same enclosure. To avoid confusion, the two animals from each camp were separated for the duration of the sample collection. The diet of the cheetah selected for the trial consisted of either horse meat or chicken. Some cheetah at the De Wildt Cheetah Centre are also fed cat pellets but were not chosen based on the fact that the methods developed in this study will eventually be applied to wild cheetah that have a diet consisting of meat.

2.1.1 Faecal sample collections from the De Wildt Cheetah Centre

The camps of the 20 chosen animals were cleaned one day before the faecal collections were to be done. The sizes of the study animals' camps are relatively large, ranging in size from 24m^2 to 2000m^2 (Table 2.1). This would have made the collection of the faecal samples difficult in most instances, therefore the cheetah were lured into their respective corner camps with a piece of meat. As can be seen in Figure 2.1, the corner camps are a smaller size (approximately 10m^2) making the task of finding the faeces easier. There are corner camps in all the enclosures at De Wildt to simplify tasks such as medicating, vaccinating or observing sick animals.

Faecal samples were collected in the early morning in order to ensure that the samples were still fresh. Samples were obtained from 19 of the 20 trial animals selected for the study. Attempts to obtain a faecal sample from the remaining individual for three days were unsuccessful and the animal was excluded from the trial. Three faecal samples were collected for each of the trial animals. One sample was immediately placed in a 50ml screw cap plastic storage container and filled with 99% ethanol for preservation. The two remaining samples from each animal were then laid out on a concreted area where they were exposed to the weather (Figure 2.2). This method of aging the samples was applied because it will not always be possible to collect fresh cheetah faeces in the wild.

Table 2.1 De Wildt Cheetah Centre camp sizes

Camp Type	Size (m^2)
Breeding camp	2000
Hospital camp	24-30
Retirement camp	600
Ambassador camp	750-1000

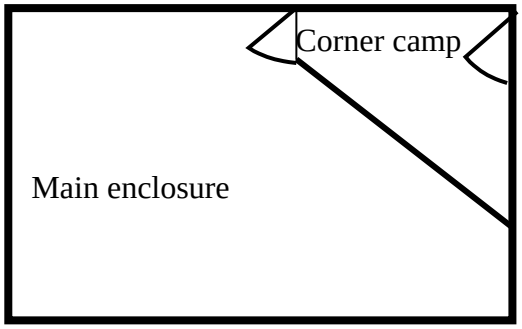


Figure 2.1 A schematic representation of the typical cheetah enclosures found at the De Wildt Cheetah Centre indicating the division of the camp into a corner camp area and a main enclosure

Figure 2.2 A photograph of cheetah faeces left on a concrete slab to age for four days to two weeks

The remaining two samples were collected after a four day period and a two week period respectively. These samples were stored in a screw cap plastic storage container with 99% ethanol and kept at room temperature. Once the initial DNA extractions were completed, all faecal samples were stored in the refrigerator at 4ºC.

Faeces was re-sampled from the camps of nine cheetah individuals that formed part of the pilot study. The samples were not stored in 99% ethanol, but were kept cool and transported to the laboratory where DNA isolations using the Qiagen mini stool kit (Southern Cross Biotechnology (Pty) Ltd, South Africa) were done on the same day. The remaining portion of the faecal sample was placed outside using the same procedure as described above and DNA isolations were done

at four days of age and again at two weeks of age. The aim was to see if the DNA quality increased if extractions were performed without storage in ethanol.

2.1.2 Blood sample collections

Blood samples from 11 of the 19 animals were obtained from the wBRC (Wildlife Biological Resource Centre) Biobank (National Zoological Gardens of South Africa, Pretoria), a national storage facility for wildlife biological samples. Samples included whole blood samples or white blood cells (buffy coat) (Table 2.2). Seven blood samples were collected from the relevant cheetah at the De Wildt Cheetah Centre during routine sampling by the veterinarian, Dr Peter Caldwell. One sample was not collected as the animal died prior to the blood collection procedure therefore the final sample size of the pilot study was 18 captive cheetah.

Table 2.2 Origin and type of blood samples from the cheetah individuals used in the pilot study

Sample number	Origin	Camp type	Type
1	Biobank	Hospital	White blood cells
2	Biobank	Breeding	White blood cells
3	Biobank	Breeding	White blood cells
4	Biobank	Breeding	White blood cells
5	Biobank	Breeding	White blood cells
6	Biobank	Breeding	White blood cells
7	De Wildt	Ambassador	Whole blood
8	De Wildt	Ambassador	Whole blood
9	De Wildt	Retirement	Whole blood
10	Biobank	Ambassador	Whole blood
11	Biobank	Ambassador	Whole blood
12	Biobank	Breeding	Whole blood
13	Biobank	Breeding	White blood cells
14	De Wildt	Hospital	Whole blood
15	Biobank	Breeding	White blood cells
16	De Wildt	Ambassador	Whole blood
17	De Wildt	Ambassador	Whole blood
18	De Wildt	Ambassador	Whole blood

The blood samples were collected in 4.5ml EDTA tubes from the vein in the hind leg of the animal. The task was easier to perform on the cheetah that are part of the De Wildt Cheetah Ambassador program as they are well trained (Figure 2.3a/b). Blood collection can be stressful for the animal especially for those that are not trained. Every measure was taken to keep the individual calm while making the sample collection as fast as possible (Figure 2.3 c). The tubes of blood were kept on ice until they were transferred into 1.5ml eppendorf tubes and stored at -20˚C (Animal Breeding and Genetics laboratory, Department of Animal and Wildlife Sciences, University of Pretoria).

(a) b) (c)

Figure 2.3 'Graca' (De Wildt Cheetah Ambassador) has been trained to lie down when the veterinarian needs to perform procedures such as drawing blood.

2.1.3 Collection of faecal samples from the Thabazimbi area

The De Wildt Wild Cheetah Project has a number of camera traps set up on farms and conservation areas surrounding Thabazimbi, a town in the Limpopo province. These cameras have proved to be a useful tool in recording animals such as cheetah, brown hyena and leopard that frequent the road or tree at which the camera is placed.

During two routine trips to the Thabazimbi district to service the cameras, scats that resembled those of cheetah were collected at the camera trap sites and stored in a screw cap plastic storage container with 99% ethanol. The faeces was collected on various farms around Thabazimbi (Table 2.3). Figure 2.4a shows faeces that was found near a 'play tree' where a camera trap had been placed. The faeces seen in Figure 2.4b was found next to a concrete block, blocks such as the one seen in the photograph are often used by cheetah as vantage points when hunting. The screw cap plastic storage containers were then labelled according to the GPS location where it was found (Table 2.3).

Table 2.3 Collection locations of presumed cheetah faeces in the Thabazimbi district.

Sample number	GPS coordinates	Farm name
1	S24˚33 27.7 E26˚39 51.9	Silent valley
2	S24˚33 27.7 E26˚39 51.9	Silent valley
3	S24˚37 23.6 E26˚31 49.6	Koedoes laagte
4	S24˚25 37.9 E26˚42 3.9	Atherstone (possibly hyena)
5	S24˚25 37.9 E26˚42 3.9	Atherstone
6	S24˚24 22.1 E26˚38 49.3	Jakkals kraal
7	S24˚24 22.1 E26˚38 49.3	Jakkals kraal (possibly monkey)
8	S24˚37 23.6 E26˚31 49.6	Koedoes laagte (fresh)
9	S24˚37 23.6 E26˚31 49.6	Koedoes laagte
10	S24˚24 22.1 E26˚38 49.3	Jakkals kraal
11	S24˚37 23.6 E26˚31 49.6	Koedoes laagte
12	S24˚24 22.1 E26˚38 49.3	Jakkals kraal
13	S24˚25 37.9 E26˚42 3.9	Atherstone
14	S24˚25 37.9 E26˚42 3.9	Atherstone
15	S24˚24 22.1 E26˚38 49.3	Jakkals kraal

(a) (b)

Figure 2.4 Faeces of suspected cheetah origin found and collected in the Atherstone Nature Reserve in the Thabazimbi district. a) Faeces found close to a cheetah 'play tree'. b) Faeces found next to a concrete block that cheetah use as a vantage point.

2.1.4 Genotypic data obtained from the NZG

Genotypic data of 45 De Wildt captive cheetah samples was obtained from the NZG (National Zoological Gardens), Pretoria. The samples had been genotyped at the 12 microsatellite loci that were used throughout this study. The aim of using this data was to see if the genotypes obtained in the scatology pilot study are comparable with the larger set of NZG samples.

2.2 Methods

2.2.1 Genomic DNA extraction from faecal and blood samples

The Qiagen® QIAamp stool mini kit (Southern Cross Biotechnology (Pty) Ltd, South Africa) was used for the faecal DNA extractions by following the manufacturers' protocol for the extraction from human faeces. DNA was isolated from blood using the Qiagen® tissue kit (Southern Cross Biotechnology (Pty) Ltd, South Africa) following the manufacturers' protocol.

Success of DNA extractions were determined by running the extracted DNA through 1% agarose gels and visualising the DNA products by ethidium bromide staining under UV light. The DNA was then quantified with the use of a Nanodrop ND-1000 UV-vis Spectrophotometer (http://www.nanodrop .com/). The ratio of absorptions at 260nm vs. 280nm is commonly used to assess the purity of DNA with respect to protein contamination, since protein tends to absorb at 280nm. A 260/280 ratio of higher than 1.7 is indicative of pure DNA whereas a low value indicates protein contamination. The ratio of absorptions at 260nm vs. 230nm is used to determine purity of DNA with regard to organic compound contamination, DNA is considered pure when the ratio is above 2, a lower value indicates contamination (Yeates *et al.*, 1998).

2.2.2 Molecular marker optimization and PCR amplification

2.2.2.1 Species-specific mitochondrial DNA markers

The diet of the captive animals in this study included chicken and horse meat thus necessitating confirmation that the DNA extracted originated from the cheetah. Furthermore, the faeces collected in the Thabazimbi area from potential wild cheetah could not be classified as originating from cheetah solely on appearance as there are a number of carnivores of similar size and diet residing in the area. A species-specific primer set was therefore used to confirm that each DNA sample was from a cheetah.

The species-specific primer set was designed for the cheetah using the mtDNA control region of cheetah (GenBank accession number NC005212, Burger *et al.*, 2004), chicken (GenBank accession number NC001323, Valverde *et al.*, 1994), horse (GenBank accession number NC001640, Xu & Arnason, 1994) and human (GenBank accession number NC001307, Ingman *et al.*, 2000). These species were chosen as chicken and horse formed part of the diet of the study animals and the human sequence was used to rule out the potential for contamination during collection and DNA extraction. These sequences were obtained from the GenBank database (http://www.ncbi. nlm.nih.gov). The sequences were aligned using Bioedit 7.0.9.0 (Hall, 1999), a biological sequence alignment editor (Addendum 1). The primer set (Addendum 1, Table 2.4) was designed by identifying areas on the mitochondrial DNA that were most unique to the cheetah and were as short as possible in order to amplify successfully when the DNA was intact or if it was degraded slightly. The potential primer set sequences were checked on an oligoanalyzer (OligoAnalyzer 3.0) to determine the GC content, length, annealing temperature, and molecular weight. When designing a primer set it is important that the primers are not complementary, there should be no 3' T, Tm temperatures of the primers should be almost the same (Tm (melting temp) = 2(A+T)+4(G+C)) and there should be a balanced distribution of G/C and A/T rich domains. A "G" or "C" is desirable at the 3' end of primers but Gs or Cs should not be added if they adversely influence the overall specificity of the primer set. When designing primer sets, it is important to have a minimum of intramolecular or intermolecular homology. This would result in either hairpins or primers binding to one another.

Once the primer set was correctly designed, The Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/) was used to find regions of local similarity between the primers and sequences on the NCBI library. The aim was to identify the primer region as coming from cheetah and to ensure that they could be distinguished from other felids. A successfully designed primer set would result in BLAST identifying that the most similar sequences to the primer set are of cheetah. The terminology used to describe the output was the 'expect' value, a 'score' and 'gaps'. The 'expect' value is the expectation value, this is the number of different alignments with scores that are equal to or higher than the scores that are expected to occur in the database search by chance. A lower 'expect' value indicates a significant score. The score for an alignment is calculated by adding the scores for each aligned position and the scores for gaps. Gaps are spaces introduced into an alignment to compensate for insertions and deletions in the sequences being aligned.

Table 2.4 Cheetah specific primer set sequences designed using the mtDNA control region of cheetah for the use of confirming that a specific faecal DNA sample was of cheetah origin.

Sequence name	**Bases (5'-3')**
Forward Primer	AAC TTG TTA CGG GAC AAT TC
Reverse Primer	GGA TTG GAT TCA AGG TAT TC

Initially PCR's were performed using the species-specific primer set to amplify DNA isolated from cheetah blood. Negative controls were included in each PCR. Once the amplification of blood DNA using the species-specific primer set was optimised, the DNA isolated from the cheetah faeces collected at De Wildt and Thabazimbi was used. The PCR's were performed using 5µl of DNA template. The concentration of the DNA was not taken into account due to the possibility that the concentration may be attributed to cheetah as well as prey DNA and bacteria, and the concentrations obtained from the extraction were very low in most instances. The final PCR reaction contained 10 ρmol of each primer, 300µM dNTP, 3 mM $MgCl_2$, and 1 unit of *Taq* DNA Polymerase (GoTaq® Flexi DNA Polymerase, Promega, Madison, WI, USA). The reactions were carried out in 15µl reactions and cycled as follows: 94˚C for 5 min followed by 25 cycles of 94˚C for 30 seconds, 56˚C for 45 seconds and 72˚C for 30 seconds and then a final extension step of 7 min at 72˚C. A reliable PCR product was obtained after 25 cycles.

PCR products were electrophoresed through 3% agarose gels and viewed by ethidium bromide staining under UV light. To confirm the successful amplification of DNA originating from cheetah, two random PCR products were chosen to be sequenced at the University of Pretoria's DNA Sequencing Facility of the Faculty of Natural and Agricultural Sciences.

To prepare the PCR product for sequencing, an ethanol/sodium acetate precipitation was done (Applied Biosystems, 2000). A volume of 12µl sterile water was added to the product and incubated for 10 minutes to return the precipitated DNA product to a 12µl volume. To prepare the cycle sequencing reaction, 3µl of the precipitated DNA product was added to two separate micro centrifuge tubes. Primers were added separately to each tube at a concentration of 3.2 ρmol and 1µl of Big Dye® v.3.1 Terminator Kit was added to the tubes. For cycle sequencing, the protocol of Dye Terminator Chemistries for Geneamp 3700, 3900 or 2400 was followed (Applied Biosystems, 2000), the thermocycler was set for a volume of 5µl. To remove any unincorporated dye terminators, the ethanol/sodium acetate precipitation was repeated.

Sequence analysis was done using the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). The 3100 separates the mixture of DNA fragments according to their lengths, provides a profile of separation and then determines the order of the four deoxyribonucleotide bases. Sequences were viewed using the program Chromas® 2.32 (Technelysium Pty Ltd, http://www.technelysium.com.au/chromas.html). The Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/) was used to ensure that the primer set had accurately amplified cheetah DNA.

An attempt was made to amplify DNA from blood samples of chicken, horse and human using the species-specific primer set. This was done as a control to exclude the possibility of the primer set amplifying species other than cheetah. It is important to note that this species-specific primer set

was mainly designed for the De Wildt pilot study group. For wild cheetah, a primer design would need to take other closely related species that occur in the area into account.

2.2.2.2 Microsatellite markers

Twelve microsatellites (Table 2.5) were chosen by the Genetics Committee of the South African Cheetah Metapopulation Management Project under the De Wildt Wild Cheetah Project for use as a standard when conducting genetic analyses on cheetah so that resulting data are comparable between laboratories. The 12 selected microsatellites included those accepted and recommended by the ISAG (International Standard for Animal Genetics) for DNA profiling. The suitability of the microsatellites was based on the fact that these loci have high polymorphic information content (PIC) values, are distributed on chromosomes throughout the genome, are internationally comparable and also comparable to two other laboratories in South Africa conducting research on cheetah. The oligonucleotide primers were synthesised and labelled with an appropriate fluorescent label (Matrix Standard Set DS-33: 6FAM, VIC, NED, PET™, LIZ™) at the 5' end of each forward primer by Applied Biosystems®. The specific dye used for each microsatellite is shown in Table 2.5. The primer sequences for the chosen microsatellite markers can be seen in Table 2.6.

Table 2.5 The chromosomal location (Menotti-Raymond *et al.*, 1999), PIC values (Driscoll *et al.*, 2002), $MgCl_2$ concentration for successful PCR and fluorescent labelling dye for each microsatellite locus typed. Group 1 and Group 2 indicate the grouping of microsatellites for fragment analysis according to range and fluorescent label.

Marker	**Chromosome**	**PIC value**	**$MgCl_2$ concentration (mM)**	**Fluorescent label** (Matrix Standard Set DS-33: 6FAM, VIC, NED, PET™, LIZ™)
FCA 8	A1	0.698	1	6FAM (Group 1)
FCA 26	D3	0.749	1	NED (Group 1)
FCA 69	B4	0.662	2	PET (Group 1)
FCA 78	D2	0.788	1	6FAM (Group 2)
FCA 96	A2	0.856	1	VIC (Group 1)
FCA 105	A2	0.76	2	NED (Group 2)
FCA 126	B1	0.698	1	6FAM (Group 2)
FCA 133	B2	0.781	2	PET (Group 2)
FCA 220	F2		2	PET (Group 2)
FCA 224	A3	0.763	2	PET (Group 1)
FCA 247	C1	0.786	1	VIC (Group 1)
FCA 310	C2	0.624	1	VIC (Group 2)

Table 2.6 Reverse and forward primer set sequences for each microsatellite locus typed (Menotti-Raymond *et al.*, 1999)

Marker	Forward primer (5' - 3')	Reverse primer(5' - 3')
FCA 8	ACTGTAAATTTCTGAGCTGGCC	TGACAGACTGTTCTGGGTATGG
FCA 26	GGAGCCCTTAGAGTCATGCA	TGTACACGCACCAAAAACAA
FCA 69	AATCACTCATGCACGAATGC	AATTTAACGTTAGGCTTTTTGCC
FCA 78	TGAACTGAAGTCAGATGCTTAACC	CGGAATCAGCTATTTTTACGG
FCA 96	CACGCCAAACTCTATGCTGA	CAATGTGCCGTCCAAGAAC
FCA 105	TTGACCCTCATACCTTCTTTGG	TGGGAGAATAAATTTGCAAAGC
FCA 126	GCCCCTGATACCCTGAATG	CTATCCTTGCTGGCTGAAGG
FCA 133	ATGAGGACCGATGACTCTTTT	GGAAATGTTGACAGAGGGACA
FCA 220	CGATGGAAATTGTATCCATGG	GAATGAAGGCAGTCACAAACTG
FCA 224	CTGGGTGCTGACAGCATAGA	TGCCAGAGTTGTATGAAAGGG
FCA 247	GGAAATTAGGAGCTCTGCCA	AAGATTTACCCAGTTGCCCC
FCA 310	TTAATTGTATCCCAAGTGGTCA	TAATGCTGCAATGTAGGGCA

The PCR reactions were performed using 5µl of DNA template. Negative controls were included in each PCR reaction to test for contamination. The final volume of the PCR reaction was 15µl containing 1 ρmol of each primer, 250 µM dNTP, 1 unit of *Taq* DNA polymerase (GoTaq® Flexi DNA Polymerase, Promega, Madison, WI, USA), 1.3mg/ml BSA solution and the $MgCl_2$ concentration varied from 1mM and 2mM depending on the microsatellite locus (Table 2.5). The reactions were carried out in a Gene Amp ® PCR system 9700 (Applied Biosystems) thermocycler under the following conditions: 93˚C for 3min, followed by 10 cycles of 94˚C for 45 seconds, 56˚C for 45 seconds and 72˚C for 90 seconds followed by 20 cycles of 89˚C for 45 seconds, 56˚C for 45 seconds and 72˚C for 90 seconds and a final extension step of 30 min at 72˚C.

Different approaches were used to improve amplification such as including bovine serum albumin (BSA), increasing the units of *Taq* (GoTaq® Flexi DNA Polymerase, Promega, Madison, WI, USA), changing the type of *Taq* (Supertherm Gold DNA polymerase, Southern Cross Biotechnology, Cape Town), lowering the concentration of $MgCl_2$ and testing alternative thermal cycling procedures (Applied Biosystems, 2000, www.med.yale.edu/genetics/ward/tavi/Trblesht. html).

