Insights into the genetics of the Ground Pangolin (*Smutsia temminckii*)

Christle de Beer

Dissertation submitted in fulfilment of the requirements for the degree of

Magister Scientiae

in the Faculty of Natural and Agricultural Sciences

Department of Genetics

University of the Free State

<u>Supervisor</u>

Prof. A. Kotzé

National Zoological Gardens of South Africa

Genetics Department, University of the Free State

<u>Co-supervisors</u>

Dr. D. Dalton

National Zoological Gardens of South Africa

Genetics Department, University of the Free State

Dr. K. Ehlers

Genetics Department, University of the Free State

Prof. R. Jansen

Department of Environmental, Water and Earth Sciences, Tshwane University of Technology

February 2013

Table of Contents

List of abbreviations and symbols

AL	Lysis Buffer
AE	Elution Buffer
ASL	Stool Lysis Buffer
ATE	Low EDTA Elution Buffer
ATL	Tissue Lysis Buffer
AW1	Wash Buffer (denatures proteins)
AW2	Wash Buffer (removes salts and purifies DNA)
bp	Base pair
°C	Degree Celsius
CITES	Convention on International Trade in Endangered Species
CTAB	Cetyl trimethylammonium bromide
cm	centimetre
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	double strand deoxyribonucleic acid
DTT	Dithiothreitol
FTA	Fast Technology for Analysis of nucleic acids
g	gram
h	hour
IUCN	International Union for Conservation of Nature
K	number of populations
kg	Kilogram
km	Kilometre
m/v	mass per volume
m	Metre
mg	miligram
ml	millilitre
mM	millimolar
min	minute
NEMBA	The National Environmental Management: Biodiversity Act no. 10 of 2004

ng	Nanogram
nm	Nanometre
PAUP	Phylogenetic Analysis Using Parsimony
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RPS	Relative priority scores
rpm	Revolutions per minute
RTD	Relative taxonomic distinctness
sp	species
s	second
ssDNA	single strand deoxyribonucleic acid
TBE	Tris/Borate/EDTA
TOPS	Threatened or Protected Species
USD	United States Dollars

List of figures

List of tables

Chapter 1: Justification

Literature suggests modest research on population genetic studies on pangolin species globally. Studies mainly focused on the ecology (Wu *et al.,* 2004), behaviour (Yang *et al.,* 2007), physiology (Webber *et al.,* 1986; Nisa *et al.,* 2005) and comparative anatomy (Ishimoto, 1983) of pangolins. Molecular studies focused mainly on the evolution and phylogeny of mammals in general and comparative genomics (Che *et al.,* 2008). With reference to molecular studies, limited data is available on the genetic diversity of pangolins (Heath, 1992; Hsieh *et al.,* 2011). In 1991, Zhang and Shi reported on the genetic diversity of *Manis pentadactyla* based on the partial sequence of the cytochrome b gene by restriction enzyme analysis, while in 2007, Luo *et al*. developed dinucleotide microsatellite markers for the Malayan Pangolin.

This study is a first attempt to determine the level of genetic variation within and between Ground Pangolin populations in South Africa using cross-species molecular markers. It is a project based on optimizing cross-species markers already available to inform on the Ground Pangolin genetic structure.

Chapter 2: Introduction

2.1 Global distribution and classification of pangolins

Pangolins are members of the only family in the order Pholidota with eight extant species in the world (Lim & Ng, 2008) assigned to Manidae. There is however, some disagreement regarding the number of genera to which species should be assigned (Figure 1) based on osteological characters, although a study by Gaudin e*t al.* (2009) shed some light on this issue. The Ground and the Giant Pangolin are grouped together within the genus *Smutsia*, while the Malayan, Chinese and Indian Pangolins are within the genus *Manis*. The African Tree Pangolin and the Arboreal Pangolin make up the third genus, *Phataginus*.

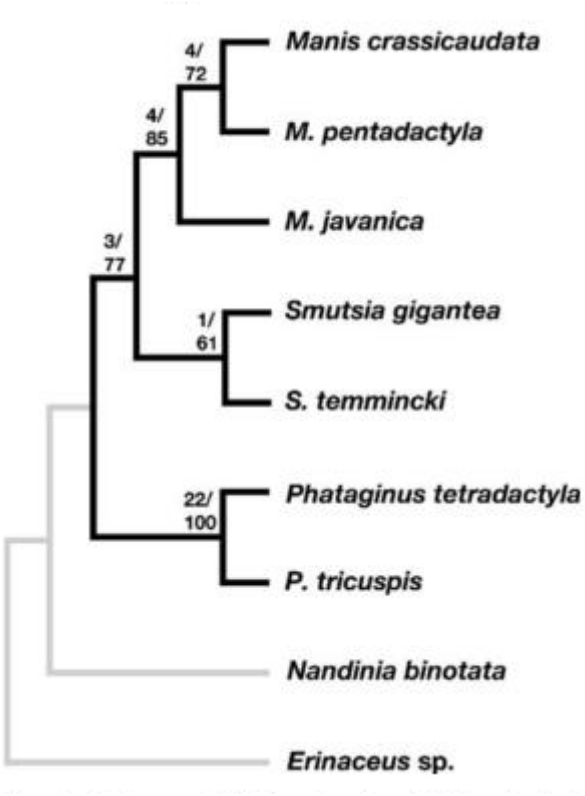


Figure 1: The phylogeny of the pangolin as given by a PAUP analysis of 395 osteological characters in nine taxa, including seven species and two outgroups, *Nandinia binotata* and an *Erinaceus* sp. The numbers in bold at each node represent Bremer support and bootstrap values (Gaudin e*t al.,* 2009).

Based on the study by Gaudin *et al.* (2009) four species of pangolin are known from Africa. The Four-toed Arboreal Pangolin, *Phataginus tetradactyla*, previously known as *Manis tetradactyla*; the Arboreal Pangolin, *Phataginus tricuspis*, previously known as *Manis tricuspis*; the Giant Pangolin, *Smutsia gigantea,* previously known as *Manis gigantea* (all co-existing in western and central Africa); and the Ground Pangolin, *Smutsia temminckii,* previously known as *Manis temminckii* (distributed in southern and eastern Africa). The four Asian species include *Manis crassicaudata* from India, *Manis pentadactyla* from China and northern South East Asia, *Manis javanica* from South East

Asia and the East Indies, and *Manis culionensis* from the Palawan islands of the Philippines (Botha & Gaudin, 2007; Gaudin e*t al.,* 2009). *Manis culionensis* was not reviewed in the study by Gaudin *et al.* (2009) as it was not widely recognized as a distinct species at the time the study was conducted. The order Pholidota is one of the smallest of the extant placental mammal orders with its modern representatives restricted to the Old World tropics. Due to the fact that these animals are toothless, may never have been speciose, and typically exist in low population densities, there is an insufficient fossil record of this group. They also prefer forested environments that have a low preservation potential in terms of fossils (Gaudin e*t al.,* 2009).

2.2 General pangolin facts

Pangolins are myrmecophagous and thus have unique specialized anatomical features which are adapted to this function (Yang *et al.,* 2006). These adaptations include a conical-shaped head, no teeth, a long sticky tongue and robust forelimbs with enlarged claws for procuring and eating ants and termites (Swart *et al.,* 1999; Botha & Gaudin, 2007). Another atypical morphological characteristic they exhibit, as compared to Old World mammalian fauna, are overlapping horny scales which are made of keratin, are yellow-brown in colour and consists of agglutinated hair (Luo e*t al.,* 2007). These scales offer protection not only against predators but also against the bites and stings of their prey, as well as protecting the skin against scratches from the underbrush or sharp rocks (Lim & Ng, 2008). Scales, however, provide little insulation or protection from external parasites (Heath, 1992). The four African species of pangolin can easily be visually distinguished from Asian pangolin species as the external ears are absent and there is no hair between the scales.

2.3 The Ground Pangolin (*Smutsia temminckii*, Smuts, 1832)

The Ground Pangolin (*Smutsia temminckii*) has its evolutionary origin in the northern hemisphere (Yang *et al.,* 2006) and lives in arid regions (Heath & Coulson, 1997). Common names for *S. temminckii* include: Temminck's Ground Pangolin (after Prof. C. J. Temminck, a Dutch zoologist); Steppen-schuppentier (German) or Steppe Pangolin; Ietermagog (Afrikaans) and Ground Pangolin (Heath, 1992). For the purpose of this study the most recent taxonomic name *S. temminkii* shall be used for the Ground Pangolin (Gaudin *et al.*, 2009).

Currently the Ground Pangolin occurs throughout most of southern and eastern Africa from the Cape northward and north into north-eastern Chad (Heath, 1992) (Figure 2) and is regarded as having the widest distribution range of any of the African pangolin species where it occurs in low densities when compared to other myrmecophagous (feed on ants and termites only) mammals (Swart *et al.,* 1999). The Ground Pangolin, reaches 1.3 m in length and is crepuscular (active during twilight), but have been known to forage into the night and even early hours of the morning. Studies have indicated that 56% of their activity is between 16h00 and 18h00 (Jacobsen e*t al.,* 1991). The same author also mentions that their tails are very powerful and to a considerable extent the tip is prehensile. Heath (1992) indicated that Ground Pangolins are capable swimmers and avid climbers.

Currently, there is insufficient biological data on the Ground Pangolin due to the difficulty of locating and observing this species. Ground Pangolins are mostly nocturnal and solitary, often meagrely distributed and only rarely seen by humans (Heath &

Coulson, 1997; Richer *et al.*, 1997). Typical behaviour exhibited by pangolins is that they promptly flee or freeze in their tracks when they detect human presence in their immediate vicinity (Lim & Ng, 2008). In addition, its spoor is small, subtle and hard to find or follow. It has been determined that Ground Pangolins can travel great distances with home ranges from 0.17 to 11.07 km^2 (Heath & Coulson, 1997). The home ranges within each sex are contiguous, but there is a conspicuous overlap of range between sexes. Individual pangolins use the same home range over multiple years (Jacobsen *et al.*, 1991; Heath & Coulson, 1997).

Figure 2: Distribution of the Ground Pangolin in Africa (IUCN Red List of Threatened Species; Version 2011.2).

2.4 Threats to the survival of pangolins

Today the existence of these mammals is threatened by several factors. If threatened or attacked, the pangolin rolls into a tight defensive ball, which often also leads to them becoming entangled in, and even killed by, electric fencing (Jacobsen *et al.*, 1991). Death by electrocution is regarded as the single largest threat to Ground Pangolins in southern Africa (Pietersen *et al*., in prep). With regards to the African Forest Pangolin (*Phataginus tetradactyla*) anthropogenic pressures and large-scale, rapid loss of forest habitat, and uncontrolled bush fires caused the population to decline (Newton *et al.,* 2008). According to Laurence *et al.* (2007) this is exacerbated in central Africa with a dramatic increase in logging, roads and hunting activities. The authors further confirmed that nocturnal species such as the pangolin are more affected by human-induced changes in forest structure. The creation of forest roads, for example, affect pangolins as they exhibited an increased proximity on road margins as compared to, their natural habitat, and forest transects. This not only makes them vulnerable targets for hunters, but also puts them at risk of being run over by vehicles. Another example is logging, as this allows poachers to reach unexploited wildlife populations and lowers the cost of transporting bush-meat to markets (Laurence e*t al.,* 2007). Indeed, mammal populations considered at greatest risk as a result of over-hunting include pangolins due to their highly valued meat, scales and thick skins. Even though Asian pangolin species appear in Appendix 2 of CITES 2007 with a zero annual export quota for wild caught individuals or those traded for commercial purposes, hunters are still paid hundreds of US dollars per kilogram meat and scales (Newton e*t al.,* 2008; Anon, 2013). The only official statistics of the trade of pangolin scales were reported by Bräutigam *et al*. (1994) which emanated from South Korea, with 7067 kg imported from China, 1850

kg from Indonesia, 1000 kg from Malaysia and 1026 kg from Vietnam in 1992. Based on the conservative assumption that a single pangolin yields 0.5 kg of scales (Anon, 2002); this translates to approximately 14000 Chinese pangolins (*M. pentadactyla*) *and* 7500 Malayan pangolins (*M. javanica*) in one year.

Trafficking seems to be one of the most harmful threats to the global population of pangolin species, specifically to the Asian meat market or for use in traditional medicine practices. According to an online encyclopaedia, Animal Life Resource (Anon, 2012a, b) traffickers sell the animals and their body parts to buyers who use the animals for food and for the muti trade, as these animals are believed to have healing properties to cure various other ailments. Pangolin meat is considered a delicacy in Vietnam and by tribal communities in India (Newton e*t al.,* 2008; Chakkaravarthy, 2012).

Within Africa, scales are used as love charms, and in the Shona culture, used as traditional medicine or as a gift to the chief (Hishin, 2011). The Yorubas of south-western Nigeria use pangolins in ethnozoological practices. It is believed that pangolin flesh gives fertility for women, is used as an aphrodisiac for men and even to incite fortune (Soewu, 2008). In Zimbabwe it is considered a good omen to see a pangolin, and it is tradition to catch and present the pangolin to a superior or chief to be consumed. In East Africa pangolin scales are burned as it is believed to repel lions. The pangolin is known as a doctor in Tanzania as each part of it is believed to have some specific healing power. Pangolins are presumed to be rare /endangered in Malawi due to its importance to traditional medicine pharmacopoeia (Heath, 1992).

Kyle, (2000) has predicted that the hunting of Ground Pangolins in Kwazulu-Natal for medicinal reasons may easily lead to their extinction. Pangolins have not been located within this province for a number of years now and they are regarded by local conservation authorities as extinct in this province (Kyle, 2000).

2.5 Current conservation status of *Smutsia temminckii*, Smuts, 1832

In 1976 the Ground Pangolin was listed as endangered by the United States Fish and Wildlife Services, but currently the status is of "Least Concern" according to the International Union for the Conservation of Nature (IUCN) Red Data Book (June, 2013). Taxa in the IUCN Red list categories are considered at risk of becoming extinct unless conservation actions are taken. The Ground Pangolin is also listed on CITES II as a lower risk/ near threatened species (June, 2013). In Table 1 below, a more detailed breakdown of the conservation status of the Ground Pangolin per province within South Africa is given. The most important legislation in terms of Ground Pangolin protection in South Africa is the NEMBA 10 of 2004, through which penalties and offences are, according to Section 57(1), R 10 million or ten years imprisonment or both. Section 57(1) states that a person may not carry out a restricted activity involving a specimen of listed threatened or protected species without a permit.

Table 1: Conservation status of the Ground Pangolin per province in Southern Africa.

Smutsia temminckii – Ground Pangolin									
National	Western Cape	North West	Mpumalanga	Northern Cape	Limpopo	Gauteng	Free State	Kwa-Zulu Natal	Eastern Cape
Listed as: Vulnerable species in terms of NEMBA	Listed as: Endangered Wild Animals (Schedule 1) in terms of the Western Cape Nature Conservation Laws Amendment Act, 3 of 2000	Listed as: Protected Game (Schedule 2) Section 15 (1) (a) in terms of the Transvaal Nature Conservation Ordinance 12 of 1983	Listed as: Protected Game (Schedule 2) Section 4 (1) (b) in terms of the Mpumalanga Nature Conservation Act, 10 of 1998	Listed as: Endangered Wild Animals (Schedule 1) in terms of the Nature and Environmental Conservation Ordinance, 19 of 1974	Listed as Specially Protected Wild Animals (Schedule 2) in terms of the Limpopo Environmental Management Act, 7 of 2003	Listed as: Protected Game (Schedule 2) Section 15 (1) (a) in terms of the Nature Conservation Ordinance, 12 of 1983	Listed as: Schedule 1 Protected Game (section 2) in terms of the Nature Conservation Ordinance, 8 of 1969	Listed as: Specially Protected Game (Schedule 3) in terms of the Nature Conservation Ordinance, 15 of 1974	Listed as: Endangered Wild Animals (Schedule 1) in terms of the Cape Nature and Environmental Conservation Ordinance, 19 of 1974

In Mpumalanga the Ground Pangolin is listed under Protected Game (Schedule 2) Section 4 (1) (b) and penalties would depend on the offence and the same applies in the Western Cape, where Ground Pangolins are listed as Endangered Wild Animals (Schedule 2).

A study by Freitag and van Jaarsveld (1997) placed the Ground Pangolin fourth in a regional national conservation concern based on national relative priority scores (RPS). The RPS technique attempts to evaluate the regional conservation importance of taxa by assigning a RPS to each taxon with reference to the extinction risk, vulnerability and conservation value or irreplaceability. Keith *et al.* (2007) places this species fifth and third respectively based on variations of RPS. The Ground Pangolin also had a very high relative taxonomic distinctness (RTD) score where RTD is a measure in which taxonomically more distinct taxa receive higher scores than speciose taxa, as it can be argued that they contribute proportionately more to regional biodiversity. However, regional provincial ordinances in South Africa vary as to the species conservation status, e.g. in Mpumalanga Province it is considered “protected game” but in the Northern

Province it is considered as “endangered” and placed on schedule 1 of the protection status.

The aforementioned statistics show that the Ground Pangolin is an important and neglected component to the biodiversity of Southern Africa and has a high priority for conservation due to its irreplaceability and distinctness. Another important factor to consider is the impact pangolins have on the ecosystem. A suitable example is the species-specific tick, *Amblyomma compressum* (Macalister, 1872). This small species of *Amblyomma* is almost exclusively found on the three African species of pangolins, *S. temminckii, P. tetradactyla,* and *P. tricuspis* (Voltzit & Keirans, 2003) and cannot survive on another host. Other than this, little is known about *A. compressum* which shows the importance of protecting the pangolin, as its extinction will also in effect lead to the extinction of this unstudied *Oxodida* ectoparasite.

2.6 Use of molecular genetics as a tool to contribute to conservation of pangolins

According to De Oliveira *et al.* (2012) conservation genetics is the application of molecular methods to preserve species as dynamic entities capable of coping with environmental change. In order to reduce the rates of extinction and preserve biodiversity, populations of species under anthropogenic impact need to be monitored. This includes management of small populations, molecular forensics and the identification of certain aspects of species' biology or social structure through the use of molecular markers (De Oliveira *et al*., 2012)

The genetic diversity in a population has to be monitored to determine whether inbreeding or other genetic factors such as drift has occurred. This is to determine whether the species needs to be managed to preserve the genetic diversity as well as the genetic integrity of the population (Grobler *et al.,* 2004). Increases in individual homozygosity, due to large losses of alleles, may reduce individual fitness through inbreeding depression (Markert *et al*., 2010). According to the same author, even modest losses of allelic diversity may negatively impact long-term population viability by reducing the capacity of populations to adapt to altered environments.

