Developmental validation of a forensic profiling system for black rhinoceros (*Diceros bicornis*) and extraction of nuclear DNA from historic rhinoceros horns

Thesis submitted in accordance with the requirements of the University of Chester for the degree of Master of Philosophy

by

Kara Dicks

November 2014

Abstract

Demand for rhinoceros horn in the late 20th century led to the population decline of black rhinoceros (*Diceros bicornis*) by more than 96 %, and the species remains critically endangered. Trade in rhinoceros and their products is prohibited under CITES legislation. Despite this, there has been a recent increase in rhinoceros horn demand for use in traditional Asian medicine (TAM) to treat a range of illnesses including cancer, childhood fevers and as a hangover cure. Consequently, the value of rhinoceros horn in illegal trade has risen dramatically, alongside a dramatic increase in poaching. Large numbers of horns exist in historic collections, and since 2011 there has been numerous thefts in both rhinoceros range states and throughout Europe. Wildlife DNA forensic analysis can be used to individually identify rhinoceros horn, matching a seized horn to a crime scene. This can provide valuable evidence in criminal proceedings. This study therefore aimed to validate a profiling system that could be used for *D. bicornis* horns and to determine an optimal DNA extraction method for historic horns.

A developmental validation of a marker panel consisting of 15 short tandem repeat (STR) and one sexing locus was carried out to determine its reliability and efficacy in singleplex reactions. The markers were not found to deviate from patterns of Mendelian inheritance, and species specificity studies found that only two loci (WR7C and ZF1) amplified human and dog DNA. Alleles could be scored at DNA concentrations of less than 1 ng/µL for all loci except WR7C. The marker panel was found to be highly reproducible with a maximum d-value of 0.545, mean heterozygous balance of 1.30 and stutter ratio of 0.35. The marker panel was also robust to reductions in annealing temperature (T_A) and increasing cycle number (up to 43 cycles), although increasing T_A reduced amplification success. The markers were then combined into two multiplexes and allele frequency data was generated for 52 *D. bicornis michaeli*, and two markers (BlRh37D and IR12) were rejected for violating assumptions of Hardy-Weinberg equilibrium and linkage equilibrium. The average probability of identity (PI_{ave}) ranged between 1.06×10^{-11} and 1.16×10^{-7} with varying levels of relatedness and population structure. The marker panel was determined suitable for use in forensic casework analysis under the defined conditions, and limitations were described.

Previous studies have shown that nuclear DNA (nDNA) can be extracted from modern horns, but how effective those methods would be on historic horns was unknown. Optimisation of a number of aspects of the extraction method was carried out, including sample preparation, chemical breakdown with dithiothreitol (DTT), mass of starting material and extraction kit.

Mechanical breakdown of rhinoceros horn material using a mixer mill was determined unnecessary, whilst the addition of DTT was found to improve digestion. Irrespective of extraction kit, 20 mg of horn material yielded sufficient quality and quantity nDNA for forensic profiling. Qiagen extraction kits using the QIAamp Mini columns were found to yield superior DNA quantity compared with QIAamp MinElute columns. Whilst KingFisher Cell and Tissue DNA Kits yielded equivalent extract quantity, mitochondrial DNA contamination between samples was detected. The Qiagen method using QIAamp MinElute columns was therefore considered the most robust extraction method for historic rhinoceros horns.

This work describes the conditions and limitations of a forensic profiling system for *D. bicornis* and an effective extraction methodology for historic rhinoceros horns. This enables this marker panel to be used with confidence in forensic casework analysis of both modern and historic horns.

Declaration

I declare that the material being presented for examination is my own work and has not been submitted for an award of this or any other Higher Education Institution.

…… ……..

Signature Date

Acknowledgements

Firstly, I am extremely grateful to Lucy Webster. Your guidance and teaching improved this project immensely, and your support was invaluable throughout. Thank you also to Ian McDowall for helping to keep this project on track and all your supervision.

Thanks to Science and Advice for Scottish Agriculture (SASA) for allowing me to initially work in the Department of Molecular Biology laboratories for two weeks and stay for a year and a half. Especially thanks to David Kenyan for facilitating this work, and also to Alex Reid and Vince Mulholland for answering my infinite list of questions. Many thanks also to Wulansari McEwing for all your help, for organising me and for all your optimism.

Kim Davie, thank you for all your support throughout. Two weeks in Edinburgh turned into two and a half years, so thank you for putting a roof over my head, laughing at all the crazy times throughout and so much more.

Many thanks to Mike Bruford (Cardiff University), Yoshan Moodley (University of Veterinary Medicine Vienna), Shadrack Muya (Jomo Kenyatta University of Agriculture and Technology) and Jane Hooper (Port Lymne Wild Animal Park) for providing access to samples to carry out the validation. Thanks especially to Isa-Rita Russo (Cardiff University) for all your guidance and help.

Lastly, thanks to Chester University and Reaseheath College for funding me throughout this work. Paul O'Donoghue developed this studentship and provided research funds.

Table of contents

Glossary of Abbreviations

bp	Base pair
CITES	Convention on International Trade in Endangered Species of Wild Fauna and Flora
DRC	Democratic Republic of Congo
DTT	Dithiothreitol
H_E	Expected heterozygosity
H_O	Observed heterozygosity
HWE	Hardy-Weinberg equilibrium
IUCN	International Union for Conservation of Nature
N	Sample number
Na	Number of alleles
NTC	No template control
PI_{ave}	probability of identity
PIC	Polymorphism information content
QA	Quality assurance
QC	Quality control
RFU	Relative fluorescence units
RMP	Random match probability
SE	Standard error
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
TA	Annealing temperature
TAM	Traditional Asian Medicine
VNTR	Variable number tandem repeat
ZF	Zinc finger protein
ZFX	Zinc finger protein X chromosome
ZFY	Zinc finger protein Y chromosome

List of Tables

List of Figures

Chapter 1 – General introduction

Humans have greatly impacted animal populations for thousands of years both indirectly, for example by altering the landscape (e.g. megafaunal extinction in Pleistocene Australia - Miller *et al.* 2005) or introducing invasive species (e.g. the decline of avifauna following mammalian introductions to oceanic islands - Blackburn *et al.* 2004), and directly, for example by overhunting (Diamond 1989). The world is currently facing a rapid loss of biodiversity (Baillie *et al.* 2004). Whilst climate change is now increasingly recognized as a major driver of recent population declines and species extinctions (Thomas *et al.* 2004), the 'evil quartet' of drivers (habitat loss, introduced species, extinction cascades and overexploitation) (Diamond & Diamond 1989) remain important causes of biodiversity loss.

Overexploitation has been cited as a primary cause of the decline and extinction of numerous species (Burney & Flannery 2005) historically (Steadman 1995; Alroy 2001; Jackson 2001; Slikas 2003; Lyons, Smith & Brown 2004; Burney & Flannery 2005) and contemporarily (Corlett 2007; Linder 2008; Jacquet 2009; Grogan *et al.* 2010; Rovero *et al.* 2012). Species are overexploited for a variety of purposes including subsistence hunting, commercial food, tourist curios, sport, pets, clothing, artistic pieces, and medicines (Broad *et al.* 2003; Warchol 2004; Wilson-Wilde 2009). The illegal international wildlife trade is worth in excess of US$6 billion annually (Warchol 2004) and, after drugs and weapons, is probably the third greatest illegal trade in the world (McGraw *et al.* 2012). Recognizing the devastating impact of overexploitation and its subsequent trade can have on threatened species, CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) was established. This international agreement between governments regulates international trade in wildlife specimens with the aim to ensure that trade is either prohibited or can be undertaken sustainably in order to safeguard wildlife species (Convention on Biological Diversity 1992).

1.1 Overexploitation of rhinoceros prior to the 21st century

Rhinoceros species have experienced massive overexploitation and are listed under CITES in an attempt to control and reduce trade in rhinoceros and their parts. The Rhinocerotidae are a single monophyletic Family within the Order Perissodactyla (Cerdeno 1995; Tougard *et al.* 2001) and at present, there are three Asian species of rhinoceros – Indian (*Rhinoceros unicornis*, Linnaeus, 1758), Javan (*Rhinoceros sondaicus*, Desmarest, 1822) and

Sumatran (*Dicerorhinus sumatrensis*, Fischer, 1814) – and two African species – white (*Ceratotherium simum*, Burchell, 1817) and black (*Diceros bicornis*, Linnaeus, 1758).

The Rhinocerotoidea probably first appeared in the Eocene (Prothero *et al.* 1989) and once included over 50 genera and hundreds of species ranging across North America, Eurasia and Africa (Kingdon & Hoffmann 2013). At least until the Holocene when humans began to impact other species, Rhinocerotoidea species evolved and went extinct according to natural processes. The role of humans versus climatic changes in the extinction of the woolly rhinoceros (*Coelodonta antiquitatis*) is much debated, though overhunting of the last remaining remnant populations may have tipped the species into extinction (Kuzmin 2010; Stuart & Lister 2012; Markova *et al.* 2013).

The conservation statuses of all five extant species of Rhinocerotoidea are listed as threatened by the International Union for Conservation of Nature (IUCN) (Table 1.1), and the primary threats to all species at present are overexploitation and habitat loss (Emslie *et al.* 2013). The Asian species are particularly restricted in their distribution (Fig. 1.1), and indeed the Javan rhinoceros now persists as a single population in the Ujung Kulon Peninsula (van Strien *et al.* 2013), although the African species have a wider distribution across southern and eastern Africa.

Overexploitation, both historical and contemporary, has nearly driven all extant species of rhinoceros to extinction. Throughout the 20th century, five subspecies of rhinoceros went extinct (Table 1.2). All five extinctions were likely due to overhunting.

The southern subspecies of white rhinoceros, *C. simum simum* was hunted as both sport and vermin, particularly by colonialists in Southern Africa, and the subspecies was believed to be extinct by the late 1800s (Emslie & Brooks 1999; Rookmaaker 2000). A small, remnant population was rediscovered in the Umfolozi region of Zululand, South Africa (now KwaZulu Natal) numbering between 20 and 150 animals, and possibly very small populations had persisted in other regions of southern Africa (Rookmaaker 2000). Through intensive conservation efforts, the southern white rhinoceros is now the most numerous of the extant species, numbering over 20 429 individuals (Table 1.2); although 90 % of the population is in South Africa (Emslie *et al.* 2013).