PCR products were run for approximately 25 minutes through 3% agarose gels and the products were visualised by ethidium bromide staining under UV light. A 100bp DNA ladder (Promega, Madison, WI, USA) was used when running PCR products on the agarose gel. Visible bands of the correct expected size based on the DNA ladder, indicated a successful PCR.

The fluorescent dye marking a particular microsatellite ensures that they can be interpreted by the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). The 3100 is a multi-colour fluorescence-based DNA analysis system with 16 capillaries operating in parallel. The 12 microsatellites were divided into two groups for fragment analysis according to range and the fluorescent label (Table 2.5). There were six markers each in group one (FCA8, FCA26, FCA69, FCA96, FCA247 and FCA224) and group two (FCA78, FCA105, FCA126, FCA133, FCA220 and

FCA310). The final volume for fragment analysis was 15 µl containing 6 µl of PCR product (1 µl of each microsatellite in the group) and 9ul of the formamide and LIZ™ 500 size standard mixture (1000µl formamide:14µl LIZ™ 500 size standard). When doing fragment analysis it separated the mixture of DNA fragments according to their lengths, provided a profile of the separation, determined the length of each fragment (bp) and estimated the concentration of each fragment in the sample. The different colours of fluorescent dye were used to distinguish between PCR products that may overlap in a single lane. A positive control was included in each analysis, this control sample was of a known animal with a known genotype. Once the results were generated, allele sizes were determined by using ABI PRISM® Genemapper software version 3.0 (Applied Biosystems) using the LIZ™ 500 size standard.

To serve as a negative control ruling out the possibility of the microsatellites amplifying contaminant DNA, the 12 microsatellite markers were used in PCR reactions with DNA of a horse, chicken and human (the human DNA used was that of the person that performed the sample collections). Positive amplifications would have indicated that the microsatellites amplify contaminant DNA in addition to the DNA of the cheetah.

2.3 Data analyses

For faecal samples forming part of the pilot study, it was important to calculate the success rate of amplification and reliable genotypes obtained from microsatellite amplification. Successful faecal DNA genotypes were determined by comparing them to the genotypes obtained from the blood DNA samples. The rates of allelic dropout and false alleles were calculated. Confidence intervals were calculated as an indication of the reliability of the estimates obtained. Finally, the data set consisting of genotypes from blood and faecal samples, genotypes from the Thabazimbi faecal samples and the 45 De Wildt captive cheetah genotypes were edited and analyzed using the Microsatellite Toolkit for MS Excel 97 or 2000 (Park, 1999) to format the data and describe the data with regard to the mean number of alleles per locus (*A*) and observed and expected heterozygosity. Probability of Identity (Peakall & Smouse, 2006) was also calculated using the NZG data in order to estimate the probability that two individuals would have the same genotype.

2.3.1 Allelic dropout rate

Due to the various methods applied for determining genotyping errors, differences in calculations make it difficult to compare the results between different studies (Broquet & Petit, 2004). In order to standardise the results, the estimation of an error, as pointed out by Creel *et al.* (2003), should be calculated as the number of errors detected, divided by the number of instances in which the error could have occurred. Allelic dropout can only occur in a heterozygous animal. The rate of

allelic dropout (ADO) was calculated for fresh, four-day-old and two-week-old samples by determining the ratio of ADO at locus *j* (Broquet & Petit, 2004). The rate of ADO is estimated as the total number of amplifications involving the loss of one allele (D_j) divided by the total number of successful amplifications for a heterozygous animal at a locus ($\mathrm{A}_{het\ j}$). The allelic dropout rate at locus *j* (p_j) is estimated by:

$$p_j = \frac{D_j}{\mathrm{A}_{het_j}} \qquad \text{(Broquet \& Petit, 2004)}$$

The average of p_j over L loci is calculated as the ratio of observed ADO over L loci as a proportion of the total number of heterozygous genotypes:

$$p = \bar{p}_w = \frac{\sum_{f=1}^{L} D_j}{\sum_{j=1}^{L} \mathrm{A}_{het_j}} \qquad \text{(Broquet \& Petit, 2004)}$$

2.3.2 False allele rate

False alleles can occur in both homozygote and heterozygote individuals. The rate of false allele (FA) occurrence was calculated for fresh, four-day-old and two-week-old samples by determining the ratio of FA at locus *j* (Broquet & Petit, 2004). The rate of FA is estimated as the total number of amplifications with one or more false alleles at locus *j* and A_j is the number of successful amplifications (homozygous and heterozygous). FA rate at locus j (f_j) is estimated as:

$$f_j = \frac{F_j}{\mathrm{A}_j} \qquad \text{(Broquet \& Petit, 2004)}$$

The average of f_j over L loci is calculated as the ratio of observed FA over L loci as a proportion of the total number of successful amplifications:

$$f = \bar{f}_w = \frac{\sum_{f=1}^{L} F_j}{\sum_{j=1}^{L} A_j} \qquad \text{(Broquet \& Petit, 2004)}$$

2.3.3 Confidence interval estimation

Confidence intervals were used to indicate the reliability of the estimates for allelic dropout and false alleles. A confidence interval gives an estimated range of values which is likely to include an unknown population parameter. The width of the confidence interval gives an indication of the certainty of the unknown parameter, for example, a very wide interval may indicate that more data should be collected before conclusions can be made regarding the parameter.

Confidence intervals were calculated for allelic dropout and false alleles using the following formulae:

$$\mathrm{P} - 1.96\sqrt{\rho\frac{1-\rho}{n}}$$ (Samuels & Witmer, 2003)

$$\mathrm{P} + 1.96\sqrt{\rho\frac{1-\rho}{n}}$$

For a 95% confidence interval, ρ is the number of allelic dropout or false allele occurrences per number of successful PCR amplifications and n is the number of successful amplifications. Calculation of the confidence interval for the occurrence of false alleles included all successful PCR amplifications (homozygous and heterozygous individuals). The confidence interval for allelic dropout is only based on successful heterozygous PCR amplifications.

2.3.4 Regressions

A study by Buchan *et al.* (2005) included the calculation of regressions to determine if there was any correlation between allelic dropout and false alleles with median allele size and amplification success. The regression is a statistical technique used to explain or predict the behaviour of a dependent variable. Generally, a regression equation takes the form of $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ (Mendenhall & Ott, 1980), where $\hat{y}$ is the dependent variable that the equation tries to predict, x is the independent variable that is being used to predict $\hat{y}$. The values of $\hat{\beta}_0$ and $\hat{\beta}_1$ b are selected so that the square of the regression residuals is minimized. The percent amplification success, percent false alleles and percent allelic dropout as a function of median allele size was calculated with the regression equation: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ (Mendenhall & Ott, 1980, Statistica software, Statsoft, 2300 East 14th street, Tulsa, OK 74104, USA). The same was done when calculating the correlation between the observed heterozygosity found in the De Wildt pilot study samples and the occurrence of allelic dropout or false alleles.

2.3.5 Analysis of data obtained from the De Wildt pilot study, Thabazimbi and the NZG

The data set consisting of genotypes obtained in the De Wildt pilot study, the Thabazimbi field trial and the genotypic data of the 45 cheetah from the NZG was edited and analyzed using the Microsatellite Toolkit for MS Excel 97 or 2000 (Park, 1999). The mean number of alleles per locus (*A*) and observed and expected heterozygosity were calculated.

2.3.6 Estimating the number of subpopulations and assignment of individuals to populations

The multilocus genotypic data for the 18 animals of the pilot study, the eight animals from Thabazimbi that had genotypic data and amplified successfully for the species-specific primer, and the data of the 45 animals obtained from the NZG (n=71 individuals) were subjected to the Bayesian clustering method implemented in the software Structure 2·0 (Pritchard *et al.,* 2000) to identify clusters of related individuals. Twenty simulations were run for each putative subpopulation, from $K = 1$ to $K = 10$, and then the statistic ΔK was computed to detect the highest rate of change in the log-likelihood between successive *K*s (Evanno *et al.,* 2005). All simulations were run setting 10^4 iterations as the burn-in period, and 10^5 iterations for Markov chain convergence. For each run, the admixture model was set and allele frequencies were allowed to be correlated. F_{ST} values were also obtained from the Structure output.

Confidence intervals were determined for the Log Likelihood (LnK) values (calculated with Structure 2.0 software) in order to determine the most accurate value with the least variation. The confidence intervals were computed using R, statistical computing software (R Development Core Team, 2008).

2.3.7 Estimating the probability of identifying individuals correctly

The Probability of Identity (P_{ID}) was calculated using GenAlEx (Peakall & Smouse, 2006) to provide an estimate of the average probability that two individuals from the same population would have the same genotype at multiple loci. The NZG data was used to determine this statistic due to the larger sample size which would be more likely to give accurate results. P_{ID} also provided an indication of the statistical power of the set of marker loci. Another advantage of determining the P_{ID} is that it gives an indication of the minimum number of loci required for reliable genetic profiling.

Chapter 3

Results

3.1 Quantification of blood and faecal DNA

The results of the quantification of the 18 DNA samples extracted from the blood collected from the De Wildt Cheetah Centre of the animals for the pilot study are shown in Table 3.1. The yield and purity varied between the DNA samples extracted from blood. These differences in quantity and quality were due to the different sources from which the samples were obtained. Despite low yields for some samples, satisfactory PCR, fragment analysis and genotyping results were obtained. During optimization of microsatellite markers, numerous repeats were done yielding consistent genotypic results. Scats contain DNA not only from the target species but also from other sources such as bacteria and prey DNA making it difficult to determine which portion is from the host species. Amplification with a species-specific primer set will be more useful in measuring the DNA quality for scats.

Table 3.1 Quantification of the 18 DNA samples extracted from blood using a Nanodrop®ND-1000 UV-vis Spectrophotometer

Lab no	Blood		
	ηg/μl*[1]	260/230*[2]	260/280*[2]
1	39.22	1.92	2.09
2	26.70	2.04	2.08
3	78.57	1.65	1.77
4	31.43	1.94	1.95
5	28.88	1.99	2.36
6	23.13	1.88	1.42
7	2.58	-0.34	1.81
8	4.97	2.38	0.60
9	15.67	1.92	0.98
10	2.10	22.14	0.35
11	1.25	1.74	0.45
12	53.09	1.88	2.32
13	54.02	1.84	1.95
14	8.29	1.70	0.74
15	12.92	2.03	1.66
16	15.61	1.95	0.72
17	10.13	2.02	0.89
18	10.07	1.76	0.39

*[1] ng/*μl* (DNA concentration)
*[2] DNA purity

The results of the quantification of fresh faecal DNA, the four-day-old faecal DNA and the two-week-old faecal DNA are shown in Tables 3.2-3.4. Due to insufficient quantities of DNA from the first extraction, DNA was re-extracted from 32 samples (extraction 2). Faeces of nine of the trial

animals were re-sampled; these samples were not stored in 99% ethanol but transported to the laboratory immediately for DNA extraction with the Qiagen mini stool kit (extraction 3).

Table 3.2 Quantification results of the 18 fresh faecal DNA samples collected from the De Wildt Cheetah Centre

Lab no	Fresh Faeces								
	Extraction 1			Extraction 2			Extraction 3		
	ng/µl*[1]	260/230*[2]	260/280*[2]	ng/µl*[1]	260/230*[2]	260/280*[2]	ng/µl*[1]	260/230*[2]	260/280*[2]
1	32.72	0.37	2.04	57.87	0.34	3.63	25.15	1.3	2.12
2	20.88	1.08	2.17	35.42	0.32	24.1			
3	42.34	1.37	1.85	13.01	0.72	1.97			
4	4.38	2.07	-0.31	27.71	1.13	1.32			
5	20.46	0.72	1.83	21.25	0.66	1.57			
6	52.75	2.2	2.12	19.1	0.67	2.48			
7	5.75	2.13	-0.83	12.76	0.37	2.09	11.63	0.61	1.64
8	13.44	1.92	0.45				275.2	0.83	1.57
9	13.46	0.62	2.25	65.9	1.31	1.44	28.39	1.38	2.04
10	20.24	1.96	2.68				17.22	1.24	1.99
11	52.2	0.67	2.15	29.36	-1.11	1.45	88.12	1.39	1.94
12	30.02	1.95	1.25						
13	154.69	2.05	2.01	57.28	1.72	1.93			
14	109.54	1.9	2.02	38.36	1.77	-6.62	32.19	1.51	2.49
15	59.82	2.17	2.51	51.48	-2.22	1.08			
16	70.81	2.15	2.19	21.04	0.57	1.71			
17	24.06	2.92	2.1				32.68	1.8	2.29
18	26.35	1.98	11.23	43	2.53	1.95	31.2	0.9	1.61

*[1] ng/*µl* (DNA concentration)
*[2] DNA purity

Table 3.3 Quantification results of the 18 four-day-old faecal DNA samples collected from the De Wildt Cheetah Centre

Lab no	Four-day-old faeces								
	Extraction 1			Extraction 2			Extraction 3		
	ng/µl*[1]	260/230*[2]	260/280*[2]	ng/µl*[1]	260/230*[2]	260/280*[2]	ng/µl*[1]	260/230*[2]	260/280*[2]
1	269.54	2.16	3.11	20.18	0.14	85.46	7.55	0.69	1.68
2	28.35	1.76	1.91	9.87	0.87	1.71			
3	90.63	2.15	2.77	39.09	0.7	1.97			
4	-10.97	1.5	0.29	22.3	0.5	1.9			
5	37.75	2.1	5.12	18.69	-13.09	1.97			
6	20.67	2.16	20.29	20.65	3.47	1.77			
7	5.61	2.04	0.34	-0.59	-0.01	-0.28	12.04	0.7	1.57
8	21.91	1.96	3.53	1.56	0.02	-0.09	15.32	1.04	2.25
9	65.61	2.11	3.07	22.4	0.53	1.92	54.94	1.36	1.75
10	10.68	1.74	-3.96	-1.03	-0.01	0.06	6.45	0.47	2
11	9.29	1.56	5.77	3.59	0.04	-0.29	60.39	1.56	2.08
12	94.92	2.1	2.76	17.97	0.15	-2.23			
13	28.88	1.84	1.91	70.63	2.72	2.22			
14	119.05	2.14	2.95	43.93	3.48	1.03	17.63	1.14	1.75
15	110.72	2.18	2.22	13.49	-1.42	0.15			
16	10.13	1.94	1.15	9.62	0.34	1.98			
17	5.01	2.03	-1.59	17.5	0.15	-3.18	12.18	1.69	1.68
18	10.02	1.93	1.17	5.7	0.85	1.72	5.73	0.47	1.42

*[1] ng/*µl* (DNA concentration)
*[2] DNA purity

Table 3.4 Quantification results of the 18 two-week-old faecal DNA samples collected from the De Wildt Cheetah Centre

Lab no	Two week old faeces					
	Extraction 1			Extraction 3		
	ηg/μl*[1]	260/230*[2]	260/280*[2]	ηg/μl*[1]	260/230*[2]	260/280*[2]
1	84.73	2.3	5.27	6.72	0.7	1.8
2	16.31	2.24	-2.45			
3	-10.31	1.48	0.35			
4	5.94	3	-3.28			
5	1.99	-15.57	-0.1			
6	67.04	2.29	22.62			
7	9.05	1.87	1.58	9.21	0.88	1.69
8	-2.88	1	0.61	7.96	0.72	2.21
9	24.24	1.92	1.1	6.75	0.71	1.78
10	19.85	1.8	2.21	7.96	0.72	2.21
11	134.61	1.97	2.4	11.55	1.31	2
12	5.18	2.53	-0.33			
13	5.58	2.5	-0.44			
14	-4.46	1.28	0.28	54.06	1.46	2.13
15	4.66	3.94	0.59			
16	30.04	2.26	0.42			
17	-11.11	4.21	0.16	3.56	0.55	2.81
18	16.05	2.16	3.42	5.69	2.05	0.39

*[1] ng/*μl* (DNA concentration)
*[2] DNA purity

Notable differences in DNA quantity and purity were seen in the faecal DNA samples, this is especially apparent in the older samples. It is also notable that quality and purity differed with later extractions sometimes being of better quantity and purity. DNA yield is highly dependent on the amount of starting material. Unexpectedly high quantity values were also obtained for certain samples such as fresh faecal sample 8, extraction 3.

3.2 Species-specific primer set

Two random PCR products (using the species-specific primer set designed for cheetah) were chosen to be sequenced to determine if the DNA originated from cheetah. The results were viewed in Chromas® 2.32 (Technelysium Pty Ltd, http://www.technelysium.com.au/chromas.html) and are shown in Addendum 2.

The result after using BLAST indicated that the sequence that amplified was 96% similar to the first five search results that were all partial *Acinonyx jubatus* sequences from the mitochondrial DNA control region (Figure 3.1a), the sixth search result was of the domestic cat.

```
>gb|AY463962.1|  Acinonyx jubatus isolate S1571 control region, partial sequen
mitochondrial
Length=441

 Score =  294 bits (159),  Expect = 4e-76
 Identities = 171/177 (96%), Gaps = 4/177 (2%)
 Strand=Plus/Plus

Query  3    TGG-CCTC-ACTATCC-AAGG-ACCTANATCNCCTAGCTTCGAGAAACCAGCAATCCTTG  58
            ||| |||| ||||||| |||| ||||| ||| ||||||||||||||||||||||||||||
Sbjct  30   TGGACCTCAACTATCCAAAGGAACCTAAATCACCTAGCTTCGAGAAACCAGCAATCCTTG  89

Query  59   CCTGAACGTGTACCTCTTCTCGCTCCGGGCCCATTTCAACGTGGGGGTTTCTATAACGGA  118
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  90   CCTGAACGTGTACCTCTTCTCGCTCCGGGCCCATTTCAACGTGGGGGTTTCTATAACGGA  149

Query  119  ACTATACCTGGCATCTGGTTCTTACCTCAGGGCCATGGAATACCTTGAATCCAATCC  175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  150  ACTATACCTGGCATCTGGTTCTTACCTCAGGGCCATGGAATACCTTGAATCCAATCC  206
```

Figure 3.1a The sequence alignment with *Acinonyx jubatus* isolate using BLAST

The BLAST search results were also used to construct a distance tree (Figure 3.1b), the distance tree graphically depicts the degree of similarity between the DNA sequenced (sample M4346 indicated by the yellow colour) and sequences of the NCBI library. The distance tree also demonstrates that the sequence was most similar to *Acinonyx jubatus* isolates proving that the species-specific primer set amplified successfully.

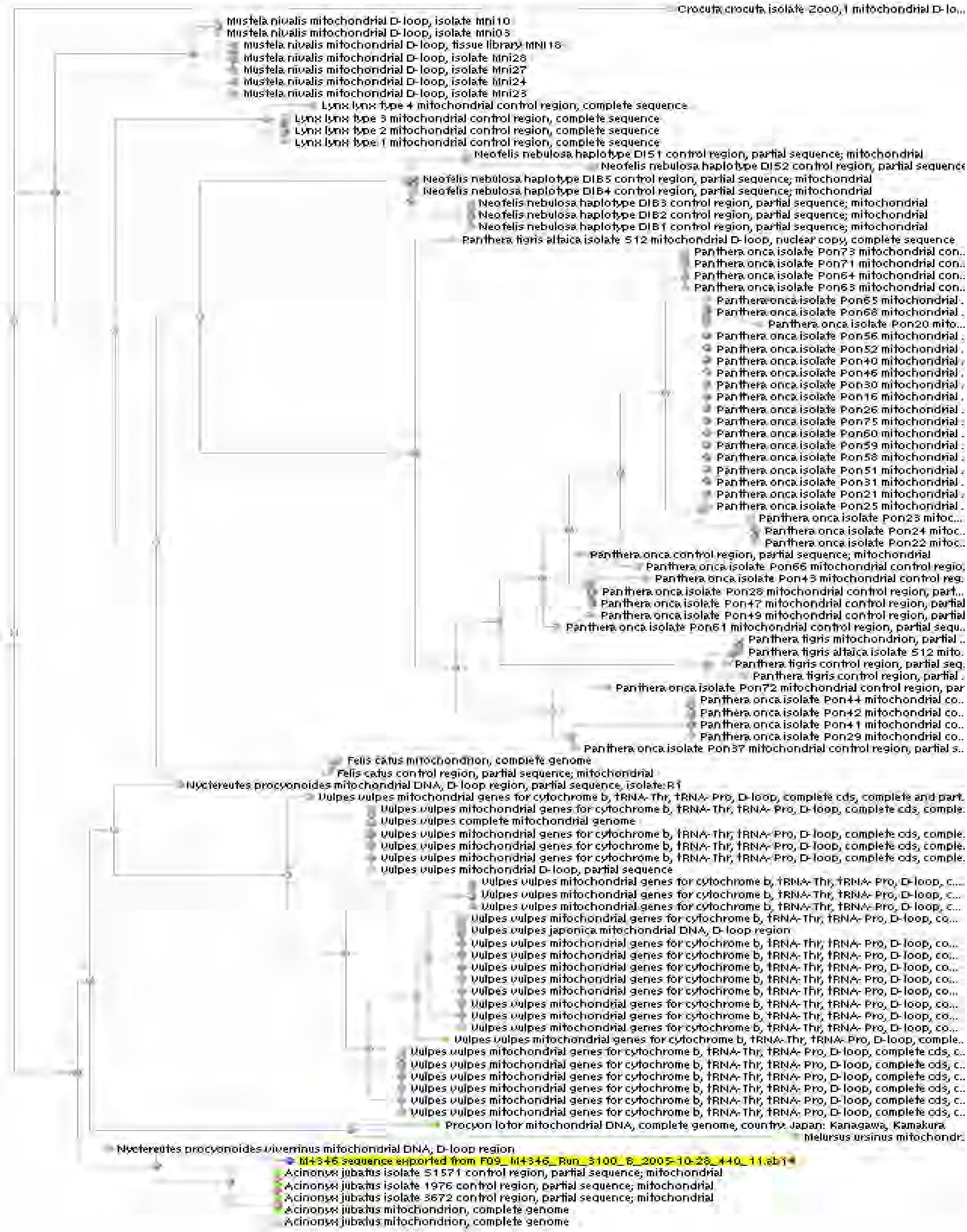


Figure 3.1b Distance tree of results obtained from a BLAST search (NCBI Blast) indicating that the sample sequenced, M4346 highlighted, was most similar to partial sequences of cheetah (*Acinonyx jubatus*)

The results of the amplifications of all faecal DNA samples using the cheetah species-specific marker are shown in Table 3.5. All fresh faecal samples had a positive result (100%), four-day-old samples were 83% positive and two-week-old samples were 94% positive.