If inbreeding or the loss of genetic diversity is found in populations then the possibility of translocation of animals from other populations should be investigated (Grobler & Van der Bank, 1993). According to Tracy *et al.*, (2011) one way in which loss of genetic diversity can be managed, is to source different individuals from genetically diverse populations. However, Jacobsen *et al.* (1991), suggests that on the grounds of their behaviour, pangolins should not be relocated or translocated, as it would be unsuccessful. Recent efforts by members of rehabilitation organisations in Zimbabwe (Tikki Haywood Trust) and in South Africa (FreeMe), have however proven successful on a number of occasions. Furthermore, their survival rate in captivity is low with 71% mortality rate in the first year of captivity in *M. Javanica* (Wilson, 1994). This shows the importance of the home range and the effectual stress that results from relocation to an unfamiliar location (Heath & Coulson, 1997).

Another important application of molecular data is in wildlife forensics, as it applies to the illegal trading of pangolin and the bush-meat industry. In Brazil this is one of the

major challenges to conservation of wildlife as poor local law enforcement and corruption allow hunting to continue unabated, even in protected areas (Sanches *et al.*, 2012). DNA forensic services for domesticated animals are well established and have expanded to include various wildlife species (Grobler *et al.*, 2005). The precise identification of bush meat is necessary as the only evidence recovered from the crime scene when hunters are captured are the remains of meat, fur, skin and bone (Sanches *et al.*, 2012). In wildlife forensics the resolution power of microsatellite markers and assignment tests have been applied to determine species identity and even identify the geographical origins of individuals (Grobler *et al*., 2007).

A recent study by Tobe and Linacre (2011) focussed on the mitochondrial genome to amplify species-specific fragments that could easily be separated using a genetic analyzer. Each fragment was of a different size and the specificity of each primer pair allowed for species identification to be made even if a mixture of several species was present. This procedure allowed the rapid and simultaneous identification of rhino, tiger, bear, leopard, pangolin, musk deer and several non-endangered mammals in traditional East Asian medication. The fact that pangolin DNA was identified in this medication serves as motivation for the further development of wildlife forensic markers for the species.

It is therefore of vital importance that molecular data be collected for current populations not only to assess their genetic variation but also to develop a set of possible microsatellite markers that could be used in future wildlife forensic cases.

2.7 Collecting molecular data from museum specimens

DNA extraction from museum specimens has, after the initial recovery of DNA from various avian museum samples by Houde and Braun (1988), been accepted as a viable source of information in molecular research. This has provided evolutionary biologists the opportunity to explore the phylogeographic relationships between and within populations and species (Reiss, 2001). For example, a genetic profile obtained from a pangolin museum specimen may provide insight into the genetic history of their populations; information about possible colonisation and divergence may be inferred; and the presence or absence of alleles from specimens can be compared to existing populations to indicate whether a loss of genetic diversity or inbreeding has occurred (Yang *et al*., 1997). Some species, such as the pangolin, are elusive and difficult to sample which make museum samples more desirable and valuable. And for some species, museum samples provide an opportunity to investigate individuals prior to a possible genetic bottleneck. By analysing museum specimens, the potential impact of an ongoing population bottleneck in the Tasmanian devil (*Sarcophilus harrisii*) was inferred (Paijmans *et al.*, 2012).

One of the apparent consequences of a genetic bottleneck within a population is the loss of genetic diversity and the related reduction in individual fitness and evolutionary potential. The effective management of populations suffering such a loss in diversity is often hindered by a lack of understanding of how adaptive genetic variation will respond to population fluctuations, given these are affected by selection as well as drift (Olivier & Piertney, 2012).

Genetic analyses of museum specimens are consumptive and require skin or bone to be removed. As such, most museums have understandably imposed restrictions on the use of their often irreplaceable collections (Rohland *et al.,* 2004). Often, only a small sample can be collected from these specimens as they are very valuable and the majority are used for display purposes. Not only has this led to difficulties in obtaining samples and extracting suitable DNA, but older samples often have degraded DNA or the precise locality from where the sample was sourced is often not available. Furthermore, it becomes increasingly difficult to extract genomic DNA from older samples and often only mitochondrial DNA can be obtained.

In addition, the treatments used to preserve the specimens often contain chemicals which can accelerate DNA degradation. Some even inhibit the reagents in DNA extraction thereby preventing any molecular analyses. A variety of treatments are available that can be used to preserve samples. Finally, the conditions under which the sample is stored may have implications as increasing temperature, for example, has an impact on the rate of DNA degradation (Paabo, 1989). Zoological samples include: bird and mammal study skins, mounted specimens, skeletal material, casts, pinned insects, dried material, animals preserved in spirit and microscope slides.

When an animal study skin is preserved for museum purposes, the skin is rolled in a mixture of sodium fluoro silicate and borax. Flat skins on the other hand are left in coarse salt and borax for 6-8 weeks and then pinned and left to dry. After 2-3 weeks the skins are sanded down on the flesh side and painted with 1-2 layers of liquid paraffin to soften the skin. From a DNA perspective, dry-salted skin harbours a lot of epithelial skin

cells, each with more or less intact nuclei and mitochondria which is why only a small piece of skin is required for DNA studies (Payne & Sorenson, 2002).

The major challenges in ancient DNA extraction are to maximize the DNA yield, to eliminate inhibitors that prevent polymerase chain reaction (PCR) amplification and to minimize the possibility of contamination (Farrugia *et al.*, 2010). Too much DNA may inhibit the PCR and too little DNA may result in the over amplification of non-specific products. If the DNA is too degraded no amplification may occur as there will be no binding sites for the primers (Westring *et al*., 2007).

A study by Yang *et al*. (1997) showed that the CTAB-based extraction (Doyle & Doyle, 1990) method consistently yielded visible DNA from air-dried skin from a frozen carcass of an extinct *Mammuthus primigenius* (Woolly mammoth) and bones from two extinct species, *M. primigenius* and *Mammut americanum* (American mastodon). The Chelex 100 method (Walsh *et al.*, 1991) has also proven to be successful in isolating DNA from museum samples in a study by Su *et al.* (1999). Pieces of dried skins cut from whole leather specimens stored for periods ranging from several years to decades from the endangered animal group, musk deer (genus *Moschus*) were used and this method yielded high recovery rates of over 70%.

2.8 Non-invasive sampling

Technical advances in areas such as high-throughput sequencing, microsatellite analysis and non-invasive DNA sampling have led to a much-expanded role for genetics in conservation (De Salle & Amato, 2004). Non-invasive sampling is increasingly valuable in studies of free-ranging mammals, like the pangolin (Frantzen *et al.,* 1998). Faecal and hair samples, as well as oral swabs, offer non-invasive options for genetic sampling of mammals. This is especially useful in the case of endangered species (Prendini *et al.,* 2002) as some animals may die of stress when invasive samples are collected which makes non-invasive sampling a safer method. In this study, pangolin scales were also tested to determine their possible use as a source of DNA.

However, non-invasive sampling has limitations. Low DNA quantity or quality or poor extract quality due to the presence of PCR inhibitors can complicate DNA analysis. Other drawbacks include the risk of contamination during the extraction and amplification process and difficulty amplifying long sequences because most DNA will be degraded into short fragments (Taberlet *et al.*, 1999).

When working with samples such as faeces, scales and oral swabs, where the total amount of DNA available for genetic typing can be very low, genotyping with nuclear DNA microsatellite loci can lead to variable results including: (1) no PCR product, (2) a PCR product and incomplete genotype, or (3) a PCR product and complete genotype. An incomplete genotype may be obtained when only one allele of a heterozygous individual is detected, in which case the error is called allelic dropout, which produces false homozygotes (Taberlet *et al.*, 1999).

2.9 Microsatellite markers

Microsatellites are tandemly repeated sequences with a unit of repetition of between one and five base pairs (bp), with di-, tri- and tetra-nucleotide repeats being used most often as markers (Jarne & Lagoda, 1996). The varying lengths of these repeats are considered as alleles, which are identified at a given locus using specific PCR amplification. Based on relative electrophoretic migration these alleles are compared to a ladder of known sizes. Due to the fact that these microsatellite markers are numerous and ubiquitous throughout the genome, show a higher degree of polymorphisms, and have a codominant inheritance they have been extensively used for estimating genetic structure, diversity, and relationships (Dávila *et al.*, 2009). Information in literature has revealed that microsatellite markers are the most accurate and efficient markers for estimating genetic diversity and relationships among populations (Takezaki & Nei, 1996; Jarne & Lagoda, 1996; Dávila *et al.*, 2009).

Microsatellite markers are useful for phylogeographical, evolutionary and population genetic studies and for helping design and implement effective conservation management plans (Luo *et al.,* 2007). There is currently no information on the population genetics of *S. temminckii* aside from the phylogenetic study on living and extinct pangolins by Gaudin *et al.* (2009). In the latter study the author reported that a comprehensive molecular phylogenetic study on modern pangolins has yet to be conducted. Due to the fact that pangolins are not well represented in zoos and museum collections worldwide, and tend to live at low population densities, obtaining fresh tissue samples for sequence analysis presents particular difficulties.

2.9.1 Microsatellite markers in pangolin studies

A single study has been undertaken where microsatellites were isolated and characterized in pangolin species (Luo *et al.,* 2007). Thirty-four polymorphic dinucleotide microsatellite loci were developed for the Malayan Pangolin, *Manis javanica*. Of the 34 markers, 32 also amplified in the Chinese Pangolin (*Manis pentadactyla*) and 18 amplified in the African Tree Pangolin (*P. tricuspis*) (Table 2).

In the Malayan Pangolin, the number of alleles for the 34 markers varied between 2 to 21, with a size range between 170 and 299 bp. For the Chinese Pangolin only 32 markers of the 34 markers were amplified with 27 markers being polymorphic with the number of alleles between 2 to 10 alleles and a base pair range of 155 to 263. As for the African Tree Pangolin in which only 18 markers of the 34 markers were amplified, the alleles ranged from 2 to 4 with the size varying from 150 to 293 base pairs. Two markers were monomorphic. The high variability across species from different continents shows the value of these markers for evolutionary and conservation genetic studies in pangolins (Luo *et al.,* 2007).

Table 2: Characteristics of microsatellite markers isolated from the Malayan Pangolin (*Manis javanica*) (MJA) and cross-species amplified in the Chinese Pangolin (*Manis pentadactyla*) (MPE) and the African Tree Pangolin (*Phataginus tricuspis*) (MTI) (Luo *et al.,* 2007).

Locus (accession no.)	Repeat motif	Primer sequence (5′–3′)	No. of alleles			Size range (bp)			H_O			H_E		
			MJA	MPE	MTI	MJA	MPE	MTI	MJA	MPE	MTI	MJA	MPE	MTI
MJA01	$(GT)_{22}$	F: CAGAAGATGGCCTAGGTGGA	13	4	0	187–219	195–201	0	1	0.778	—	0.895	0.624	—
(DQ886446)		R: CTTGGGGCAGAGCTATCTGA												
MJA02	$(GT)_{11}$	F: GAGGGTACATCCCACAAAGG	7	4	2	223–237	229–237	229–235	0.739	0.600	1	0.780	0.700	0.500
(DQ886447)		R: GGGTACTTCCGAAGGAAATG												
MJA03	$(CA)_{40}$	F: TAGGTGGCAGACGATTTGCT	21	1	3	175–237	189	185–199	0.565	0	0.500	0.942	0	0.625
(DQ886448)		R: CTGAGTGAGGCTGGCTTTCT												
MJA05	$(GT)_{12}$	F: GTGGAAGGCAGGAAAAACAA	13	2	3	261–299	259–263	277–293	0.750	0	0.500	0.846	0.153	0.625
(DQ886450)		R: CCCTTTGGGAAGAGTGTGAA												
MJA06	$(CA)_{26}$	F: CTGGCAGATTCCATCTTGCT	19	4	3	220–288	190–230	230–238	0.667	0.500	1	0.895	0.462	0.625
(DQ886451)		R: GGATGATGAAATACGGCTGAA												
MJA07	$(GT)_{21}$	F: CAGCCCAGGTAACAGACTGG	11	3	3	238–270	198–228	192–206	0.792	0.0833	0.500	0.862	0.226	0.625
(DQ886452)		R: TTCCATCTGGGTGTCCTACAG												
MJA08	$(GT)_{11}$	F: CACCCACATTATTGCAAACG	3	2	2	178–184	156–172	172–178	0.375	0.0833	0.500	0.312	0.219	0.375
(DQ886453)		R: AAAGATATTGCCACCCACTTG												
MJA09	$(GT)_{21}$	F: TCTGCATAAGGTTGAAGAGCAA	8	6	4	200–216	194–208	184–204	0.667	0.417	1	0.773	0.667	0.750
(DQ886454)		R: GACAAGGCAGTGTTGCTGAA												
MJA10	$(CA)_{18}$	F: CTAGGGTTGGGTCCTTCCTC	13	7	2	211–245	233–263	205–207	0.792	0.333	0	0.828	0.427	0.500
(DQ886455)		R: CTCAGGTGCTTTGGACTTAGG												
MJA11	$(CA)_{29}$	F: CTCACCGTGACAGCAGAGAC	14	5	0	188–224	176–208	0	0.909	0.250	—	0.897	0.420	—
(DQ886456)		R: GCTTATCCTGGTTTCAATCATTC												
MJA12	$(CA)_{20}$	F: GGAGTGCTGAACTTGGGTGT	5	1	2	178–186	166	180–184	0.250	0	0.500	0.631	0	0.375
(DQ886457)		R: TGGAGGGAAGTCTACCCAAA												
MJA13	$(GT)_{16}$	F: CTGGGGATGCCCTAATTTCT	10	4	3	204–226	216–222	214–264	0.652	0.700	0.500	0.701	0.715	0.625
(DQ886458)		R: CACAGCACAGTTGGGATTGT												
MJA14	$(CA)_{20}$	F: CTTGGGGCAGAGCTATCTGA	13	4	1	184–216	192–200	196	1	0.333	0	0.895	0.615	0
(DQ886459)		R: CAGAAGATGGCCTAGGTGGA												
MJA15	$(GT)_{18}$	F: TTTCGAAGATGGCCTAGGTG	13	2	0	173–205	185–187	0	1	0.250	—	0.895	0.219	—
(DQ886460)		R: TCTGACCCCTGTTCTCCACT												
MJA16	$(CA)_{17}$	F: TTCCCCATCTTCTCCTTCCT	10	4	2	170–208	164–204	174–178	0.750	0.167	0.500	0.833	0.514	0.375
(DQ886461)		R: TGAATGTTGTAAAGAGGTAAAAACCA												
MJA17	$(GT)_{22}$	F: AAAAAGGAGGGAGCCTTCTG	12	5	3	173–211	161–199	191–199	0.750	0.222	0.500	0.891	0.642	0.625
(DQ886462)		R: AGCCGCTGCTTTATCACACT												
MJA18	$(CA)_{11}$	F: GATCCTCGAAACCAAGCAG	5	7	2	183–201	163–191	189–191	0.375	0.583	0	0.532	0.580	0.500
(DQ886463)		R: AGGCTCTAGGCTTCGTCCTT												
MJA19	$(CA)_{13}$–	F: CCAAGAGCTGGAGAGTGCAT	16	1	1	230–280	222	236	0.875	0	0	0.893	0	0
(DQ886464)	$(CA)_{19}$	R: TCAGTTGTATTTCCACGTTCTGA												
MJA20	$(CA)_{11}$	F: GGTTAGTGAGCCACCCTGAA	5	0	0	261–271	0	0	0.263	—	—	0.601	—	—
(DQ886465)		R: ATCAGCGCCCCTAATACTTG												
MJA21	$(CA)_{16}$	F: GAACCTGGGTTGGGGTAACT	13	10	4	218–254	212–256	220–236	0.286	0.833	1	0.883	0.879	0.750
(DQ886466)		R: GCAGGGTTTCTCAACTTTGG												
MJA22	$(CA)_{15}$	F: GGATGTGGGTATCCTTGTGG	13	5	4	196–226	182–198	150–176	0.870	0.400	1	0.925	0.735	0.750
(DQ886467)		R: CCTCTCAGTGGGTGGGAGTA												
MJA23	$(CA)_{30}$	F: CCACCTCACTCACACCACTG	19	1	0	173–217	155	0	0.917	0	—	0.936	0	—
(DQ886468)		R: TGAATGTCGGGGAGAGAAAC												
MJA24	$(GT)_{11}$	F: GAGGGGTCAGAAGTGTCCAA	7	4	0	199–229	185–205	0	0.708	0.100	—	0.822	0.415	—
(DQ886469)		R: GCAGCCTGCTTCCATTTTAT												
MJA26	$(GT)_{15}$	F: TGTGAAAGCAGAATGCAAAAG	9	7	0	235–251	233–253	0	0.917	0.833	—	0.882	0.813	—
(DQ886471)		R: ACTTGCCTGAAGTGGACACC												
MJA27	$(GT)_{16}$	F: GGTGACTTTGGGCAATTCAT	8	6	0	249–269	239–251	0	0.696	0.600	—	0.796	0.705	—
(DQ886472)		R: CCCTCTTTGGAGGCATCATA												
MJA28	$(GT)_{24}$	F: GCCTTCAAGTGTGCCTGTCT	9	5	3	241–265	237–253	207–217	0.875	0.429	0.500	0.885	0.725	0.625
(DQ886473)		R: CAGGCAAAATTTGGGCTAGA												
MJA29	$(GT)_{22}$	F: GGAGGCAAGGATGATTCTGA	10	1	0	227–251	223	0	0.833	0	—	0.842	0	—
(DQ886474)		R: TCACACATAAGCAGAGCTCCA												
MJA30	$(GT)_{20}$	F: GAAGCCCTAACCCCAGTCTC	6	5	0	242–256	236–254	0	0.625	0.167	—	0.780	0.361	—
(DQ886475)		R: CCTGGAGCAGTAAGGGAACA												
MJA31	$(GT)_{16}$	F: CCATGTGGGCCTGTATTAGG	11	0	0	192–216	0	0	0.875	—	—	0.884	—	—
(DQ886476)		R: TCATGTGAGCCAGCACCTTA												
MJA32	$(GT)_{8}$	F: CCCGGTCAGCTTCTTTCAAT	2	0	0	176–178	0	0	0.250	—	—	0.330	—	—
(DQ886477)		R: ACAGGAGATGGAGTGCAGGT												
MJA33	$(GT)_{23}$	F: GGAAGTGAGCAGCAAACACA	10	0	0	192–224	0	0	0.6667	—	—	0.872	—	—
(DQ886478)		R: CTCTCAGTCCCTTTCCGTAGA												
MJA34	$(GT)_{24}$	F: AGCTGCACATTCTCAGCAAA	6	0	0	217–231	0	0	0.385	—	—	0.864	—	—
(DQ886479)		R: CCCATGTGTCCTCATTCTCA												
Average									0.708	0.32	0.553	0.805	0.437	0.513

H_O, observed heterozygosity; H_E, expected heterozygosity.