Black rhinoceros were abundant across Africa in the 19th century, but sport hunting greatly reduced population sizes by the early 20th century (Emslie & Brooks 1999). Hunting pressures changed from sport to poaching for horns, and by the 1960s, the horn trade began to place heavy pressure on black rhinoceros populations. Between 1960 and 1995, the black

rhinoceros declined by over 95 % (Emslie & Brooks 1999; Milliken & Shaw 2012) to a total population size of 4 880 individuals composed of four subspecies (Emslie *et al.* 2013).

The Sumatran and Javan rhinoceros are currently at the brink of extinction (van Strien *et al.* 2008, 2013). Population numbers are critically low and decreasing for the Javan (< 50 individuals) and Sumatran rhinos (< 200 individuals) (Table 1.1) (Emslie *et al.* 2013). Both species have been hunted historically (Foose & van Strien 1997). The Indian rhinoceros has also been heavily hunted historically (Wiliam Andrew Laurie 1978), particularly for sport (Foose & van Strien 1997); although hunting was rare in its Nepalese range until the 1950s (Thapa *et al.* 2013). Habitat loss has also been a major driver of population decline in all three Asian rhinoceros species (Talukdar *et al.* 2008; van Strien *et al.* 2008, 2013).

1.2 The rhinoceros horn trade

1.2.1 Use of rhinoceros horn

Rhinoceros have been hunted for centuries (Amin *et al.* 2006), particularly the African species, to extirpate them from land claimed for human purposes or as for sport (Cumming *et al.* 1987; Emslie & Brooks 1999). They were hunted for their horns in the 1800s and early 1900s as both hunting trophies, typically by European travellers, and to export their horns into trade. Between 1840 and 1900, up to 11 000 kg of rhinoceros horn was exported from East Africa each year, which would equate to approximately 170 000 rhinoceros (primarily black) hunted in 60 years (Martin & Martin 1982).

1.2.2 Yemeni jambiya handles

Jambiyas (daggers) are part of the Yemeni men's traditional dress, and the most prestigious have rhinoceros horn handles (Anonymous 1983; Leader-Williams 1992; Martin *et al.* 1997). A market was created for rhinoceros horn to produce such handles after carrying a jambiya became legal following South Yemen's independence from Britain in 1967 and the economy improved (Martin 1979). A ban on the import of rhinoceros horn to Yemen, a downward trend in the economy and an Islamic edict against killing for reasons other than food or protection against predators eventually reduced the demand for rhinoceros horn jambiyas by the early 1990s (Martin *et al.* 1997). Recent assessments of the Yemeni jambiya culture suggest that most handles are now made from water buffalo horn or other alternatives (Vigne & Martin 2008). During the 20th Century, however, Yemen was probably the greatest end-use market for rhinoceros horn (Anonymous 1983; Leader-Williams 1992; Martin *et al.* 1997).

Table 1.1. Details of rhinoceros species status, population trends and population estimates.

Species	Common name	IUCN status	Population trend	Most recent total population estimate
Diceros bicornis	Black rhino	Critically endangered	Increasing	5 081 (Dec. 2012 - Emslie & Knight, 2013)
Ceratotherium simum	White rhino	Near threatened	Increasing	20 429 (Dec. 2012 - Emslie & Knight, 2013)
Rhinoceros unicornis	Indian rhino	Vulnerable	Increasing	3 264 (Jun. 2012 - Emslie, Milliken, & Talukdar, 2013)
Rhinoceros sondaicus	Javan rhino	Critically endangered	Unknown	35 - 45 (Jun. 2012 - Emslie, Milliken, & Talukdar, 2013)
Dicerorhinus sumatrensis	Sumatran rhino	Critically endangered	Decreasing	140 - 210 (Jun. 2012 - Emslie, Milliken, & Talukdar, 2013)

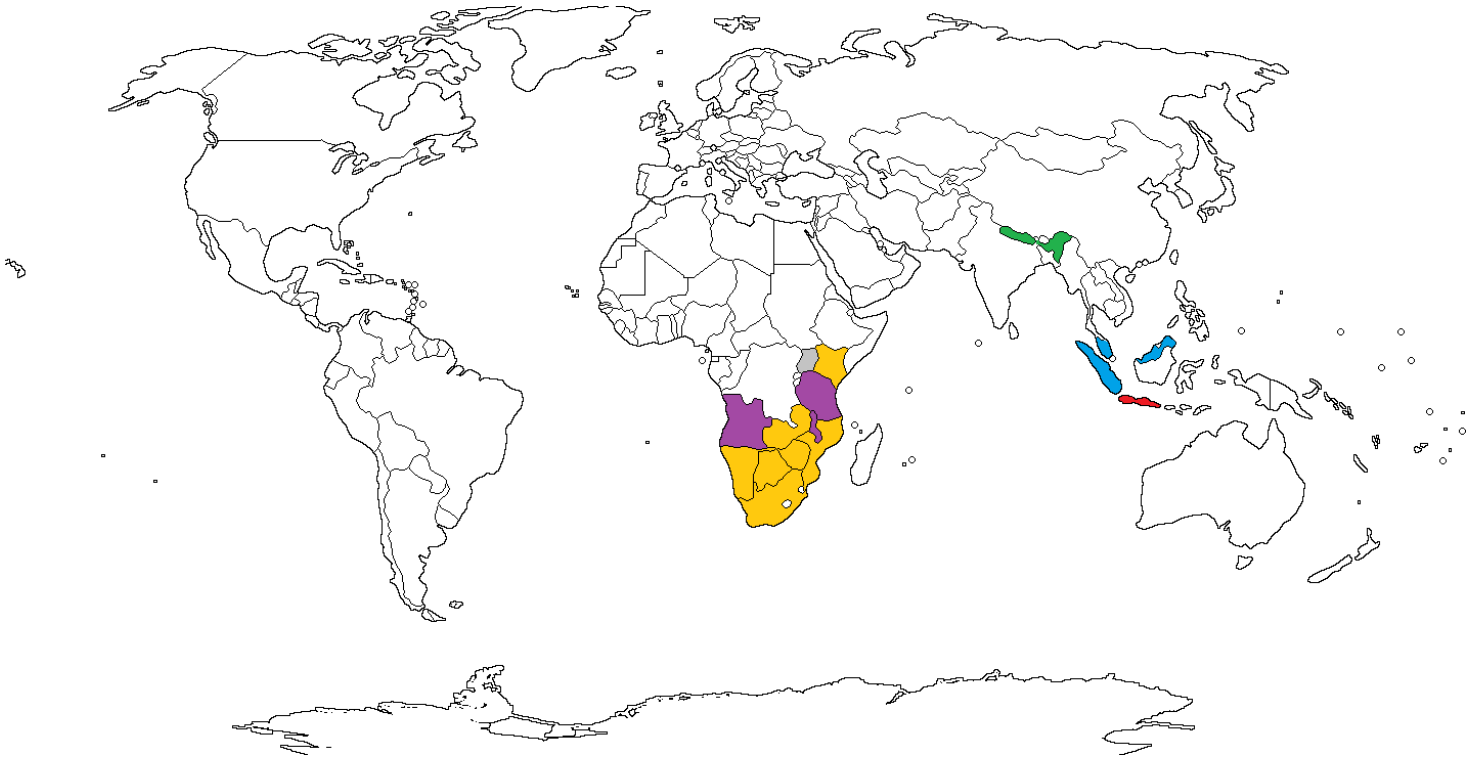


Figure 1.1. Distribution of all five rhinoceros species (*Rhinoceros unicornis* green, *Rhinoceros sondaicus* red, *Dicerorhinus sumatrensis* blue, *Ceratotherium simum* only grey, *Diceros bicornis* only purple, *Ceratotherium simum* and *Diceros bicornis* yellow). Range is shown to political country scale, or island where appropriate.

Table 1.2. Details of extinctions of rhinoceros subspecies

Subspecies	Historic distribution	Approx. date of extinction	Details of extinction	References
Rhinoceros sondaicus inermis	Sunderbans region of India and Bangladesh	Early 20th century	Probably poached	(Vanleeuwe *et al.* 1997)
Rhinoceros sondaicus animaticus	Viet Nam	Declared by IUCN in 2011	Thought extinct by 1970 then rediscovered in 1989. Faecal DNA evidence shows last remaining individual was shot in 2010	(Brook *et al.* 2011, 2012)
Dicerorhinus sumatrensis lasiotis	India, Bhutan, Bangladesh, Myanmar	Unknown	Unknown but likely poaching and habitat loss	(van Strien *et al.* 2008)
Diceros bicornis longipes	Sub-Saharan Africa	Nearly extinct by 1930 but remained in Cameroon throughout the 20th century. Declared extinct by IUCN in 2012	Heavy poaching greatly reduced population and restricted it to Cameroon by 1930, and poaching likely to have driven subspecies to complete extinction by 2006	(Lagrot *et al.* 2008; Emslie 2012)
Ceratotherium simum cottoni	Central Africa	Extinct in the wild by 2006, declared by IUCN 2012	Declined to 2 250 individuals by 1960. Civil unrest in strongholds of Sudan and Democratic Republic of Congo (DRC) led to overexploitation for meat and horns. Last known stronghold in Garamba National Park, DRC failed to protect the last few individuals and by 2006 all were poached. Four captive bred individuals have been translocated to Kenya in an attempt to interbreed with *C. simum simum* to maintain some of the genetic variation.	(Groves 1972; Hillman Smith *et al.* 1986; Emslie & Brooks 1999; IUCN SSC African Rhino Specialist Group 2008)

1.2.3 Traditional Asian Medicine (TAM)

Rhinoceros horn has been used in TAM for centuries, listed as a medicinal in a 2 000 year old pharmacology text 'The Divine Peasant's Herbal' (Nowell *et al.* 1992)**,** and used in the Tang Dynasty (approximately 7th to 9th Century) as both a medicine and a luxury item (Heller 2011). TAM considers the body as a whole (Ellis 2005) and health is maintained by keeping the body in 'balance' as opposed to 'imbalanced' (diseased) (Van & Tap 2008), differing from western medicine which treats disease and illness in parts of the body. Rhinoceros horn has traditionally been used to a wide range of illnesses, but primarily to reduce "heat" (Graham-Rowe 2011; Milliken & Shaw 2012; Nowell 2012), a condition not directly recognised by Western medicine but often presents with fevers, headaches and other symptoms caused by infections (Martin & Martin 1982; Leader-Williams 1992; Graham-Rowe 2011; Milliken & Shaw 2012; Nowell 2012). Rhinoceros horns from African and Asian rhinoceros are thought differ in effectiveness, with Asian "fire" horns being more potent than African "water" horns (Martin 1979, 1981; Nowell *et al.* 1992; Emslie & Brooks 1999).