Table 3.5 Amplification results using the cheetah species-specific marker for all faecal DNA samples collected at the De Wildt Cheetah Centre. Positive results are indicated by 'Y' and 'N' indicates a negative result.

Animal number	Fresh faecal DNA sample	Four-day-old faecal DNA	Two-week-old faecal DNA
1	Y	Y	Y
2	Y	Y	Y
3	Y	Y	Y
4	Y	Y	N
5	Y	Y	Y
6	Y	Y	Y
7	Y	N	Y
8	Y	Y	Y
9	Y	Y	Y
10	Y	Y	Y
11	Y	Y	Y
12	Y	N	Y
13	Y	Y	Y
14	Y	Y	Y
15	Y	N	Y
16	Y	Y	Y
17	Y	Y	Y
18	Y	Y	Y
Average	100%	83%	94%

Chicken, horse and human DNA were used during PCR amplifications with the cheetah specific marker as a control to exclude the possibility of the marker amplifying DNA other than that belonging to the cheetah. All attempts failed, suggesting that the marker was specific for cheetah in the presence of horse, human and chicken DNA.

3.3 Microsatellite markers

The results of the 12 cat microsatellite markers that were used to genotype the 18 blood and faeces samples of the pilot study are shown in Addendum 3. The fragment ranges of the alleles for each locus (Figure 3.2) was similar to and included the ranges found by Menotti-Raymond *et al.* (1999).

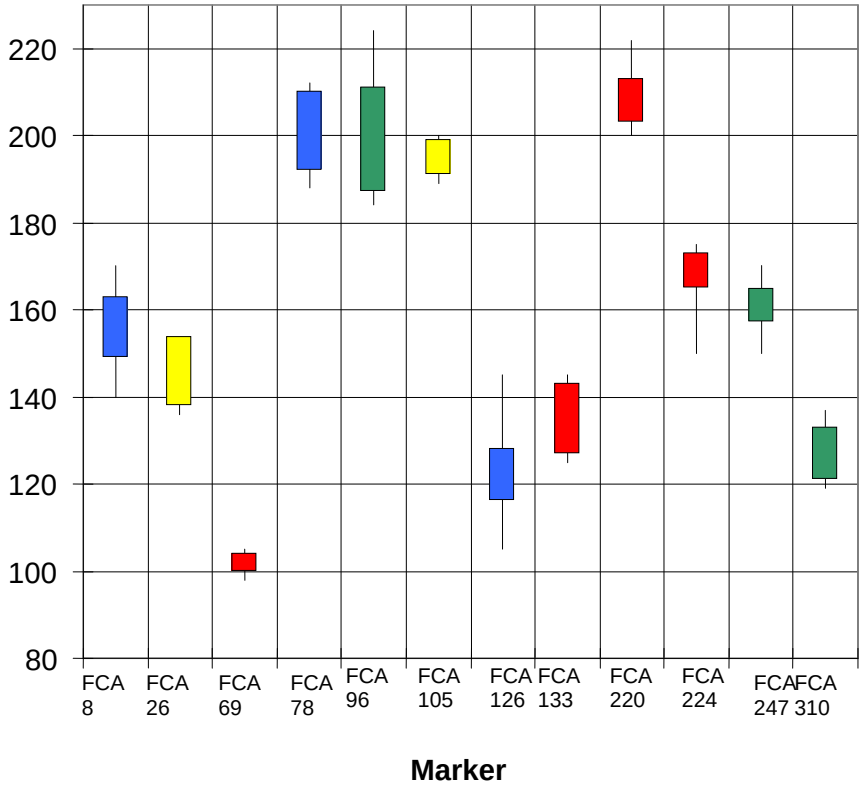


Figure 3.2 Fragment ranges of the alleles and fluorescent label of each microsatellite locus analysed based on blood DNA results. The coloured areas indicate the range of fragment sizes found in this study per allele. The black line extending on either side of the coloured area indicates the range (exceeding current study) found by Menotti-Raymond *et al.* (1999).

3.4 Amplification success, allelic dropout and false allele occurrences

Amplification results comparing the difference of success of faecal samples stored in ethanol versus those extracted immediately are shown in Table 3.6. The table indicates that storage in ethanol appeared to be more effective than immediate extraction with a clear difference between fresh faeces stored in ethanol (60.48%) versus fresh faeces that was immediately extracted (41.4%). It is important to note that more repetitions were performed (1-5 repeats per sample per marker) for the faeces stored in ethanol compared to the one amplification performed for the faeces that was immediately extracted. The reason for only one repeat (one amplification per sample per marker) was that it was a preliminary trial to test if there would be a large improvement in amplification success if the samples weren't stored in ethanol. Table 3.6 also shows that fresh faecal samples had the best overall amplification success (50.54%) when compared to the older samples. The average amplification success for the four-day-old samples (20.06%) was substantially lower than for fresh samples but the lowest amplification success was for the two-week-old samples (2.47%).

Table 3.6 Amplification success of fresh, four-day-old and two-week-old faecal samples stored in ethanol versus those that were immediately analyzed.

Sample type	Amplification success after faecal storage in 99% ethanol	Amplification success after direct extraction without storage in 99% ethanol	Average amplification success
Fresh faeces	60.48%	41.4%	50.54%
Four-day-old faeces	24.44%	10.10%	20.06%
Two-week-old faeces	3.56%	0%	2.47%
Average	39.63%	17.17%	24.37%

Summaries indicating the repetitions done for the faecal DNA samples, amplification success, allelic dropout and the occurrence of false alleles are shown in Tables 3.7-3.10. Amplification results are summarized per marker in Tables 3.7 and 3.8 and summarized per animal in Tables 3.9 and 3.10.

Allelic dropout is calculated based on individuals that are heterozygous at specific loci, for example, 14 cases of allelic dropout were found at locus FCA 8 and the number of heterozygous samples that were amplified successfully was 35 resulting in a 40% rate of allelic dropout. False alleles are based on all individuals, for example, eight cases of false alleles were found at locus FCA 8 and the number of successful amplifications was 58, this means that the frequency of false alleles at the locus FCA 8 was 13.79%. Allelic dropout was found to be very high for a number of markers; ranging from 0% to 37.21%. False allele frequency ranged from 0% to 16.13%.

Table 3.7 Results obtained per microsatellite marker from PCR reactions and genotyping results of fresh and four-day-old faecal samples for all trial animals*

MARKER	FRESH FAECAL SAMPLES									FOUR - DAY – OLD FAECAL SAMPLES								
	REPEATS	*TOTAL SUCCESSFUL AMPLIFICATIONS*	*NO. OF FAILED AMPLIFI-CATIONS*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*[1]	*NO. CONSISTENTLY SCORED*[2]	*ALLELIC DROPOUT*	*ALLELIC DROPOUT RATE*	*FALSE ALLELE*	FALSE ALLELE RATE	*REPEATS*	*TOTAL SUCCESSFUL AMPLIFICATIONS*	*NO. OF FAILED AMPLIFI-CATIONS*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*[1]	*NO. CONSISTENTLY SCORED*[2]	*ALLELIC DROPOUT*	*ALLELIC DROPOUT RATE*	*FALSE ALLELE*	FALSE ALLELE RATE
FCA 8	81	58	23	35	36	14	40.00%	8	13.79%	27	11	16	7	8	1	14.29%	2	18.18%
FCA 26	81	53	28	34	46	4	11.76%	3	5.66%	27	9	18	5	9	0	0.00%	0	0.00%
FCA 69	81	72	9	56	66	2	3.57%	4	5.56%	27	9	18	8	7	2	25.00%	0	0.00%
FCA 78	81	15	66	13	15	0	0.00%	0	0.00%	27	0	27	0	0	0	0.00%	0	0.00%
FCA 96	81	35	46	29	25	8	27.59%	2	5.71%	27	2	25	1	2	0	0.00%	0	0.00%
FCA 105	81	68	13	41	58	1	2.44%	9	13.24%	27	9	18	5	9	0	0.00%	0	0.00%
FCA 126	81	56	25	21	43	3	14.29%	10	17.86%	27	6	21	2	6	0	0.00%	0	0.00%
FCA 133	81	53	28	39	41	9	23.08%	3	5.66%	27	8	19	7	6	2	28.57%	0	0.00%
FCA 220	81	35	46	26	29	5	19.23%	1	2.86%	27	4	23	3	4	0	0.00%	0	0.00%
FCA 224	81	48	33	39	36	10	25.64%	2	4.17%	27	3	24	1	3	0	0.00%	0	0.00%
FCA 247	81	35	46	23	30	4	17.39%	1	2.86%	27	1	26	1	1	0	0.00%	0	0.00%
FCA 310	81	41	40	23	38	3	13.04%	0	0.00%	27	3	24	1	3	0	0.00%	0	0.00%
Total	972									324								

* This table is representative of the results for all the trial animals (from De Wildt Cheetah Centre) at a specific locus
*[1] Successful amplifications of heterozygous individuals at specific loci for calculation of dropout rate (Chapter 2.3.1)
*[2] Instances where the genotype obtained corresponded with the genotype obtained for the blood sample of the same animal

Table 3.8 Results obtained per microsatellite marker from PCR reactions and genotyping results of two-week-old faecal samples and a summary of all the samples for all the trial animals*

MARKER	TWO – WEEK – OLD FAECAL SAMPLES									TOTAL (All faecal samples)					
	REPEATS	*TOTAL SUCCESSFUL AMPLIFICATIONS*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*[*1]	*NO. OF FAILED AMPLIFI-CATIONS*	*NO. CONSISTENTLY SCORED*[*2]	*ALLELIC DROPOUT*	*ALLELIC DROPOUT RATE*	*FALSE ALLELE*	FALSE ALLELE RATE	TOTAL SUCCESSFUL AMPLIFICATIONS	SUCCESSFUL AMPLIFICATIONS (heterozygous)	TOTAL ALLELIC DROPOUT	TOTAL ALLELIC DROPOUT RATE	TOTAL FALSE ALLELES	TOTAL FALSE ALLELE RATE
FCA 8	27	2	1	25	1	1	100.00%	0	0.00%	71	43	16	37.21%	10	14.08%
FCA 26	27	1	0	26	1	0	0.00%	0	0.00%	63	39	4	10.26%	3	4.76%
FCA 69	27	0	0	27	0	0	0.00%	0	0.00%	81	64	4	6.25%	4	4.94%
FCA 78	27	0	0	27	0	0	0.00%	0	0.00%	15	13	0	0.00%	0	0.00%
FCA 96	27	1	1	26	1	0	0.00%	0	0.00%	38	31	8	25.81%	2	5.26%
FCA 105	27	2	0	25	1	0	0.00%	1	50.00%	79	46	1	2.17%	10	12.66%
FCA 126	27	0	0	27	0	0	0.00%	0	0.00%	62	23	3	13.04%	10	16.13%
FCA 133	27	1	0	26	1	0	0.00%	0	0.00%	62	46	11	23.91%	3	4.84%
FCA 220	27	0	0	27	0	0	0.00%	0	0.00%	39	29	5	17.24%	1	2.56%
FCA 224	27	1	1	26	1	0	0.00%	0	0.00%	52	41	10	24.39%	2	3.85%
FCA 247	27	0	0	27	0	0	0.00%	0	0.00%	36	24	4	16.67%	1	2.78%
FCA 310	27	0	0	27	0	0	0.00%	0	0.00%	44	24	3	12.50%	0	0.00%
Total	324														

* This table is representative of the results for all the trial animals (from De Wildt Cheetah Centre) at a specific locus
*[1] Successful amplifications of heterozygous individuals at specific loci for calculation of dropout rate (Chapter 2.3.1)
*[2] Instances where the genotype obtained corresponded with the genotype obtained for the blood sample of the same animal

Table 3.9 Results obtained per animal from PCR reactions and genotyping results of fresh and four-day-old faecal samples for all microsatellite markers*

ANIMAL	FRESH FAECAL SAMPLES									FOUR - DAY – OLD FAECAL SAMPLES								
	REPEATS	*TOTAL SUCCESSFUL AMPLIFICATIONS*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*1*	*NO. OF FAILED AMPLIFI-CATIONS*	*NO. CONSISTENTLY SCORED*2*	*ALLELIC DROPOUT*	*ALLELIC DROPOUT RATE*	*FALSE ALLELE*	FALSE ALLELE RATE	*REPEATS*	*TOTAL SUCCESSFUL AMPLIFICATIONS*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*1*	*NO. OF FAILED AMPLIFI-CATIONS*	*NO. CONSISTENTLY SCORED*2*	*ALLELIC DROPOUT*	*ALLELIC DROPOUT RATE*	*FALSE ALLELE*	FALSE ALLELE RATE
1	60	47	42	13	45	2	4.76%	0	0.00%	24	0	0	24	0	0	0.00%	0	0.00%
2	48	43	29	5	38	3	10.34%	2	4.65%	12	2	1	10	1	0	0.00%	1	50.00%
3	48	41	30	7	40	1	3.33%	0	0.00%	12	7	4	5	7	0	0.00%	0	0.00%
4	48	24	19	24	13	8	42.11%	3	12.50%	12	1	0	11	1	0	0.00%	0	0.00%
5	48	29	16	19	20	1	6.25%	8	27.59%	12	0	0	12	0	0	0.00%	0	0.00%
6	48	29	24	19	20	7	29.17%	2	6.90%	12	7	7	5	6	1	14.29%	0	0.00%
7	60	29	16	31	25	3	18.75%	1	3.45%	24	4	1	20	3	1	100.00%	0	0.00%
8	60	35	18	25	28	4	22.22%	3	8.57%	24	6	3	18	6	0	0.00%	0	0.00%
9	60	37	18	23	36	1	5.56%	0	0.00%	24	2	2	22	1	1	50.00%	0	0.00%
10	60	34	27	26	27	5	18.52%	2	5.88%	24	8	8	16	7	1	12.50%	0	0.00%
11	60	35	25	25	31	1	4.00%	3	8.57%	24	6	3	18	6	0	0.00%	0	0.00%
12	48	23	20	25	15	5	25.00%	3	13.04%	12	1	1	11	0	0	0.00%	1	100.00%
13	48	17	15	31	5	9	60.00%	3	17.65%	12	2	1	10	1	1	100.00%	0	0.00%
14	60	28	16	32	24	2	12.50%	2	7.14%	24	7	4	17	7	0	0.00%	0	0.00%
15	48	19	7	29	16	0	0.00%	3	15.79%	12	3	1	9	3	0	0.00%	0	0.00%
16	48	28	15	20	23	2	13.33%	3	10.71%	12	2	1	10	2	0	0.00%	0	0.00%
17	60	27	13	33	18	6	46.15%	3	11.11%	24	7	4	17	7	0	0.00%	0	0.00%
18	60	44	29	16	39	3	10.34%	2	4.55%	24	0	0	24	0	0	0.00%	0	0.00%
Total	972	569								324	65							

* This table is representative of the results of the 12 microsatellite markers for a specific animal
*[1] Successful amplifications of heterozygous individuals at specific loci for calculation of dropout rate (Chapter 2.3.1)
*[2] Instances where the genotype obtained corresponded with the genotype obtained for the blood sample of the same animal

Table 3.10 Results obtained per animal from PCR reactions and genotyping results of two-week-old faecal samples and a summary of all the samples for all the microsatellite markers*

ANIMAL	TWO – WEEK – OLD FAECAL SAMPLES									TOTAL (All faecal samples)					
	REPEATS	*TOTAL SUCCESSFUL AMPLIFICATIONS per animal for all markers*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*[1]	*NO. OF FAILED AMPLIFI-CATIONS*	*NO. CONSISTENTLY SCORED*[2]	*ALLELIC DROPOUT*	*ALLELIC DROPOUT RATE*	*FALSE ALLELE*	FALSE ALLELE RATE	*TOTAL SUCCESSFUL AMPLIFICATIONS per animal for all markers*	*SUCCESSFUL AMPLIFICATIONS (Heterozygous)*[1]	*TOTAL ALLELIC DROPOUT*	*TOTLA ALLELIC DROPOUT RATE*	*TOTAL FALSE ALLELES*	TOTAL FALSE ALLELE RATE
1	24	0	0	24	0	0	0.00%	0	0.00%	47	42	2	4.76%	0	0.00%
2	12	0	0	12	0	0	0.00%	0	0.00%	45	30	3	10.00%	3	6.67%
3	12	0	0	12	0	0	0.00%	0	0.00%	48	34	1	2.94%	0	0.00%
4	12	0	0	12	0	0	0.00%	0	0.00%	25	19	8	42.11%	3	12.00%
5	12	0	0	12	0	0	0.00%	0	0.00%	29	16	1	6.25%	8	27.59%
6	12	1	1	11	1	0	0.00%	0	0.00%	37	32	8	25.00%	2	5.41%
7	24	2	1	22	2	0	0.00%	0	0.00%	35	18	4	22.22%	1	2.86%
8	24	0	0	24	0	0	0.00%	0	0.00%	41	21	4	19.05%	3	7.32%
9	24	3	0	21	3	0	0.00%	0	0.00%	42	20	2	10.00%	0	0.00%
10	24	0	0	24	0	0	0.00%	0	0.00%	42	35	6	17.14%	2	4.76%
11	24	0	0	24	0	0	0.00%	0	0.00%	41	28	1	3.57%	3	7.32%
12	12	0	0	12	0	0	0.00%	0	0.00%	24	21	5	23.81%	4	16.67%
13	12	0	0	12	0	0	0.00%	0	0.00%	19	16	10	62.50%	3	15.79%
14	24	0	0	24	0	0	0.00%	0	0.00%	35	20	2	10.00%	2	5.71%
15	12	0	0	12	0	0	0.00%	0	0.00%	22	8	0	0.00%	3	13.64%
16	12	0	0	12	0	0	0.00%	0	0.00%	30	16	2	12.50%	3	10.00%
17	24	2	1	22	0	1	100.00%	1	50.00%	36	18	7	38.89%	4	11.11%
18	24	0	0	24	0	0	0.00%	0	0.00%	44	29	3	10.34%	2	4.55%
Total	324	8													

* This table is representative of the results of the 12 microsatellite markers for a specific animal
*[1] Successful amplifications of heterozygous individuals at a specific loci for calculation of dropout rate (Chapter 2.3.1)
*[2] Instances where the genotype obtained corresponded with the genotype obtained for the blood sample of the same animal

The average rate of allelic dropout and frequency of false alleles per marker are displayed in Figure 3.3. The graph indicates that the highest rate of allelic dropout was found for marker FCA 8 and the lowest was found for marker FCA 78. The highest occurrence of false alleles was found for FCA 126 and the lowest was for FCA310 and FCA 78. It is important to note that the amplification success for FCA 78 was very low and the results regarding allelic dropout and false allele occurrence are thus unreliable.