Although cross-species markers cannot generally be used universally, microsatellite amplification between closely related species is possible (Primmer *et al.*, 2005). The success rate of cross-species microsatellite amplification has been shown to be directly related to the evolutionary divergence between the species from which the microsatellite loci have been isolated (the source species in this case the Malayan Pangolin) and the species to which the heterologous loci are being applied (the target species in this case the Ground Pangolin). Cross-species microsatellite amplification has been proved to be sufficiently successful for a diverse range of evolutionary genetic studies (e.g. Brunner *et al.*, 1998; Johnsen *et al.*, 1998; Palo *et al.*, 2001). In various species, especially of the avian variety, research has been conducted based solely on cross-amplified microsatellites (Craig *et al.,* 2005).

2.10 Optimization

During optimization of PCRs containing cross-species markers, several difficulties have to be overcome. These include poor sensitivity and specificity and/or preferential amplification of certain targets. The most important one of these limitations is the primer-to-template ratio, too high and primer-dimers are formed. Whereas if the ratio is too low, product will not accumulate exponentially, since newly synthesized target strands will re-nature after denaturation subsequently reducing the yield considerably or inhibiting the formation of PCR product (Markoulatos *et al.,* 2002).

It is therefore necessary to adjust the primer-to-template ratio to minimize these non-specific interactions. Hot start PCR often eliminates non-specific reactions caused by primer annealing at low temperature before commencement of thermocycling. The

alteration of other PCR components such as PCR buffer constituents, dNTPs, $MgCl_2$, and enzyme concentrations usually also result in considerable improvement in the sensitivity and/or specificity of the reaction (Markoulatos *et al.,* 2002).

2.11 Objectives and aims of the study

The overall objective of the study is to determine the genetic structure of the Ground Pangolin using cross-species microsatellite markers and to assess the level of genetic diversity within current South African populations as compared to historic populations. The suitability of these markers for wildlife forensic cases will also be assessed.

2.11.1 Specific aims

1. Optimise DNA isolation protocols for invasive, non-invasive and museum samples.
2. Optimise PCR protocols and microsatellites designed for the Malayan Pangolin on the Ground Pangolin.
3. Assess the viability of the microsatellites to measure the level of genetic diversity of the Ground Pangolin.
4. Determine whether the optimised microsatellites can be applied in illegal wildlife trafficking cases.

Dissertation outline

This dissertation is presented as a number of chapters focussed on addressing the specific aims mentioned above. Chapter 2 will focus on the materials and methods used during the course of this study and it is hoped that this information will be useful in future studies to overcome the number of difficulties faced when working with non-invasive samples. The results of optimization of protocols and the evaluation of non-invasive samples for downstream applications (aims 1 and 2) will be addressed in Chapter 3 to enable the collection of viable non-invasive samples that yield optimal results. Chapter 4 focuses on the genotypes produced and the various results from molecular analyses employed to address aim 3 and 4 while also looking at the application of these markers in wildlife forensics. Chapter 5 discusses the results from the previously mentioned chapters and how the results address the specific aims outlined above, as well as the shortcomings of this study and possible solutions to the problems encountered. This final chapter concludes the study and includes suggestions for future studies.

Chapter 3: Materials and Methods

3.1 Sample collection

Seventy three samples were collected from various locations in South Africa (Figure 3) ranging from Upington and surrounding arid areas in the west, the Waterberg (Vaalwater) in central South Africa to the Hoedspruit/Phalaborwa area in the eastern lowveld. Samples were either collected from pangolins that had become entangled in electric fencing and had died (2 whole pangolins); confiscated live specimens (9 blood samples); specimens located in the wild (1 faecal, 14 scales and 8 nasal/oral swab samples); 4 specimens caught in snares and from road mortalities (4 epidermal, 2 muscle and 29 tissue samples). Five additional outgroup samples were collected from animals that were to be sold as bush-meat in Ghana (3 blood samples collected on FTA paper), with single scale samples collected from Namibia and Zimbabwe as well. Two museum samples were also collected as an additional source of DNA for comparison from the National Museum in Bloemfontein and the McGregor Museum located in Kimberly (see Appendix 1).

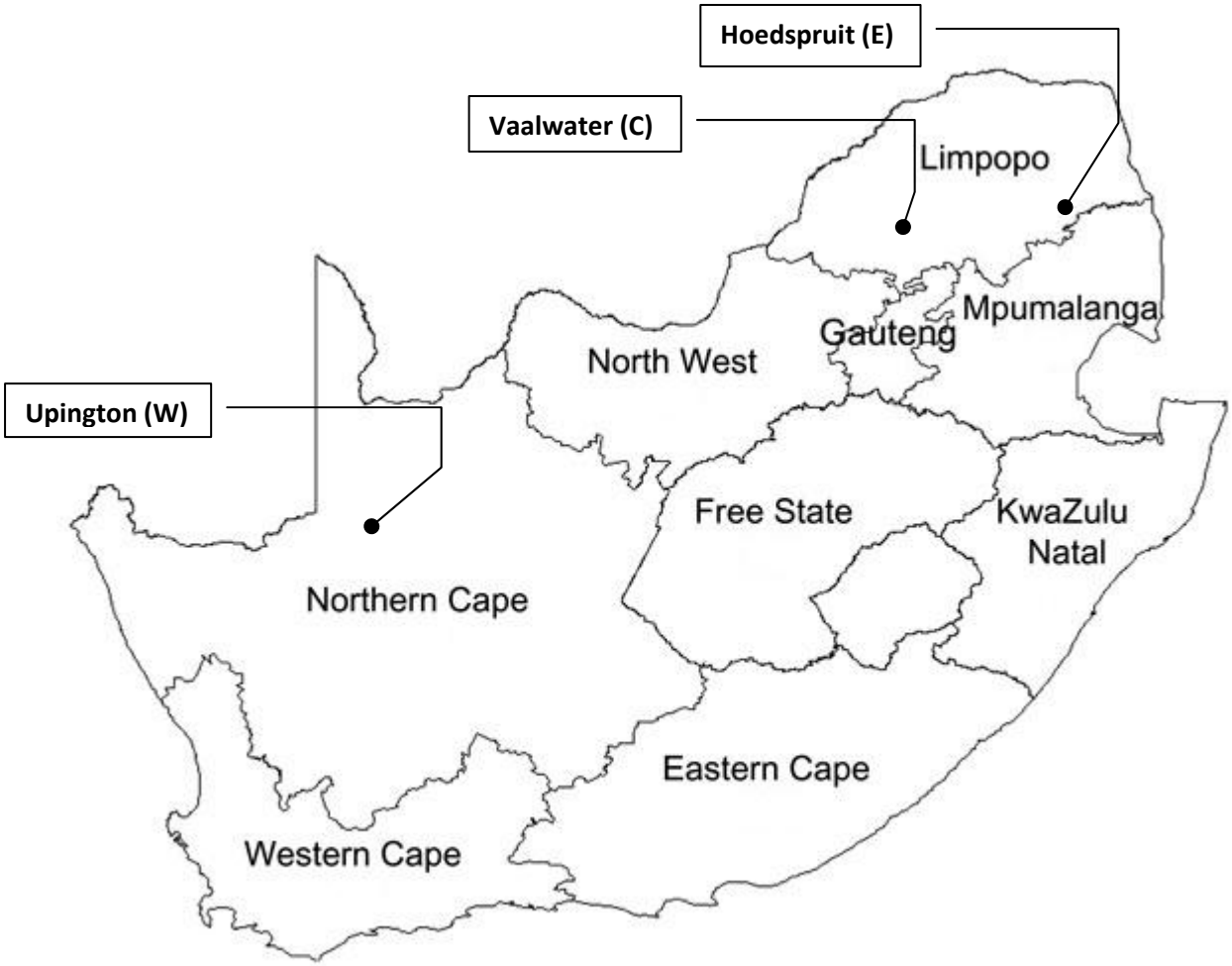


Figure 3: Map of South Africa indicating, from left to right, the 3 sampling localities representing western (42), central (7) and eastern (24) regions.

3.2 DNA isolation

Various samples (blood, tissue, blood on FTA paper, scales, oral swabs, faecal and museum samples) had been collected from live, dead or museum specimens, and therefore various extraction techniques had to be employed and optimised. In every isolation protocol the incubation period in the heating block; and the amount of proteinase K added; was increased to ensure that optimal lysis takes place.

3.2.1 Scale and tissue

The Qiagen isolation of total DNA from body fluid stains protocol (according to the manufacturer) was used with the following changes: A section of 0.5 – 1 centimetre^2 tissue or scale was lysed in 300 µl Buffer ATL, 20 µl of 20mg/ml proteinase K and 20 µl 1 M DTT and vortexed for 10 s. The tubes were placed in a heating block at 56°C for 6h while being vortexed every hour after which it was briefly centrifuged at 2000 rpm. Thereafter 300 µl Buffer AL and 1 µl carrier RNA was added and the mixture was vortexed again for 10 s. The samples were incubated at 70°C for 10 min while being vortexed every 3 min and briefly centrifuged at 2000 rpm. After which 150 µl of 96% ethanol was added and the samples were vortexed for 15 s and briefly centrifuged at 2000 rpm. The mixture was transferred to a QIAamp MinElute spin column and centrifuged at 8000 rpm for 1 min. After 500 µl Buffer AW1 was added and the column centrifuged at 8000 rpm for 1 min, 700 µl Buffer AW2 was added and centrifuged at 8000 rpm for another minute. Thereafter 700 µl of 96% ethanol was added and again centrifuged at 8000 rpm for 1 min and again for 3 min at full speed. For a final time the spin column was transferred to a clean 1.5 ml microcentrifuge tube and the lid opened and left to incubate at room temperature for 10 min. Finally 50 µl of Buffer ATE was applied to the centre of the membrane and the lid closed and left to incubate at room temperature for 1 min after which it was centrifuged at full speed for 1 min. The eluted samples were diluted 1:4 with distilled water.

3.2.2 Blood samples

The Qiagen isolation of total DNA from blood protocol from the QIAamp DNA investigator handbook was used with the following changes: Firstly 20 µl of 20mg/ml

proteinase K was pipetted into a 1.5 ml microcentrifuge tube to which 100 µl blood was added. Following that 120 µl PBS and 200 µl Buffer AL was added and immediately vortexed for 15 s. The sample was then incubated at 70°C for 10 min after which 200 µl of -20°C 96% ethanol was added and vortexed. The tube was then placed in the freezer for 5 min and transferred to a DNeasy spin column. It was centrifuged at 8000 rpm for 1 min, 500 µl Buffer AW1 added and the column centrifuged again at 2000 rpm for 1 min. Thereafter 500 µl Buffer AW2 was added and the column centrifuged at full speed for 3 min after which 200 µl Buffer AE was added and incubated at room temperature for 5 min. The sample was centrifuged at 8000 rpm for 1 min and used as is without dilution.

3.2.3 Nasal /Oral swab

The Qiagen isolation of total DNA from swabs protocol from the QIAamp DNA investigator handbook was used with the following changes: The swab was placed in a 2 ml microcentrifuge tube to which 20 µl of 20mg/ml proteinase K and 400 µl Buffer ATL was added and then vortexed for 10 s. The tube was placed in a heating block at 56°C and left overnight. The tube was briefly centrifuged and 400 µl Buffer AL and 1 µl carrier RNA was added. The tube was again incubated in a heating block at 70°C for 10 min while being vortexed every 3 min. The tube was briefly centrifuged and the lysate transferred to a QIAamp MinElute column and centrifuged for 1 min at 8000 rpm. Thereafter 500 µl Buffer AW1 was added to the column and then centrifuged for another minute at 8000 rpm after which 700 µl Buffer AW2 was added and again centrifuged for 1 min at 8000 rpm. Finally 700 µl 96% ethanol was added and the column centrifuged for 1 min at 8000 rpm and then again at full speed for 3 min. The column was opened and incubated at room temperature for 10 min after which 50 µl

Buffer ATE was added and incubated at room temperature for another minute. The column was centrifuged at full speed for 1 min and the sample was used as is without dilution.

3.2.4 Faecal

The Qiagen isolation of total DNA from stool protocol from the QIAamp DNA investigator handbook was used with the following changes: Three grams of faecal sample was weighed out in a 15 ml conical tube to which Buffer ASL was added to bring the final volume to 10 ml. The samples were vortexed for 1 min and 2 ml of the lysate was pipetted into a 2 ml collection tube. The tubes were centrifuged at full speed for 1 min and 1.4 ml of the supernatant pipeted into a new 2 ml collection tube. One inhibitEX tablet was added to each tube and immediately vortexed for 1 min and then incubated at room temperature for 1 min. The samples were then centrifuged at full speed for 3 min and all the supernatant was pipetted into a new 1.5 ml tube and again centrifuged at full speed for another 3 min. Following this 20 µl of 20mg/ml proteinase K was added to a new 2 ml tube and 600 µl of the supernatant added to that. Then 600 µl Buffer AL was added and vortexed for 15 s and then incubated at 70°C for 10 min. The tubes were then briefly centrifuged at 2000 rpm and 600 µl of 98% ethanol added to the lysate which was then mixed by vortexing. Thereafter 600 µl of the lysate was applied to a QIAamp spin column and centrifuged at full speed for 1 min. A second 600 µl of lysate was then applied to the same column and centrifuged at full speed for 1 min. Another 600 µl aliquot of lysate was again applied to the same column and centrifuged at full speed for 1 min. Following this 500 µl Buffer AW1 was added and centrifuged at full speed for 1 min and then 500 µl Buffer AW2 was added and centrifuged at full speed for

3 min. Finally 100 µl Buffer AE was applied directly to the membrane and incubated at room temperature for 1 min and centrifuged at full speed for 1 min. The sample was then diluted 1:10 with distilled water.

3.2.5 Blood on FTA paper

The QIAamp DNA investigator kit protocol was used with the following changes: A 3mm section of the FTA paper was punched and placed into a 1.5 ml microcentrifuge tube to which 180 µl Buffer ATL was added. This was incubated at 85°C for 10 min and briefly centrifuged at 2000 rpm. Thereafter 20 µl of 20 mg/ml of proteinase K was added, vortexed and incubated at 56°C for 1 hour. After again briefly centrifuging at 2000 rpm the samples 200 µl buffer AL was added, vortexed and incubated at 70°C for 10 min. The samples were once more briefly centrifuged at 2000 rpm and 200 µl 96% ethanol was added and vortexed. After again briefly centrifuging at 2000 rpm the samples the entire lysate was transferred to a DNeasy mini Spin Column, and centrifuged at 8000 rpm for 1 min. Following this 500 µl Buffer AW1 was added and centrifuged at 8000 rpm for 1 min. Another 500 µl Buffer AW2 was added and centrifuged at full speed for 3 min. The spin column was placed into a 1.5 ml microcentrifuge tube, and again centrifuged at full speed for 1 min. Finally 100 µl Buffer AE was added and left to incubate at room temperature for 1 min, then centrifuged at 8000 rpm for 1min and used as is without dilution.

3.2.6 Museum Samples

The Qiagen isolation of total DNA from nail clippings and hair from the QIAamp DNA investigator handbook was used with the following changes: A 0.5 – 1 cm^2 section of the

museum sample was lysed in 300 µl Buffer ATL, 20 µl of 20 mg/ml proteinase K and 20 µl 1 M DTT and then vortexed for 10 s. It was then placed in a heating block at 56°C for 6h while being vortexed every hour after which it was briefly centrifuged at 2000 rpm. Then 300 µl Buffer AL and 1 µl carrier RNA was added and then it was vortexed again for 10 s. It was then incubated at 70°C for 10 min while being vortexed every 3 min and briefly centrifuged at 2000 rpm. Following this 150 µl 96% ethanol was then added and then vortexed for 15 s and briefly centrifuged at 2000 rpm. The mixture was then transferred to a QIAamp MinElute spin column and centrifuged at 8000 rpm for 1 min. Thereafter 500 µl Buffer AW1 was added and then the column centrifuged at 8000 rpm for 1 min after which 700 µl Buffer AW2 was added and the column centrifuged at 8000 rpm for 1 min. Then 700 µl of 96% ethanol was added and again centrifuged at 8000 rpm for 1 min and again for 3 min at full speed. The spin column was then transferred to a clean 1.5 ml microcentrifuge tube and the lid opened and left to incubate at room temperature for 10 min. Finally 50 µl of Buffer ATE was applied to the centre of the membrane and the lid closed and left to incubate at room temperature for 1 min after which it was centrifuged at full speed for 1 min. The samples were then diluted 1:4 with distilled water.

The samples from blood (sample as well as FTA paper) and swabs (nasal/oral) were not diluted as to ensure optimum DNA concentration detection. The scales, tissue, museum and faecal samples, however, were diluted due to high amounts of inhibitors present as shown by the NanoDrop results.

3.3 DNA quantification

The DNA concentration was determined with the use of a Thermo Fisher Scientific NanoDrop 1000 Spectrophotometer. Following the user manual the subsequent procedure was performed. A total of 2 µl of distilled water was used to clean the measurement surfaces. The 'Nucleic Acid' spectral measurement was initiated using the operation system on the computer. With the sampling arm open, 1 µl of Buffer was pipetted onto the lower measurement pedestal. The sampling arm was closed and the 'Blank' button was clicked on. After opening the arm, the upper and lower pedestals were wiped clean using a soft laboratory wipe. Once the machine was blanked samples of DNA could be read. With the sampling arm open, 1 µl of DNA sample was pipetted onto the lower measurement pedestal. The sampling arm was closed and the 'Measure' button was clicked on. The readings on the computer screen for the ratio A260/280 (absorbance ratio) and the concentration (ng/µl) was recorded. Absorbance measurements made on the spectrophotometer include the absorbance of all molecules in the sample that absorb at the wavelength of interest. Since ssDNA and dsDNA both absorb at 260 nm, they will contribute to the total absorbance of the sample and therefore the ratio of absorbance at 260 nm and 280 nm is used to assess the purity of DNA. A ratio of ~1.8 is generally accepted as "pure" for DNA but if the ratio is appreciably lower, it may indicate the presence of protein, phenol or other contaminants that absorb strongly at or near 280 nm (Thermo Fisher Scientific, 1975). Following three consecutive sample readings, the blanking process was followed to ensure accurate readings. The results were used to determine which samples would be used for downstream applications and which sample types would need dilution, as

based on the A260/280 measurements for sample purity and concentration (see Appendix 2).