The end use market for rhinoceros horn in TAM has changed from primarily China, South Korea and Taiwan in the 20th century (Martin 1981; Emslie & Brooks 1999) to Viet Nam in the 21st century (Milliken & Shaw 2012; Emslie *et al.* 2013). Although rhinoceros horn has been a medicinal in Vietnamese TAM for traditional uses for many years, a booming economy with 65% of the population under 30 has created a young affluent population with disposable income and a desire for rare and luxury products, especially rhinoceros horn (Milliken & Shaw 2012). Three principle user groups have been identified: i) cancer patients (and possibly those with other incurable diseases), ii) the young and affluent who use rhinoceros horn to treat overconsumption of recreational drugs including alcohol, and iii) young affluent mothers who stock rhinoceros horn to treat a range of childhood illnesses (Milliken & Shaw 2012). This new and changing end-use market in Viet Nam differ from historical East Asian rhinoceros horn users and tackling the illegal trade will require a multifaceted approach both within the consumer and source countries.

Very little scientific data has been generated to determine the efficacy of rhinoceros horn as a medicinal. The scientific community has shown little interest in the medicinal value of keratin (Nowell 2012), which is the primary component of rhinoceros horn (Geerinckx *et al.* 2007). A number of studies have attempted to demonstrate the efficacy of rhinoceros horn (for review see Nowell 2012b) but with little concordance amongst results. There has been a lack of well designed, peer-reviewed clinical trials (Milliken & Shaw 2012), but little

agreement can be made between TAM and Western medicine as to how such trials should be carried out due to the differences in the pathology of disease (Nowell 2012).

1.3 Impact of the resurgence in the rhinoceros horn trade

1.3.1 Wild rhinoceros populations

The demographics of both rhinoceros horn sources and users has changed in the recent resurgence of the rhinoceros horn trade. The vast majority of horns entering the illegal market until 1995 originated from poached wild rhinoceros, primarily in scattered and unprotected regions of Africa (Emslie & Brooks 1999). South Africa is now the major range state for African rhinoceros (Emslie *et al.* 2013) and has become the primary source of rhinoceros horn to the illegal markets with poaching incidents continuing to rise (Milliken & Shaw 2012; Emslie *et al.* 2013). Since 2006, a minimum of 3 926 poached black and white rhinoceros have been reported (Figure 1.2) (Emslie & Knight 2014) and between 1999 and 2012, at least 226 Indian rhinoceros were poached (Rhino *et al.* 2009; Emslie *et al.* 2013).

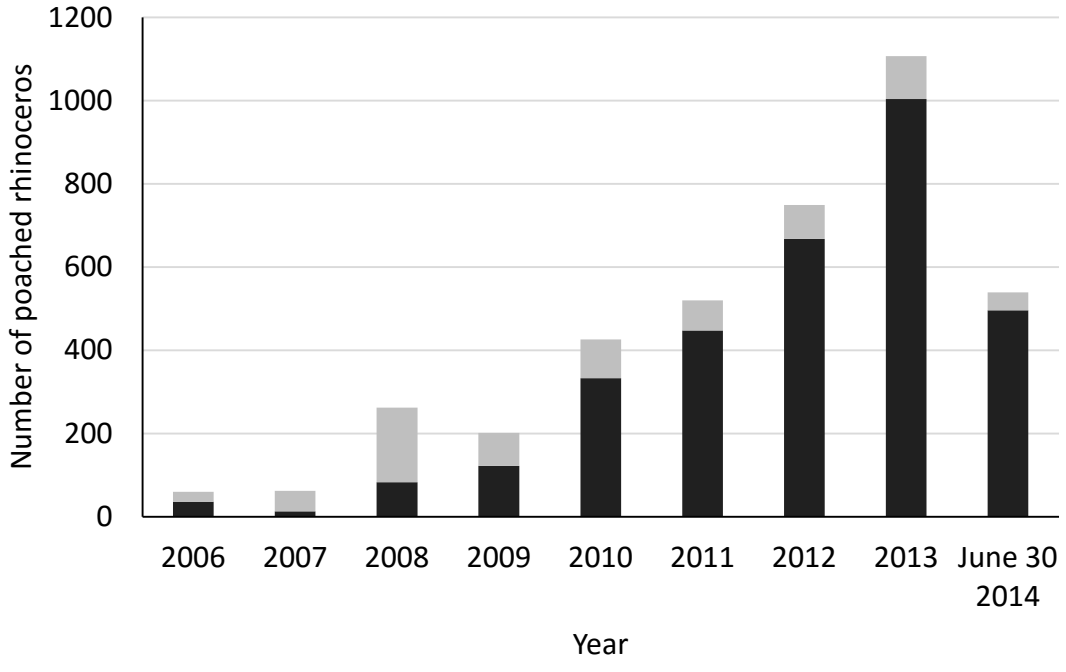


Figure 1.2. Reported numbers of poached black and white rhinoceroses between January 1, 2006 and June 30, 2014 in South Africa (black) and all other African range States (grey). This represents the minimum number of poaching incidences as many may not be detected. Figures from Emslie & Knight (2014).

1.3.2 Historical and zoological collections

Since 2009, increasing numbers of horns have been entering the illegal market from outside of rhinoceros range States, many of which were from private collections around the world (Milliken & Shaw 2012). Unknown numbers of horns are held in museums, zoological exhibitions, art galleries, antiquaries and private family collections. Horns in such collections are historical, scientific, decorative or educational but are rarely valuable in monetary terms. Within South Africa, vast and unknown numbers of horns have been stolen from government stockpiles, private collections and museums (Milliken & Shaw 2012). There has been an unprecedented trend in thefts of horns from European collections. Europol data indicates that 82 horns were stolen across Europe in 2011, and further thefts were reported in the US and Argentina (Emslie *et al.* 2013). Thefts have not only been limited to complete horns, but have also included carvings, cups and a number of imitation horns (Emslie *et al.* 2013).

Rhinoceros in zoological collections around the world could be targeted for their horns. There are over 1 100 captive rhinoceros in zoological collections around the world (Foose & Wiese 2006). In 2007, the horns of a deceased rhinoceros from Colchester Zoo were seized at Manchester Airport where they were being smuggled out of the UK in a fake antique sculpture by an antiques dealer after being taken from the abattoir (Bhattacharya 2010). No captive rhinoceros has yet been poached, but intelligence suggested that rhinoceros at several European wildlife parks were being targeted by poachers, although no attempt has been made following increased security measures (McCarthy 2012; The Aspinall Foundation 2013).

1.3.3 Involvement of organised crime

There is increasing evidence that organised crime groups are heavily involved in the rhinoceros horn trade (Rhino *et al.* 2009; Milliken & Shaw 2012; Montesh 2012; UNODC 2012). Organised international criminal networks are known to be linked to wildlife crime, using drug trafficking routes to smuggle illegal wildlife goods (Alacs *et al.* 2010). Wildlife goods, rhinoceros horn in particular, are often highly profitable, carry low risks of detection and penalties are minimal (Alacs *et al.* 2010), especially compared to other illegal activities carried out by these groups such as drug and weapons smuggling, money laundering and human trafficking (Milliken & Shaw 2012).

Historically, poachers have been locals or skilled trackers and shooters including former military and game scouts (Milliken & Shaw 2012). However, poaching has become

increasingly sophisticated, and guns are often being replaced by 'silent' methods including the use of immobilising drugs, poisons and cross-bows, requiring large teams and the expertise of trained professionals (Thomas 2010; Milliken & Shaw 2012; CITES Secretariat 2013). Furthermore, South African populations of white rhinoceros are listed under CITES Appendix II (Cites 2011), which allows hunting under license as a sport and strictly not for profit. At least five different Vietnamese-run syndicates were known to be involved in pseudo-hunting in 2007, with the horns exported as trophies by Czech and Vietnamese citizens before ending up in the illegal trade in South East Asia (Rhino *et al.* 2009; Milliken & Shaw 2012). South African law was changed in 2012 to prevent this practice (Milliken & Shaw 2012; CITES Secretariat 2013).

Horns have been smuggled through countries which have not previously been linked to the horn trade. Individuals from the Czech Republic are believed to have been involved in pseudo hunts (Milliken & Shaw 2012; CITES Secretariat 2013). Seizures have been made in the USA, Hong Kong SAR, Australia and the Philippines (CITES Secretariat 2013) suggesting that horns are being smuggled out of, or through, these countries.

Rhinoceros horn syndicates are organised at different levels (Figure 1.3) within the illegal trade network in South Africa as identified by the National Wildlife Crime Reaction Unit (Milliken & Shaw 2012). At the lowest levels are the typical poachers. The rhinoceros horn moves up the chain to national buyers and exporters to international buyers and exporters. Enforcement activities are typically targeted at individuals operating at the lowest levels (CITES Secretariat 2013). These are numerous and easily replaceable, therefore their removal from the trade chain has minimal impact. Removal of individuals at the top of the trade chain operating internationally is more likely to disrupt the entire chain and have greater impact (CITES Secretariat 2013). Furthermore, many of these individuals may be involved in organising other criminal activities, thus enforcement may have wide reaching consequences.

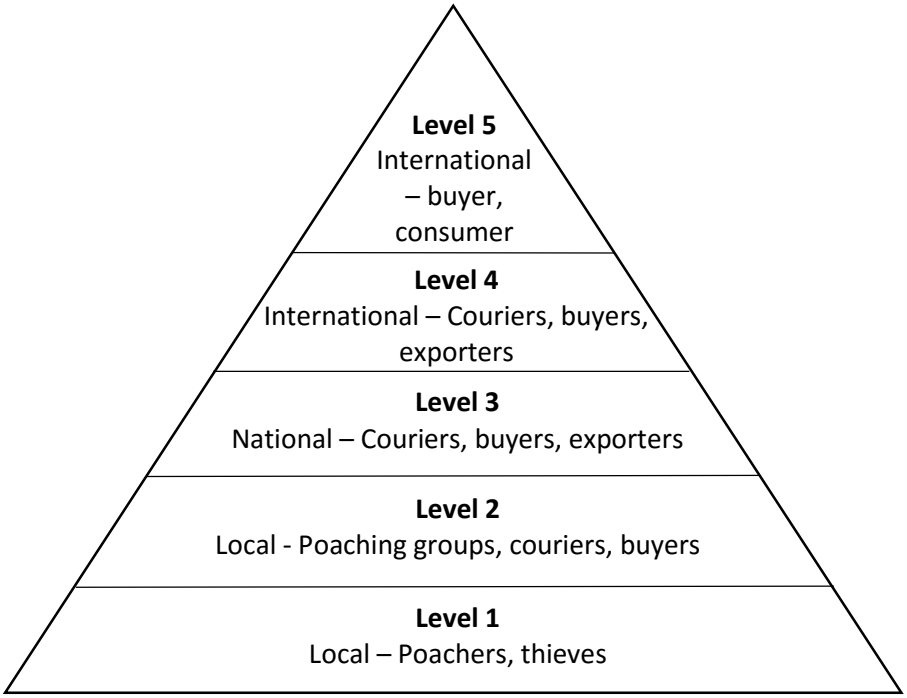


Figure 1.3. The five levels of organised crime, from the lowest (Level 1) to the highest (Level 5), involved in the rhinoceros horn trade in South Africa (Adapted from Milliken & Shaw 2012). This hierarchy is likely to be similar if level one involves a poaching incident or a theft, and in other countries, although the same individuals may encompass levels 1 and 2.