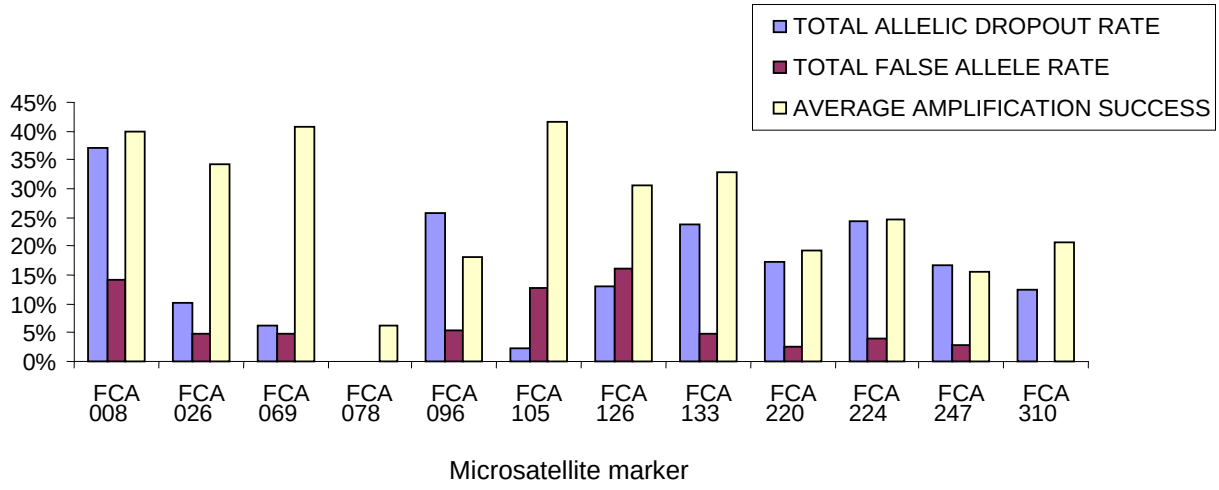


Figure 3.3 Rate of allelic dropout, rate of false allele occurrence and average amplification success per microsatellite marker based on genotypes from fresh, four-day-old and two-week-old faecal DNA.

Examples of the two types of genotyping errors found, allelic dropout and false alleles, are shown in Figure 3.4 and Figure 3.5 where the genotypes of the blood samples are also shown as a reference. In Figure 3.4 two DNA extractions done from fresh faeces (b & c) are shown. The allelic dropout differed between the two extractions. It is interesting to note that the four-day-old sample (d) did not have allelic dropout and the correct genotype could be distinguished clearly. The peak depicted before the 207bp allele is recognized as stutter, peaks occurring after the first allele were only recognised as a second allele if they were larger than 50% of the initial allele. Figure 3.5a shows the genotype obtained for the blood sample of a particular animal at locus FCA 8, Figure 3.5b-d show the consistent false allele occurrences when the PCR was repeated using the same DNA.

a.

b.

c.

d.

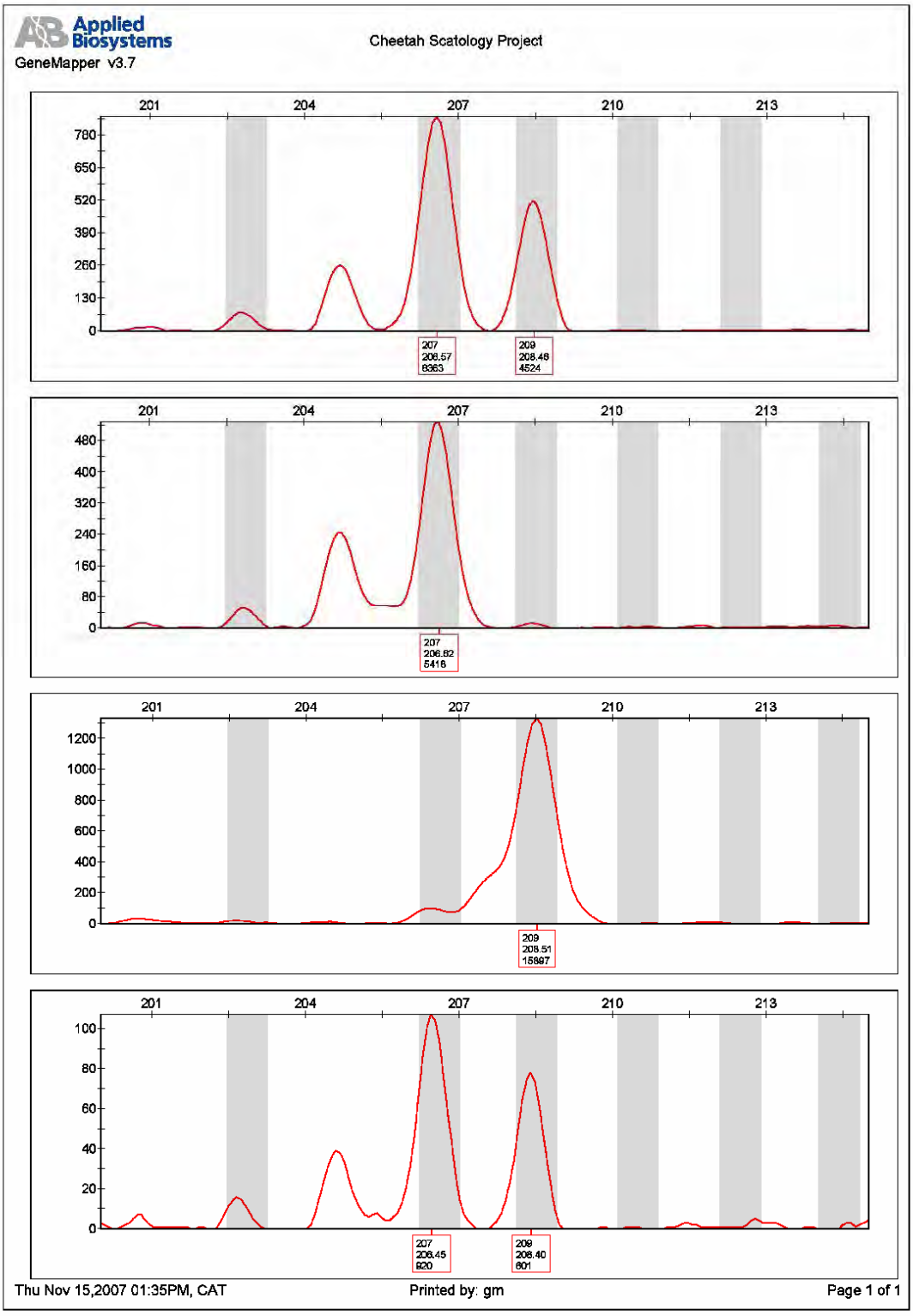


Figure 3.4 Genemapper electropherograms illustrating an example of allelic dropout at locus FCA 220 using DNA from different sample types from animal 10 forming part of the pilot study.
a. Blood sample, b. Fresh faecal sample (extraction 1), c. Fresh faecal sample (extraction 2), d. four-day-old faecal sample

a.

b.

c.

d.

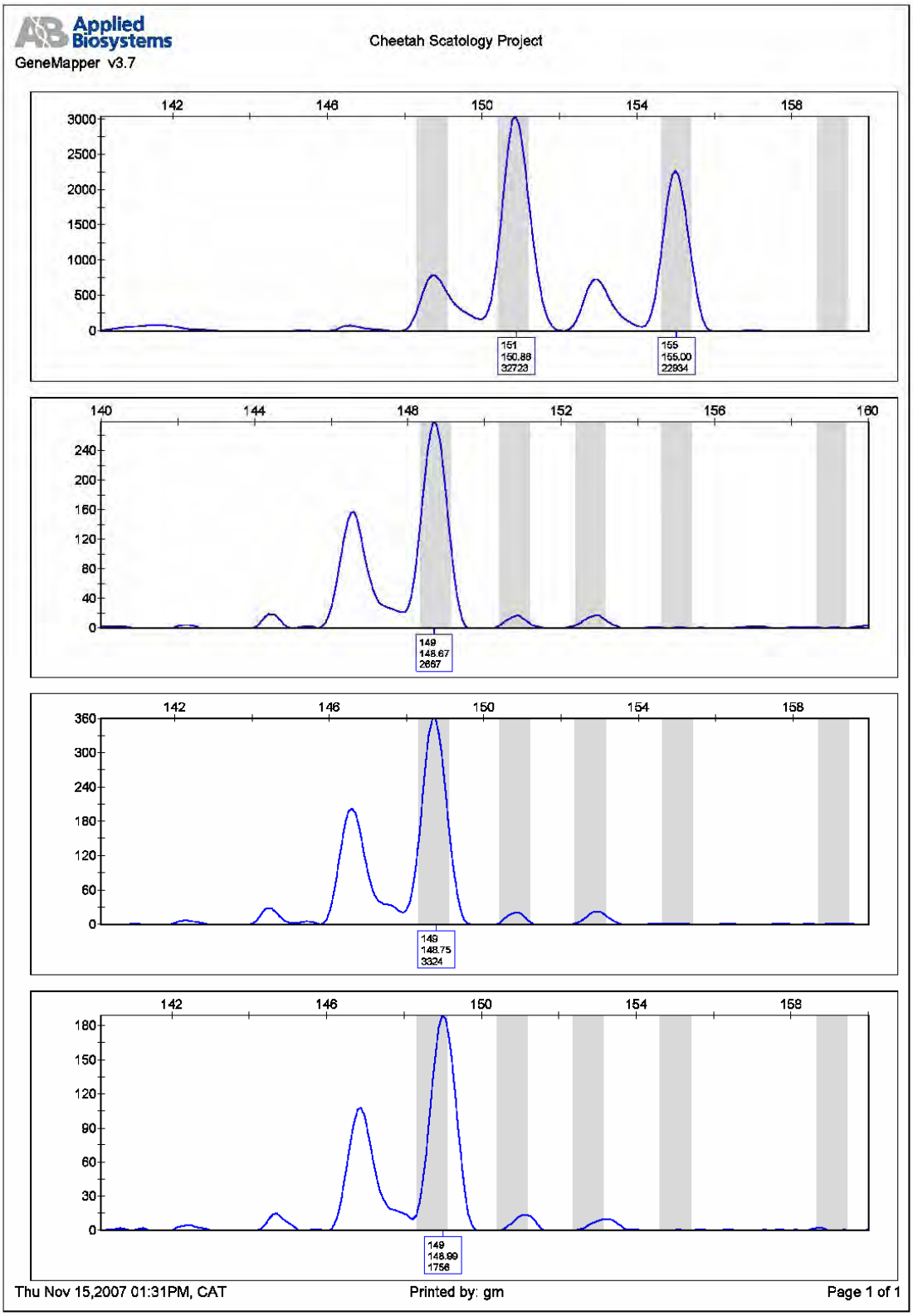


Figure 3.5 Genemapper electropherograms illustrating an example of false allele occurrence at locus FCA 8 using DNA from different sample types from animal six forming part of the pilot study.
a. Blood sample, b-d. Separate repetitions of a fresh faecal sample.

3.4.1 Confidence intervals

Confidence intervals were calculated to indicate the reliability of the estimates for allelic dropout and false alleles. The results are summarized in Table 3.11.

Table 3.11 Confidence intervals for allelic dropout and false alleles of the fresh, four-day-old and two-week-old samples

	Fresh Samples	Four-day-old samples	Two-week-old samples	Total
Total Allelic dropout	63	5	1	69
Total False Alleles	43	2	1	46
Total Successful amplifications	569	65	8	642
Total Successful heterozygous amplifications	379	41	3	432
Dropout rate	0.17	0.12	0.33	0.16
Allelic dropout – Lower CI	0.13	0.02	-0.20	0.12
Allelic dropout – Upper CI	0.20	0.22	0.87	0.19
False Allele rate	0.08	0.03	0.13	0.07
False allele – Lower CI	0.05	-0.01	-0.10	0.05
False allele – Upper CI	0.10	0.07	0.35	0.09

3.4.2 Regression analyses

Regression lines were calculated and drawn for amplification success (Figure 3.6), the false allele occurrences (Figure 3.7) and rate of allelic dropout (Figure 3.8) as functions of median allele size.

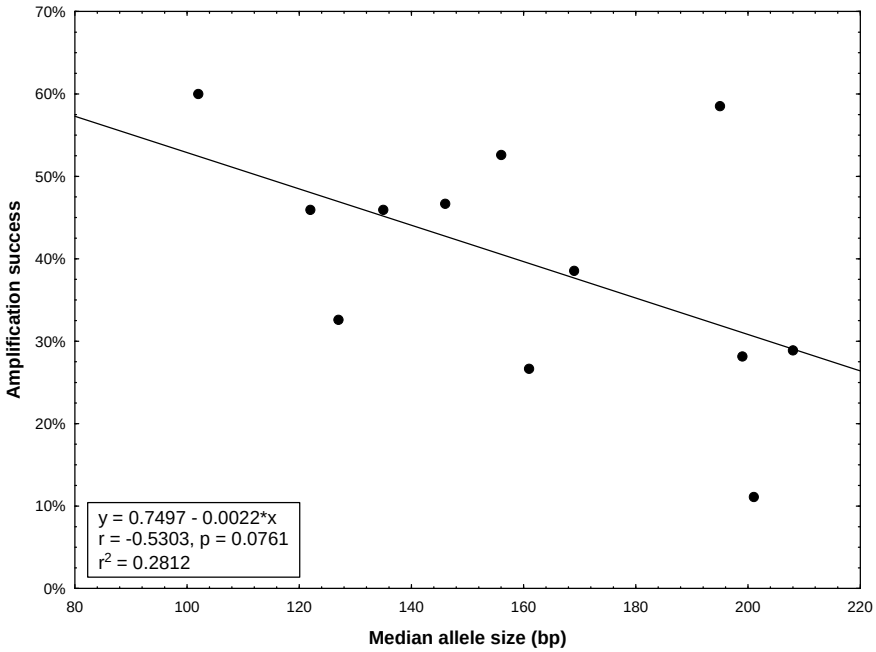


Figure 3.6 Amplification success as a function of median allele size

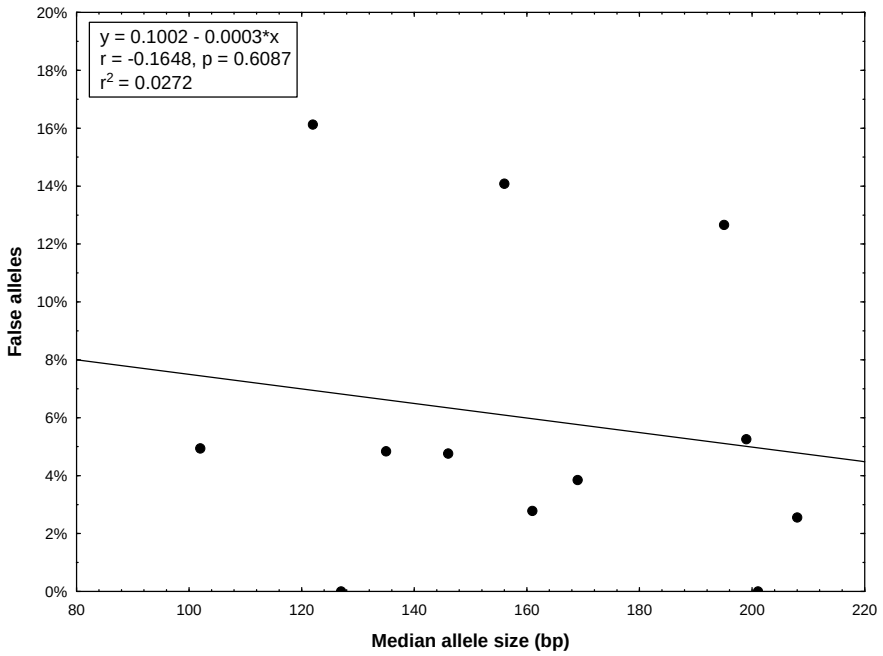


Figure 3.7 Occurrence of false alleles as a function of median allele size

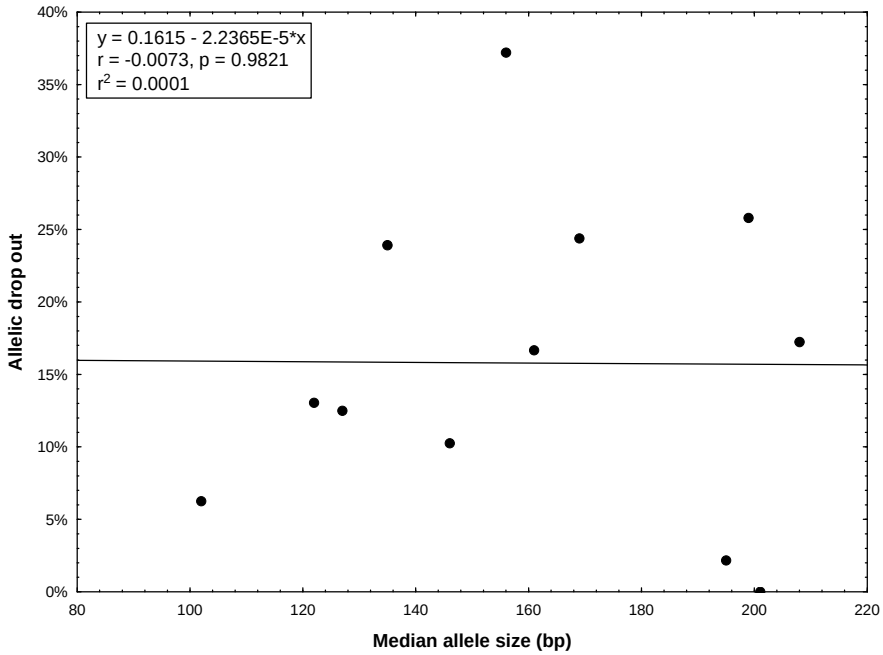


Figure 3.8 Allelic dropout as a function of median allele size

Although the r^2 values are low for all the graphs, indicating that the linear regression is not the best fit to the data, there is indeed a trend that can be seen by the slope of the line. Amplification success tends to be better when median allele sizes are smaller and the occurrence of false alleles tends to decrease when the median allele size increases. The graph indicating the correlation between median allele size and allelic dropout seems inconclusive as there is no clear trend (Figure 3.8) but if it weren't for the two outliers around 195 and 200bp there would be a strong correlation showing that allelic dropout is positively correlated with allele size.

3.5 Analysis of faecal samples collected in the Thabazimbi area

DNA extracted from the Thabazimbi faecal samples was quantified and the results are shown in Table 3.12. The quantities obtained from these samples were mostly found to be low but this was expected based on the results obtained from the pilot study. The results of Thabazimbi faecal DNA amplifications using the species-specific primer set are shown in Table 3.12. Of the 15 samples analysed, 12 had a positive result indicating the presence of DNA originating from cheetah.

The results of the Thabazimbi faecal DNA amplification using the 12 cat microsatellites after genotyping are shown in Table 3.13. Marker FCA 78 did not yield any successful amplifications and was excluded from the results.

Table 3.12 Quantification of all DNA from samples collected in the Thabazimbi area and the success of the species-specific (SS) primer

Lab no	Faeces collected in Thabazimbi area			
	Extraction 1			
	ηg/μl*[1]	260/230*[2]	260/280*[2]	SS*[3]
W01	8.03	0.60	1.55	Y
W02	48.40	1.01	0.77	Y
W03	15.74	0.62	1.40	Y
W04	18.67	0.86	1.25	N
W05	18.48	0.46	1.50	Y
W06	25.88	0.55	1.69	Y
W07	14.14	1.08	1.46	N
W08	16.35	0.50	1.65	Y
W09	79.50	1.46	2.02	Y
W10	13.95	0.70	1.89	Y
W11	9.32	0.44	1.45	Y
W12	29.12	1.39	2.15	Y
W13	11.79	0.30	1.66	Y
W14	13.71	0.52	1.62	Y
W15	24.83	1.10	1.85	N

*[1] ng/μl (DNA concentration)
*[2] DNA purity
*[3] SS (Species-specific primer set): Y (Yes) indicates that the species-specific primer set amplified successfully, N (No) indicates that the amplification was unsuccessful

Table 3.13 Genotypes from wild cheetah samples collected in the Thabazimbi Area

Marker/ Animal	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 8	FCA 96	Successful Amplifications (for the 11 loci)
W01				127 127	213 213	159 159						3/11
W02			116 118					138 150	121 131	149 149		4/11
W03			116 118									1/11
W04*[1]	102 102											1/11
W05												0
W06				127 127		157 157				149 149		3/11
W07*[1]												0
W08												0
W09	102 104	191 193	118 126	127 143	207 207	169 169	163 163	138 154		153 153	199 199	10/11
W10		195 197	116 118	127 127	213 213			152 154	121 123	153 155	201 201	8/11
W11												0
W12	104 104	195 194		127 127	213 213	167 169	155 155	152 154		153 155	201 201	9/11
W13	102 104	191 193		127 143	207 207	169 169					199 199	6/11
W14												0
W15*[1]	114 114					157 157						2/11

*[1]Samples that did not amplify when the species-specific primer set was used

One of the applications of scatology is to use the genotypic data obtained from the molecular analysis together with the information of where the sample was collected in order to estimate the home range of the animals being studied. Matches were found between a number of samples for the loci analyzed. For example, animal 9 and animal 13 had 12 alleles (6 markers) that could be compared and all of the alleles matched meaning that there is a possibility that it is the same animal. Close relatives may also have close matches and be physically near one another in a given area, for this reason it is important that a number of loci are used in order to distinguish between individuals.