3.4 DNA profiling and optimization

The extracted DNA was amplified using PCR with the 14 selected microsatellites. These microsatellites (Table 3) were selected for use in the current study as they were the most polymorphic in the three (Malayan, Chinese and African-Tree) tested pangolin species (Luo *et al.*, 2007).

Table 3: Fourteen *Manis javanica* microsatellite markers (Luo *et al.*, 2007) utilized in the current study.

Name	Size range in Base pairs	Sequence	Plexes and Dyes
MJA02F	223–237	GAG GGT ACA TCC CAC AAA GG	
MJA02R	223–237	GGG TAC TTC CGA AGG AAA TG	
MJA28F	241–265	GCC TTC AAG TGT GCC TGT CT	
MJA28R	241–265	CAG GCA AAA TTT GGG CTA GA	
MJA05F	261–299	GTG GAA GGC AGG AAA AAC AA	
MJA05R	261–299	CCC TTT GGG AAG AGT GTG AA	
MJA21F	218-254	GAA CCT GGG TTG GGG TAA CT	**Plex 1 NED**
MJA21R	218-254	GCA GGG TTT CTC AAC TTT GG	
MJA22F	196-226	GGA TGT GGG TAT CCT TGT GG	**Plex 1 6-FAM**
MJA22R	196-226	CCT CYC AGT GGG TGG GAG TA	
MJA09F	200-216	TCT GCA TAA GGT TGA AGA GCA A	**Plex 1 HEX**
MJA09R	200-216	GAC AAG GCA GTG TTG CTG AA	
MJA07F	238-270	CAG CCC AGG TAA CAG ACT GG	**Plex 2 HEX**
MJA07R	238-270	TTC CAT CT GGG TGT CCT ACA G	
MJA08F	178-184	CAC CCA CAT TAT TGC AAA CG	**Plex 2 6-FAM**
MJA08R	178-184	AAA GAT ATT GCC ACC CAC TTG	
MJA18F	183-201	GAT CCT CGA AAC CAA GCA G	**Plex 2 NED**
MJA18R	183-201	AGG CTC TAG GCT TCG TCC TT	
MJA16F	170-208	TTC CCC ATC TTC TCC TTC CT	**Plex 3 6-FAM**
MJA16R	170-208	TGA ATG TTG TAA AGA GGT AAA AAC CA	
MJA13F	204-226	CTG GGG ATG CCC TAA TTT CT	**Plex 3 HEX**
MJA13R	204-226	CAC AGC ACA GTT GGG ATT GT	
MJA12F	178-186	GGA GTG CTG AAC TTG GGT GT	**Plex 4 ROX**
MJA12R	178-186	TGG AGG GAA GTC TAC CCA AA	
MJA03F	175-237	TAG GTG GCA GAC GAT TTG CT	**Plex 4 NED**
MJA03R	175-237	CTG AGT GAG GCT GGC TTT CT	
MJA14F	184-216	CTT GGG GCA GAG CTA TCT GA	**Plex 4 HEX**
MJA14R	184-216	CAG AAG ATG GCC TAG GTG GA	

The PCR protocol was then optimized using samples of different quality, as determined by the NanoDrop results, to determine whether the non-invasive samples yielded usable results for downstream applications. Furthermore, due to the fact that cross-species markers were used, it was necessary to optimise the PCR protocol as the conditions published by Luo *et al.* (2007), did not yield successful amplification of the samples in

the study presented here. Therefore, the protocol was optimized by three modifications: Varying $MgCl_2$ concentrations and an annealing temperature gradient using various *Taq* DNA polymerase.

3.4.1 Varying $MgCl_2$ concentrations

Invasive samples:

Invasive samples (2 blood samples and 3 tissue samples) with an A260/280 value range of 1.6 – 2.0 were amplified with primer sets MJA09 and MJA16. A total of 5 µl of 5 x QIAGEN PCR buffer, 2.5 µl of 0.2 mM dNTPs, 200 nM each of forward and reverse primer, 0.2 µl QIAGEN *Taq* DNA polymerase and 5.3 µl of either 2mM, 2.5 mM, 3 mM or 3.5 mM $MgCl_2$ was added and made up to 25 µl with water. Then 6 µl of 20 ng DNA of blood and 6 µl of 20 ng DNA (diluted 1:4) of tissue was added to separate tubes. The PCR was performed with one 10 min cycle at 95°C; 35 cycles of 30 s at 94°C, 30 s at 50°C and 30 s at 74°C; and one 30 min cycle at 72°C. A total of 10 µl of undiluted product was visualized on a 2% (m/v) agarose gel.

Non-invasive samples vs. Invasive samples:

Non-invasive (3 scales and 2 museum specimens) with an A260/280 value range of 1.6 – 2.0 was used and compared to invasive samples (3 tissue samples) with an A260/280 value range of 1.6 – 2.0 with primer sets MJA09 and MJA16. Then 5 µl of 5 x QIAGEN PCR buffer, 2.5 µl of 0.2 mM dNTPs, 200 nM each of forward and reverse primer, 0.2 µl QIAGEN *Taq* DNA polymerase and 5.3 µl of either 2.5 mM, 3 mM or 3.5 mM $MgCl_2$ was added and made up to 25 µl with water. A total of 6 µl of 20 ng (diluted 1:4) DNA of each sample type was added. The PCR was performed with one 10 min cycle at 95°C; 35

cycles of 30 s at 94°C, 30 s at 45°C and 30 s at 74°C; and one 30 min cycle at 72°C. Finally 10 µl of undiluted product was visualized on a 2% (m/v) agarose gel.

3.4.2 PCR with a temperature gradient using SuperTherm *Taq*

Invasive samples:

Invasive samples (5 tissue) with an A260/280 value range of 1.6 – 2.0 were used with primer sets MJA09, MJA16, MJA21 and MJA22. To each tube the following was added: 2.5 µl 10 x QIAGEN Buffer with 15 mM $MgCl_2$, 2.5 µl of 0.2 mM dNTPs, 200 nM each of forward and reverse primer, 0.1 µl SuperTherm *Taq* and 7.9 µl water. Then 6 µl of 20 ng (diluted 1:4) DNA was added. The PCR was performed with the following annealing temperatures: 45°C, 45.3°C, 46°C, 47.3°C, 49°C, 53.6°C, 57.4°C and 59.6°C. The PCR was performed with one 10 min cycle at 95°C; 35 cycles of 30 s at 94°C, 30 s at various annealing temperatures as indicated and 30 s at 74°C; and one 30 min cycle at 72°C. Finally 10 µl of undiluted product was visualized on a 2% (m/v) agarose gel.

Non-invasive samples:

Non-invasive samples (2 scales and 1 faecal) with an A260/280 value range of 1.6 – 2.0 and an A260/280 value of 2.5 respectively were used with primer sets MJA09, MJA16, MJA21 and MJA22. Even though the NanoDrop results showed very poor purity, as determined by the A260/280 measurements, for the faecal sample, the sample were diluted 1:10 to attempt amplification. A PCR was done using 2.5 µl 10 x QIAGEN Buffer with 15 mM $MgCl_2$, 2.5 µl of 0.2 mM dNTPs, 200 nM each of forward and reverse primer, 0.1 µl SuperTherm *Taq* and 7.9 µl. Finally 6 µl of 20 ng (diluted 1:4) scale sample and 6 µl of 20 ng (diluted 1:10) faecal sample DNA was added. The PCR was performed with

the following annealing temperatures: 45°C, 45.3°C, 46°C, 47.3°C, 49°C, 53.6°C, 57.4°C and 59.6°C. The PCR was performed with one 10 min cycle at 95°C; 35 cycles of 30 s at 94°C, 30 s at various annealing temperatures as indicated and 30 s at 74°C; and one 30 min cycle at 72°C. Finally 10 µl of undiluted product was visualized on a 2% (m/v) agarose gel.

3.4.3 Optimized PCR protocol

Once each primer set's optimal reaction conditions had been determined, capillary electrophoresis was used to determine the genotype of each individual for all loci. Markers that displayed polymorphism in the subset of *Smutsia temminckii* samples were selected for further analysis. Polymerase chain reaction products of different sizes and labelled with different fluorescent labels were plexed together (Table 3). A plex consisted of 5 µl of each PCR product (per sample).

The PCR was done with 1.25 µl of 10 x QIAGEN PCR buffer, 2.5 µl 5 x Q-Solution, 1.25 µl of 0.2 mM dNTPs, 200nM each of forward and reverse primer, 0.125 µl QIAGEN *Taq* polymerase and 1.5 µl 1 mM $MgCl_2$ was added and made up to 12.5 µl with distilled water. Finally 2.5 µl of 20 ng DNA of each sample type was added. The PCR was performed with one 3 min cycle at 94°C; 10 cycles of 30 s at 94°C, 30 s at 50°C and 30 s at 72°C; 10 cycles of 30 s at 94°C, 30 s at 45° and 30 s at 72°C; 20 cycles of 30 s at 94°C, 30 s at 40° and 30 s at 72°C and one 40 min cycle at 72°C.

The various PCR products were visualized on agarose gels as the size ranges of the amplified PCR products were large enough that polyacrylamide gels were unnecessary.

A total 2 µl loading dye and 8 µl amplicon was loaded onto a 2% (m/v) agarose (consisting of 2 g agarose, 100 ml 1x TBE and 2.5 µl Gel Red™) gel and run for 30 min at 100 volts. Thereafter the gel was placed on a UV light to enable visualisation.

3.5 Fragment Analysis

A volume of 0.4 µl of a fluorescent internal standard GeneScan™-500 Liz® was added to a total of 1 µl PCR product and the sample was denatured by using 8.6 µl of Hi-Di™ Formamide (genetic analysis grade). The ABI PRISM® 3130 DNA genetic analyser was calibrated with Applied Biosystems™ five-dye chemistry, the DS-33 Dye set. The ABI PRISM® 3130 DNA genetic analyser was used for electrophoresis of the samples. The PCR product was detected by a laser that illuminates the incorporated fluorescent dyes. GeneMapper™ Software (Applied Biosystems™) was used to analyse the wavelengths that are characteristic of the light emitted by particular dyes and which was collected throughout the run and for allele scoring (see Appendix 1).

3.6 Molecular Analysis

Of the 80 samples collected, 79 samples were analysed as the single faecal sample (collected in the western region) did not amplify during PCR. Of the 14 loci selected, 6 loci were used for further molecular analysis as they amplified successfully across ≥ 50% of the samples. To determine whether genotyping errors had occurred the profiles were analysed using MICRO-CHECKER V2.2.3 (van Oosterhout *et al.,* 2004).

Chakraborty *et al.* (1992) and Brookfield (1996) discuss several methods to estimate the frequency of null alleles from an apparent heterozygote deficiency. The choice of

method varies depending on whether some samples failed to amplify, and on whether these non-amplified samples represent null homozygotes. MICROCHECKER employs these methods for determining null alleles as well as a method developed by van Oosterhout *et al.* (2004) in the results produced by the software (van Oosterhout *et al.,* 2004).

The three sampling localities (west, central and east South Africa) and outgroup profiles were then run through the STRUCTURE (Pritchard *et al.,* 2000) programme V2.1 to indicate population differentiation. Firstly a new parameter set was created with a burn in period of 1000 at 1000 repetitions using a no admixture model. The allele frequencies were selected as independent and the program was used to compute the probability of data thereby estimating K. Ten individual repetitions of K = (1-6) were run in order to check for consistency in the results.

The profiles were analysed using ARLEQUIN (Excoffier *et al.*, 2005). ARLEQUIN is a software package integrating several basic and advanced methods for population genetic data analysis, like the computation of standard genetic diversity indices and the estimation of allele and haplotype frequencies amongst many others. Analysis can occur at both intra-population and inter-population levels (Excoffier *et al.,* 2005).

Due to the incompleteness of the profiles, some of the analyses could not be completed using ARLEQUIN. As such POPGENE V1.32 (Yeh & Yang, 1999) was employed. The data format was set as variable as column and the check all function was selected to test for, among others, allele frequencies, polymorphism, Hardy Weinberg equilibrium, linkage

disequilibrium and F-statistics. The significance of p for linkage disequilibrium test was 0.5 and the number of simulations run for the neutrality test was 1000.

To determine the possibility of applying the optimised microsatellites for individual identification in forensic cases, the software programme GIMLET V1.3.3 (Valerie, 2002) was used. In wildlife forensics it is important to be able to identify an individual, but also to be able to assign this individual to a specific locality. Due to the lack of resources and often a limited amount of sample in wildlife forensic cases, it is vital that the number of markers needed to successfully identify an individual is known. GIMLET (Valière 2002) was used estimate the probability of identity, whereby the probability that two individuals in the populations share the same genotype (PI) is computed using the equations of biased PI, unbiased PI (with sample size correction) and PI for sibs (Valière 2002).

Chapter 4: Results of the optimization of protocols and evaluation of non-invasive samples

4.1 Sample collection and DNA isolation

Due to the variety of sample types and conditions under which the samples were collected it was necessary to optimize the DNA isolation protocols. It was also attempted to limit the amount of inhibitors present in the sample by diluting the extracted DNA from scale, tissue, museum and faecal samples, thereby ensuring good quality DNA. In this study, any problems regarding contamination were overcome by observing strict laboratory guidelines, such as using aseptic techniques, to minimize contamination to nearly negligible.

4.2 DNA quantification

The quality and quantity of the DNA was assessed using a Thermo Fisher Scientific NanoDrop 1000 Spectrophotometer. The results were used to determine which samples will be used for downstream applications and which sample types needed dilution, as based on the A260/280 measurements for sample purity (see Appendix 2). From these measurements it was found that the faecal sample and the scale, museum and tissue samples would need to be diluted for downstream application, whereas the blood samples (and on FTA paper) needed no dilution. The quantity of DNA extracted varied from 0.6 ng/µl to 246 ng/µl, with most samples being well out of the recommended 1.8 A260/280 value range.

4.3 DNA Profiling and optimization

Once optimal quantities of DNA was isolated from the samples, the PCR protocol was optimized using samples of different quality, as determined by the NanoDrop results, to determine whether a) the samples, especially those collected non-invasively, yielded usable results for downstream applications and b) to determine if a so-called universal protocol could be adopted. The PCR protocol published by Luo *et al.* (2007) did not yield successful amplification of the samples and therefore it was necessary to optimise the protocol. Therefore the protocol was optimized by varying three factors: Varying $MgCl_2$ concentrations and a PCR with an annealing temperature gradient using various *Taq* DNA polymerase.

4.3.1 Varying $MgCl_2$ concentrations:

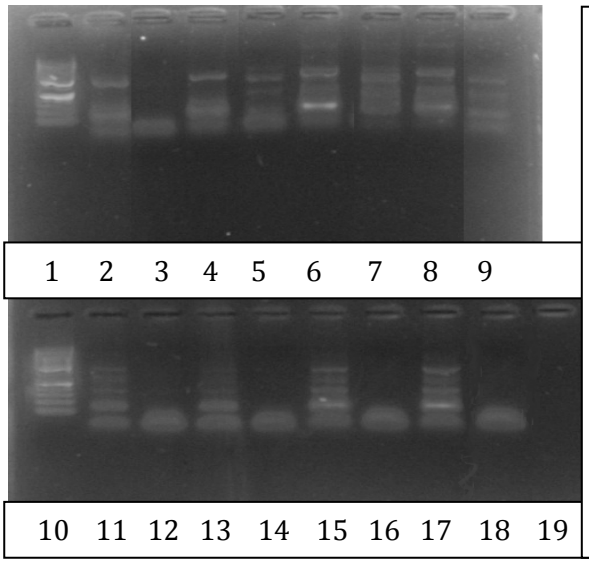
Optimization of magnesium is critical since *Taq* DNA polymerase is an Mg-dependent enzyme and the template DNA primers and dNTPs bind Mg. Too much Mg stabilizes the DNA double strand and prevents complete denaturation of DNA, which reduces yield. Too little Mg reduces the amount of product (Markoulatos *et al*., 2002). Because template DNA primers are Mg concentration dependent, it was attempted to optimize the PCR by varying these concentrations.

QIAGEN *Taq* DNA polymerase is a high-quality recombinant enzyme that is suitable for general and specialized PCR applications. The QIAGEN PCR kit is supplied with the unique QIAGEN PCR Buffer that minimizes the requirement for optimization as it contains both KCl and $(NH_4)_2SO_4$ providing important advantages over conventional buffers, such as a wider temperature range or specific annealing and a greater tolerance

to variable Mg concentration. Q-Solution, also included in the kit, is a novel additive that enables efficient amplification of GC-rich templates by modifying the melting behaviour of DNA. Use of this unique reagent will often enable or improve sub-optimal PCR (Anon, 2012c). For these reasons it was decided to use QIAGEN *Taq* DNA polymerase to optimize the protocol with varying Mg concentrations.

Invasive samples:

Invasive samples (blood and tissue) with an A260/280 value in the range of 1.6 – 2.0 were tested with primer sets MJA09 and MJA16 and the results are shown in Figure 4 below.



Samples were loaded as follows:

1 – Ladder (50-2k)	10 – Ladder (50-2k)
Primer set MJA16	**Primer set MJA09**
2 – Blood (2 mM $MgCl_2$)	11 – Blood (2 mM $MgCl_2$)
3 – Tissue (2 mM $MgCl_2$)	12 – Tissue (2 mM $MgCl_2$)
4 – Blood (2.5 mM $MgCl_2$)	13 – Blood (2.5 mM $MgCl_2$)
5 – Tissue (2.5 mM $MgCl_2$)	14 – Tissue (2.5 mM $MgCl_2$)
6 – Blood (3 mM $MgCl_2$)	15 – Blood (3 mM $MgCl_2$)
7 – Tissue (3 mM $MgCl_2$)	16 – Tissue (3 mM $MgCl_2$)
8 – Blood (3.5 mM $MgCl_2$)	17 – Blood (3.5 mM $MgCl_2$)
9 – Tissue (3.5 mM $MgCl_2$)	18 – Tissue (3.5 mM $MgCl_2$)
	19 – Negative control

Figure 4: PCR products from blood and tissue samples amplified with primer sets MJA09 and MJA16 using varying $MgCl_2$ concentrations.