1.4 Improved methods to combat the horn trade

With the increasing sophistication of the methods used by those involved in the rhinoceros horn trade and increasing awareness of the involvement of organised crime, it is apparent that methods used to combat the trade must improve. Enforcement at levels 1 – 3 of the organised crime hierarchy (Fig. 1.3) is relatively effective in range States, but must become more effective at levels 4 and 5 in order to impact the driving forces behind the trade (CITES Secretariat 2013). As such, efforts must be carried out on an international scale.

Law enforcement measures must improve and adapt to the increasingly sophisticated and international trade in order to reduce the illegal trade in rhinoceros horns. This must be a multifaceted approach involving securing live rhinoceros and horns, gathering intelligence and applying effective policy and legislation. Wildlife DNA forensic approaches can be extremely valuable in providing intelligence in wildlife crime cases, including cases of poaching, theft and illegal sales.

1.4.1 Wildlife DNA forensic analysis

The field of human DNA forensic analysis, whilst having a wide range of uses and technical capabilities, is primarily concerned with individual and familial identification of a single species. Non-human DNA forensic tests can, however, be developed for and applied to any species (Butler 2005); although wildlife forensic priorities are typically plants and animals. There are four primary applications of DNA forensic analyses to wildlife crime investigations: species identification, geographical origin, familial identification and individual identification (Ogden *et al.* 2009).

- **Species identification** methods are frequently necessary where morphological characteristics are lacking to determine that a seized product is from a protected species. Processing of wildlife products, trace evidence on a suspect's clothing or possession (Tobe & Linacre 2008; Ogden *et al.* 2009), or decomposed carcasses (McGraw *et al.* 2012) often yield DNA evidence that can be used to identify the species (Ogden *et al.* 2009).
- **The geographic origin** of wildlife products can be carried out using assignment based methods which utilise extensive population genetic data to determine the likely population of reproductive origin or exclude an individual as originating from a specific protected population (Ogden *et al.* 2009; McGraw *et al.* 2012). This may be necessary to determine whether a crime has occurred as CITES and other wildlife legislation typically operates at a national or regional level (Ogden *et al.* 2009), protecting only specific populations of a species, which may not differ morphologically.
- **Familial identification** compares the DNA profiles of individuals to determine their relatedness. This can be used to determine whether an individual is the offspring of its supposed captive-bred parents (Ogden *et al.* 2009). Failure to identify captive-bred familial relations may suggest that the individual was actually illegally wild-caught (Ogden *et al.* 2009; McGraw *et al.* 2012). Captive breeding programmes for highly prized species allow trade in highly threatened species to be maintained without impacting wild populations, but where species still occur in the wild, there is the possibility to take animals and plants from the wild and trade them as captive-bred (Ogden *et al.* 2009).
- **Individual identification** is carried out using short tandem repeat (STR) or single-nucleotide polymorphism (SNP) profiles to determine whether two pieces of DNA evidence originated from the same source (Butler 2005). These techniques can be

used to match trace evidence, such as blood stains on a poacher's knife (e.g. Lorenzini *et al.* 2011), and seized wildlife products back to a crime scene (e.g. White *et al.* 2012).

1.4.1.1 Use of individual identification profiling systems for rhinoceros crimes

DNA profiling uses nuclear DNA markers which are bi-parentally inherited (Butler 2005). Nuclear DNA can be extracted from rhinoceros horn (Peppin *et al.* 2010; Harper *et al.* 2013) as well as other tissues, and can be used to match a horns or other DNA evidence to a carcass in poaching cases. A rhinoceros DNA profiling system has been established in South Africa, the Rhino DNA Index System RhoDIS, which has been useful in providing evidence in a number of South African cases (CITES Secretariat 2013).

Following the spate of thefts of horns from collections and the potential threat to captive rhinoceros, a DNA profiling technique may also be useful in non-rhinoceros range States. Being able to match a horn back to a crime may provide useful evidence to build a case. Horn thefts leave no remaining DNA, unlike poaching cases where the horn is removed and the carcass is left. It is therefore necessary to generate and store the DNA profiles of historic and contemporary horns in collections prior to a theft. This may act as a deterrent to reduce such crimes occurring and would help to improve prosecutions if they occurred.

1.5 Aims

There is now a need to be able to forensically identify both contemporary and historic rhinoceros horns. A panel of markers has been developed and is in use in South Africa to identify live African rhinoceroses (*C. simum* and *D. bicornis*) and for use in poaching crimes (Harper *et al.* 2013); although the panel of markers has not yet, to our knowledge, been tested under a developmental validation to ensure it is fit for purpose following guidelines developed for animal forensic systems (Budowle *et al.* 2005; Linacre *et al.* 2011; SWGWILD 2012). Furthermore, whilst it has been demonstrated that nuclear DNA can be extracted from contemporary horns (Peppin 2009; Peppin *et al.* 2010; Harper *et al.* 2013) and that they can be identified using STR markers (Harper *et al.* 2013), the application of such methods to historic horns has not previously been investigated. The aims of this thesis were therefore to describe a forensic developmental validation of a panel of STR markers for *D. bicornis* and to identify an optimal DNA extraction method for historic horns.

Chapter 2 – Validation of a forensic profiling system for black rhinoceros

2.1 Introduction

Identification of a sample at an individual level, rather than to species or population level, is carried out using DNA profiling (see section 1.4.1.1). DNA profiling, also known as DNA fingerprinting, was developed in 1985 for humans (Jeffreys *et al.* 1985a; b), and later that same year its potential for use in forensic analysis was recognised (Gill *et al.* 1985). The methods used for human DNA profiling were adapted for use in birds (Burke & Bruford 1987; Wetton *et al.* 1987) and have since become a standard method in both forensic (Ogden *et al.* 2009; Roewer 2013) and wildlife DNA analysis (Chambers *et al.* 2014; Nybom *et al.* 2014).

Since the method's conception, the markers used in DNA profiling, and the methods used to genotype them, have changed as technological advances have been made (for reviews see Roewer 2013; Chambers *et al.* 2014; Nybom *et al.* 2014), and will continue to do so. A good panel of markers for use in DNA profiling has the following qualities: high variability among individuals enabling high discriminatory power, ability to amplify from degraded DNA samples, and have a rapid and scalable genotyping method that is robust, reproducible and accurate (Butler 2005). STRs are able to meet these criteria, and, at present, are the gold standard for forensic DNA profiling (Butler 2005; Wictum *et al.* 2013).

2.1.1 STRs

STRs are a type of variable number tandem repeat (VNTR), which are hypervariable and formed of repetitive nucleotide sequences (Parker *et al.* 1998; Ellegren 2004). STRs (also known as microsatellites) are formed of the shortest repeat units, typically between one and six bp (Parker *et al.* 1998); larger repeat units are called minisatellites. The number of times a sequence must be repeated in order to be classified as an STR is not well defined (Ellegren 2004). However, in practice, fragment length analysis via PCR amplification and capillary electrophoresis of sequences shorter than 50 – 80 nucleotides becomes impractical due to background noise generated by the fluorescent tags and primers. Therefore, STRs used in population studies are rarely short enough to raise debate over their classification as an STR.

An STR locus is composed of the STR repeat and the flanking regions of DNA on either side (Figure 2.1). Whilst the STR repeat is variable, the flanking regions are conserved among individuals. This enables a single primer pair to be designed within the upstream and

downstream flanking regions which will amplify all alleles within a population or species. The amplified fragment will therefore show size polymorphism due to the varying lengths of the STR repeats. PCR amplification using fluorescently labelled primers and subsequent capillary electrophoresis enables fragment size polymorphism at a locus to be detected.

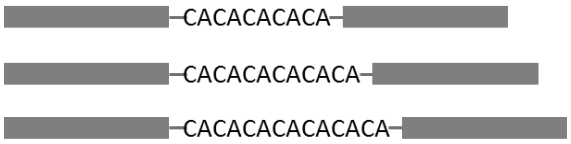


Figure 2.1. Schematic of STR alleles of a dinucleotide CA repeat. Each line represents a different STR allele that could exist within a population, formed of 5, 6 or 7 repeats of the CA dinucleotide. Grey bars on either end represent the flanking regions.

STR loci are typically described according to the characteristics of the repeat unit (Table 2.1). Firstly, they are described by the type of repeat they contain (Urquhart *et al.* 1994). Secondly, they are named for the number of base pairs within the repeat, for example a dinucleotide repeat contains two nucletodides.

Table 2.1. Characteristics of STR loci that can occur. STRs are typically named after the number of nucleotides included in a repeat unit, e.g. a dinucleotide contains two bp repeat units and is denoted by the prefix "di", or whether they are complex or compound. The first repeat unit is in bold.

Type of STR*	Number base pairs in repeat units	Prefix-	Example	Sequence repeat designation	Allelic nomenclature
Simple	1	Mono-	**A**AAAAAA	$(A)_7$	7
	2	Di-	**CA**CACACA	$(CA)_4$	4
	3	Tri-	**CAG**CAGCAG	$(CAG)_3$	3
	4	Tetra-	**TACA**TACATACA	$(TACA)_4$	4
	5	Penta-	**GAAAA**GAAAAGAAAA	$(GAAAA)_3$	3
	6	Hexa-	**AGGAAT**AGGAATAGGAAT	$(AGGAAT)_3$	3
Compound	1-6		**AC**ACACGCACGC**AC**GCAC	$(AC)_3(GCAC)_3$	6
Complex	1-6		**AC**ACACAGGACACACTGCATGCA	$(AC)_3AGG(AC)_3(TGCA)_2$	8

* A further designation of these STR types exists, simple/compound/complex with non-consensus allele, in which one or more of the alleles contains a sequence variant (e.g. incomplete repeat, transition, indel - Urquhart *et al.* 1994).