3.6 Analysis of data obtained from the De Wildt pilot study and the NZG

Heterozygosity values and mean number of alleles were estimated for the De Wildt pilot study data and the data provided by the NZG, these values are summarized in Table 3.14. Genotypic data obtained from the NZG are attached in Addendum 4. An additional calculation using 18 random NZG samples was done to determine if the heterozygosity values and mean number of alleles would differ when using a smaller sample size. As can be seen in Table 3.14, the heterozygosity values of the smaller NZG sample size decreased slightly, as did the mean number of alleles due to the smaller sample size.

Table 3.14 Heterozygosity values and mean number of alleles calculated from the genotypic data obtained from the blood and faecal samples of the pilot study and the genotypic data from the NZG

Source	Sample size	Loci typed	H_E	SD	H_O	SD	Mean no. alleles	SD
Blood (pilot study)	18	12	0.71	±0.03	0.69	±0.03	4.92	±1.24
Faeces (pilot study	18	12	0.64	±0.04	0.58	±0.02	4.83	±1.19
NZG	45	12	0.7692	0.0132	0.7093	0.0195	7.08	1.62
NZG	18	12	0.7550	0.0207	0.6806	0.0317	6.17	1.11

Allele frequencies per locus are shown in Table 3.15. This table was constructed to compare the results from a small sample size such as the pilot study of 18 animals with the larger sample size of the NZG. The values obtained for the allele frequencies for the NZG samples give a clearer indication of what can be expected in the cheetah population because it is a larger sample size. It is also important to ensure that the alleles found in the pilot study are also found in the NZG samples confirming that similar alleles are obtained with the use of scatology.

Table 3.15 Allele frequencies (%) for blood and NZG samples by locus

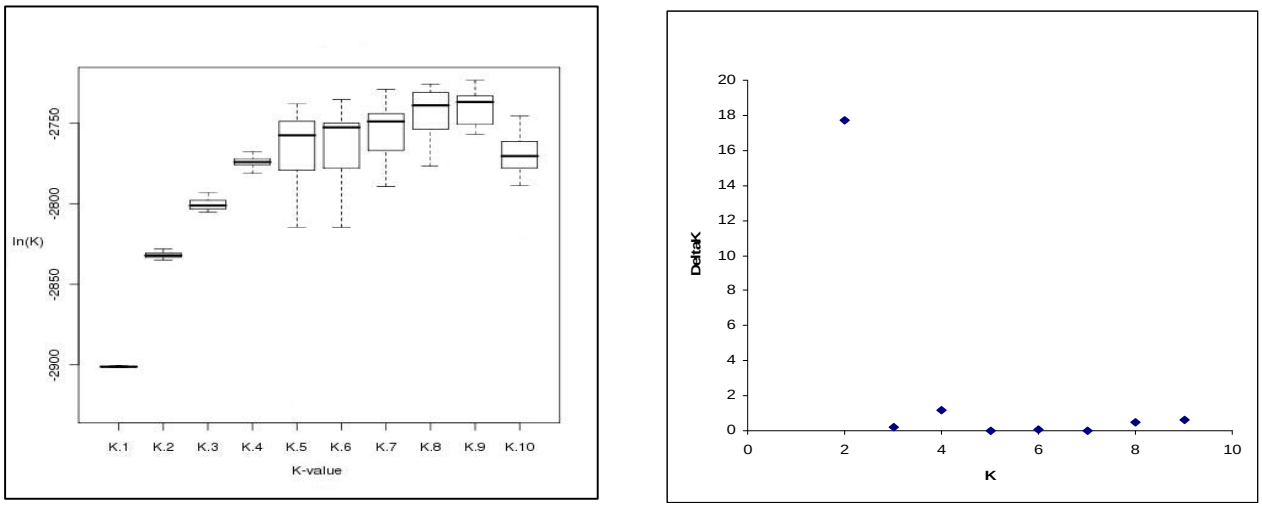
Locus	Blood	NZG
FCA 69a		
100	8.33	12.22
102	44.44	35.56
104	47.22	25.56
106		8.89
108		17.78
FCA 105a		
191	38.89	32.22
193	41.67	24.44
195	16.67	22.22
197		17.78
199	2.78	2.22
201		1.11
FCA 126a		
114		1.11
116	2.78	2.22
118	69.44	34.44
120		11.11
122		24.44
124		3.33
128	27.78	16.67
132		6.67
FCA 133a		
117		1.11
125		11.11
127	36.11	24.44
129		3.33
133	19.44	24.44
135	8.33	16.67
137	19.44	8.89
141	2.78	2.22
143	13.89	4.44
145		1.11
147		2.22
FCA 220a		
203	25.00	24.44
207	11.11	10.00
209	27.78	20.00
211	27.78	26.67
213	8.33	15.56
215		3.33
FCA 224a		
165	30.56	10.00
167	25.00	27.78
169	30.56	31.11
171		15.56
173	13.89	7.78
175		7.78
FCA 247a		
149		1.11
153		2.22
155		37.78
157	25.00	3.33
159	8.33	6.67
161	8.33	33.33
163	47.22	15.56
165	11.11	

Locus	Blood	NZG
FCA 26a		
136		17.78
138	19.44	
142	2.78	7.78
146		
148		51.11
150	55.56	6.67
152	5.56	2.22
154	16.67	14.44
FCA 310a		
121	27.78	37.78
123	13.89	17.78
125		1.11
127	13.89	15.56
129		1.11
131	41.67	24.44
133	2.78	2.22
FCA 78a		
190		1.11
192	19.44	16.67
198		6.67
200	22.22	15.56
202	16.67	26.67
204	13.89	16.67
206	13.89	10.00
208		2.22
210	13.89	4.44
FCA 8a		
149	47.22	31.11
151	25.00	32.22
153		14.44
155	16.67	15.56
159	5.56	3.33
161	2.78	2.22
163	2.78	1.11
FCA 96a		
187	11.11	2.22
199	33.33	46.67
201	16.67	7.78
205	2.78	10.00
207	22.22	12.22
209	5.56	6.67
211	8.33	14.44

3.7 Estimating the number of subpopulations using the data from the De Wildt pilot study, Thabazimbi and the NZG

The approach used to infer the most likely number of populations showed that K = 2 was the optimal number of clusters when using ΔK. The values that were calculated per population for LnK and ΔK are represented in the graphs of Figure 3.9.

The most appropriate K value was determined by calculating the confidence intervals for the LnK values. As shown in Figure 3.9 (a), the most accurate K value according to the confidence intervals for LnK was 2. Therefore it is assumed that the most likely number of populations for the data in this study is 2 as this corresponds with the ΔK result.



(a) (b)

Figure 3.9 (a) Graphic representation of LnK values with confidence intervals for each population (K), (b) Graphic representation of ΔK values calculated for each population (K)

The assignment of individuals of the three sample groups (De Wildt pilot study, Thabazimbi and the NZG) to two inferred clusters are shown in Figure 3.10. The De Wildt pilot study animals are represented by '1', the Thabazimbi animals are represented by '2' and the NZG samples are represented by '3' on the x-axis. If an individual showed a probability of assignment of ≥80% it is assigned to that cluster as can be seen in cluster 1 (red) where the De Wildt pilot study group contains 15 individuals assigned to cluster 1 and three individuals that were admixed. Individuals that fell into the range between 20% and 80% were referred to as admixed. Of the Thabazimbi field trial samples, one individual was assigned to cluster 1, two assigned to cluster 2 (green) and five individuals were admixed. The NZG samples assigned four individuals to cluster 1, 28 individuals to cluster 2 and 13 were admixed. The F_{ST} values calculated by Structure were estimated as 0.1457 for $F_{ST}1$ and 0.0126 for $F_{ST}2$, this indicates that the variation within cluster 1 is more than for cluster 2.

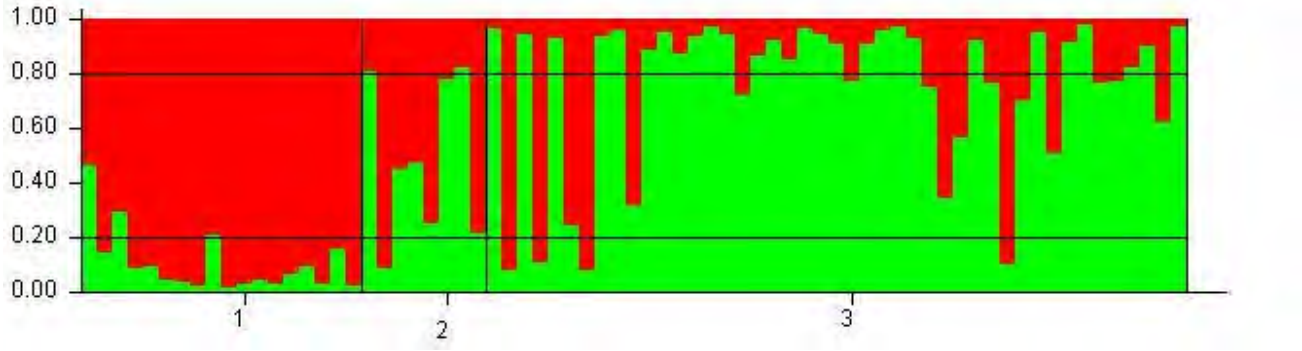


Figure 3.10 Assignment of individuals to two inferred clusters 1=De Wildt, 2=Thabazimbi and 3=NZG (De Wildt captive)

3.8 Estimating the probability of identity for increasing locus combinations using NZG data

The NZG data was used to obtain an indication of the probability that two individuals will have the same genotype. The 12 microsatellite loci allowed a high degree of discrimination among genotypes within the NZG sample set, as indicated by the probability of identity ranging from 0.103 to less than 0.002 when three loci are included (Figure 3.11). The expected number of individuals that might display the same multi-locus genotype within each population was therefore very low.

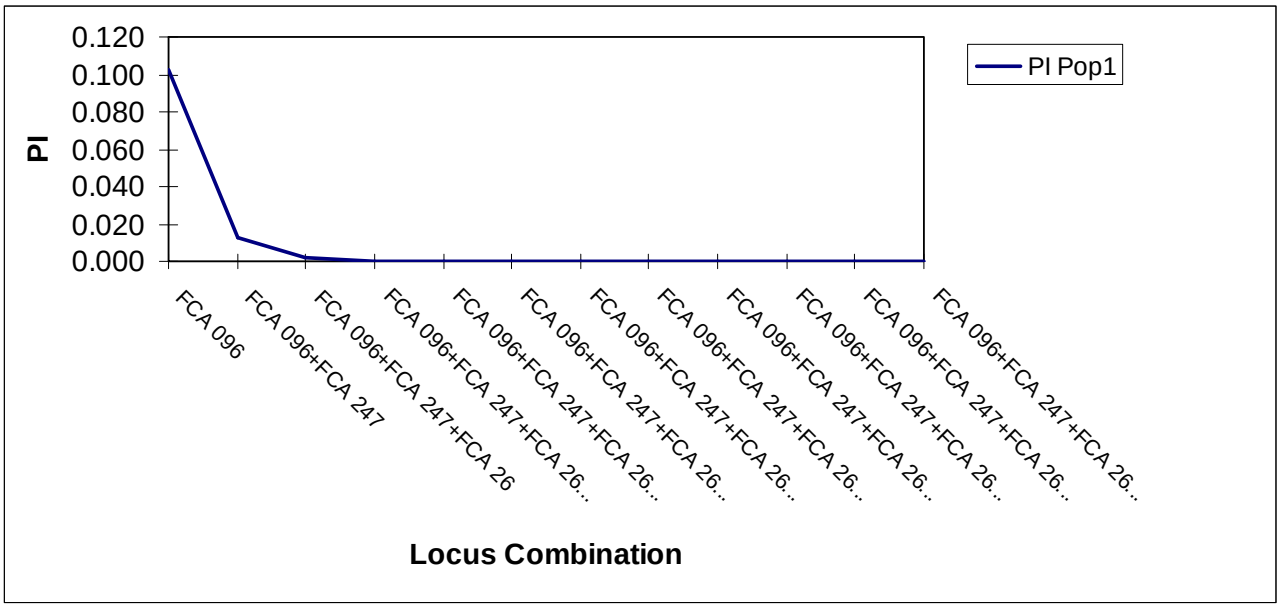


Figure 3.11 The probability of identity for increasing locus combinations using NZG genotypic data

Chapter 4

Discussion

Conventionally, blood or tissue samples are taken from animals for genetic profiling. For an elusive animal, such as a cheetah, the application of this method is not practical in the wild (game reserves and farms). In a number of areas in South Africa such as Thabazimbi, for example, it is extremely difficult to track the cheetah on foot as the vegetation is very thick in certain areas; tracking will prove to be a time consuming and labour intensive method of finding the animal and obtaining the genetic material. Additionally, if a cheetah is caught, the procedure of obtaining the sample will be stressful for these wild animals. An alternative method should be employed to obtain samples for accurate genetic analysis, these samples should be relatively easy to obtain and should not place any unnecessary stress on the cheetah. Many studies have shown that the DNA can be obtained from the faeces but the method used to obtain DNA and the quality of this DNA varies and therefore it is important to verify the extent to which scatology can be used (Kohn & Wayne, 1997, Ernest, 2000, Farrell, 2001, Creel, 2003, Dallas *et al.*, 2003, Kurose *et al.*, 2005). Varying and inaccurate genotyping results are an important concern with scatology as important decisions need to be made based on the genotypic information obtained.

The aim of this study was to determine if scatology could be used as a non-invasive method for genetic analysis of cheetah in South Africa. The ex situ component of the study included determining the accuracy of identifying a species using faecal samples and the genotyping of faecal DNA samples relative to a blood standard to determine how the results will differ in accuracy. The in situ study was the field trial in Thabazimbi where potential wild cheetah faecal samples were collected and then analysed.

Quality and quantity of DNA from faecal material

Two faecal storage methods were used in this study namely storage in 99% ethanol and direct extraction without storage. By comparing the amplification success of samples stored in 99% ethanol and those that were extracted immediately, it would seem that storage in 99% ethanol results in better amplification success than extracting the samples immediately. This is, however, not conclusive because a number of repetitions of the PCR were done for the samples stored in ethanol as opposed to the one PCR done per sample that were not placed in a storage agent. A number of trials have been performed comparing different storage methods for example Frantzen *et al.* (1998) suggested

that storage in a DMSO/EDTA/Tris/salt solution (DETS) was most effective in preserving faecal samples. Another option is the drying of faecal DNA as a preservation method (Taberlet *et al.,* 1999). Other comparisons by Frantz et al. (2003) included freezing, storage in ethanol and storage in buffered DETS where the highest amplification success was also obtained from samples stored in ethanol. In the study by Fernando *et al*. (2000), samples stored in ethanol for up to four years provided DNA of sufficient integrity indicating no detectable degradation over long periods of time (Buchan *et al.*, 2005). Storage in ethanol does appear to be the accepted method (Fernando *et al.*, 2000, Lucchini *et al.,* 2002, Adams *et al.*, 2003, Frantz *et al.*, 2003).

For DNA extraction of faecal samples two methods are available. These were compared in the study by Frantz *et al.* (2003), namely the GuSCN/silica method and the QIAamp DNA stool mini kit which is an adaptation of the GuSCN/silica method. For both blood and hair the GuSCN/silica method yielded the best results. In this study the Qiagen® QIAamp stool mini kit (Southern Cross Biotechnology (Pty) Ltd, South Africa) was used for the extractions of all faecal samples. It is especially designed for the extraction of faecal DNA and includes an Inhibitex (Qiagen^{TM}) tablet which aids in the removal of inhibitors from the final DNA product. Inhibitors have mostly been found in plant secondary compounds in herbivore dung but there are also suspicions that inhibitors may be present in carnivore faeces, this may influence the success of DNA isolation and inhibit PCR. Although the occurrence of plant compounds in the cheetah faeces seems unlikely, grass and other plant remnants were found, further supporting the use of the QIAamp kit. The QIAamp kit has also been reported on in scatology studies of a similar nature such as on baboons and elephants (Buchan *et al.,* 2005).

The success of DNA extractions based on the purity of the DNA varied, this was expected as similar results have frequently been reported in other scatology studies (Taberlet *et al.,* 1999). There is general consensus between a number of researchers that changing environmental conditions, especially differences in humidity or exposure to sun or shade, have an effect on the quality of faecal DNA, therefore a variation in extraction success is expected (Frantz *et al.*, 2003). The probability of finding fresh faeces in the wild is low and therefore aged samples were included in this study. Samples included fresh faeces as well as four-day-old and two-week-old faecal samples that were aged in the sun. The results indicate that the quality of the DNA deteriorates rapidly over time. A similar observation was made by Frantz *et al*. (2003) that the time interval between defecation and collection has an effect on the DNA extraction success. Fernando *et al.* (2000), however, reported that sufficient DNA was extracted from elephant dung sampled up to eight days after defecation. In the study by Gottelli *et al*. (2007), faecal samples were collected directly after the cheetah had defecated and moved on, similarly, many other studies using faecal samples for genetic analysis used

fresh samples and found relatively good results when analyzing these samples (Ernest *et al.,* 2000, Fernando *et al*., 2000, Bhagavatula & Singh, 2006).

The faecal samples that were collected in the Thabazimbi area were mostly found in the shade except sample four which was lying approximately ten meters from the camera trap, in full sunlight. All of the samples collected from the wild were relatively old and dry; the only sample that was fresh was sample eight. Upon analyzing the results, there did not appear to be a difference between the fresh sample (number eight) and the remaining samples, but what is important to note is that even some of the very old dry samples did amplify successfully but not for all loci. Therefore, the time of exposure to the elements is an important factor when using scatology and fresh samples for analysis will most likely yield the best results.

Accuracy of species identification from faecal samples

The De Wildt Wild Cheetah Project has a trained sniffer dog for the detection of cheetah faeces in the Limpopo province. This Staffordshire terrier has been trained to find cheetah faeces exclusively and it has been proven on numerous occasions that a dog is more accurate than humans. For the pilot study, the need to prove that the scats obtained were from cheetah was not necessary as they were taken from cheetah enclosures at the De Wildt Cheetah and Wildlife Centre but collecting samples from wild areas is far more complicated as there are a number of carnivores of similar size and diet that are found in those areas, making accurate identification of cheetah faeces difficult. The samples at De Wildt served as a control to optimise the analyses of faecal samples in cases where scats collected where of uncertain origin. When faecal samples were collected from the Thabazimbi area for the field trial, the sample collectors were not always certain that the samples were in fact of cheetah origin, therefore notes were made for all the samples. Of special interest was sample four, possibly being of monkey origin as it contained seeds and sample seven possibly being of hyena origin as it was white in colour, a characteristic often associated with hyena faeces as their diet is rich in calcium. Another concern comes with the amplification of DNA from the faeces because it can contain DNA from the target species but also DNA from prey items, other members of the same species or bacteria (Adams *et al*., 2003). Whether or not the samples are located by a scat sniffing dog or by humans, a method to confirm that the scat did originate from the target animal is necessary.

A species-specific primer set was designed for high copy mtDNA to detect cheetah DNA and the primer set was optimised on DNA from cheetah blood. The results of the species-specific primer used with fresh faeces from De Wildt indicated a 100% positive identification. DNA from faeces that was four days old and two-week-old samples amplified with a lower but adequate success rate. There may

also be increased accuracy when using fresh samples as the DNA is presumably of better quality. The success rate obtained for the mitochondrial DNA is comparable with other studies such as 56-90% in brown bears (Murphy *et al.*, 2002 & Wasser *et al*., 1997), 91% in Eurasian otters (Hung *et al.*, 2004), 66.5% in Taiwanese macaque's (Chu *et al.*, 2006) and 83-89% in red wolves and coyotes (Adams *et al*., 2003).