Non-invasive samples vs. Invasive samples:

Non-invasive (scales and museum) specimens with an A260/280 value in the range of 1.6 – 2.0 was used and compared to invasive samples (Tissue) with an A260/280 value

in the range of 1.6 – 2.0 by amplification with primer sets MJA09 and MJA16 and the results are shown in Figure 5 below.

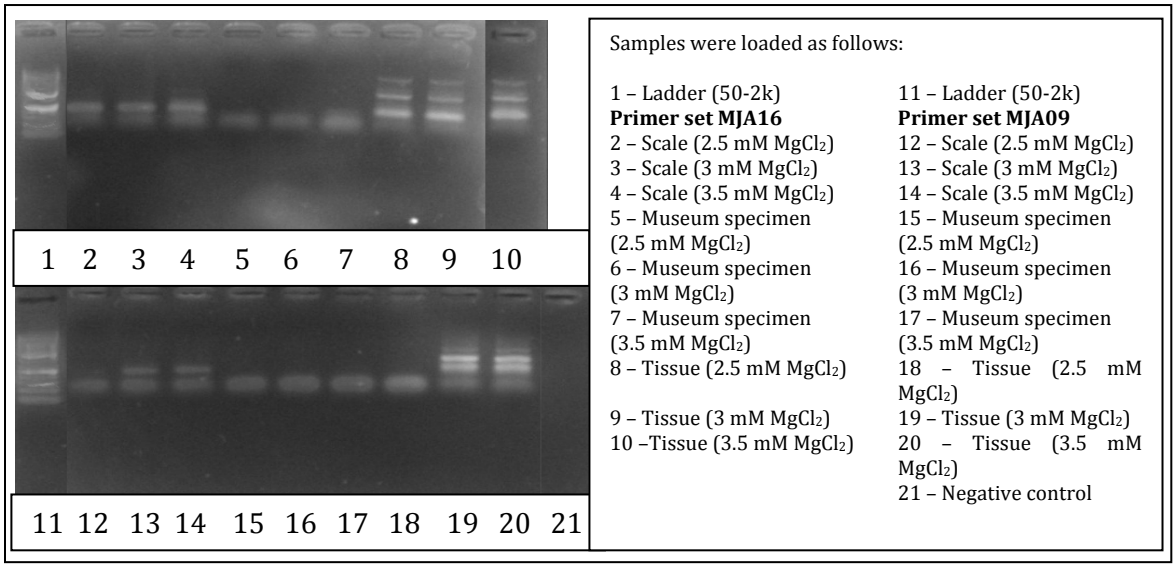


Figure 5: PCR products from scale, museum and tissue samples amplified with primer sets MJA09 and MJA16 using varying $MgCl_2$ concentrations.

4.3.2 PCR with a temperature gradient using SuperTherm *Taq*

Annealing temperature on the other hand is critical when co-amplifying the same loci in a multiplex mixture. A study by Markoulatos *et al*. (2002) showed that although many individual loci could be specifically amplified at 56–60°C, lowering the annealing temperature by 4–6°C was required for the same loci to be co-amplified in multiplex mixtures. When many specific loci are simultaneously amplified, the more efficiently amplified loci will negatively influence the yield of product from the less efficient loci (Markoulatos *et al*., 2002). Therefore the annealing temperature gradient was chosen as a second variable to optimize.

SuperTherm *Taq* has improved fidelity as shown in the lower error frequency than standard *Taq* DNA polymerase. With increased thermal stability, SuperTherm *Taq* has a 2½ hour half life at 96°C and is better suited to cope with the high temperatures often required to overcome the problems associated with G-C rich regions or secondary structure (Anon, 2012c). It was for these reasons that SuperTherm *Taq* was used to optimize the protocol with an annealing temperature gradient.

Invasive samples:

Invasive samples (tissue) with an A260/280 value in the range of 1.6 – 2.0 were amplified with primer sets MJA09, MJA16, MJA21 and MJA22 and the results are shown in Figure 6 below.

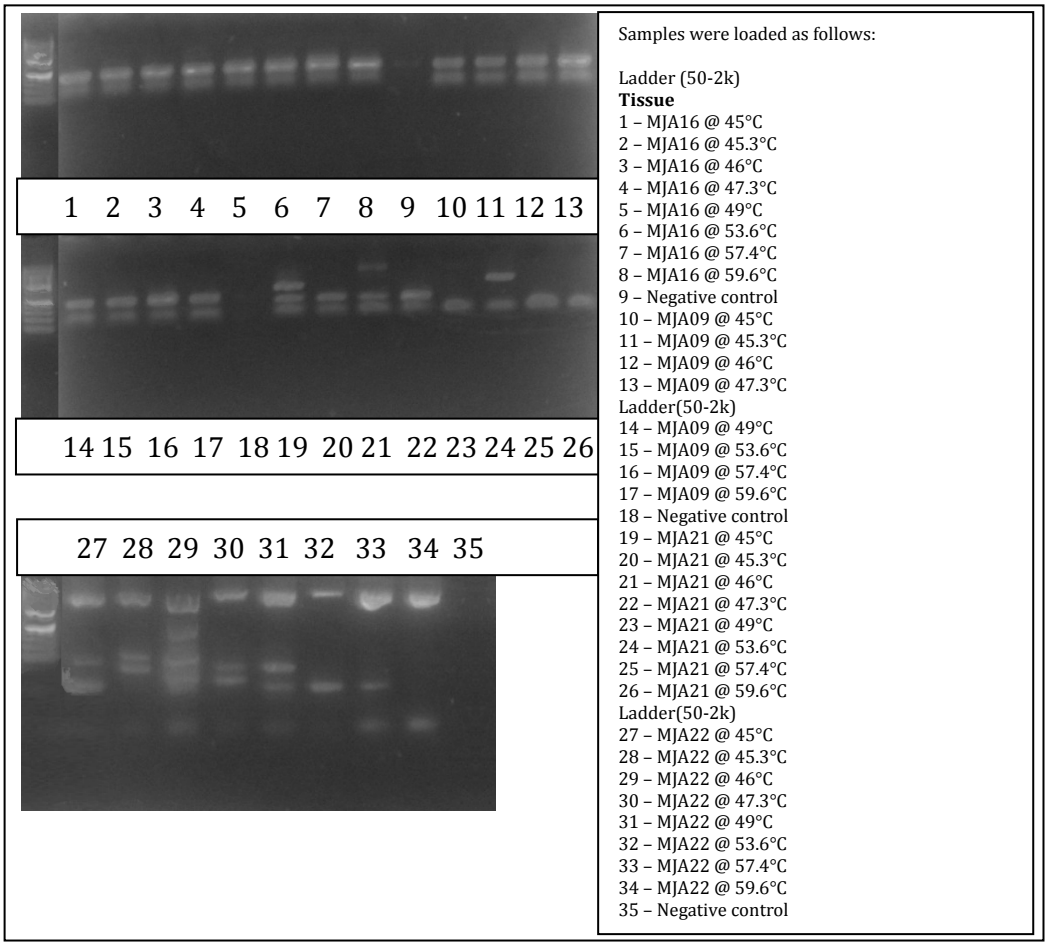
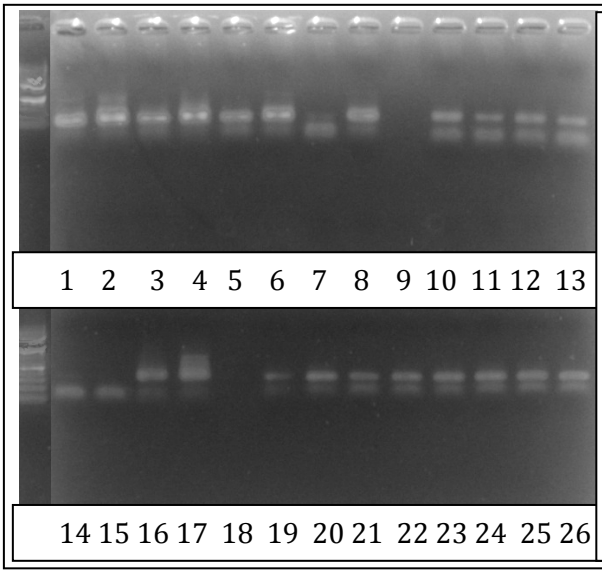


Figure 6: PCR products from tissue samples amplified with primer sets MJA09, MJA16, MJA21 and MJA22 using an annealing temperature gradient.

Non-invasive samples:

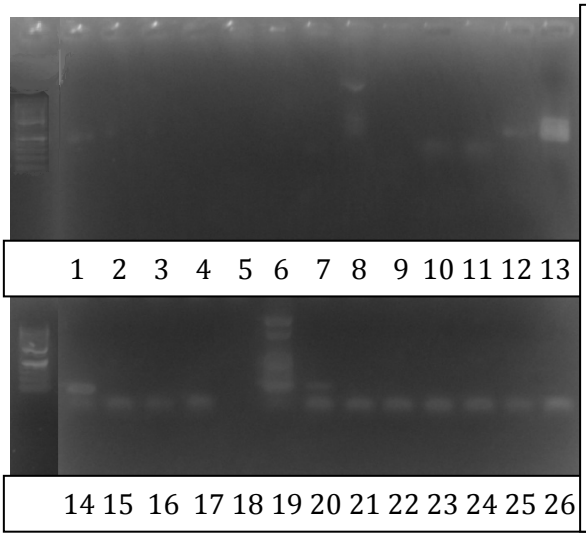
Non-invasive samples (scales and faecal) with an A260/280 value in the range of 1.6 – 2.0 and an A260/280 value of 2.5 respectively were tested with primer sets MJA09, MJA16, MJA21 and MJA22 and the results are shown in Figure 7 below.



Samples were loaded as follows:

Ladder (50-2k)	Ladder (50-2k)
Scale	
1 – MJA16 @ 45°C	14 – MJA09 @ 49°C
2 – MJA16 @ 45.3°C	15 – MJA09 @ 53.6°C
3 – MJA16 @ 46°C	16 – MJA09 @ 57.4°C
4 – MJA16 @ 47.3°C	17 – MJA09 @ 59.6°C
5 – MJA16 @ 49°C	18– Negative control
6 – MJA16 @ 53.6°C	19 – MJA21 @ 45°C
7 – MJA16 @ 57.4°C	20 – MJA21 @ 45.3°C
8 – MJA16 @ 59.6°C	21 – MJA21 @ 46°C
9 –Negative control	22 – MJA21 @ 47.3°C
10 – MJA09 @ 45°C	23 – MJA21 @ 49°C
11 – MJA09 @ 45.3°C	24 – MJA21 @ 53.6°C
12 – MJA09 @ 46°C	25 – MJA21 @ 57.4°C
13 – MJA09 @ 47.3°C	26 – MJA21 @ 59.6°C

Figure 7a: PCR products from scale and faecal samples amplified with primer sets MJA09, MJA16, MJA21 and MJA22 using an annealing temperature gradient.



Samples were loaded as follows:

Ladder (50-2k)	Ladder (50-2k)
Scale	
1 – MJA22 @ 45°C	14 – MJA16 @ 49°C
2 – MJA22 @ 45.3°C	15 – MJA16 @ 53.6°C
3 – MJA22 @ 46°C	16 – MJA16 @ 57.4°C
4 – MJA22 @ 47.3°C	17 – MJA16 @ 59.6°C
5 – MJA22 @ 49°C	18 –Negative control
6 – MJA22 @ 53.6°C	19 – MJA09 @ 45°C
7 – MJA22 @ 57.4°C	20 – MJA09 @ 45.3°C
8 – MJA22 @ 59.6°C	21 – MJA09 @ 46°C
Faecal	22 – MJA09 @ 47.3°C
9 – Negative control	23 – MJA09 @ 49°C
10 – MJA16 @ 45°C	24 – MJA09 @ 53.6°C
11 – MJA16 @ 45.3°C	25 – MJA09 @ 57.4°C
12 – MJA16 @ 46°C	26 – MJA09 @ 59.6°C
13 – MJA16 @ 47.3°C	

Figure 7 b: PCR products from scale and faecal samples amplified with primer sets MJA09, MJA16, MJA21 and MJA22 using an annealing temperature gradient.

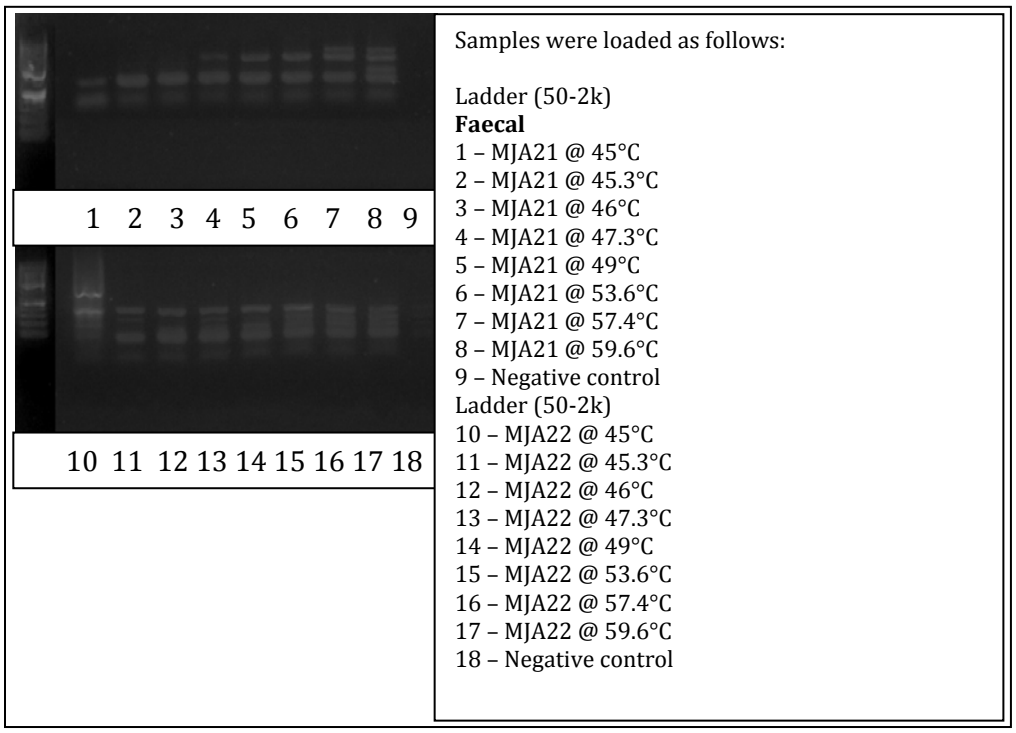


Figure 7 c: PCR products from scale and faecal samples amplified with primer sets MJA09, MJA16, MJA21 and MJA22 using an annealing temperature gradient.

From these primary optimization results if was found that non-invasive samples as compared to invasive samples amplified equally well during PCR, based on amplification as seen in the quality and quantity of bands visualized on the gels. When comparing the results of tissue samples (invasive) in Figure 6 and scales (non-invasive) in Figure 7a, it can be seen that the bands are of equal width and clarity.

Furthermore it was found that non-invasive samples, such as scales, yielded usable DNA for downstream applications such as PCR, as shown by successful amplification in Figures 5, 7a and 7b. Even though faecal samples amplified with some success as shown in Figure 7b and 7c, it was concluded that Ground Pangolin faecal samples were in fact not viable for downstream applications due to the quality and quantity of bands visualized on the gels.

The NanoDrop results (see Appendix 2) however, indicated that not all samples with the recommended A260/280 value of 1.8 amplified successfully. For example, sample 4 (an invasive sample) had an A260/280 value of 1.82 and a DNA concentration of 4.8ng/µl, but only amplified with two microsatellites (see Appendix 1). Sample 33 (a non-invasive sample) on the other hand amplified successfully with all six microsatellites, but had an A260/280 value of 1.16, which is well out of range, and a DNA concentration of 9ng/µl. Ideally 2ng/µl of DNA is used for PCR reactions, and samples are diluted, which would mean that the A260/280 value of sample 33 decreased even more upon dilution. After comparing the NanoDrop results from invasive (blood and tissue) and non-invasive (scales, nasal/oral swabs and museum) samples, it was found that invasive samples constantly yielded better results, as based on the A260/280 value. When comparing,

however, the samples that amplified successfully across all six loci, it is found that on average, more non-invasive samples were successful than invasive samples. This may be an indication that, in the case of this study at least, non-invasive samples perform better.

4.4 Fragment Analysis

The samples were amplified using the optimized PCR protocol and the following microsatellites from the study by Luo *et al.* (2007): MJA05, MJA02 and MJA28 yielded no results across all the samples. MJA03, MJA14, MJA12 (Plex4) and MJA21 and MJA18 did not amplify successfully across 50% of the samples irrespective of whether the samples were invasive of non-invasive. The reason for non- or partial amplification of samples with these microsatellites may very well be due to the fact that the markers used were cross-species markers. The remaining loci amplified successfully across more than half of the samples, therefore allowing sufficient data to work with. These loci were MJA07, MJA08, MJA09, MJA13, MJA16 and MJA22.

Chapter 5: Results of the molecular and population analyses

Seventy nine profiles were generated but only those microsatellites that had successfully amplified across ≥50% of the samples were analysed therefore only 6 loci were evaluated (Appendix 1).

5.1 Molecular Analysis

DNA degradation, low DNA concentrations and primer-site mutations may result in the incorrect assignment of microsatellite genotypes, caused by such occurrences as allelic drop out, potentially biasing population genetic analyses (Van Oosterhout *et al.*, 2004). Many instances of low DNA concentrations, averaging in the range of 25ng/µl but some as low as 0.6ng/µl, in most samples were present in the current study. Therefore there was cause for concern that allelic drop-out had occurred.

Genotyping errors are also known to occur, according to Zhan *et al.* (2010), when using DNA isolated from non-invasive samples, as was the case with most samples collected in the study presented here. To screen the dataset for possible genotyping errors and also to estimate the frequency of null alleles, data was imported into the software programme MICRO-CHECKER (Van Oosterhout *et al.*, 2004). MICRO-CHECKER can help identify genotyping and scoring errors and where multi-locus genotypes are available, it can also discriminate between inbreeding and Wahlund effects (a reduction in heterozygosity), and Hardy–Weinberg deviations caused by null alleles. Furthermore, it can identify possible typographic errors and if null alleles are present, it can estimate the null allele frequency, and adjust the observed allele and genotype frequencies

accordingly. These adjusted allele frequencies can be used subsequently for further population genetic analysis (van Oosterhout *et al.,* 2004).