Allelic nomenclature can be assigned following guidelines for forensics (Gill *et al.* 1994, 1997; Bär *et al.* 1997; Hellmann *et al.* 2006). Nomenclature is assigned according to the number of times each unit is repeated within the STR for simple repeats (Table 2.1). In compound and

complex repeats, it is the sum of all the repeats. An allele may contain complete or incomplete repeat units, for example:

1. 5' – AGT GGA **CAC CAC CAC CAC CAC** TCT – 3'
2. 5' – AGT GGA **CAC CAC CAC CAC CAC CA**T CTG – 3'

The first example contains five complete copies of CAC. In the second example, however, the sixth repeat unit is incomplete and is missing the final C. Incomplete alleles are designated according to the number of full repeats and the number of nucleotides in the partial repeat, separated by a decimal point (Bär *et al.* 1997)– example 1 is allele 5 and example 2 is allele 5.2.

The highly variable nature of STRs make them well suited to identity testing as they provide high discriminatory power between individuals (Butler 2005). Due to the repetitive nature of STRs, slippage and proofreading errors frequently occur during replication (Eisen 2001), producing alleles which are typically an entire repeat unit larger or smaller than the template allele; although replication errors of less than a whole repeat unit do occur. Highly variable STRs are usually located in non-coding regions of the genome (Ellegren 2004) and are therefore not under selective pressure to maintain advantageous or eliminate deleterious alleles, allowing mutations to build up. The STR mutation rate is not constant and may differ between alleles, loci or species; however the greater the number of repeat units, the greater the rate of mutation (Ellegren 2004).

Offspring inherit one copy of the STR from each parent, and can thus can inherit two different alleles or two copies of the same allele. At a single locus with many alleles, there will be lots of variation amongst individuals within a population, but, barring new mutations, no individual will be unique. Whereas across all STRs within an individual's genome, the allelic complement is unique. It is not, however, feasible to genotype all STRs within the genome and instead a subsample are genotyped. With increasing numbers of loci, the specific combination of alleles within an individual becomes increasingly rarer and eventually unique. Thus, by amplifying a panel of loci which contains an adequate number of loci with sufficient variability, it is possible to generate unique individual profiles (Ogden *et al.* 2009; McGraw *et al.* 2012).

Forensic casework that requires identity testing is often typified by trace samples with very low DNA quantity and degraded DNA. PCR amplification of STRs is very sensitive, thus the amount of template DNA required is very low (Butler 2005; Cassidy & Gonzales 2005). Providing a few intact strands can be bound by the primers, successful amplification and

genotype scoring can occur. Degraded DNA is typified by random shearing into small fragments, often only a few hundred bp long (Butler 2005). An inverse relationship has been found for human DNA between the size of a locus and its amplification success from degraded DNA (Whitaker *et al.* 1995; Sparkes *et al.* 1996; Takahashi *et al.* 1997; Schneider 2004; Butler 2005), and therefore the shorter the amplified fragment, the greater the success rate. STRs themselves are typically 20 – 200 bp, and thus, including the flanking regions required to bind the primers, amplified fragments are often in the size range of 100 – 500 bp (Butler 2005; Cassidy & Gonzales 2005). STRs are therefore typically short enough to be recovered from degraded DNA (Butler 2005; Cassidy & Gonzales 2005).

2.1.2 Validations

Unlike typical research, the results of each and every test in criminal casework has serious implications. Wrongful accusation (and particularly subsequent wrongful conviction) of a crime can be extremely detrimental to an individual both personally and professionally (Moore & Kornfield 2012). On the other hand, it would be detrimental for evidence to be missing or rejected simply because a laboratory or a method is not up to standard. The field of human forensic analysis has been developing for many years and is accompanied by standards and guidelines to ensure that methods are robust for use. Forensic evidence cannot afford to be inaccurate. Quality assurance (QA) and quality control (QC) allow confidence to be placed on a test result that it is accurate, reliable and reproducible (Butler 2005; Moore & Kornfield 2012).

QA is a system of verification that ensures tests generate reliable results (Budowle *et al.* 2005; Butler 2005; Moore & Kornfield 2012; Ansell 2013), including procedures such as staff training, maintenance of reliable equipment, traceability of reagents, standard operating procedures, proficiency testing and auditing of procedures. QC involves the daily activities that ensure that methods are functioning as expected and generating appropriate results (Budowle *et al.* 2005; Butler 2005; Moore & Kornfield 2012; Ansell 2013). QC activities include the use of negative controls, positive standards, dual person checking of sample transfer, use of separate pre- and post-PCR facilities, and traceability of data analyses. QA and QC systems allow confidence to be placed in the results generated by DNA forensic tests.

A developmental validation is a QA process which defines the limitations of a new or novel technology applied in a forensic context (Moore & Kornfield 2012). Ultimately, the developmental validation aims to demonstrate that the method is reproducible, reliable and

accurate in trained hands (Butler 2005). As wildlife DNA forensic testing has become established, there have been many calls for such tests to be carried out in adequate laboratories and for QA/QC practices to be followed, including developmentally validation new techniques (Budowle *et al.* 2005; Cassidy & Gonzales 2005; Ogden 2010; Linacre *et al.* 2011; Moore & Kornfield 2012).

The DNA Advisory Board (DNA Advisory Board 2000) has established firm guidelines for validating human profiling systems. Standards and guidelines have also been proposed for wildlife species (Budowle *et al.* 2005; Linacre *et al.* 2011) and minimum standards were formalised by SWGWILD (2012). Wildlife validations cannot always meet all guidelines established for human systems, for example it is rarely possible to obtain sufficient pedigree data to establish mutation rates of STRs (Dawnay *et al.* 2008). Comprehensive validations of STR marker panels have been carried out for a number of domestic and commercial species including dogs (*Canis lupis familiaris* - Kun *et al.* 2013; Wictum *et al.* 2013), cats (*Felis catus* - Menotti-Raymond *et al.* 1997; Coomber *et al.* 2007; Berger *et al.* 2014) and pigs (*Sus scorfa* - Lin *et al.* 2014). Developmental validations of STR panels which meet many of the standards and guidelines described by Budowle *et al.* (2005) and Linacre *et al.* (2011) have been carried out for non-domestic species including Eurasian badgers (*Meles meles* - Dawnay *et al.* 2008) and European brown bears (*Ursus arctos* - Andreassen *et al.* 2012). Many studies, however, have only considered parts of the developmental validation process, either aspects of marker characterisation (Rütten *et al.* 2001; Anderson *et al.* 2002; Harper *et al.* 2013) or population studies (Jobin *et al.* 2008; Dawnay *et al.* 2009; Caratti *et al.* 2010; Goor 2011), but remain limited in their knowledge of the system's limitations.

This study aims to describe a developmental validation for a panel of 16 markers for black rhinoceros *D. bicornis*. A number of challenges are faced when conducting such tests on a wildlife species. Firstly, access to high quality and quantity DNA in order to characterise the markers and their specificity, reproducibility and accuracy. Captive individuals may provide such samples; however due to ethical and practical reasons, these are only available opportunistically from veterinary procedures. Secondly, large numbers of samples are required for population studies to assess the markers for their discriminatory power, Hardy-Weinberg equilibrium, null alleles and linkage disequilibrium. For large species such as rhinoceros, obtaining such numbers is not possible from captive individuals and from wild individuals is expensive, time consuming and ethically challenging as the process requires anaesthetising the animals. As such, the black rhinoceros, *D. bicornis*, was selected for developmentally validating an STR marker panel as stored samples from captive rhinoceros

were kindly provided by Port Lympne Wild Animal Park and population samples had previously been collected for a previous study (Muya *et al.* 2011).

2.2 Materials and methods

2.2.1 Sample collection

For the developmental validation studies, 21 *D. bicornis michaeli* blood and tissue samples were obtained from captive rhinoceros (n=5; Port Lympne Wild Animal Park) and wild Kenyan rhinoceros (n=16; Muya *et al.* 2011). A further 36 wild rhinoceros samples (Muya *et al.* 2011) from reserves across Kenya were genotyped to generate population data. Briefly, Kenyan samples were collected during routine procedures and stored in 70 % ethanol or 25 % dimethylsulphide (DMSO), but see Muya *et al.* (2011) for further details.

Horn material was obtained from one deceased *D. bicornis michaeli* individual (Port Lympne Wild Animal Park) and from one deceased *C. simum* individual. From a cross-section of the *D. bicornis michaeli* horn, a standard drill with a sterile 5 mm drill bit was used to drill slowly into the base of the horn, pausing frequently to prevent the drill bit heating and damaging the DNA. Only fragments of the *C. simum* horn were available and the position on the horn from which they originated was unknown. Fragments and shavings were obtained by drilling as before from the surface of the fragments. Shavings of horn were collected on clean paper and transferred to a sterile 50 mL tube for storage at – 20 °C.

2.2.2 DNA extraction, PCR amplification and STR analysis

DNA was extracted from all samples using either Qiagen® DNEasy® Blood and Tissue or QIAamp® DNA Investigator kits. Blood and tissue samples were extracted following standard protocols. Horn material was extracted following a modified protocol by which 100 mg of material was ground using a Retsch mixer mill (2 mins at 25 revolutions/s) and digested for 24 hours in extraction buffers at three times volume, 40 µL proteinase K and 10 µL dithiothreitol (DTT), and completed under standard protocol. Extraction controls were included throughout following standard forensic protocols (SWGWILD 2012).

Sixteen loci were chosen for the validation study consisting of fifteen STR markers and one sexing marker (Table 2.2). Markers were selected for their polymorphism in *D. bicornis* (Cunningham *et al.* 1999; Scott 2008; Nielsen *et al.* 2008), potential for multiplex amplification and potential cross-amplification in other rhinoceros species (Scott 2008).

Furthermore, all loci have been included on the panel used by Harper *et al.* (2013), thus there is potential for cross comparability of data between laboratories. Sex determination was achieved by using a single primer pair to co-amplify a polymorphic region of a zinc-finger protein (ZF) intron which exhibits a size difference of 7 bp between female (ZFX) and male (ZFY) homologs (Peppin *et al.* 2010).