DNA from the faeces found in the Thabazimbi area was also analysed using the species-specific primer set. This was useful not only for identifying cheetah faeces samples but also for determining the accuracy of species identification by sample collectors. The primer set identified 12 of the 15 samples as those belonging to cheetah. This emphasized the need not only for knowledge of scat morphology to verify that the scat was not from a sympatric carnivore, but also for a species-specific primer set. Based on the results, the sampling accuracy of the human collectors was 80%, enforcing the need for a species-specific primer set in these types of studies especially when the species of interest shares its environment with species of similar type, size and diet, this will ensure that time is not wasted and that costs are limited. Three of the samples did not amplify, two of these samples were already suspected to be monkey and hyena (four and seven).

Accuracy of genotyping faecal DNA samples

Success of PCR amplification of the 12 microsatellite markers loci varied between the faecal samples due to a lower quantity and quality of DNA compared to the blood. The PCR protocol was adapted by using a lower annealing temperature and a lower concentration of magnesium chloride, this resulted in a higher amplification success. Repetitions of PCR were performed up to five times per animal per loci.

A total of 972 amplifications were performed for all 12 microsatellite markers using the fresh faecal DNA (Table 3.6 and Table 3.8). Four repeats were performed per microsatellite locus per animal and an additional repeat was performed for animals 1, 7 to 11, 14, 17 and 18 at all 12 microsatellite markers (Table 3.8). These additional repeats were as a result of re-sampling and re-extraction of the samples to verify if storage length had a notable effect on DNA quality and quantity.

The 59% amplification success of the **fresh samples** is comparable with the success rates of other similar studies amplifying DNA from scats such as 59% in felids (Farrell *et al*., 2000) and 53% in wolves (Lucchini *et al.,* 2002), higher amplification rates were also reported, 63% in mountain lions (Ernest *et al*., 2000), 79% in coyotes (Kohn *et al*., 1999) and 70% in savannah baboons (Bayes *et al.*, 2000). Very low amplification results (19.1%) were found in seals (Reed *et al.,* 1997). Of the

successful faecal amplifications achieved in this study, 81% corresponded with the genotypes obtained from the matching blood samples (Table 3.6). The microsatellite markers used were designed to amplify domestic cat DNA and one could expect higher amplification success if microsatellites are designed to amplify cheetah DNA specifically.

When comparing the results of the **four-day-old** faecal DNA and **two-week-old** faecal DNA with those obtained from the fresh samples, it is clear that the age of the sample has an effect on PCR amplification. For the two-week-old faecal DNA, eight of the 324 amplifications were of sufficient quality and six of these genotypes corresponded with that of the blood (Table 3.7 and Table 3.9). The proportion of amplifications for the faecal DNA samples from which genotypes could be scored was lower than obtained with blood. This is reported in a number of studies using scats, due to degradation in faecal samples the template DNA contains very few copies of the target DNA sequence (Taberlet *et al*, 1996).

Genotype inconsistencies between blood and faecal DNA were due to allelic dropout and false alleles. This is also observed in other non-invasive studies (Creel *et al*., 2003, Buchan *et al*., 2005) as the DNA extracted from the samples is often degraded. Allelic dropout and other genotyping errors such as allelic false alleles, errors due to electrophoresis artefacts, incorrect scoring of allele banding patterns and data entry errors can also occur in blood samples but it is more commonly found in faecal samples or low quality DNA (Creel *et al*., 2003, Buchan *et al*., 2005). On average, 2-3 repeats were necessary to obtain the genotype of the reference blood sample. It is important that the reference genotype is reliable and that the genotyping results are clear for interpretation.

The **overall** allelic dropout rate in this study was found to be 16% and the occurrence of false alleles was 7% (Figure 3.3 and Table 3.8). The overall rate of allelic dropout and false alleles was calculated using the results from the fresh, four-day-old and two-week-old samples but the amplification success of the older samples was poor and therefore the allelic dropout and false allele rates are not reliable. There are insufficient results from the two-week-old samples and therefore no conclusions can be drawn from this study.

The rate of allelic dropout for **fresh samples** was 17% and the occurrence of false alleles was found to be 8% with notable differences between markers (Figure 3.3 and Table 3.7). Locus FCA 8 amplified with a good success rate but the allelic dropout observed at this locus was high (40%) and the occurrence of false alleles was 13.79%. Allelic dropout of as high as 51% and a rate of 13% for false alleles were reported by Chu *et al.* (2006) on Taiwanese Macaques. The allelic dropout was calculated as the number of allelic dropout occurrences over total successful amplifications resulting in

a lower value than if it had been calculated in the same manner as this study. Similar to this study, they also found great differences in the performance and results of different markers ranging from 13.3% to 63.8% for allelic dropout and 2.6% to 21% for occurrence of false alleles. Allelic dropout rates were found in other studies such as baboons 8-19% (Bayes *et al.*, 2000; Buchan *et al.* 2005), wolves 25% (Lucchini *et al.*, 2002), tigers 44-66% (Bhagavatula & Singh, 2006) Eurasian badgers 27% (Frantz *et al.*, 2003) and elephants 21% (Buchan *et al.*, 2005). From these results it can be seen that the allelic dropout of 17% found in this study is well within the range of similar studies using scatology.

In this study, FCA 78 had a zero rate of allelic dropout and no false alleles, but the amplification success of the locus was very poor with few successful PCRs after numerous changes in protocol to optimize the microsatellite. This microsatellite marker proved very difficult to use and gave unsatisfactory results.

The FCA 69 locus yielded very positive results with a high amplification success (60%) and low rates of allelic dropout and false alleles. This can possibly be attributed to the small fragment range of FCA 69. A regression for the influence of loci size on amplification success was calculated (Figure 3.6), it clearly shows how amplification success decreases as the range of the loci become larger. The study by Frantzen *et al.* (1998) supports this observation, where better amplification success for microsatellite markers with shorter alleles was found. Further studies using alternative microsatellite markers with smaller ranges may yield better amplification results.

In the study by Buchan *et al.* (2005) a clear relationship was found between median allele size and allelic dropout. They found that there was a trend for allelic dropout to be lower with smaller allele sizes. In this pilot study, median allele size did not seem to influence the rate of allelic dropout but this was due to two outliers at 195 and 200. If they were removed from the data set a clear trend can be seen that increasing median allele size negatively affects allelic dropout. Additionally, the markers used in this pilot study all had a relatively small allele range with the maximum locus size being 213bp. The relationship between median allele size and the occurrence of false alleles was also determined and the trend indicated that false allele occurrence decreased as the median allele size increased (Figure 3.7). With the use of additional microsatellite markers of more diverse ranges these results could be challenged.

The average occurrence of false alleles across all loci for fresh samples was found to be 8%. False alleles in other studies include wolves 5.6% (Creel, 2003), Eurasian badgers 8% (Frantz *et al.*, 2003), Barbary macaques 15.3% (Lathuilliere *et al.*, 2001) and Taiwanese macaques 12.7% (Chu *et al.*,

2006). Considering the values found for false alleles in other studies using scatology, a false allele rate of 8% is well within an acceptable range.

Notable differences in allelic dropout and false alleles were also seen in specific animals. For example, animal one performed relatively well compared to the other animals with an allelic dropout of 4.76% and 0% false alleles whereas animal 13 had an allelic dropout of 60% and 17.65% false alleles (Table 3.10). There could be numerous reasons for results such as this but the most important reason being the purity of the extracted DNA. A possible reason for the good results from animal one is that according to the quantification results obtained from the Nanodrop spectrophotometer (Table 3.1), the quantity of DNA was acceptable as was the purity but for animal 13 the quantity of DNA is suspiciously high from the first extraction from which most of the results were obtained, this high value could be due to an impure DNA extraction or bacterial DNA which is likely to cause poor genotyping results and increased frequency of errors such as the ones obtained for this animal.

Figure 3.4 shows how allelic dropout can vary between DNA extractions. In the example shown, there is allelic dropout of one allele (209) from one extraction from a fresh faecal sample and allelic dropout of the other allele (207) after another DNA extraction of the same fresh faecal sample. The four-day-old sample amplified correctly and its genotype corresponded with that of the blood sample. This example demonstrates the importance of DNA quality in the process of genetic profiling. In samples with low DNA quality such as faeces, it is important to perform numerous repeats of the PCR and DNA extraction if necessary, to determine the correct genotype especially when no blood controls are available.

Confidence intervals for allelic dropout and false alleles were determined. The range of a confidence interval is indicative of the certainty of the unknown parameter which in this case is the unwanted chance of experiencing allelic dropout or false alleles. The average range of the confidence interval found for allelic dropout is larger than the interval obtained for false alleles, indicating that more information is needed to make sound genotyping decisions to avoid allelic dropout than is needed for false alleles. For this reason a number of repeats to confirm a genotype is suggested by other researchers such as Taberlet & Luikart (1999). An interesting point to note is the difference between the confidence intervals found for fresh, four-day-old and two-week-old samples because the confidence interval ranges increase as the age of the samples increased for both allelic dropout and false alleles. This indicates that the older samples are more likely to be susceptible to allelic dropout and false alleles and that more repeats will render the samples more reliable.

The application of scatology in the wild

The locations of the faecal samples that were collected in the Thabazimbi area are shown on the map in Figure 4.1. The coloured circles indicate where certain samples were collected. After genotyping the Thabazimbi samples at the 12 microsatellite markers, matches were found between four animals at six loci (9 & 13 - red, 10 & 12 - blue). The remaining six loci did not amplify and could not be compared. These matches indicate a possibility that two animals had been sampled twice. By applying the method of scatology in the wild cheetah population, clear maps can be constructed giving an indication of each individual's territory. Samples 1 and 10, 2 and 3, 2 and 6 and 3 and 10 also matched but they could only be compared at 1-2 loci therefore it would not be accurate to assume that they are matching individuals. According to the probability of identity values obtained for increasing loci combinations (figure 3.10), three markers are sufficient. Although a match at six loci will give a very good idea of whether or not two samples are from the same animal, the ideal would be to be able to compare the genotypes at all 12 loci. With more frequent sampling and repetitions of amplifications at the 12 loci, more accurate data could possibly have been generated to determine if the matches were true.

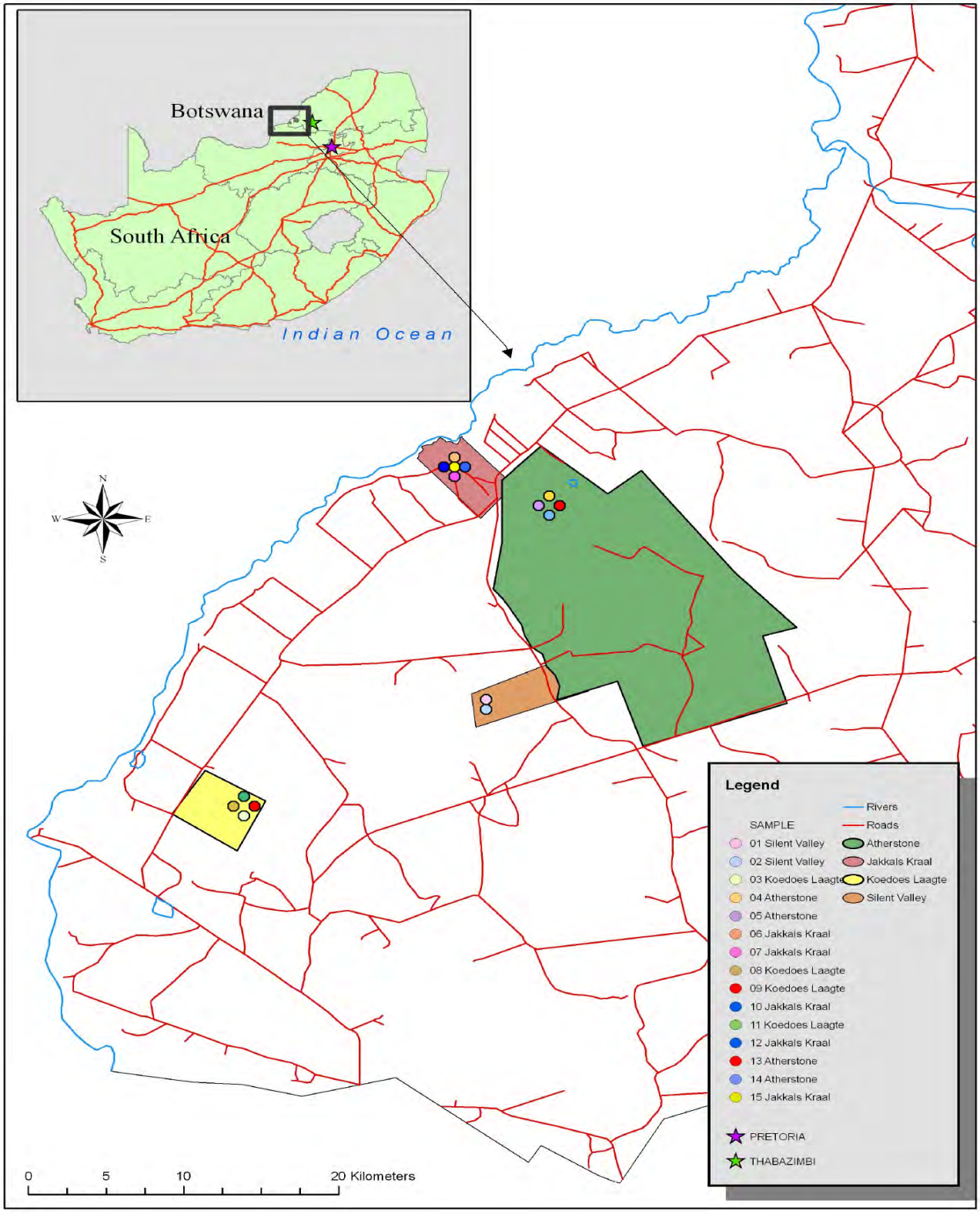


Figure 4.1 Map of Thabazimbi district (and location within South Africa) depicting sample collection areas

Genetic diversity parameters

The expected and observed heterozygosity values for the blood samples and the faecal samples are not a good indication of the trend in the general population because of the small sample size and because all of the samples come from De Wildt captive bred cheetah. The reason for the calculation of the observed and expected heterozygosities was to show that the values obtained for the faecal samples were lower than the blood samples. The reason for this is the genotyping errors (i.e. allelic dropout and false alleles) that occur when using non-invasive sources of DNA such as faeces. This decrease in heterozygosity is a clear indication of one of the dangers when using scatology because it gives an inaccurate indication of the genetic variation within a population.

The number of alleles per locus varied from 4.92 for the pilot study blood samples to 7.08 for the NZG samples. The lower value obtained for the blood samples is most likely due to the smaller number of samples. With the use of the genotypic information obtained for the 45 samples from the NZG, the results from the faeces could be verified by comparing the alleles obtained. Although the results from the NZG indicate additional alleles, this was expected because of the increased sample size. There is also an instance the pilot study has an additional allele at locus FCA 247. This allele will most likely be identified in the metapopulation study performed on approximately 400 cheetah.

The low F_{ST} values calculated by Structure indicate that the variation within cluster 1 (0.1457) is more than for cluster 2 (0.0126) signifying that there is more variation within the samples from the De Wildt pilot study animals than from the NZG samples. The Structure analysis indicated admixture occurring between the Thabazimbi animals whereas the De Wildt and NZG animals formed separate clusters (Figure 3.10). This is most likely due to sampling techniques as both sets of data originate from De Wildt.

Future prospects for the application of scatology for cheetah conservation

This pilot study concludes that molecular scatology will provide an important tool in studying the elusive South African cheetah population by using microsatellite markers to identify individuals. This genetic approach can contribute to determining if cheetah are present in a given area, produce a genetic identity for each individual in a population, estimate cheetah population size and it can be used to estimate genetic diversity and inbreeding and it can be used to trace pedigrees. With this information, decisions can be made by cheetah population managers to preserve the remaining gene pool and decrease the risk of inbreeding by translocating cheetah between reserves and conservation centres in a more scientific manner.

For a species that is threatened, wide-ranging, elusive and difficult to capture, this non-invasive method will be useful in gathering information without capturing or handling the animal. Future developments in the field of molecular scatology will undoubtedly further facilitate genetic analysis of cheetah populations and make a significant contribution to the conservation of this threatened species.

Chapter 5

Conclusion and Recommendations

Molecular scatology is still a developing field and holds much potential for the application of genetic analysis of wild, free-ranging animal populations. The genotypic data obtained can be useful in determining the presence of a species in an area, individual identification, sex, home range, kinship and providing information for determining the genetic variation within a population. It is extremely beneficial for species that are shy, elusive, threatened and wide-ranging. The results from this study demonstrate the potential use of molecular scatology for the conservation of cheetah in South Africa.

This study provides valuable guidelines for the future studies of cheetah using the non-invasive technique of scatology. The genotypic information from the 12 microsatellites can now be used as a standard when analyzing future samples from South African cheetah. Using this information, additional markers should be tested for improved amplification success, decreased rates of allelic dropout and occurrence of false alleles. It is also recommended that as stated in other studies, repetitions of DNA extractions and three to four amplifications should be done so that conclusive and validated decisions can be made as to the genotype of an individual at a specific locus. This has the disadvantage of significantly increasing costs and laboratory work.

The storage of samples in ethanol in this study and in others simplifies the task of sample collection and facilitates a reduction in sampling costs. It is also a practical method that can be employed by sample collectors in the field where no refrigeration is possible. In order to avoid unnecessary costs, species identification by species-specific primers should be performed first to avoid analyses of incorrect samples. There is general agreement between studies that PCR success and accurate genotyping are not only influenced by DNA quality and quantity but also other factors such as diet, preservation methods and freshness of samples. Freshness of samples played a role in this pilot study with fresh samples performing far better for species identification and for genotyping using microsatellite markers. The age of the samples and location could have an effect on the DNA extraction success and PCR amplification, it is therefore suggested that this is taken into account when collecting samples for analyses. Based on the results from the field trial, it is recommended that as many samples as possible are collected in order to obtain enough useful data. Factors to consider regarding the collection of these samples include the time taken for sample collections to done and the storage facilities that would be required.

Although there are many advantages to using non-invasive sampling, faecal samples yield inferior quality DNA when compared to blood or skin biopsies. The main limitation found in this study was the inferior amplification success, the occurrence of allelic dropout and occurrence of false alleles. These errors, when applied in a wild population could cause failure to distinguish individuals which would result in an under or over estimation of the population size. Inefficient markers could be replaced in future analyses with markers that are equally informative but show better amplification success, less allelic dropout and fewer false alleles in low quality DNA amplification. Repeated extractions to obtain a good DNA sample and a minimum of three to four microsatellite amplifications are recommended per sample for validating of the results.

The results indicate the potential use of cheetah faeces as a source of DNA for population studies of cheetah in South Africa in addition to the methods currently employed such as camera trapping and telemetry to provide a better understanding of the spatial distribution and levels and structure of genetic variation among wild cheetah in South Africa.

Summary

Cheetah are difficult to census and as a result there is only a limited amount of information regarding the status of cheetah populations across South Africa making it difficult to identify and address potential threats and to plan effective conservation strategies. The aim of this study was to determine if scatology could be used as a non-invasive method for genetic analysis of cheetah in South Africa.

Corresponding blood and faecal samples were collected from cheetah at the De Wildt Cheetah Centre. The faecal samples were aged and then collected and stored in ethanol after four days and after two weeks. Faecal samples of suspected wild cheetah were also collected in the Thabazimbi area in the Limpopo Province. DNA was than isolated from all the blood and faecal samples.

A species-specific primer set was developed in the mitochondrial control region to reliably identify cheetah faecal samples. The PCR amplification success of the species-specific marker was high, all fresh faecal samples had a positive result (100%), four-day-old samples were 83% positive and two-week-old samples were 94% positive. The species-specific primer was found to be effective in accurate cheetah identification and had greater amplification success on fresh samples. Twelve microsatellite markers were used to profile the blood and faecal samples. Multiple PCRs were performed per faecal sample per marker to ensure that the genotype obtained was correct. Amplification success of the 12 microsatellites was 59% for fresh samples, 20% for four-day-old samples and 2.4% for two-week-old samples.

Difficulties associated with the technique included poor DNA quality and quantity and genotyping errors, resulting in allelic dropout and false alleles. Genotyping errors were determined by comparing the genotype of the faeces with that of the matching blood sample. Overall, allelic dropout experienced was 16% and 7% for false alleles. These values correspond to the results of similar studies.

Samples of suspected wild cheetah were collected in the Thabazimbi area and were analysed. The alleles obtained were similar to those from the pilot study. Data of 45 De Wildt cheetah samples obtained from the NZG were analysed and compared with the results of the pilot study, as expected, all alleles in the De Wildt population occurred in the NZG samples. The alleles found in this study and in the NZG samples can be used as the standard when using molecular scatology techniques.