The profiles listed in Appendix 1 were analysed as one population and included all samples and the outgroup. Firstly the profiles were checked for unusual observation and it was found that 235 values were outside the expected range, and 3 values had an inconsistent modulus. The profiles were analyzed at a confidence interval of 95% at 1000 repeats and all suspect data was included. There was no evidence for large allele dropout but stuttering might have resulted in scoring errors, as was indicated by the highly significant shortage of heterozygote genotypes with alleles of one repeat unit difference. Further to this, more than 50% of the alleles at each locus were of one allele size class and therefore binomial analysis could not be performed. Also, due to the general excess of homozygotes for most allele size classes all loci show evidence for null alleles. The results given by the programme state that this population is also not in Hardy Weinberg equilibrium, but this may be because the populations were all assessed as one and differentiated based on the 6 loci evaluated (table 4).

Table 4: Summary of results from four methods (van Oosterhout; Chakraborty; Brookfield 1; Brookfield 2) for determining null alleles across six loci.

Locus	Null Allele Present	Van Oosterhout	Chakraborty	Brookfield 1	Brookfield 1
22	Yes	0.2278	0.3118	0.182	0.2689
9	Yes	0.2367	0.3127	0.2007	0.311
8	Yes	0.2276	0.3148	0.2172	0.449
7	Yes	0.2502	0.3514	0.2	0.7796
16	Yes	0.366	0.616	0.3484	0.7539
13	Yes	0.2802	0.398	0.2633	0.6804

The three sampling localities and outgroup profiles were then run through the STRUCTURE (Pritchard *et al.,* 2000) programme to confirm possible population differentiation. STRUCTURE is a software programme that is used to assess the genotypes of individuals in a fully Bayesian approach using a clustering method. It involves the placing of individuals into K (number) of populations of origin and it simultaneously assigns individuals to the population with explicit estimates of their 90% confidence intervals. The results provide a probability value that can be interpreted directly as a probability of origin for each individual tested in each population. This software attempts to assign individuals to populations on the basis of their genotypes (Pritchard *et al.,* 2000).

A study by Evanno *et al.* (2005) found that the intensity of sampling, and as such the number of samples collected in one location as compared to another location, has an impact on the correct detection of the number of groups, and that microsatellites perform better than other types of markers in population structure detection because they are both codominant and highly polymorphic. Even though microsatellites were used in this study, the intensity of sampling may impact the results. There were 42 samples collected in the west, but only 24 in the east and 7 in the central sampling localities.

Six runs with different K values were run at a random configuration, as suggested by Pritchard *et al.* (2000), with a burn in period of 1000 at 1000 replications using a no admixture model. The results showed a clear plateau in –Ln(Probability) values at K=4 (Figure 8), with this also being the highest probability value overall. Since there is a clear plateau of probability estimates from K=4 onwards, it was regarded as true K and

it was not considered necessary to use other methods to elucidate true K, such as the ΔK approach by Evanno *et al.* (2005).

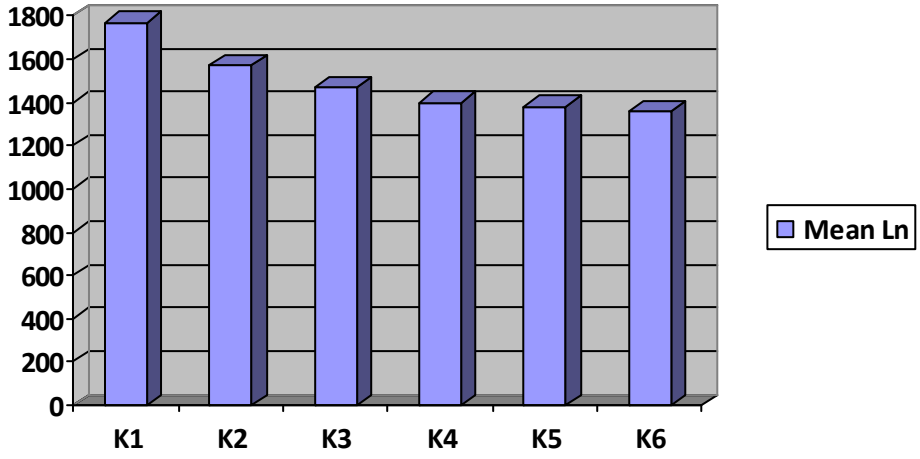


Figure 8: –Ln(Probability) for K = 1 – 6 in the Ground Pangolin population.

A no admixture model assumes that individuals come purely from one of the K populations and this model is appropriate for studying discrete populations, as is the case of the pangolins due to their distinct geographical isolation. This model also detects subtle population structure. The alpha values for degree of admixture were at K =3 (0.1369); K = 4 (0.0940) and K = 5 (0.0835). This would suggest that there is a small degree of admixture within the populations indicating that this may once have been one population and that geographic dispersal took place (Figure 9).

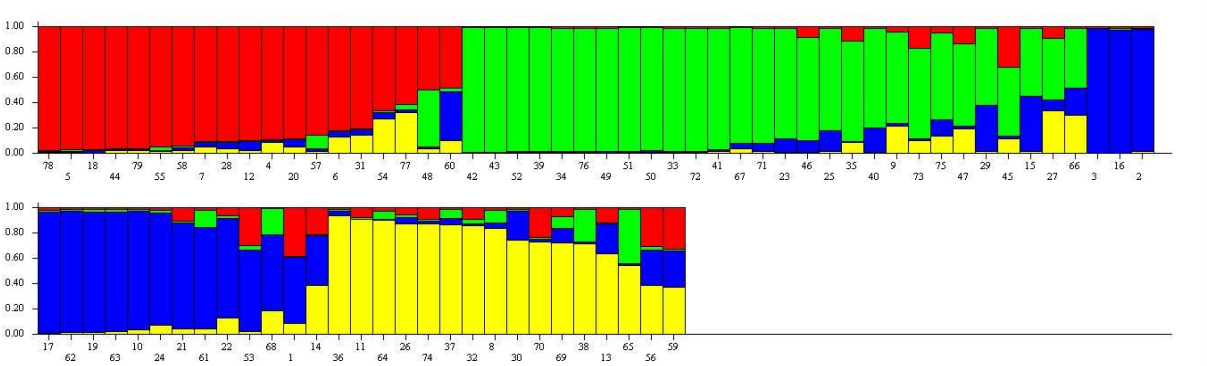


Figure 9: STRUCTURE analysis with samples representing west (yellow); east (green); central (red) sampling localities and an outgroup (blue) with K = 4.

Samples were input into STRUCTURE as follows:

1 - Outgroup	23 – East	45 - East	67 - East
2 - Outgroup	24 - West	46 - East	68 - West
3 - Outgroup	25 - West	47 - East	69 - East
4 - West	26 - West	48 - East	70 - West
5 - West	27 - West	49 - East	71 - East
6 - West	28 - Central	50 - Central	72 - East
7 - Museum	29 - West	51 – East	73 - West
8 - West	30 - West	52 – East	74 - West
9 - West	31 - West	53 - Central	75 - East
10 - West	32 - West	54 - Central	76 - East
11 - West	33 – West	55 - Central	77 – Central
12 - West	34 - West	56 - West	78 - Outgroup
13 - West	35 - West	57 - West	79 - Museum
14 - West	36 - West	58 - West	
15 - West	37 - West	59 - East	
16 - West	38 - West	60 - East	
17 - West	39 – East	61 - West	
18 - West	40 – East	62 - East	
19 - West	41 – East	63 - East	
20 - West	42 – East	64 - Outgroup	
21 - West	43 – East	65 - West	
22 - West	44 – Central	66 - East	

The genetic structure of populations is not always reflected in the geographical proximity of individuals (Evanno *et al.*, 2005). It is interesting to note that the localities where the samples were collected did not necessarily coincide with the results in Figure 9. This may be an indication that some individuals in a current population may be historically genetically related to other populations, as confirmed by the alpha values for degree of admixture. For example, sample 60 was collected in the east, but the results show a higher similarity with the central sampling locality and outgroup, than with the eastern sampling locality. Based on the assignment of individuals to specific populations by STRUCTURE, the data was grouped accordingly for intra-population analysis.

Population analysis includes the computation of different standard genetic indices such as the number of different alleles, standard deviation and the expected and observed heterozygosity. In Table 5 a comparison is made between the results found in this study and that of Luo *et al.* (2007) of the six loci that were used for analysis on the basis of the number of alleles and size ranges of the alleles from the Malayan Pangolin. It should be taken into consideration that for the study on the Malayan Pangolin 38 samples were assessed, whereas in the current study 79 samples were assessed.

Table 5: Comparison of the number of alleles and allele size ranges at 6 loci for the Ground Pangolin (current study) and the Malayan Pangolin (Luo *et al.,* 2007).

Locus	Malayan Pangolin		Ground Pangolin	
	Number of Alleles	**Allele size range**	**Number of Alleles**	**Allele size range**
MJA07	11	238-270	4	158-182
MJA08	3	178-184	12	116-190
MJA09	8	200-216	10	132-174
MJA13	10	204-226	12	224-256
MJA16	10	170-208	10	126-164
MJA22	13	196-226	11	127-145

The results in Table 5 show that both populations had similar numbers of alleles for MJA 09, 13, 16 and 22. MJA07 and 08 however do not share similar numbers of alleles, and it would appear that for MJA08 the Ground Pangolin populations have greater allelic diversity than the Malayan Pangolin population, and the converse is true for MJA07. This indicates that the DNA region in which the microsatellite MJA08 and MJA07 is located is a site for possible differentiation between the Ground and Malayan Pangolin.

Table 6 shows the expected and observed heterozygosities for each locus across all Ground Pangolin populations as compared to the Malayan Pangolin.

Table 6: Observed and expected heterozygosity values for Malayan and Ground Pangolin populations

Locus	**Malayan Pangolin (Luo *et al.,* 2007)**		**Ground Pangolin (This Study)**	
	Expected Heterozygosity	**Observed Heterozygosity**	**Expected Heterozygosity**	**Observed Heterozygosity**
MJA07	0.862	0.792	0.680	0.462
MJA08	0.312	0.375	0.814	0.387
MJA09	0.773	0.667	0.746	0.406
MJA13	0.701	0.652	0.837	0.544
MJA16	0.833	0.750	0.827	0.300
MJA22	0.925	0.570	0.611	0.400
Total	0.734	0.634	0.753	0.416

The observed heterozygosities for the Ground Pangolin populations were almost always consistently lower than that of the Malayan Pangolin populations except for MJA08, due to the fact that the Ground Pangolin populations have a greater allelic diversity for that microsatellite compared to the Malayan Pangolin populations. When comparing observed and expected heterozygosities it is generally accepted that the closer these values are, the better the genetic diversity within the population. Within the Ground Pangolin populations there is a significant difference between the observed and expected heterozygosities, which indicates diminished genetic diversity.

When looking at the three sampling localities, the observed heterozygosity was markedly lower than the expected heterozygosity, as shown in Table 7. This may be an indication that, as found in the MICROCHECKER (Van Oosterhout *et al.*, 2004) results, that the population is not in Hardy Weinberg equilibrium. The current populations

could not be compared to the historical populations because, as indicated in Table 7, only two samples could be typed and only across three loci.

When comparing the sampling localities to each other the expected heterozygosities differ significantly from the observed heterozygosities. It would appear that the eastern sampling locality has higher genetic diversity than the central or western sampling localities.

Table 7: A comparison of expected and observed heterozygosities at six loci in 5 sampling localities: central (C), east (E), outgroup (OG), west (W) and historic (H) populations.

Population	**Sample size**	**Loci typed**	**Expected Heterozygosity**	**Observed Heterozygosity**
C	7	6	0.6903	0.4504
E	24	6	0.7084	0.5214
H	2	3	0.3333	0.3333
OG	5	6	0.6350	0.4695
W	41	6	0.7525	0.3321

POPGENE V1.32 (Yeh & Yang, 1999) was then used to test for Hardy-Weinberg equilibrium using the Chi-square method and the Likelihood-ratio. In both tests, the probability was less than 0.05 and as such the hypothesis that the population was in Hardy-Weinberg equilibrium was rejected.

When making use of F-statistics, a few factors can be addressed that are of concern such as genetic distance and gene flow. Genetic distance is a comparison of the variance of allele frequencies between populations, where if the distance is zero the populations are identical and if it is one the populations are completely different. POPGENE V1.32 (Yeh

& Yang, 1999) was used to determine the F-statistics in the Ground Pangolin population. It was found that the central, western and eastern sampling localities had genetic distances between each other of less than 0.5. The lowest genetic distance, however, was between the central sampling locality and the western sampling locality, at 0.1510. Conversely all of the sampling localities had genetic distances with the outgroup of more than 0.5. The highest genetic distance being between the eastern sampling locality and the outgroup at 1.0410. There was one exception though, in that the genetic distance between the central sampling locality and the outgroup was only 0.3557. This may be due to the fact that in the central sampling locality only 7 samples were assessed as compared to the larger sample sizes in the eastern and western sampling localities at 24 and 42 samples respectively.

Gene flow can also be determined using F-statistics where Fst is seen as a reduction in heterozygosity and Nm the number of migrants per generation. Therefore if there are zero migrants per generation the reduction in heterozygosity will be one, and vice versa. According to POPGENE V1.32 (Yeh & Yang, 1999) MJA16 had 1 migrant per generation, MJA08 and MJA13 had 2 migrants per generation, MJA22 had 3 migrants per generation; and MJA09 had 4 migrants per generation. This indicates a high level of gene flow between the sampling localities as the Fst values for these loci were proportionately low and for all 5 loci mentioned below 0.1. Locus MJA07 however, had an Nm value of 0.6544 and a Fst value of 0.2764 which is not as high as the other loci, but not necessarily a negative indication.

POPGENE V1.32 (Yeh & Yang, 1999) was finally used to test for linkage disequilibrium between the loci. This measure is important, especially in wildlife forensics, where a marker needs to be discriminatory to test for identity. The results of the linkage disequilibrium testing indicated that the markers MJA22, MJA16, MJA13 MJA09 and MJA08 had a probability of less than 0.05 and as such were in linkage disequilibrium. MJA07 was the only marker that had a probability of more than 0.05. This may be because MJA07 only has 4 alleles, which is markedly less than the other markers.

Another application of the data gathered is the identification of a species, population or individual based on their genotype. GIMLET (Valière, 2002) estimates the probability of identity, whereby the probability that two individuals in the populations share the same genotype (Probability of identity; PI) is computed using the equations of biased PI, unbiased PI (with sample size correction) and PI for sibs (Valière, 2002). In wildlife forensics it is important to be able to identify an individual, but also to be able to assign this individual to a specific locality. Due to the lack of resources and often a limited amount of sample in wildlife forensic cases, it is vital that the number of markers needed to successfully identify an individual is known. The results from GIMLET (Valière, 2002) are shown in Table 8.

Table 8: Probability of identity per locus for the Ground Pangolin arranged from most to least informative markers, using unbiased probability of identity (PIunb), biased probability of identity (PIb) and probability of identity for sibs (PIsibs).

Locus Name	PIunb	PIb	PIsibs
MJA22	3.942e-02	4.521e-02	3.437e-01
MJA09	1.796e-03	2.323e-03	1.202e-01
MJA08	9.002e-05	1.307e-04	4.279e-02
MJA07	7.224e-06	1.157e-05	1.699e-02
MJA16	1.002e-06	1.960e-06	7.768e-03
MJA13	1.627e-07	2.960e-07	3.647e-03

When plotting these probabilities (Figure 10), it is shown that identity can, with a high probability, be determined after the use of 4 loci (PI = 7.224e-06), with MJA22 being the most informative locus for Ground Pangolin populations. Unfortunately, due to the limited information available on the current status of Ground Pangolin populations, and estimated population size cannot be given.

Furthermore, due to the fact that all the markers except MJA07 are in linkage disequilibrium these markers will not be sufficient to use in forensic genetics for the purpose of identifying an individual.

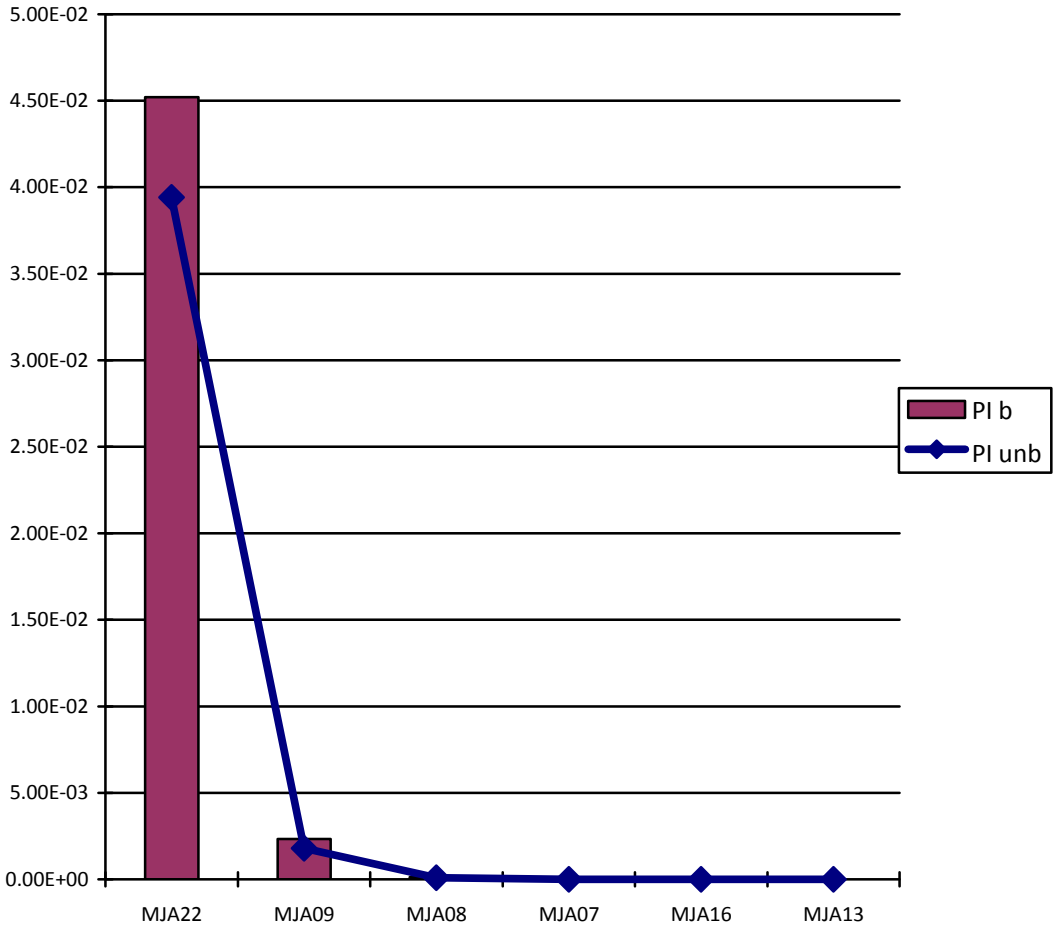


Figure 10: Probability of identity per locus arranged from most to least informative locus.