PCRs were performed in 10 µL reactions containing 1x Qiagen Type-It Multiplex PCR master mix (includes 3mM $MgCl_2$), primers, water and approximately 10 ng of template DNA. Primers were included at 0.2 µM in singleplex reactions or according to Table 2.2 for multiplex reactions. PCR conditions were 5 min at 95 °C, 28 (blood/tissue samples) or 33 (horn samples) cycles of 95 °C for 30 s, T_A (60 °C or 65 °C) for 90 s, 72 °C for 30 s, followed by final extension at 60 °C for 30 min.

Capillary electrophoresis was carried out on the AB 3130xL genetic analyser using 1 µL PCR product (undiluted or up to 1/500 dilution) added to 10 µL Hi-Di Formamide mixed with 0.05 µL Genescan™ 500LIZ® as an internal ladder. Fragment lengths were binned and scored using GeneMapper® 4 (Applied Biosystems). Allele calls were checked by eye using a conservative analytical peak height threshold of 100 relative fluorescence units (RFU) and a stochastic peak height threshold of 200 RFU for homozygote calls.

2.2.3 Developmental validation studies

Chromosomal locations of the 15 STR loci were determined on the horse (*Equus caballus*) genome, using the Ensemble BlastN search. As genomic sequences were not yet available for any rhinoceros species, the horse was the most closely related species to *D. bicornis* with extensively mapped genomic sequence. Samples from two known parent-offspring pairs (one father–daughter and one mother–daughter) were used to assess for consistency of allelic Mendelian inheritance patterns. One horn sample and corresponding blood sample (blood card) were analysed for one *D. b. michaeli* individual to confirm that different stating materials generate concordant genotypes, as well as horn and blood samples from a *C. simum* individual. DNA samples from one female domestic dog and one female human were tested against the marker set to assess cross-species amplification, alongside DNA samples from four *C. simum* individuals.

Samples from 13 individuals were amplified and genotyped three times independently to assess the reproducibility of genotyping results. Following the third round of amplification and genotyping, two samples per locus were transferred to a new plate and re-analysed by

the AB 3130xL to confirm reproducibility of results between capillaries. Accuracy and precision were assessed by calculating d-values, stutter ratios and heterozygote balance for each locus. D-value is calculated between repeat runs for every allele per individual as:

$$\text{D-value} = \text{maximum allele size (bp)} - \text{minimum allele size (bp)}$$

D-values therefore measure the drift of alleles between runs (schematic shown in Appendix I). Heterozygote balance is calculated as:

$$\text{Heterozygote balance} = \frac{\text{peak height of short allele}}{\text{peak height of long allele}}$$

There have been many debates regarding the best way to calculate and assess heterozygote balance (Gill & Buckleon 2005), and many human forensic studies calculate the reverse (long allele/short allele). Wildlife DNA profiling studies rarely assess this, but this calculation is used by those that do (Howard *et al.* 2008; Andreassen *et al.* 2012; Berger *et al.* 2014) and it provides more information regarding the direction of imbalance for dinucleotides. A heterozygote balance of 1 indicates that both peaks are of equal heights, >1 that the short allele is taller than the long allele and <1 that the short allele has a lower peak height than the long allele. Heterozygote balance was not calculated for adjacent alleles as the short allele RFU value will be a measure of the intensity of both the true short allele and the stutter of the long allele.

Finally, stutter ratio is calculated as:

$$\text{Stutter ratio} = \frac{\text{peak height of stutter peak}}{\text{peak height of true allele}}$$

The stutter ratio is calculated for the stutter peak at -1 repeat unit; all loci are dinucleotides, thus ratios are calculated using the stutter peak two bp smaller than the true allele. Stutter ratios were only calculated for non-adjacent alleles.

Sensitivity to starting DNA template concentration was assessed by amplifying DNA from five individuals serially diluted to concentrations of 5, 2.5, 1.25, 0.62, 0.31 and 0.15 ng/μL. Amplification success and heterozygote balance were assessed at each concentration.

The effects of annealing temperature (±2 °C and ±4 °C of standard T_A of 60 °C or 65 °C) and cycle number (+5, +10 and +15 cycles above the standard 28 cycles) were assessed for three samples. $MgCl_2$ concentration was not varied as this is standardized at 3 mM in the Type-It microsatellite mastermix.

Table 2.2. Details of primer pairs validated in this study.

Primer Pair	Accession number (GenBank)	Forward primer sequence	Reverse primer sequence	T_a (°C)	Fluorescent Tag	Multiplex	Multiplex concentration (uM)*	Reference
BlRh1B	AY606078	GATCAGTAACACCAAAGTCC	AGTGAAGACAGAAGGATCAC	65	VIC	1	0.2	Nielsen *et al*. 2008
BlRh1C	AY606079	AGATTCTTGGAAAGGTCACT	AACATTGGGTTTCACCTC	60	6FAM	2	0.2	Nielsen *et al*. 2008
BlRh37D	AY606083	ACATGTGTAAACTTGGGAAC	TGGTTCATTGATCTCTTCTC	60	NED	2	0.2	Nielsen *et al*. 2008
BR6		TCATTTCTTTGTTCCCCATAGCAC	AGCAATATCCCACGATATGTGAAGG	65	NED	1	0.2	Cunningham, Harley, & O'Ryan 1999
DB1	AF129724	TAAGTCACAGGGACTAATCTG	GAGGGTTTATTGTGAATGAG	65	6FAM	1	0.2	Nielsen *et al*. 2008
DB23	AF129734	ATCTTCCTCAGCAATAAGG	ATCATCAGAGTTTCCAGTTC	60	6FAM	2	0.15	Nielsen *et al*. 2008
DB44	AF129730	AGGGTGGAATGTCAAGTAG	CTTCTAGAGGGAGACTAGGAG	65	6FAM	1	0.3	Nielsen *et al*. 2008
DB52	AF129732	CATGTGAAATGGACCGTCAGG	ATTTCTGGGAAGGGGCAGG	60	PET	2	0.2	Brown & Houlden 2000
DB66	AF129733	CCAGGTGAAGGGTCTTATTATTAGC	GGATTGGCATGGATGTTACC	60	VIC	2	0.3	Brown & Houlden 2000
IR12		GAATGCTGATCATTTAGTGAC	GGGTCCAGTTGAGATATCAC	65	PET	1	0.4	Scott 2008
IR22		ATGGTGGAAGAAGTGCAGCC	ACTTCTGTGTCTCTAGCGCC	65	PET	1	0.05	Scott 2008
SR63	AY427965	CTTGAGCAGAGTAGAATTTGG	CTCTGTATCCACCTCATTCC	65	NED	1	0.2	Scott *et al*. 2004
WR32A	AY138541	CTAGCAAAATCTCAAAGAGG	TTACTAAGGGAATCACCAAG	60	6FAM	2	0.3	Nielsen *et al*. 2008
WR7B	AY138544	AACCAACTTGTAATGAGAGG	AATGAACAGGAAGGAAGAC	60	NED	2	0.15	Nielsen *et al*. 2008
WR7C	AY138543	GTCAGTTCAAGTTTTTGCTC	CTCATCCATGCTTCTTCTAC	60	PET	2	0.35	Nielsen *et al*. 2008
ZF1	DQ519375	GATTTGGAASCTAGGCATTTCC	GCCATGATACTCATGAATGACA	65	6FAM	1	0.15	Peppin *et al*. 2009

2.2.4 Allele sequencing

Common alleles were sequenced for each locus to confirm the repeat motif and number of repeats. Homozygous samples were amplified in a final reaction volume of 20 µL containing 0.2 µM of each primer, 1x HotStarTaq Plus Master Mix (contains 1.5mM $MgCl_2$), water and approximately 2 ng of template DNA. PCR conditions were 5 min at 95 °C, 30 or 35 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s, followed by final extension at 72 °C for 10 min. BIGDYE version 1.1 chemistries were used to sequence the alleles on an AB 3130xL. Sequences were edited using Geneious 5.5.7. Allelic nomenclature was assigned following established guidelines (Gill *et al.* 1994, 1997; Bär *et al.* 1997).

2.2.5 Multiplex, population studies and data analysis

Markers were combined to form two multiplex reactions (containing eight primer pairs each; allelic size ranges shown in figure 2.2) using the same thermocycling conditions as previously for singleplex reactions, with adjusted primer concentrations (Table 1). The multiplex was tested on 13 individuals previously genotyped in singleplex in order to confirm that both singleplex and multiplex reactions generated consistent profiles. This test was repeated for DNA extracted from *D. b. michaeli* and *C. simum* horn material. Genotyping error was assessed by re-amplifying 10 samples in multiplex which were scored independently by an experienced DNA forensic analysis with no prior experience of scoring this panel of markers.

DNA profiles were generated for a total of 52 wild rhinoceros using the two multiplex reactions to determine allele frequencies across the Kenyan population of *D. b. michaeli*. Allele frequency data and descriptive statistics were generated in GenAlEx 6.5 (Peakall & Smouse 2006), and polymorphism information content (PIC) scores were calculated using Microsatellite Toolkit 3.1.1 (Park 2001). All loci were checked for the presence of null alleles using Micro-checker 2.2.3 (Van Oosterhout *et al.* 2004). GenePop on the Web 4.2 (Raymond & Rousset 1995) was used to evaluate Hardy-Weinberg equilibrium (HWE) and linkage equilibrium. Sequential Bonferroni correction was applied to determine significance thresholds for HWE and linkage equilibrium in order to account for multiple comparisons.

2.2.6 Match probabilities

Average probability of identity (PI_{ave}) was calculated in API-Calc (Ayres & Overall 2004) for the wild *D. b. michaeli* population. The impact of population structure (F_{st}) was modelled assuming the population is composed of entirely unrelated individuals, entirely related as cousins and entirely related as full siblings.