The results of this study provide valuable guidelines for future studies of cheetah using the non-invasive technique of scatology. The genotypic information from the 12 microsatellites can now be used as a standard when analyzing samples of cheetah origin. Using this information, additional markers can now be tested for improved amplification success and decreased rates of allelic dropout and occurrences of false alleles.

The results prove the feasibility of using cheetah faeces as a potential source of DNA for population studies of cheetah in South Africa in addition to the methods currently employed.

References

Adams, J.R., Kelly, B.T. & Waits, L.P., 2003. Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). Molecular Ecology 12: 2175–2186.

Amos, A. & Harwood, J., 1998. Factors affecting levels of genetic diversity in natural populations. Philosophical Transactions of the Royal Society of London B 353: 177-186.

Applied Biosystems, 2000. Automated DNA Sequencing Chemistry Guide. www.appliedbiosystems/techsupport.com.

Avise, J.C., 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York, NY 10003, USA.

Bayes, M.K., Smith, K.L., Alberts, S.C. & Bruford, M.W., 2000. Testing the reliability of microsatellite typing from faecal DNA in savannah baboon. Conservation Genetics 1: 173-176.

Bhagavatula, J. & Singh, L., 2006. Genotyping faecal samples of Bengal tiger *Panthera tigris tigris* for population estimation: A pilot study. BMC Genetics 7:48 doi: 10.1186/1471-2156-7-48.

Bottriell, L.G. 1987. King cheetah: the story of the quest. EJ Brill, Leiden.

Bouwer, V., 2003. Letting cheetah roam free. Farmers Weekly 10:36.

Broquet, T. & Petit, E., 2004. Quantifying genotyping errors in non-invasive population genetics. Molecular Ecology 13: 3601-3608.

Bruford, M.W. & Wayne, R.K., 1993. Microsatellites and their application in population genetic studies. Current opinion in Genetics and Development 3: 939-943.

Buchan, J.C., Archie, E.A., van Horn, R.C., Moss, C.J. & Alberts, S.C., 2005. Locus effects and sources of error in non-invasive genotyping. Molecular Ecology Notes 5: 680-683.

Burger, P.A., Steinborn, R., Walzer, C., Petit, T., Mueller, M. & Schwarzenberger, F., 2004. Analysis of the mitochondrial genome of cheetah (*Acinonyx jubatus*) with neurodegenerative disease. Gene 338 (1): 111-119.

Cat Specialist Group, 2002. *Acinonyx jubatus*. In: IUCN 2004. 2004 IUCN Red List of Threatened Species.

Chu, J., Lin, Y. & Wu, H., 2006. Applicability of Non-Invasive Sampling in Population Genetic Study of Taiwanese Macaques (*Macaca cyclopis*). Taiwania 51 (4): 258-265.

Convention on International Trade in Endangered Species of Wild Flora and Fauna. 1973. Endangered Species Act. United Nations International Law 93-205, Title 50: Part 23.

Cooper, A.L., 1926. A curious skin. Field 147: 690.

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 non-invasive microsatellite genotypes. Molecular Ecology 12: 2003-2009.

Dallas, J.F., Coxon, K.E., Sykes, T., Chanin, P.R.F., Marshall, F., Carss, D.N., Bacon, P.J., Piertney, S.B. & Racey, P.A., 2003. Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter *Lutra lutra*. Molecular Ecology 12 (1): 275.

Driscoll, C.A., Menotti-Raymond, M., Nelson, G., Goldstein, D. & O'Brien, S.J., 2002. Genomic microsatellites as evolutionary chronometers: A test in wild cats. Genome Research 12: 414-423.

Durant, S.M., 1998. Competitive refuges and coexistence: an example from Serengeti carnivores. Journal of Animal Ecology 67: 370-386.

Durant, S.M., 2000. Predator avoidance, breeding experience and reproductive success in endangered cheetah, Acinonyx jubatus. Animal Behaviour 60: 121–130.

Ernest, H.B., Penedo, C.T., May, B.P., Syvanen, M. & Boyce, W.M., 2000. Molecular tracking of mountain lions in the Yosemite Valley region in California: genetic analysis using microsatellites and faecal DNA. Molecular Ecology 9: 433-441.

Ennis, S. & Gallagher, T.F., 1994. A PCR-based sex determination assay in cattle based on the bovine amelogenin locus. Animal Genetics 25: 425–427.

Evanno, G., Regnaut, S. & Gaudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Farrell, L.E, Roman, J. & Sunquist. M.E., 2000. Dietary separation of sympatric carnivores identified by molecular analysis of scats. Molecular Ecology 9: 1583-1590.

Farrell, L.E, 2001. Molecular scatology as a conservation tool. Endangered species update 18: 4.

Fernando, P., Pfrender, M.E., Encalada, S.E. & Lande, R., 2000. Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. Heredity 84: 362–372.

Frantz, A.C., Pope, L.C., Carpenter, P.J., Roper, T.J., Wilson, G.J., Delahay, R.J. & Burke, T., 2003. Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Molecular Ecology 12 (6): 1649.

Frantzen, M.A.J., Silk., J.B., Ferguson, J.W.H., Wayne, R.K. & Kohn, M.H., 1998. Empirical evaluation of preservation methods for faecal DNA. Molecular Ecology 7: 1423-1428.

Glémin, S., Vimond, L., Ronfort, J., Bataillon, T. & Mignot, A., 2006. Marker-based investigation of inbreeding depression in the endangered species *Brassica insularis*. Heredity 97: 304-311.

Gottelli, D., Wang, J., Bashir, S. & Durant, S., 2007. Genetic analysis reveals promiscuity among female cheetah. Proceeding of the Royal Society B doi:10.1098/rspb.2007.0502.

Hall, T.A., 1999. Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Ser. 41: 95-98.

Hansen, M.M. & Jacobsen, L., 1999. Identification of mustelid species: Otter (*Lutra lutra*), American mink *(Mustela vison*) and polecat (*Mustela putorius*), by analysis of DNA from faecal samples. Journal of Zoology 247 (2):177-181.

Hedrick, P.W., 1996. Bottleneck(s) or Metapopulation in Cheetahs. Conservation Biology 10:3.

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

Höss, M., Kohn, M. & Pääbo, S., 1992. Excrement analysis by PCR. Nature 359:199.

Hunter, L.T.B. & Hamman, D., 2003. Cheetah. Cape Town: Struik Publishers.

Hung, C.M., Li, S.H. & Lee, L.L., 2004. Faecal DNA typing to determine the abundance and spatial organisation of otters (*Lutra lutra*) along two stream systems in Kinmen. Animal Conservation 7: 301-311.

Ingman, M., Kaessmann, H., Paabo, S. & Gyllensten, U., 2000. Mitochondrial genome variation and the origin of modern humans. Nature 408 (6813): 708-713.

Jarne, P. & Lagoda, P.J.L., 1996. "Microsatellites, from molecules and populations and back." Trends in Evolution and Ecology 11 (10): 424-429.

Jiminez, J.A., Hughes, K.A., Alaks, G, Graham, L. & Lacy R.C., 1994. An experimental study of inbreeding depression in a natural habitat. Science 266: 271-73.

Karp, K. & Edwards, K.J., 1998. DNA markers: a global overview. G. Caeteno-Annole's and Gresshoff, P.M. (Eds). DNA markers: protocols, applications and overviews. Wiley, VCH, New York.

Kohn, M.H. & Wayne, R.K., 1997. Facts from faeces revisited. Tree 12: 223-227.

Kohn, M.H., York, E.C. & Kamradt, D.A, 1999. Estimating population size by genotyping faeces. Proceedings of the Royal Society of London B 266: 657-663.

Kurose, N., Masuda, R. & Tatara, M., 2005. Faecal DNA Analysis for Identifying Species and Sex of Sympatric Carnivores: A Non-invasive Method for Conservation on the Tsushima Islands, Japan**.** Journal of Heredity 96 (6):688-697; doi:10.1093/ jhered/esi124.

Lathuilliere, M., Menard, N. Gautier-Hion, A. & Crouau-Roy, B., 2001. Testing the reliability of non-invasive genetic sampling by comparing analyses of blood and faecal samples in Barbary macaques (*Macaca sylvanus*). American Journal of Primatology 55 (3): 151-158.

Liu, Z.J. & Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238:1-37.

Lucchini, V., Fabbri, E., Marucco, F., Ricci, S., Boitani, B. & Randi, E., 2002. Non-invasive molecular tracking of colonizing wolf (Canis lupus) packs in the western Italian Alps. Molecular Ecology 11:857-868.

Luikart, G. & England, P.R., 1999. Statistical analysis of microsatellite DNA data. Tree 14 (7):253-256.

Marker, L.L., 2002. Aspects of cheetah *Acinonyx jubatus* biology, ecology and conservation strategies on Namibian farmlands Ph.D. thesis, University of Oxford, Oxford.

Marnewick, K., Cilliers, D., 2006. Range use of two coalitions of male cheetah, *Acinonyx jubatus*, in the Thabazimbi district of the Limpopo Province, South Africa. South African Journal of Wildlife Research 36 (2): 147-151.

Marnewick, K., Beckhelling, A., Cilliers, D., Mills, G., Herring, K., Caldwell., Hall, R. & Meintjes, S., 2007. The status of the cheetah in South Africa. Submitted to the IUCN Cat Specialist Group. To appear in a special issue of Cat News: The Status of the Cheetah in Southern Africa.

Massicot, P., 2005. Animal info – Cheetah. © 1999 - 2005 Animal Info. http://www.Animalinfo.org/species/carnivor/acinjuba.htm.

May, R.M, 1995. The cheetah controversy. Nature 374: 309-310.

Mendenhall, W. & Ott, L., 1980. Understanding Statistics, Third Edition. Duxbury Press, A division of Wadsworth, Inc., Belmont, California, 94002.

Menotti-Raymond, M. & . O'Brien, S. J., 1993. Dating the genetic bottleneck of the African cheetah. Proceedings of the National Academy of Science USA 90: 3172-3176.

Menotti-Raymond, M., David, V.A., Lyons, L. A., Schäffer, A. A., Tomlin, J. F. and O'Brien, S. J., 1999. A Genetic Linkage Map of Microsatellites in the Domestic Cat (*Felis catus*). Genomics 57: 9-23.

Menotti-Raymond, M. & O'Brien, S.J., 1999. Evolutionary Conservation of ten microsatellite markers in Felidae. Journal of Heredity 86(4): 319-321.

Miller, C.R., Joyce, P. & Waits, L.P., 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160: 357-366.

Morin, P.A., Chambers, K.E., Boesch, C. & Vigilant, L., 2001. Quantitative polymerase chain reaction analysis of DNA from non-invasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Molecular Ecology 10: 1835-1844.

Murphy, M.A., Waits, L.P., Kendall, K.C., Wasser, S.K., Higbee, J.A. and Bogden, R., 2002. An evaluation of long-term preservation methods for brown bear (*Ursus arctos*). Molecular Ecology 12: 2261-2265.

Nowell, K. & Jackson, P., 1996. Wild cats: Status survey and conservation action plan. Burlington Press, Cambridge.

O'Brien, S.J., Wildt, D.E., Goldman, D., Merril, C.R .& Bush, M., 1983. The cheetah is depauperate in genetic variation. Science 221:459.

O'Brien, S.J., Roelke, M.E., Marker, L., Newman, A., Winkler, C.A., Meltzer, D., Colly, L., Evermann, J.F., Bush, M. & Wildt, D.E., 1985. Genetic basis for species vulnerability in the cheetah. Science: New series 227(4693): 1428.

O'Brien, S.J., Wildt, D.E., Bush, M., Caro, T. M., Fitzgibbon, C. , Aggundey, I. & Leakey, R.E., 1987. East African cheetahs: evidence for two bottlenecks? Proceedings of the National Academy of Science USA 84:508-511.

O'Brien, S.J., 1991. Mammalian genome mapping: Lessons and prospects. Current opinion in Genetics and development 1:105-111.

O'Brien, S.J., 1994. A role for molecular genetics in biological conservation. Proceeding of the National Academy of Sciences of the United States of America 91(13): 5748-5755.

O'Regan, H.J, 2002. Defining cheetah, a multivariate analysis of skull shape in big cats. Mammal Review 32(1): 58.

Park, S., 1999. The microsatellite toolkit for MS Excel 97 or 2000. Molecular population genetics laboratory, Smurfit Institute of Genetics, Trinity College Dublin 2, Ireland.

Paxinos, E., McIntosh, C., Ralls, K. & Fleischer, R., 1997. A non-invasive method for distinguishing among canid species: amplification and enzyme restriction of DNA from dung. Molecular Ecology 6:483-486.

Peakall, R. & Smouse, P.E., 2006. GENALEX6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295.

Pfeiffer, I. & Brenig, B., 2005. X- and Y-chromosome specific variants of the amelogenin gene allow sex determination in sheep (*Ovis aries*) and European red deer (*Cervus elaphus*). BMC Genetics 6: 16.

Pocock, R.I., 1927. Description of a new species of cheetah (*Acinonyx*). Proceedings of the Zoological Society of London 1: 245-252.

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

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Reed, J.Z., Tollit, D.J., Thompson, P.M. & Amos, W., 1997. Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. Molecular Ecology 6:225-234.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K. B., Horn, U.T., Erlich, H. A. & Arnheim, N.,1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. Science 230: 1350-1354.

Sainudiin, R., Durrett, R.T., Aquadro, C.F. & Nielsen, R., 2004. Microsatellite mutation models: Insights from a comparison of humans and chimpanzees. Genetics 168: 383-395.

Samuels, M.R. & Witmer, J.A., 2003. Statistics for the Life Sciences (3rd Edition), Prentice Hall, NJ, Upper Saddle River.

Skinner, J.C., & Chimimaba, C.T., 2005. The Mammals of the Southern African Subregion. Cambridge University Press, The Water Club, Beach Road, Granger Bay, Cape Town, 8005, South Africa.

Smith, A., Ralls, K., Hurt, A., Adams, B., Parker, M., Davenport, B., Smith, M.C. & Maldonado, J.E., 2003. Detection and accuracy rates of dogs trained to find scats of San Joaquin kit foxes (*Vulpes macrotis mutica*). Animal conservation 6: 339-346.

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

Sullivan, K.M., Mannucci, A., Kimpton, C.P. & Gill, P., 1993. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 15: 636–641.

Sunnucks, P., 2000. Efficient genetic markers for population biology. Tree 15(5):199-203.

Taberlet, P., Griffin, S. & Goossens, B., 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research 24: 3189-3194.

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

Taberlet, P., Waits, L.P. & Luikart, G., 1999. Non-invasive genetic sampling: look before you leap. Trends in Ecological Evolution 14: 323-327.

Tautz, D., 1989. Hypervariability of Simple Sequences as a General Source for Polymorphic DNA Markers. Nucleic Acids Research 17:6463-6471.

Templeton, A.R. & Read, B., 1983. Elimination of inbreeding depression from a captive population of Speke's gazelle: validity of the original statistical analysis and confirmation by permutation testing. Zoo Biology 17: 77-98.

Valverde, J.R., Marco, R. & Garesse, R., 1994. A conserved heptamer motif for ribosomal RNA transcription termination in animal mitochondria. Proceedings of the National Academy of Science U.S.A. 91(12): 5368-5371.

Van Aarde, R.J. & van Dyk, A., 1986. Inheritance of the king coat colour pattern in cheetahs Acinonyx jubatus. Journal of Zoology, London 209: 573-578.

van Dyk, A., 1991. Inheritance of the king coat colour patterns in cheetahs *Acinonyx jubatus*. Journal of Zoology. 209:573-578.

Wasser, S.K., Houston, C.S., Koehler, G.M., Cadd, G.G., Fain, S.R., 1997. Techniques for application of faecal DNA methods to field studies of Ursids. Molecular Ecology 6: 1091-1097.

Wayne, R.K., Modi, W.S. & O'Brien, S.J., 1986. Morphological variability in the cheetah (*Acinonyx jubatus*), a genetically uniform species Evolution 40: 78-85.

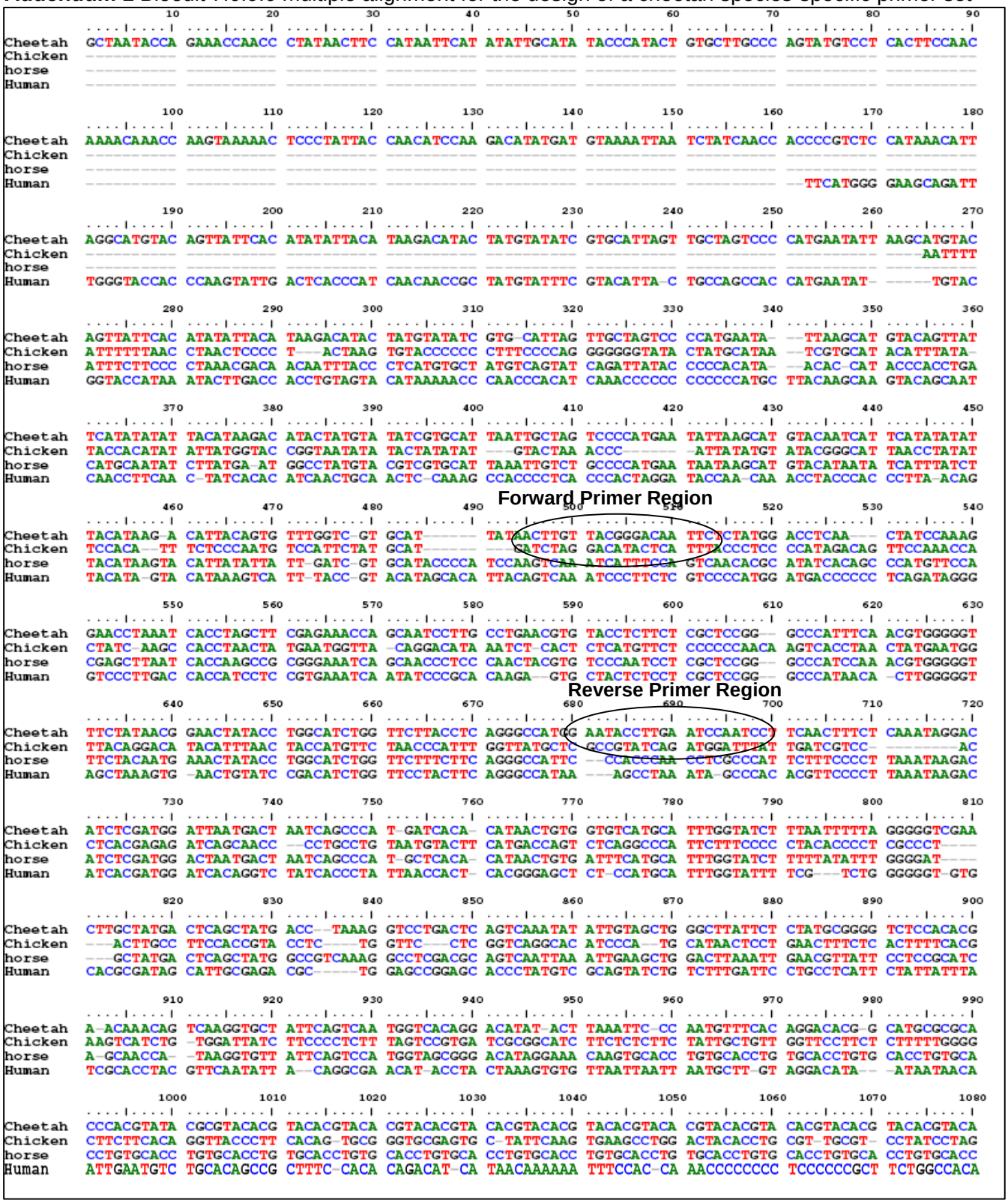
Xu, X. & Arnason,U., 1994. The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. Gene 148 (2): 357-362.

Yalden, D.W., Largen, M.J. & Kock, D., 1980. Catalogue of the mammals of Ethiopia. 4. Carnivora. Monitore Zool. Ital. (Supplement) 19: 67-221.