Chapter 6: Discussion

The aim of this study was to explore the genetic variation of the Ground Pangolin using optimised cross-species microsatellite markers to assess the level of genetic diversity within current South African populations as compared to historic populations and to determine if these markers are suitable to possibly be applied in wildlife forensic investigations. Results of this study should indicate if variations within the DNA profile of *Smutsia temminckii* exist, particularly between subpopulations and involved sampling existing populations and, in regions where the species is very rare or now absent, museum specimens. The specific aims will now be addressed.

6.1 DNA isolation protocols for invasive, non-invasive and museum samples

As discussed in Chapter 2 the various protocols for isolating DNA from invasive, non-invasive and museum samples were screened to test which gave the best results. The optimised protocols are adaptations of various Qiagen kit protocols and are discussed in Chapter 2. It should be mentioned that DNA isolation from non-invasive samples is becoming increasingly more successful with the implementation and development of new kits, especially for forensic purposes, such as those used in this study. It was found that non-invasive samples and invasive samples had similar quality and quantity of DNA isolated as based on the spectrophotometric results.

It was further determined whether samples collected non-invasively such as scats, oral swabs and scales yield viable DNA based on the spectrophotometric results. The A260/280 measurement was in the range of the recommended 1.8 for 46 of the 80

samples excluding the faecal sample. This was to be expected as DNA isolated from faecal samples is often degraded, which would make genotyping erroneous, if not impossible (Bubb *et al.,* 2010). This was also proven when, during the optimisation of the microsatellites, a faecal sample was included and produced no usable results. Furthermore, the results produced during the optimization of the PCR protocol indicated that DNA extracted from samples collected non-invasively were viable for downstream applications except for faecal samples.

6.2 PCR protocols and microsatellites designed for the Malayan Pangolin used with the Ground Pangolin

MJA05, MJA02 and MJA28 yielded no PCR product, but this is not unusual as described in the study by Taberlet *et al.* (1999). In the study mentioned, it was found that during genotyping with samples collected non-invasively, that no product was often obtained due to the fact that the DNA available was often in the picogram range and thus very low. This is not only the case with non-invasive samples however, as invasive samples that are not collected or stored correctly often degrade due to environmental factors such as temperature, UV exposure and time from collection until storage (Paabo, 1989). Every precaution was taken during the current study to ensure that these factors did not influence the invasively collected samples.

Five microsatellites did not amplify successfully across ≥ 50% of the samples. The success rate of cross-species microsatellite amplification has been shown to be directly related to the evolutionary divergence between the species from which the microsatellite loci have been isolated (the source species) and the species to which the

heterologous loci are being applied (Primmer *et al.,* 2005). It may very well be that the Ground Pangolin has sufficiently diverged, on an evolutionary level, from the Malayan Pangolin (source species), to influence the success rate of the cross-species microsatellite amplification as suggested in the study by Primmer *et al.* (2005).

6.3 Viability of the microsatellites to measure the level of genetic diversity in the Ground Pangolin

There are often problems when using cross-species markers for amplification, one of which is homoplasy, where PCR products of microsatellite loci with the same fragment length, but different sequence can arise from mutational events (deletion or insertion) in the flanking regions of the repeats. It may also be produced by interruptions in perfect repeat producing alleles of the same size, which however are not identical by descent (Gen-Hua *et al.,* 2010). Size homoplasy also increases with time since divergence among populations and taxa, which is/ or may be the case when looking at the phylogeny of the pangolin (see Figure 1).

Therefore in the current study only data generated for the 79 samples with the 6 loci most successfully amplified was analysed. MICRO-CHECKER (Van Oosterhout *et al.*, 2004) was used to analyse the data and found there was no evidence for large allele dropout but stuttering might have resulted in scoring errors, as was indicated by the highly significant shortage of heterozygote genotypes with alleles of one repeat unit difference. Further to this, more than 50% of the alleles at each locus were of one allele size class and therefore binomial analysis could not be performed. Also, due to the

general excess of homozygotes for most allele size classes all loci showed evidence for null alleles.

There may be several reasons for genotyping errors, such as null alleles, to occur. It may be due to low quantity and quality of the DNA template which causes samples not to amplify successfully which can lead to allele dropout. Genotyping errors can also occur due to the preferential amplification of small alleles (i.e. large allele dropout or short allele dominance, Wattier *et al*., 1998), where the larger allele specifically fails to amplify. Errors may also occur when the results are misinterpreted because of biochemical artefacts such as pull-up peaks, split peaks and stutter peaks, but the MICROCHECKER (van Oosterhout *et al.*, 2004) results suggest that this is not the case. Contamination could also have occurred or even human error such as mislabelling of samples or data which is entered into the system incorrectly. Finally a mutation in the primer-binding area may cause null alleles and thus incorrect genotypes (Pompanon *et al.,* 2005; Selkoe & Toonen 2006) which would result in a genotyping error, but not due to incorrect typing or technical errors as in the other cases. These deviations from the correct genotype are often very similar to those caused by inbreeding, assortative mating or Wahlund effects. A heterozygote will thus be interpreted as a homozygote. Other possible causes of null alleles include the preferential amplification of short alleles or slippage during PCR amplification (Chapuis & Estoup, 2007).

Therefore great care should be taken when genotypes are analysed from profiles that were obtained from low quantity or quality DNA, since there is an increase in allelic dropout and stutter peaks (Butler, 2005). Even though the NanoDrop results suggest

that the DNA isolated is, in most cases, of good quality, the quantity was very low for most samples.

After the profiles were screened by MICRO-CHECKER (Van Oosterhout *et al.*, 2004) the data was imported into the software programme STRUCTURE (Pritchard *et al.*, 2000). The alpha values for degree of admixture suggested a small degree of admixture within the populations indicating that this may once have been a single population and that geographic dispersal took place. An intra-population analysis was then performed using ARLEQUIN (Excoffier *et al.,* 2005) and the results are shown in Tables 5, 6 and 7. Preliminary results indicate that the observed heterozygosities are lower than the expected heterozygosities which may suggest that the populations are not in Hardy Weinberg equilibrium, which is also suggested by the results from the MICRO-CHECKER (Van Oosterhout *et al.*, 2004) programme. POPGENE V1.32 (Yeh & Yang, 1999) was used, and confirmed that the population was indeed not in Hardy Weinberg equilibrium. The results further suggested that there was a significant genetic distance between the sampling localities and outgroup, and conversely that the distance between the different sampling localities was less significant. Another measure, gene flow, suggests that there is at least 1 migrant per generation, but could be as many as 4. This information may indicate that the sampling localities aren't as geographically isolated as it may appear.

Due to limited informative markers and the lack of complete profiles for all samples analysed, it cannot be said with certainty what the current status of the genetic diversity is between the various Ground Pangolin populations within South Africa. Preliminary results indicate, however that the Ground Pangolin population has less allelic diversity

when compared to the Malayan Pangolin population based on the 6 loci investigated. The study on the Malayan Pangolin population assessed 38 samples whereas the current study assessed 79 samples, which may also influence the results. Low allelic diversity and highly skewed frequency distributions at microsatellite loci indicates strong genetic drift due to a strong founder effect and/or a previous population bottleneck. Preliminary results from intra-population analysis indicate that the observed heterozygosity was markedly lower than the expected heterozygosity which indicates that the population is not in Hardy Weinberg equilibrium and therefore has diminished genetic diversity.

Limited informative markers and the lack of complete profiles is not unexpected however (Primmer *et al.,* 2005). Cross-species markers were initially selected as they had been used successfully on several Asian species including the Tree Pangolin, but unfortunately the results obtained were not optimal. This serves as motivation for the development of species-specific markers for the Ground Pangolin in future studies.

During sample collection, it was attempted to collect historic samples, such as museum samples, to compare with current populations. Unfortunately this was unsuccessful with only two museum samples obtained, as no other museum held Ground Pangolins in their collections. Both these samples proved to be too degraded to yield a sufficient profile for comparison, as amplification only occurred with three microsatellites.

6.4 Application of microsatellites in illegal wildlife trafficking cases

The six loci were arranged by GIMLET (Valière, 2002) from most informative to least informative, and it was found that MJA22 was most informative. When compounding the probability of identity across all six loci, it was found as indicated in the graph in Figure 10, that 4 loci would be sufficient to determine the identity of an individual in the current Ground Pangolin population. MJA22 was considered most informative due to the fact that it had a low probability of identity. This means that it has the ability to differentiate between two individuals of the same population.

Unfortunately when testing for linkage disequilibrium, POPGENE V1.32 (Yeh & Yang, 1999) results showed that all loci, except MJA07 where in linkage disequilibrium. Therefore these markers cannot be used to test for identity.

Chapter 7: Conclusion

Little is known about the molecular genetic variation of Ground Pangolin populations in South Africa. In this study it was attempted to assess the genetic diversity of populations representing different areas, but this could not be achieved due to the limited amplification success of cross-species markers that were described in the literature.

Isolation protocols have been optimized, and it is shown that non-invasive samples yield good quality and quantity DNA that is usable for down-stream applications and perform as well as invasive samples. The PCR protocol was also optimized, and the results from the optimization chapter will be of assistance when species-specific markers are optimized. This study has shown the need for the development of species-specific markers, based on the genetic diversity results, as it appears that the diversity within Ground Pangolin populations is much lower with the markers tested than in Malayan Pangolin populations (Luo *et. al.*, 2007). This will however have to be confirmed with species specific markers. Furthermore, it appears that MJA22 is the most informative maker tested and may be applied in wildlife forensic cases, in combination with an additional three loci to successfully identity samples of Ground Pangolin, but due to linkage disequilibrium this may not be a viable option.

Summary

Little is known about the molecular genetic variation of Ground Pangolin populations in South Africa. In this study it was attempted to assess the genetic diversity of the populations, but this could not be achieved due to insufficient cross-species markers amplification. It should, however, be emphasized that the molecular work done in this study is novel, and that the results found during this research are key foundations for future studies. During sample collection it was found that three main populations of Ground Pangolins exist in South Africa in the Eastern, Western and Central parts of the country. Isolation protocols have been optimized, and it has been shown that non-invasive samples yield good quality and quantity DNA that is usable for down-stream applications and perform as well as invasive samples. The PCR protocol was optimized, and the results from the optimization chapter will be of assistance when species-specific markers are optimized. This study has also shown the need for the development of species-specific markers, and the use of said markers will give a better indication of the genetic diversity of the Ground Pangolin populations in South Africa. From the statistical analysis it would seem that there are some correlations between the three sampling localities which may indicate a population divergence at some point. Based on the genetic diversity results, it appears that the diversity within Ground Pangolin populations is much lower based on the markers tested than in Malayan Pangolin populations. This will however have to be confirmed with species-specific markers.

Keywords: Ground Pangolin; non-invasive sampling; genotyping errors; population; wildlife forensics; cross-species markers

Opsomming

Min is bekend oor die molekulêre genetiese variasie van die Grond Ietermagog bevolking in Suid-Afrika. In hierdie studie is gepoog om die genetiese diversiteit van die bevolking te bepaal, maar dit kon nie bereik word nie as gevolg van onsuksesvolle kruis-spesies merker amplifisering. Dit moet egter beklemtoon word dat die molekulêre werk wat gedoen is in hierdie studie nuut is, en dat die resultate wat tydens hierdie navorsing gevind is die basis vir toekomstige studies vorm. Gedurende monster versameling is daar belsuit om in drie verskillende areas bevolkings van Grond Ietermagö te versamel wat die Oos-, Wes- en Sentrale dele van Suid Afrika verteenwoordig. Die DNS isolasie protokol is geoptimiseer, en het daarop gewys dat nie-indringende monsters net sulke goeie gehalte en hoeveelhede DNS lewer wat bruikbaar vir verdere toepassing, as indringende monsters. Die PKR-protokol is ook geoptimiseer, en die resultate van die optimiserings hoofstuk sal van hulp wees wanneer spesies-spesifieke merkers geoptimiseer word. Hierdie studie het ook getoon dat daar die behoefte is vir die ontwikkeling van spesie-spesifieke merkers, en gebruik van die merkers sal 'n beter aanduiding gee van die genetiese diversiteit van die Grond Ietermagog-bevolkings in Suid-Afrika. Uit die statistiese analise sou dit blyk dat daar ooreenstemming asook duidelike verskille is tussen die drie bevolkings wat kan dui op 'n bevolkings divergensie op 'n sekere punt. Gegrond op die genetiese diversiteit, blyk dit dat die diversiteit binne Grond Ietermagog bevolkings heelwat laer is gebasseer op die merkers wat getoets is as in die Maleise Ietermagog bevolkings. Dit sal egter bevestig moet word met spesie-spesifieke merkers.

References

Anon. 2002. Seizures and prosecutions - Asia. *TRAFFIC Bulletin,* 19(2): 2.

Anon. 2012a. (http://animals.jrank.org/pages/3343/Pangolins-Pholidota.html) accessed on 24 April 2012.

Anon. 2012b. (http://www.jainbiologicals.com/PDF/dnapoly.pdf) accessed on 2 May 2012.

Anon. 2012c. (http://www.qiagen.com/products/pcr/taqsystem/taqdnapolymerase.aspx#Tabs=t1) accessed on 2 May 2012.

Anon. 2013. (http://blog.arkive.org/2013/04/in-the-news-chinese-ship-crash-reveals-illegal-cargo-of-pangolin-meat/) accessed on 11 June 2013.

Botha, J. & Gaudin, T. 2007. An early Pliocene Pangolin (Mammalia; Pholidota) from Langebaanweg, South Africa. *Journal of Vertebrate Paleontology,* 27(2): 484-491.

Bräutigam, A., Howes, J., Humphreys, T. & Hutton, J. 1994. Recent information on the status and utilization of African pangolins. *TRAFFIC Bulletin,* 15(1): 15-22.

Brunner, P. C., Douglas, M. R. and Bernatchez, L. 1998. Microsatellite and mitochondrial DNA assessment of population structure and stocking effects in Arctic charr, *Salvelinus alpinus* (Teleostei: Salmonidae) from central alpine lakes. *Molecular Ecology,* 7: 209-223.

Bubb, A., Ehlers, K., Kotze, A. & Grobler, J. P. 2010. The effect of sample age and storage method on DNA yield and microsatellite amplification from baboon (*Papio Ursinus*) faecal samples. *European Journal of Wildlife Research,* 57(4): 971-975.

Butler, J. M. 2005. Constructing STR multiplex assays. *Methods in Molecular Biology,* 297: 53-66.

Chakkaravarthy, Q. A. 2012. Research and conservation needs of the Indian Pangolin (*Manis crassicaudata*). *Proceedings of Third Seminar on Small Mammals Conservation Issues,* pp 50-55.

Chapuis, M. P. & Estoup, A. 2006. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution,* 24(3): 621-631.

Che, J., Wang, J., Su, W., Ye, J., Wang, Y., Nie, W., & Yang, F. 2008. Construction, characterization and FISH mapping of a bacterial artificial chromosome library of Chinese pangolin (*Manis pentadactyla*). *Cytogenetic and Genome Research,* 122: 55-60.

Craig, R., Primmer, J. N., Painter, M. K., Jukka, U. P. & Juha, M. 2005. Factors affecting avian cross-species microsatellite amplification. *Journal of Avian Biology,* 36(4): 348-360.

Dávila S. G., Gil M. G., Resino-Talaván P. & Campo J. L. 2009. Evaluation of diversity between different Spanish chicken breeds, a tester line, and a White Leghorn population based on microsatellite markers. *Poultry Science,* 88: 2518-2525.

De Oliveira, L. R., Loizaga de Castro, R., Cardenas-Alayza, S., & Bonatto, S. L. 2012. Conservation genetics of South American aquatic mammals: an overview of gene diversity, population structure, phylogeography, non-invasive methods and forensics. *Mammal Review,* 42(4): 275-303.

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

Doyle, J. J. & Doyle, J. L. 1990. Isolation of plant DNA from fresh tissue. *Focus,* 12: 13-15.

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

Excoffier, L., Laval, G. & Schneide, S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* 1: 47-50.

Farrugia, A., Keyser, C. & Ludes, B. 2010. Efficiency evaluation of a DNA extraction and purification protocol on archival formalin-fixed and paraffin-embedded tissue. *Forensic Science International,* 194: 25-28.

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.

Freitag, S. & van Jaarsveld, A. S. 1997. Relative occupancy, endemism, taxonomic distinctiveness and vulnerability: prioritising regional conservation actions. *Biodiversity and Conservation,* 6(2): 211-232.

Gaudin, T. J., Emry, R. J. & Wible, J. R. 2009. The phylogeny of living and extinct pangolins (Mammalia, Pholidota) and associated taxa: a morphology based analysis. *Journal of Mammalian Evolution,* 16: 235-305.

Gen-Hua, Y., Kovacs, B. & Orban, L. 2010. A new problem with cross-species amplification of microsatellites: Generation of non-homologous products. *Zoological Research,* 31(2): 131-140.

Grobler, J. P. & Van der Bank, F. H. 1993. Genetic variability in South African blue wildebeest (*Connochaetes taurinus*). *Comparative Biochemistry and Physiology,* 106(3): 755-762.

Grobler, J. P., Hartl, G. B., Grobler, N., Kotze, A., Botha, K. & Tiedemann, R. 2004. The genetic status of an isolated black wildebeest (*Connochaetes gnou*) population from the Abe Bailey Nature Reserve, South Africa: Microsatellite data on a putative past hybridization with blue wildebeest (*C. taurinus*). *Mammalian Biology,* 70(1): 35-45.

Grobler, J. P., Kotze, A., Swart, H. & Hallerman, E. M. 2005. The application of microsatellite DNA markers for forensic analysis of koi carp (*Cyprinus carpio*). *South African Journal of Science,* 101: 19-21.

Heath, M. E. & Coulson, I. M. 1997. Home range size and distribution in a wild population of Cape pangolins, *Manis temminckii*, in north-west Zimbabwe. *African Journal of Ecology,* 35: 94-109.

Heath, M. E. 1992. *Manis temminckii. Mammalian Species,* 415: 1-5.

Hishin, J. 2011. Singita Pamushana. *Malilangwe Wildlife Reserve Wildlife Journal,* pp 1-2.