Cumulative likelihood random match probabilities (RMPs) (Weir 2012) were calculated for four of the five samples profiled during the developmental validation to assess sensitivity. The fifth sample (KA) originated from a captive rhino and thus the RMP could not be calculated as allele frequencies were unknown. RMPs were also calculated for each profile generated at decreasing concentrations of template DNA. The precise geographic origins of samples was unknown, thus a default value of θ=0.1 was used (Dawnay *et al.* 2009) to account for population structure. This should provide a conservative estimate based on previous calculations for Kenyan black rhinoceros subpopulations (Muya *et al.* 2011), which estimated that pairwise θ (F_{ST}) values ranged between -0.034 and 0.075 for all population pairs except comparisons with the Masai Mara population which may contain multiple subspecies. Furthermore, the inbreeding fixation index (F_{IS}) did not exceed 0.127 for any subpopulation and the mean was 0.046 ± 0.091 (Muya *et al.* 2011), indicating that within population relatedness was minimal. A 99 % source attrition threshold value (P_X), the value above which the RMP is considered to confer 99 % confidence the profile is unique within the population, was calculated as

$$P_X = \frac{1}{1-(1-\alpha)^{1/N}}$$

where α = 0.01 (corresponding to a 99% threshold) (Budowle *et al.* 2000; Butler 2005) and N=740, the approximate size of the wild Kenyan population (Emslie 2012).

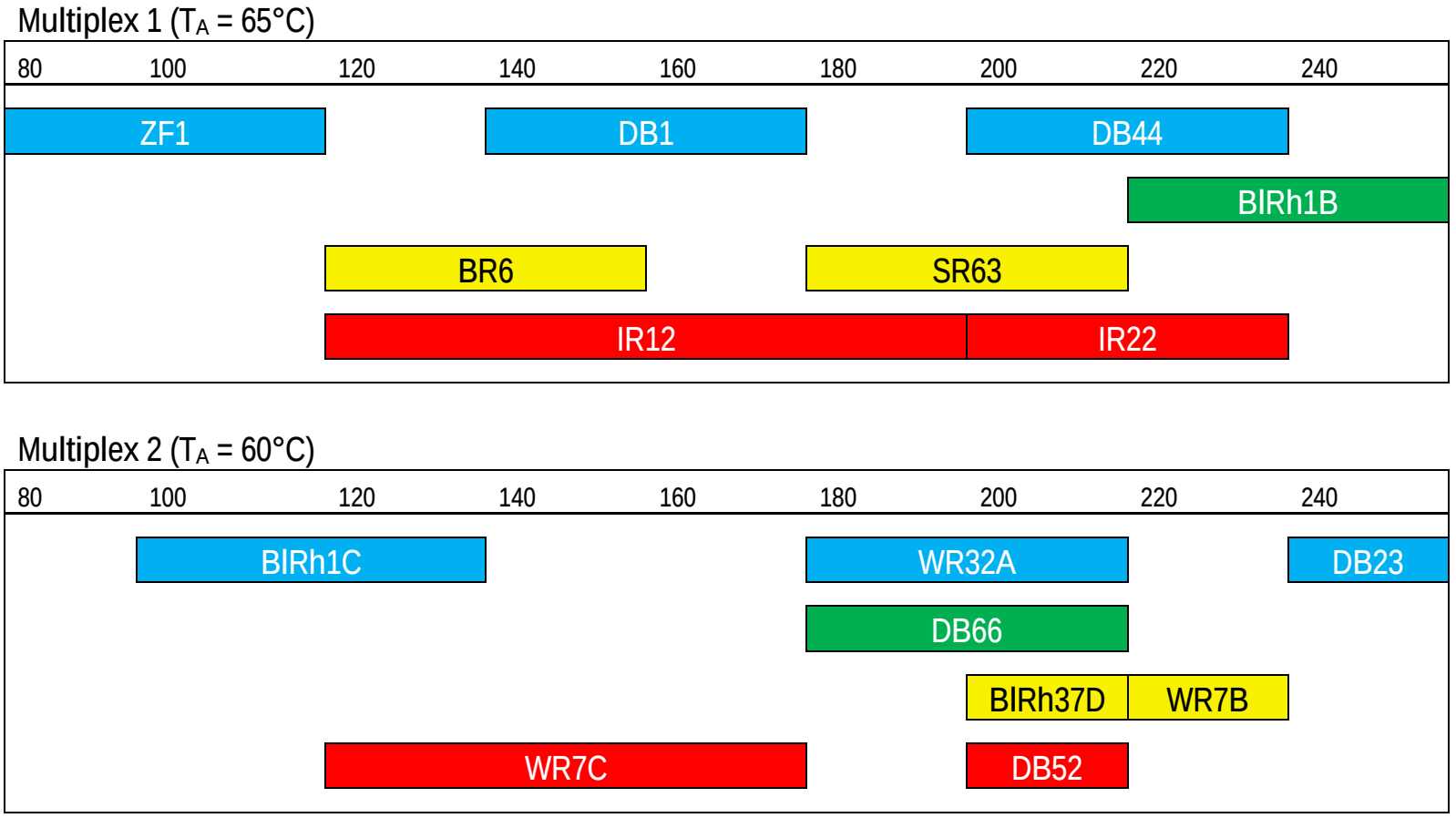


Figure 2.2. Schematic of allele size ranges of markers as included in two multiplexes. Colours indicate fluorescent labels (blue – FAM-6, green – VIC, yellow – NED, red – PET).

2.3 Results

2.3.1 Developmental validation

2.3.1.1 Mendelian inheritance

Genotypes of two parent–offspring groups (one father–daughter and one mother–daughter group) confirmed that loci inheritance was consistent with patterns of Mendelian inheritance. Genotypes are shown in Table 2.3. At least one allele was present in both parent and offspring genotypes with maximum d-values of 0.31 bp for the father–daughter pair and 0.33 bp for the mother–daughter pair.

Table 2.3. Alleles observed in two parent-offspring pairs for all 16 alleles. At least one allele was observed in both the parent and the offspring genotypes. All samples were obtained from captive rhinos from Port Lymnpe Wild Animal Park.

		BlRh1B		BlRh1C		BlRh37D		BR6	
Family	Individual	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2
1	Mother	18	20	25.1	28.1			20	20
	Daughter	18	20	23.1	28.1	23.1	23.1	20	21
2	Father	16	18	22.1	22.1	23.1	23.1	19	20
	Daughter	18	18	22.1	25.1	23.1	23.1	20	21

		DB1		DB23		DB44		DB52	
Family	Individual	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2
1	Mother	14	16	12	12	24.1	24.1	18.1	23.1
	Daughter	14	14	12	12	24.1	24.1	23.1	23.1
2	Father	12	18	12	15	18.1	22.1	16.1	17.1
	Daughter	14	18	12	12	18.1	22.1	17.1	17.1

		DB66		IR12		IR22		SR63	
Family	Individual	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2
1	Mother	27	27			20	20	19	20
	Daughter	24	27	25	25	20	20	19	19
2	Father	24	27	19	19	19	20	20	20
	Daughter	24	27	19	25	19	20	19	20

		WR32A		WR7B		WR7B		ZF1	
Family	Individual	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2	Allele1	Allele2
1	Mother	22	22	18.1	18.1	18.1	18.1	98	98
	Daughter	22	23	18.1	19.1	18.1	19.1	98	98
2	Father	21	21	18.1	19.1	18.1	19.1	98	105
	Daughter	21	21	18.1	19.1	18.1	19.1	98	98

2.3.1.2 Different starting materials

Blood and horn samples generated identical profiles for each of the *D. bicornis* and *C. simum* individuals (Appendix II). The *D. bicornis* blood sample was obtained from a blood card, generating low quantity DNA and this has resulted in amplification failure for BlRh37D, IR12 and WR7C (these genotypes were obtained from the horn sample). WR7C also failed to amplify from the *C. simum* horns in either single- or multiplex.

2.3.1.3 Chromosome mapping

The chromosomal locations of 12 of the 15 STR loci were identified on the *E. caballus* genome. The best alignment matches are shown in Table 2.4. Sequences for the flanking regions of BR6, IR12 and IR22 were not available and thus chromosomal locations could not be determined. Loci BlRh1C, DB66 and WR7C were all located on *E. caballus* chromosome 5. DB1 and SR63 were located on *E. caballus* chromosome 14. For DB44, the best alignment match was only 30 bp long, representing only 8% of the DRB44 GenBank sequence, and did not contain a microsatellite. All other STR loci were located on separate chromosomes. The sexing marker, ZF1, amplifies homologous fragments on the X and Y chromosomes (Peppin *et al.* 2010).

Table 2.4. Chromosomal location of each locus on the genome of the domestic horse *Equus caballus.*

Marker	Accession number	Chromosome on *Equus caballus* genome	Length of best alignment match (bp)*	Length of GenBank sequence (bp)	E-value (Identity match %)
DB52	AF129732	Chromosome 1	220 + 84 + 27	508	1e-40 (86.0 %) + 7e-08 (86.0 %) + 7e-05 (100 %)
BlRh37D	AY606083	Chromosome 3	258	511	5e-52 (85.3 %)
BlRh1B	AY606078	Chromosome 4	483	526	4e-102 (86.0 %)
BlRh1C	AY606079	Chromosome 5	69 + 63	165	6e-21 (94.0 %) + 8e-08 (88.9 %)
DB66	AF129733	Chromosome 5	369 + 176	714	2e-121 (90.0 %) + 1e-39 (85.2 %)
WR7C	AY138543	Chromosome 5	111+188	389	1e-21 (87.4 %) +2e-20 (83.5 %)
WR32A	AY138541	Chromosome 6	331	325	7e-78 (85.5 %)
WR7B	AY138544	Chromosome 7	315	315	7e-90 (87.6 %)
DB44	AF129730	Chromosome 9	29	379	0.19 (93.1 %)
DB1	AF129724	Chromosome 14	831	811	0.0 (86.0 %)
SR63	AY427965	Chromosome 14	204	220	1e-35 (85.3 %)
DB23	AF129734	Chromosome 15	182	342	5e-47 (89.0 %)

Note: No sequence data was available for BR6, IR12 or IR22 on GenBank or in the literature.
* Consecutive sequence in *D. bicornis* but not in *E. caballus* shown as +

2.3.1.4 Species specificity

Species specificity tests demonstrated that WR7B and the sexing marker ZF1 produced amplicons for both dog and human DNA (Figure 2.3). Locus WR7B generated homozygous alleles for both the human and dog DNA amplified, producing fragment sizes of 227.89 bp and 215 bp respectively. Both amplicons are close to the size range of observed *D. b. michaeli* alleles (222.86 bp – 225.41 bp). ZF1 also cross-amplified female human and dog DNA. Homozygous ZFX fragments were produced at 101.24 bp for human DNA and 99.8 bp for dog DNA. These sizes differ from those detected for black rhinos (ZFX = 98.06 ± 0.20 bp, ZFY = 105.32 ± 0.11 bp; see 2.3.1.5).