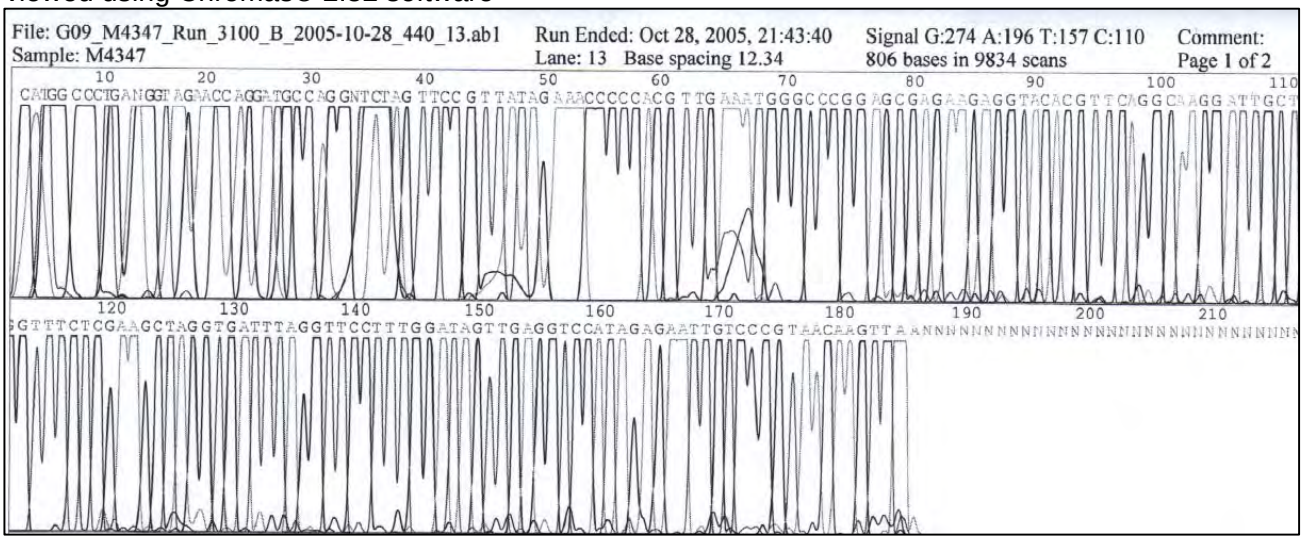
Yeates, C., Gillings, M.R., Davison, A.D., Altavilla, N. & Veal, D.A., 1998. Methods for microbial DNA extraction from soil for PCR amplification. Biological Procedures Online 1 (1).

Addenda

Addendum 1 Bioedit 7.0.9.0 multiple alignment for the design of a cheetah species-specific primer set



Addendum 2 Electropherogram of a random sample amplified using the cheetah specific marker viewed using Chromas® 2.32 software



Addendum 3 Genotypic information for blood and faecal samples of cheetah individuals from De Wildt Cheetah Centre *

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F01B	102	191	118	127	203	167	159	152	121	192	149	205
	104	195	118	135	207	173	163	154	123	210	161	207
F01R1	102	191	118	127	203	167		152	121		149	205
	104	195	118	135	207	173		154	123		161	207
F01R2	102	191	118	127	203	167	159	152			149	205
	104	195	118	135	207	173	163	154			161	207
F01R3	102	191	118	127	203	167			121			
	104	195	118	135	207	173			123			
F01R4	102	191	118	127	203	167	159	152		192	149	207
	104	195	118	135	207	173	163	154		210	161	207
F01R5		191	118	127	203	167	159	152			149	207
		195	118	135	207	173	163	154			161	207
F01D4R1												
F01D4R2												
F01D14R1												
F01D14R2												
F02B	102	193	118	141	211	167	163	138	121	202	151	199
	102	199	118	143	211	169	163	150	131	204	155	207
F02R1	102	193	118	141	211	167	163	138	121		149	199
	102	193	118	143	211	167	163	150	131		151	207
F02R2	102	193	118	141	211	167	163	138	121	202	149	199
	102	199	118	141	211	169	163	150	131	204	151	207
F02R3		193	118	141	211	167		138		202		199
		199	118	143	211	169		150		204		207
F02R4	102	193	118	141	211	167	163	138	121	202	151	199
	102	199	118	143	211	169	163	150	131	204	155	207
F02D4R1	102										149	
	102										149	
F02D14R1												
F03B	102	191	118	133	209	169	159	142	121	200	149	199
	104	195	118	133	213	173	163	150	131	204	149	201
F03R1	102	191	118	133	209	169	159	142		200		199
	104	195	118	133	213	173	163	150		204		201
F03R2	102	191	118	133	209	169	159	142	121	200	149	
	104	195	118	133	213	173	163	150	131	204	149	
F03R3	102	191	118	133	209			142		200	149	199
	104	195	118	133	213			150		204	149	199

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F03R4	102	191	118	133	209	169	159	142	121	200	149	
	104	195	118	133	213	173	163	150	131	204	149	
F03D4R1	102	191	118	133			159	142			149	
	104	195	118	133			163	150			149	
F03D14R1												
F04B	102	191	118	127	203	167	157	154	127	202	151	207
	104	191	128	133	213	169	165	154	131	204	155	211
F04R1	102	191					157		127	202	149	
	104	191					157		131	204	149	
F04R2	102	191	116			167			131	202	149	
	104	191	118			167			131	204	149	
F04R3	102	191	118	127		167						207
	104	191	118	133		167						207
F04R4		191	118				157	154	127			
		191	128				157	154	127			
F04D4R1		191										
		191										
F04D14R1												
F05B	102	191	128	127	203	165	157	150	121	202	151	199
	104	191	128	133	211	167	163	150	133	204	151	207
F05R1	102	191	116	133		165	157	150			149	201
	104	191	118	133		167	163	150			149	201
F05R2	102	191	118	137		165	157	150	121		149	
	104	191	128	143		167	163	150	133		149	
F05R3	102	191	116	127	203	165						
	104	191	128	133	211	167						
F05R4	102	191	128					138	121			
	104	191	128					150	133			
F05D4R1												
F05D14R1												
F06B	100	191	118	137	207	169	163	150	123	202	149	201
	104	193	118	143	209	173	163	154	131	210	151	209
F06R1	100	191	118	137		169			123		149	
	104	193	118	137		173			131		149	
F06R2	100	191	118	137				150	123		149	
	104	193	118	143				154	131		149	
F06R3	100	191	118		207	169					149	
	104	193	118		207	173					149	
F06R4	102	191	118		209	173	163	150			149	
	104	193	118		209	173	163	154			151	
F06D4R1	100	191		137	207	169		150			149	
	104	193		137	209	173		154			151	

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F06D14R1						169						
						173						
F07B	102	193	118	135	203	167	163	150	127	192	149	187
	104	193	118	143	211	169	165	150	131	200	149	201
F07R1	102	193	118	135		167		150	127	192	149	187
	104	193	118	143		169		150	131	200	149	201
F07R2	102		118			167		150	127		149	187
	104		118			169		150	131		149	187
F07R3	102	193	116	135		169					149	
	104	193	118	135		169					149	
F07R4	102	193		135	203			150			149	
	104	193		143	211			150			149	
F07R5												
F07D4R1	102	193						150			149	
	102	193						150			149	
F07D4R2												
F07D14R1								150				187
								150				201
F07D14R2												
F08B	100	193	118	127	203	165	159	150	131	192	149	187
	104	193	128	143	209	169	163	150	131	206	149	207
F08R1	100	191	118					150	131		149	
	104	191	128					150	131		149	
F08R2	100	191				165	159	150	131		149	187
	104	193				165	163	150	131		149	187
F08R3	100	193	118	127	203	165			131		149	
	104	193	128	127	209	169			131		149	
F08R4	100	193	118	127		173	159	150	131		149	
	104	193	128	127		173	163	150	131		149	
F08R5	100		118						131		149	
	104		128						131		149	
F08D4R1												
F08D4R2	100	193	118	127				150	131			
	104	193	128	143				150	131			
F08D14R1												
F08D14R2												
F09B	102	191	116	127	211	165	157	138	123	200	149	199
	104	191	118	127	211	167	161	152	123	200	149	201
F09R1	102	191	116	127		165	157	138	123	200	149	199
	104	191	118	127		167	161	152	123	200	149	201

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F09R2	102	191	116	127			157	138	123	200	149	199
	104	191	118	127			161	152	123	200	149	201
F09R3	102	191	116	127	211						149	
	104	191	118	127	211						149	
F09R4	102	191	116	127	211		157	138	123		149	199
	104	191	118	127	211		161	152	123		149	199
F09R5												
F09D4R1	102							138				
	102							152				
F09D4R2												
F09D14R1		191		127							149	
		191		127							149	
F09D14R2												
F10B	102	191	118	127	203	169	163	150	131	192	149	187
	104	193	128	143	211	169	163	154	131	210	155	201
F10R1	102	191					163	150	131		149	187
	104	193					163	150	131		149	201
F10R2	102	191	116	135					131		149	187
	104	193	118	137					131		149	201
F10R3	102	191	118	127	203						149	
	104	193	128	127	211						149	
F10R4	102		118				163	150	131			
	104		128				163	154	131			
F10R5	102	191	118	127	203	169		150	131			187
	104	193	128	143	211	169		154	131			201
F10D4R1	102	191									149	187
	104	193									149	201
F10D4R2		191	118	127							149	
		193	128	143							155	
F10D14R1												
F10D14R2												
F11B	102	191	118	133	209	165	161	138	131	192	149	199
	104	193	118	137	209	173	165	150	131	200	149	207
F11R1	102	191	118	133	203	165	161	138	131		149	199
	104	193	118	137	209	173	165	150	131		149	207
F11R2	102	191	118	133		165	161	138	131		149	199
	104	193	118	137		173	165	150	131		149	207
F11R3	102	193		133		165					149	
	104	195		137		165					149	
F11R4	102	191	116	133		165		138		192	149	199
	104	193	118	137		173		150		200	149	207

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F11R5												
F11D4R1	102	191	118					138	131		149	
	104	193	118					150	131		149	
F11D4R2												
F11D14R1												
F11D14R2												
F12B	102	191	118	127	203	165	157	138	121	200	151	199
	104	193	118	137	209	167	157	150	127	206	159	199
F12R1				137		165	157	150			149	
				137		167	157	150			149	
F12R2				137		165		150			149	
				143		167		150			149	
F12R3	102	191		127	203	165					151	
	104	193		137	209	165					159	
F12R4	102	191		127	203	165	157				151	199
	104	193		137	209	165	157				159	199
F12D4R1											149	
											149	
F12D14R1												
F13B	102	191	118	127	203	165	157	150	121	200	149	199
	104	193	128	137	209	165	163	150	127	206	159	209
F13R1	102	191		137				150				
	104	193		137				150				
F13R2	102	191										
	104	193										
F13R3	102	193	118		209							
	102	195	118		209							
F13R4	102	193	118		209		163	138				209
	102	195	118		209		163	150				209
F13D4R1				127		165						
				127		165						
F13D14R1												
F14B	102	193	118	127	209	169	163	138	121	202	155	199
	102	195	118	133	211	173	163	150	131	204	155	199
F14R1	102	193					163	138	121			
	102	195					163	150	131			
F14R2	102						163	138	121			199
	102						163	150	131			199
F14R3	104			127		169						
	104			133		173						

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F14R4	104	193	118	127		173		150			155	
	104	195	118	133		173		150			155	
F14R5	102	193	118	127	209	169		138				199
	102	195	118	133	211	173		150				199
F14D4R1			118	127	209			138	121		155	199
			118	133	211			150	131		155	199
F14D4R2												
F14D14R1												
F14D14R2												
F15B	104	193	118	133	207	165	157	150	121	192	151	207
	104	195	128	137	207	165	157	150	127	206	163	211
F15R1	102	191										
	102	193										
F15R2	104					165		150				
	104					165		150				
F15R3	104	191		133	207	165					151	
	104	193		137	207	165					163	
F15R4	104	193		133	207	165	157	150			151	
	104	195		137	207	165	157	150			163	
F15D4R1				133	207			150				
				137	207			150				
F15D14R1												
F16B	104	191	128	127	203	165	157	138	121	202	149	207
	104	193	128	127	211	165	161	138	131	206	151	211
F16R1	104	191		127	203	165	157	138	121		149	
	104	193		127	211	165	161	150	131		149	
F16R2	104	191			203	165	157		121		149	
	104	193			211	165	161		131		149	
F16R3	104	191	118	127		165			121			
	104	193	128	127		165			131			
F16R4	104	191	118					138	121		149	
	104	193	118					138	131		151	
F16D4R1								138			149	
								138			151	
F16D14R1												
F17B	102	195	118	127	209	167	163	150	121	200	149	199
	104	195	118	135	213	167	163	150	123	210	151	199
F17R1	102						163	150			149	199
	104						163	150			149	199
F17R2	102		118					150			149	199
	104		118					150			149	199

Laboratory Number	FCA 69	FCA 105	FCA 126	FCA 133	FCA 220	FCA 224	FCA 247	FCA 26	FCA 310	FCA 78	FCA 8	FCA 96
F17R3	102	191	118	127	209						149	
	104	193	118	135	209						149	
F17R4	102	191	118	127				150	121		149	199
	104	193	128	135				150	121		149	199
F17R5		195	118								149	
		195	118								151	
F17D4R1	102	195	118	127	209	167					149	
	104	195	118	135	213	167					151	
F17D4R2												
F17D14R1		191									149	
		193									149	
F17D14R2												
F18B	100	193	118	137	209	169	163	150	131	192	149	187
	102	193	128	137	211	169	165	154	131	210	155	201
F18R1	100	193	118	137			163	150	131		149	187
	102	193	128	137			163	154	131		149	201
F18R2	100	193	118				161	150	131		149	187
	102	193	128				163	154	131		149	201
F18R3	100	193	118	137	209	169		150			149	
	102	193	128	137	211	169		154			155	
F18R4	100	193		137	209		163	150	131	192	149	185
	102	193		137	211		165	154	131	210	155	199
F18R5	100	193	118	137		169		150	131		149	187
	102	193	128	137		169		154	131		155	201
F18D4R1												
F18D4R2												
F18D14R1												
F18D14R2												

* B=Blood sample, R1-Rx=repeat number of microsatellite amplifications of fresh faecal sample, D4=4-day-old faecal sample, D4R1-D4Rx= repeat number of microsatellite amplifications of 4-day-old faecal sample, D14=2-week-old sample, D14R1-D14Rx= repeat number of microsatellite amplifications of 2-week-old faecal sample, cells without data indicate inability to genotype the sample

Addendum 4 Genotypic information obtained from the NZG of 45 De Wildt cheetah samples.

Laboratory Number	FCA 96	FCA 247	FCA 26	FCA 78	FCA 105	FCA 69	FCA 310	FCA 133	FCA 220	FCA 8	FCA 224	FCA 126
GACH 140/05	199	161	148	190	191	100	121	135	211	153	167	122
	201	161	148	200	193	102	127	135	211	153	167	122
GACH 151/05	207	157	148	192	195	104	131	127	203	149	167	118
	209	163	154	200	195	108	131	127	211	149	171	128
GACH 152/05	199	155	148	192	195	104	121	135	213	149	169	118
	199	161	148	200	195	108	123	141	215	155	173	118
GACH 153/05	207	155	136	200	191	102	121	133	203	151	169	118
	207	163	154	202	193	102	131	143	209	155	171	128
GACH 142/05	199	155	136	192	191	102	121	129	211	149	165	122
	207	155	136	200	193	104	121	133	215	163	167	122
GACH 155/05	209	155	154	200	191	102	121	127	203	149	169	118
	211	163	154	202	191	106	131	133	209	155	175	118
GACH 156/05	207	161	148	206	191	100	121	127	203	149	167	118
	209	163	152	210	193	102	131	127	203	155	171	128
GACH 158/05	199	155	148	206	195	108	121	127	207	149	165	118
	205	161	154	206	195	108	123	135	211	151	171	120
GACH 160/05	205	155	146	204	191	108	123	133	207	151	169	128
	211	161	148	208	197	108	123	137	211	151	171	128
GACH 161/05	199	159	136	202	193	102	129	127	203	149	167	118
	207	161	154	210	195	106	131	143	211	151	169	128
GACH 162/05	199	155	136	200	191	102	121	127	209	153	167	118
	205	155	136	202	193	106	121	127	209	159	175	118
GACH 163/05	205	159	146	202	195	106	123	133	207	151	171	118
	211	161	154	210	195	108	123	137	211	161	175	120
GACH 169/05	199	161	148	200	193	104	121	127	203	149	167	122
	199	161	150	204	193	106	127	133	213	151	167	122
GACH 170/05	199	155	136	200	191	102	123	129	209	153	167	120
	207	155	150	202	197	102	123	135	209	159	173	120
GACH 171/05	199	155	136	200	191	102	123	129	209	153	167	120
	199	155	148	202	191	106	123	135	209	159	173	124
GACH 172/05	199	155	148	204	193	108	121	117	209	149	165	122
	199	163	148	206	197	108	123	125	211	149	169	122
GACH 173/05	201	155	148	198	193	100	131	137	203	149	165	122
	205	163	148	202	195	102	131	143	209	153	169	132
GACH 179/05	199	155	136	192	193	102	121	125	203	151	167	122
	207	161	154	198	193	104	131	133	209	151	173	132
GACH 180/05	199	155	136	202	191	102	127	125	203	151	167	122
	211	159	154	204	197	102	131	133	209	155	173	122
GACH 181/05	199	155	136	192	191	102	121	133	203	151	167	122
	207	159	148	204	191	104	131	133	209	155	173	122
GACH 182/05	199	155	148	202	191	100	123	127	211	151	165	132
	199	161	148	202	191	104	133	147	211	153	169	132
GACH 183/05	199	159	136	202	197	102	121	125	207	151	169	122
	199	161	148	202	197	104	131	143	209	151	173	122
GACH 203/05	199	149	136	200	191	104	121	133	207	151	169	118
	211	155	148	202	193	106	123	137	213	153	169	128

	FCA 96	FCA 247	FCA 26	FCA 78	FCA 105	FCA 69	FCA 310	FCA 133	FCA 220	FCA 8	FCA 224	FCA 126
GACH 204/05	199	155	148	204	195	102	123	127	211	149	171	118
	209	163	148	204	197	104	131	135	211	149	171	128
GACH 205/05	199	155	148	202	197	104	121	127	213	151	169	116
	199	161	150	202	199	104	127	133	213	155	175	118
GACH 206/05	199	155	148	198	191	108	127	135	203	149	169	120
	205	161	148	204	193	108	131	135	207	149	171	120
GACH 208/05	205	155	146	192	197	104	121	127	211	149	167	120
	211	161	152	198	197	108	121	135	213	151	171	120
GACH 209/05	199	153	148	202	195	104	121	133	203	151	171	118
	201	161	150	208	199	104	127	133	213	155	175	122
GACH 210/05	199	163	148	202	191	104	121	125	213	149	169	116
	199	163	150	202	197	106	127	133	213	155	169	118
GACH 211/05	187	155	148	192	195	102	121	137	203	149	167	118
	211	163	148	200	195	104	131	145	209	151	169	128
GACH 144/05	199	155	146	206	191	100	121	127	203	151	169	118
	201	157	148	206	195	104	131	133	213	153	169	118
GACH 145/05	199	155	148	192	193	100	125	133	211	149	167	118
	211	161	148	198	193	102	127	133	213	151	169	118
GACH 147/05	199	161	136	192	191	100	121	127	211	151	169	128
	199	161	148	202	197	104	131	127	213	155	169	128
GACH 149/05	201	161	148	202	191	102	121	127	203	151	167	128
	207	163	154	210	193	108	131	127	209	151	169	128
GACH 154/05	199	155	148	192	191	102	121	137	203	155	167	118
	211	161	154	206	195	104	127	141	213	155	171	118
GACH 166/05	199	153	148	204	191	104	121	135	211	149	167	118
	211	161	148	204	195	108	133	147	211	151	167	128
GACH 167/05	209	155	148	204	191	102	121	127	203	149	169	118
	211	161	148	206	191	102	127	133	203	155	171	118
GACH 175/05	187	155	146	192	191	102	121	125	211	149	165	122
	207	163	148	204	197	104	127	125	215	153	169	122
GACH 176/05	199	155	146	202	193	100	121	125	211	151	169	124
	211	161	148	202	197	108	123	135	213	153	169	132
GACH 185/05	199	157	148	192	191	100	121	125	203	149	167	122
	201	161	150	192	193	100	131	135	211	153	169	122
GACH 186/05	201	161	148	192	193	100	121	133	209	149	165	114
	205	161	154	192	195	102	131	137	211	151	165	122
GACH 187/05	199	163	146	198	193	102	121	125	203	149	165	124
	209	163	148	206	195	108	127	135	211	149	169	132
GACH 196/05	199	159	136	202	195	102	127	127	207	151	167	118
	199	161	148	204	201	102	131	133	211	155	175	118
GACH 197/05	205	155	136	200	191	102	121	127	203	151	167	118
	211	155	148	200	193	102	131	137	209	153	167	128
GACH 199/05	199	155	148	204	197	102	123	133	207	149	171	118
	199	161	154	204	197	104	127	135	207	161	175	120