Houde, P. & Braun, M. J. 1988. Museum collections as a source of DNA for studies of avian phylogeny. *Auk,* 105: 773-776.

Hsieh, H. M., Lee, J. C., Wud, J. H., Chen, C. A., Chen, Y. J., Wang, G. B., Chin, S. C., Wang, L. C., Linacre, A. & Tsai. L. C. 2011. Establishing the pangolin mitochondrial D-loop sequences from the confiscated scales. *Forensic Science International: Genetics,* 5: 303-307.

Ishimoto, Y. 1983. Comparative anatomical studies on the cerebellar nuclei of the pangolis. *Journal für Hirnforschung*, 24(5): 575-589.

Jacobsen, N. H. G., Newberry, R. E., de Wet, M. J., Viljoen, P. C. & Pietersen, E. 1991. A contribution of the ecology of the Steppe pangolin, *Manis temminckii,* in the Transvaal. *International Journal of Mammalian Biology,* 56: 94-100.

Jarne, P. & Lagoda, P. J. L. 1996. Microsatellites, from molecules to populations and back. *Tree,* 11(10): 424-429.

Johnsen, A., Lifjeld, J. T., Rohde, P. A., Primmer, C. R. & Ellegren, H. 1998. Sexual conflict over fertilizations: female bluethroats escape male paternity guards. *Behavioural Ecology and Sociobiology,* 43: 401-408.

Keith, M., Chimimba, C. T., Reyers, B. & van Jaarsveld, A. S. 2007. A comparative analysis of components incorporated in conservation priority assessments: a case study based on South African species of terrestrial mammals. *African Zoology,* 42(1): 97-111.

Kyle, R. 2000. Some notes on the occurrence and conservation status of *Manis temminckii*, the pangolin, in Maputaland, Kwazulu-Natal. *Koedoe,* 43(1): 1-3.

Laurence, W. F., Croes, B. M., Guissouegou, N., Buij, R., Detier, M. & Alonso, A. 2007. Impacts of roads, hunting and habitat alteration on nocturnal mammals in African rainforests. *Conservation Biology,* 22(3): 721-732.

Lim, N. T. L. & Ng, P. K. L. 2008. Home range, activity cycle and natal den usage of a female Sunda pangolin, *Manis javanica* (Mammalia: Pholidota). *Endangered Species Research,* 4: 233-240.

Luo, S-J., Cai, Q-X., David, V. A., Zhang, L., Martelli, P., Lim, N. T-L., Ferrand, N., Chin, S-C., Gaubert, P., Ramos, M. J., O'Brien, S. J., Antunes, A. & Johnson, W. E. 2007. Isolation and characterization of microsatellite markers in pangolins (Mammalia, Pholidota, Manis spp.). *Molecular Ecology Notes,* 7: 269-272.

Markert, J. A., D. M. Champlin, R. Gutjahr-Gobell, Grear, J. S., Kuhn, A., McGreevy T. J., Roth, A., Bagley, M. J. & Nacci, D. E. 2010. Population genetic diversity and fitness in multiple environments. *BMC Evolutionary Biology,* 10: 205-213.

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

Newton, P., Thai, N. V., Robertson, S. & Bell, D. 2008. Pangolins in peril: using local hunters' knowledge to conserve elusive species in Vietnam. *Endangered Species Research,* 67: 41-53.

Nisa, C., Kitamura, N., Sasaki, M., Agungpriyono, S., Choliq, C., Budipitojo, T., Yamada, J., and Sigit, K. 2005. Immunohistochemical study on the distribution and relative frequency of endocrine cells in the stomach of the Malayan Pangolin, *Manis javanica*. *Anatomy Histology Embryology*, 34(6): 373-378.

Oliver, M. K. & Piertney, S. B. 2012. Selection maintains MHC diversity through a natural population bottleneck. *Molecular Biology and Evolution,* 29(7): 1713-1720.

Paabo, S. 1989. Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the United States of America,* 86: 1939-1943.

Paijmans, J. L., Gilbert M.T. & Hofreiter M. 2012. Mitogenomic analyses from ancient DNA. *Molecular Phylogenetics and Evolution*. In Press.

Palo, J. U., Mäkinen, H. S., Helle, E., Stenman, O. & Väinölä, R. 2001. Microsatellite variation in ringed seals (*Phoca hispida*): genetic structure and history of the Baltic Sea population. *Heredity,* 86: 609-617.

Payne, R. B. & Sorenson, D. 2002. Museum collections as sources of genetic data. *Bonner zoologische Beitrage,* 51(4): 97-104.

Pietersen, D. W., McKechnie, A. E. & Jansen, R. (in prep). Review of threats faced by ground pangolins *Smutsia temminckii* in Southern Africa.

Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics,* 6: 847-856.

Prendini L., Hanner R., DeSalle R. Obtaining, storing and archiving specimens for molecular genetic research. In: Methods and tools in biosciences and medicine (MTBM) techniques in molecular systematics and evolution - DeSalle R., Giribet G., Wheeler W., eds. (2002) Basel: Birkhauser. Pages 176-248.

Primmer, C.R., Painter, J. N., Koskinen, M. T., Palo, J. U. & Merila, J. 2005. Factors affecting avian cross-species microsatellite amplification. *Journal of Avian Biology,* 36: 348-360.

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

Reiss, K. Z. 2001. Using phylogenies to study convergence: The case of the ant-eating mammals. *American Zoology,* 41: 507-525.

Richer, R. A., Coulson, I. M. & Heath, M. E. 1997. Foraging behaviour and ecology of the Cape pangolin (*Manis temminckii*) in north-western Zimbabwe. *African Journal of Ecology,* 35(4): 361-369.

Rohland, N., Siedel, H. & Hofreiter, M. 2004. Non-destructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *BioTechniques,* 36: 814-821.

Sanches, A., Tokumoto, P. M., Peres, W. A. M., Nunes, F. L., Gotardi, M. S. T., Carvalho, C. S., Pelizzon, C., Godoi, T. G. & Galetti, M. 2012. Illegal hunting cases detected with molecular forensics in Brazil. *Investigative Genetics,* 3: 17-22.

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

Soewu, D. A. 2008. Wild animals in ethnozoological practices among the Yorubas of south western Nigeria and the implications for biodiversity conservation. *African Journal of Agricultural Research,* 3(6): 421-427.

Su, B., Wang, Y., Lan, H., Wang, W. & Zhang, Y. 1999. Phylogenetic study of complete cytochrome b genes in musk deer (genus *Moschus*) using museum samples. *Molecular Phylogenetics and Evolution,* 12(3): 241-249.

Swart, J. M., Richardson, P. R. K. & Ferguson, J. W. H. 1999. Ecological factors affecting the feeding behaviour of pangolins (*Manis temminckii*). *Journal of Zoology,* 247: 281-292.

Taberlet, P., Waits, L. P. & Luikart, G. 1999. Non-invasive genetic sampling: look before you leap. *Trends in Ecology and Evolution,* 14(8): 323-327.

Takezaki, N., & M. Nei. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics,* 144: 389-399.

Thermo Fisher Scientific – NanoDrop products. 1975. *T009-Technical Bulletin NanoDrop 1000 & 8000*, pp 1-2.

Tobe, S. & Linacre, A. 2011. A new assay for identifying endangered species in Traditional East Asian Medicine. *Forensic Science International: Genetics Supplement Series,* 3: e232–233.

Tracy, L. N., Wallis, G. P., Efford, M. G. & Jamieson, I. G. 2011. Preserving genetic diversity in threatened species reintroductions: how many individuals should be released? *Animal Conservation,* 14: 439-446.

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

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

Voltzit, O. V. & Keirans, J. E. 2003. A review of African *Amblyomma* species (Acari, Ixodida, Ixodidae). *Acarina,* 11(2): 135-214.

Walsh, P. S., Metzger, D. A. & Higuchi, R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques,* 10: 506-513.

Wattier, R., Engel, C. R., Saumitou-Laprade, P. & Valero, M. 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Molecular Ecology,* 7: 1569-1573.

Weber, R. E., Heath, M. E., & White, F. N. 1986. Oxygen binding functions of blood and hemoglobin from the Chinese pangolin, *Manis pentadactyla*: possible implications of burrowing and low body temperature. *Respiratory Physiology & Neurobiology,* 64: 103-112.

Westring, C. G., Kristinsson, R., Gilbert, D. M. & Danielson, P. B. 2007. Validation of reduced-scale reactions for the Quantifiler ™ human DNA kit. *Journal of Forensic Science,* 52 (5): 1035-1043.

Wilson, A. E. 1994. Husbandry of pangolins. *International Zoo Yearbook,* 33: 248-251.

Wu, S., Ma, G., Chen, H., Xu, Z., Li, Y. & Liu, N. 2004. A preliminary study on burrow ecology of *Manis pentadactyla*. *Ying Yong Sheng Tai Xue Bao,* 15: 401-407.

Yamada, J., & Sigit, K. 2005. Immunohistochemical study on the distribution and relative frequency of endocrine cells in the stomach of the Malayan Pangolin, *Manis javanica*. *Anatomy Histology Embryology,* 34: 373-378.

Yang, C. W., Chen, S., Chang, C. Y., Lin, M. F., Block, E., Lorentsen, R., Chin, J. S. & Dierenfeld, E. S. 2007. History and dietary husbandry of pangolins in captivity. *Zoo Biology,* 26: 223-230.

Yang, F., Graphodatsky, A. S., Li, T., Fu, B., Dobigny, G., Wang, J., Perelman, P. L., Serdukova, N. A., Su, W., O'Brien, P. C. M., Wang, Y., Ferguson-Smith, M. A., Volobouev, V. & Nie, W. 2006. Comparative genome maps of the pangolin, hedgehog, sloth, anteater and human revealed by cross-species chromosome painting: further insight into the ancestral karyotype and genome evolution of eutherian mammals. *Chromosome Research,* 14: 283-296.

Yang, H., Golenberg, E. M. & Shoshani, J. 1997. Proboscidean DNA from museum and fossil specimens: An assessment of ancient DNA extraction and amplification techniques. *Biochemical Genetics,* 35: 165-180.

Yeh, F. C. & Yang, R. 1999. Popgene Version 1.31. Microsoft window based freeware for population genetic analysis. University of Alberta and Tim Boyle, Centre for International Forestry Research, Canada.

Zhan, X., Zheng, X., Bruford, M. W., Wei, F. & Tao, Y. 2010. A new method for quantifying genotyping errors for non-invasive genetic studies. *Conservation Genetics,* 11: 1567-1571.

Zhang, Y. P., & Shi, L. M. 1991. Genetic diversity in the Chinese pangolin (*Manis pentadactyla*): inferred from restriction enzyme analysis of mitochondrial DNAs. *Biochemical Genetics,* 29: 501-508.

Appendix 1: Raw data: DNA profiles of samples

Population	MJA22		MJA09		MJA08		MJA07		MJA16		MJA13	
Central	133	133	160	160	-	-	174	174	152	152	-	-
Central	131	131	-	-	-	-	174	174	-	-	-	-
Central	131	141	158	166	190	190	158	164	162	164	226	256
Central	131	141	160	166	-	-	-	-	152	152	250	254
Central	131	131	168	168	138	138	-	-	-	-	-	-
Central	131	131	158	160	-	-	174	174	-	-	-	-
Central	131	131	160	160	186	188	-	-	-	-	-	-
East	131	131	160	160	186	190	164	182	162	162	244	250
East	131	135	158	158	188	188	158	164	164	164	226	226
East	131	135	166	166	188	188	164	164	136	164	242	246
East	131	143	164	164	190	190	158	164	162	162	226	240
East	131	135	164	164	190	190	164	182	162	162	226	256
East	131	135	158	164	190	190	164	182	162	162	256	256
East	131	131	160	160	188	190	-	-	162	164	-	-
East	131	139	160	160	190	190	174	174	164	164	256	256
East	131	135	158	164	190	190	-	-	162	164	-	-
East	131	131	160	168	188	190	174	174	162	162	-	-
East	131	131	158	166	190	190	164	182	162	164	226	256
East	135	141	160	166	188	188	158	164	164	164	226	256
East	131	135	158	164	190	190	164	164	162	162	226	256
East	133	133	160	166	186	188	-	-	-	-	224	254
East	131	131	160	166	122	156	-	-	-	-	244	254
East	129	129	160	168	138	138	158	164	148	150	250	254
East	129	129	168	168	190	190	174	174	150	150	224	224
East	127	129	160	160	190	190	-	-	152	162	226	256
East	127	131	164	166	188	188	164	182	144	144	226	256
East	127	129	160	168	142	142	164	182	-	-	-	-
East	131	135	160	160	188	190	164	182	152	164	226	242
East	131	131	160	160	186	188	164	182	138	162	226	256
East	131	131	160	160	188	188	-	-	152	162	226	226
East	131	131	160	160	186	190	164	164	162	162	256	256
Museum	131	131	140	168	-	-	-	-	-	-	-	-

Museum	-	-	-	-	-	-	174	174	-	-	-	-
Outgroup (G)	134	144	140	160	154	156	-	-	-	-	-	-
Outgroup (G)	129	129	174	174	154	156	-	-	152	152	228	240
Outgroup (G)	129	141	146	146	154	156	174	174	152	156	228	240
Outgroup (N)	127	129	-	-	186	190	-	-	-	-	256	256
Outgroup (Z)	131	131	160	160	122	122	-	-	-	-	-	-
West	131	142	168	168	-	-	-	-	-	-	-	-
West	131	131	160	160	-	-	-	-	-	-	-	-
West	131	131	160	160	138	138	-	-	-	-	-	-
West	131	131	168	168	186	188	-	-	152	152	256	256
West	131	131	160	160	186	186	-	-	138	138	256	256
West	131	131	150	168	172	172	-	-	152	152	224	234
West	127	131	-	-	186	186	-	-	-	-	-	-
West	131	131	174	174	-	-	-	-	-	-	-	-
West	-	-	168	168	188	188	-	-	154	154	224	232
West	131	131	132	132	138	138	-	-	148	148	-	-
West	131	139	160	160	188	190	158	164	152	152	250	254
West	131	135	150	160	172	172	174	174	136	136	232	232
West	131	137	140	160	138	138	174	174	152	152	250	254
West	131	131	160	160	132	132	174	174	-	-	-	-
West	131	131	160	174	138	142	-	-	136	136	232	232
West	131	131	160	160	-	-	-	-	-	-	224	224
West	131	137	140	144	188	188	174	174	-	-	236	236
West	131	131	160	160	138	138	-	-	152	152	226	232
West	131	135	160	160	138	138	-	-	148	152	232	232
West	131	131	160	160	188	190	164	182	138	138	254	254
West	133	137	158	168	186	186	-	-	152	152	-	-
West	139	139	158	166	188	190	-	-	-	-	256	256
West	135	135	160	166	190	190	-	-	126	126	240	240
West	129	129	168	168	116	150	-	-	152	152	226	226
West	-	-	168	168	122	122	-	-	152	152	-	-
West	131	131	-	-	188	190	-	-	152	152	-	-
West	131	131	160	160	186	188	164	182	162	164	226	256
West	131	131	164	164	186	188	164	182	162	162	240	256
West	131	131	164	164	186	190	164	164	162	164	-	-

West	143	145	158	168	116	116	-	-	154	154	-	-
West	131	135	150	168	186	186	-	-	154	154	226	226
West	131	131	-	-	184	186	164	164	154	162	-	-
West	133	133	160	160	186	190	-	-	-	-	254	254
West	-	-	160	160	-	-	164	164	-	-	-	-
West	133	133	160	166	-	-	174	174	-	-	-	-
West	133	133	160	166	186	190	158	164	150	150	250	254
West	131	131	168	168	186	186	-	-	164	164	256	256
West	127	129	164	168	138	138	164	164	148	152	226	226
West	127	129	168	168	186	190	174	174	-	-	-	-
West	131	131	160	160	186	190	164	164	-	-	226	256
West	131	135	168	168	116	116	-	-	-	-	-	-
West (Faecal)	-	-	-	-	-	-	-	-	-	-	-	-

***** G – samples collected in Ghana, N – sample from Namibia, Z – sample from Zimbabwe.

Appendix 2: Raw data: NanoDrop Measurements

Lab number	A260/280 value	Conc/Ng	Lab number	A260/280 value	Conc/Ng
PNG1	1.90	197.4	PNG43	1.04	11.1
PNG2	3.81	7.4	PNG44	1.78	27.4
PNG3	0.35	30.7	PNG45	5.67	7.7
PNG4	1.82	4.8	PNG47	4.91	6.5
PNG5	0.33	35.7	PNG48	3.42	10.9
PNG6	0.7	18.7	PNG49	2.61	20.9
PNG7	1.90	52.9	PNG50	1.21	34.8
PNG8	-0.40	0.6	PNG51	1.29	15.6
PNG9	2.10	5.8	PNG52	1.57	8.7
PNG10	1.17	20	PNG53	1.54	1.9
PNG11	1.31	45.8	PNG55	1.37	11.5
PNG13	2.07	5.9	PNG56	1.24	24.5
PNG14	2.13	28.0	PNG58	-0.21	-0.3
PNG15	1.99	4.5	PNG59	0.80	1.6
PNG16	1.99	3.8	PNG60	0.57	3.2
PNG17	0.99	1.4	PNG61	0.71	3.5
PNG19	1.56	7.8	PNG62	0.57	4.3
PNG20	1.77	1.8	PNG63	1.16	19.8
PNG21	1.71	20.6	PNG64	1.65	21.2
PNG22	1.60	9.6	PNG65	2.09	55.9
PNG23	1.47	9.4	PNG66	0.46	10.8
PNG24	1.44	2.6	PNG 67	1.94	245.9
PNG25	1.53	8.7	PNG 69	1.61	7.4
PNG26	1.68	4.9	PNG 70	1.07	12.7
PNG27	1.40	5.2	PNG 71	2.07	2.2
PNG28	1.38	4.6	PNG 72	1.57	2.9
PNG29	1.34	2.8	PNG 73	1.93	276
PNG30	1.21	1.8	PNG 74	1.83	114.7
PNG31	1.70	10.6	PNG 75	2.22	6.4
PNG32	2.05	3.9	PNG 76	1.09	5.9
PNG33	1.01	40	PNG 77	2.09	5.6

PNG34	1.16	9	PNG 78	1.9	132.5
PNG35	2.16	42.9	PNG 79	1.94	224
PNG36	1.13	7.7	PNG 80	1.1	1.8
PNG37	1.30	8.6			
PNG38	1.30	3.6			
PNG39	1.12	2.4			
PNG40	0.99	3.4			
PNG41	4.43	5.9			
PNG42	0.92	26.2			