Fourteen STR markers and ZF1 amplified DNA from four *C. simum* samples; no amplification was observed from BlRh37D. BR6, DB23 and IR22 generated identical, homozygous profiles in all four individuals assessed. Fragment sizes for the remaining 11 loci were within the same range as those of *D. bicornis*.

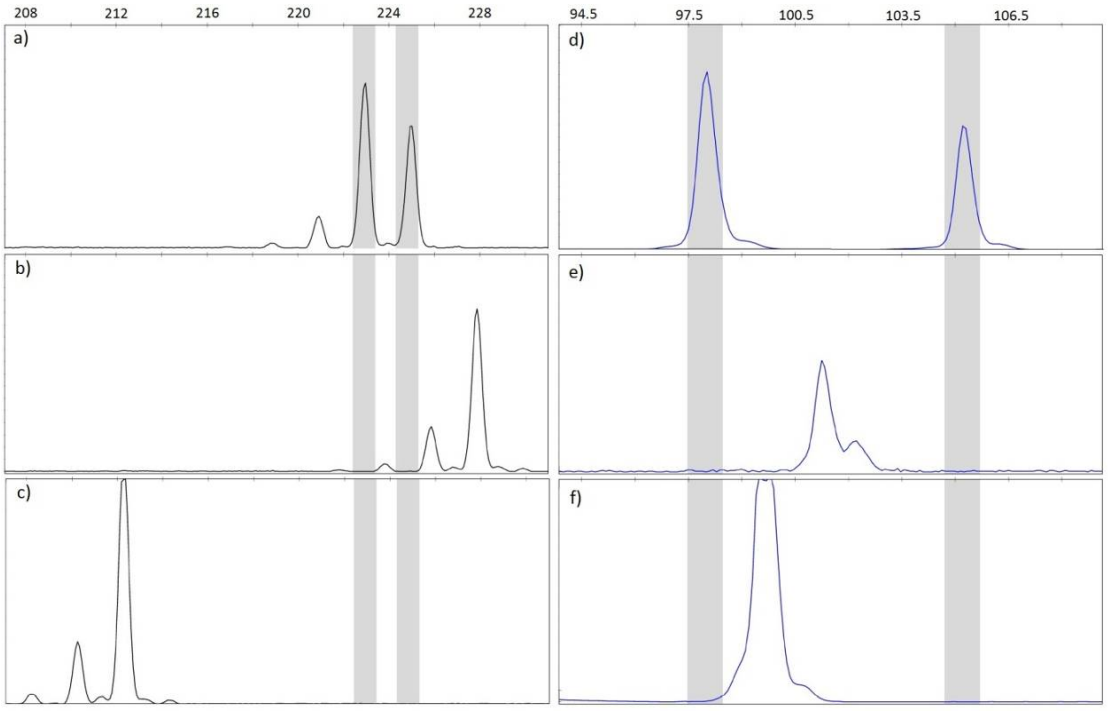


Figure 2.3. Electropherograms for locus WR7B in a) heterozygous *D. bicornis*, b) human and c) domestic dog and locus ZF1 in d) male *D. bicornis*, e) female human and f) female domestic dog. Size in bp is shown above electropherograms. Grey bars indicate allele bins for *D. b. michaeli*.

2.3.1.5 Reproducibility, precision and accuracy

DNA profiles obtained for 13 individuals to demonstrate reproducibility were identical across three separate amplifications and random injection positions for capillary electrophoresis. Mean and maximum d-values calculated for each locus are shown in Table 2.5. Mean d-values range from 0.067 for DB44 up to 0.348 for BlRh1B. The maximum d-value observed for any individual allele was 0.545 for BlRh1B.

Table 2.5. Mean and maximum d-values, mean heterozygote balance and mean stutter ratio for each locus. D-values were calculated for each set of three replicated amplifications per individual for each allele as the largest fragment size (bp) – smallest fragment size (bp). Heterozygote balance was calculated as peak height of short allele / peak height of long allele, thus values above 1 indicate that the shorter allele had a greater peak height than the longer allele. Stutter ratio was calculated as peak height of stutter at –1 repeat / peak height of true allele.

Locus	Mean d-value	Maximum d-value	Mean heterozygote balance	Mean stutter ratio
BlRh1B	0.348	0.545	1.255	0.31
BlRh1C	0.108	0.217	1.389	0.49
BlRh37D	0.108	0.162	1.220	0.22
BR6	0.277	0.385	1.411	0.44
DB1	0.185	0.329	1.222	0.33
DB23	0.195	0.370	1.241	0.15
DB44	0.067	0.142	1.153	0.34
DB52	0.068	0.149	1.311	0.42
DB66	0.242	0.409	1.344	0.37
IR12	0.086	0.167	1.468	0.48
IR22	0.081	0.163	1.436	0.47
SR63	0.202	0.331	1.414	0.28
WR32A	0.156	0.356	1.234	0.37
WR7B	0.315	0.461	1.269	0.21
WR7C	0.198	0.359	1.428	0.34
ZF1	0.119	0.403	1.001	

Heterozygote balance ranged from 0.72 (short allele has lower peak than long allele) to 2.74 (short allele has taller peak than long allele). The short allele was taller than the long allele, heterozygote balance greater than 1.0, in 90.8 % of amplifications, and heterozygote balance

ranged between 1.2 and 1.5 for 57 % of alleles (Figure 2.4). Mean heterozygote balance (Table 2.5) ranged from 1.00 for ZF1 to 1.47 for IR12 (Figure 2.4). The sexing marker ZF1 displayed heterozygote imbalance on 15 of 21 occasions, such that the shorter ZFX allele had a lower peak height than the longer ZFY allele (Figure 2.6). Heterozygote balance for this locus ranged between 0.72 and 1.72, although was typically less than 1.1 (Figure 2.4). This variation in heterozygote balance did not hinder allele scoring. Locus DB44 also showed heterozygote imbalance on 13 of 33 occasions. Heterozygote balance for this locus ranged from 0.88 to 1.62. It only dropped below 0.95, however, on four occasions and allele scoring remained unambiguous. On two occasions heterozygote balance was greater than 1.8 for locus WR7C (Figure 2.5). On both occasions, these occurred in heterozygous individuals for which the shorter allele was 29 bp smaller than the large allele, and thus the presence of two alleles remained apparent.

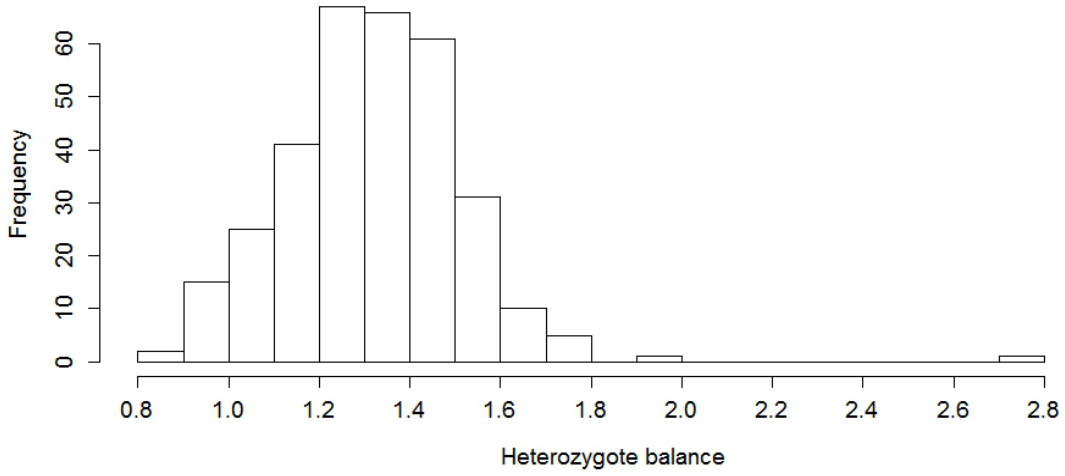


Figure 2.4. Histogram of heterozygote balance across all STR loci (ZF1 is excluded). For most amplifications, this ranged between 1.2 and 1.5.

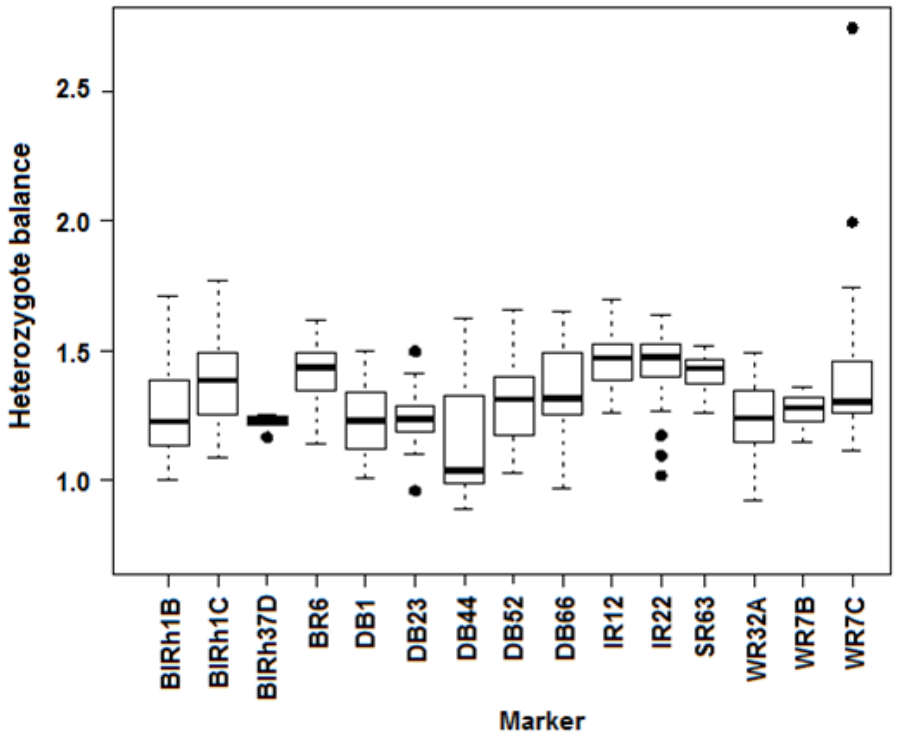


Figure 2.5. Box plots showing heterozygote balance for all markers. Dark lines represent the mean for each locus and the boxes represent the first and third quartiles. Whiskers illustrate the limits of the data, to a limit of 1.5 times the interquartile range. Circles represent outliers.

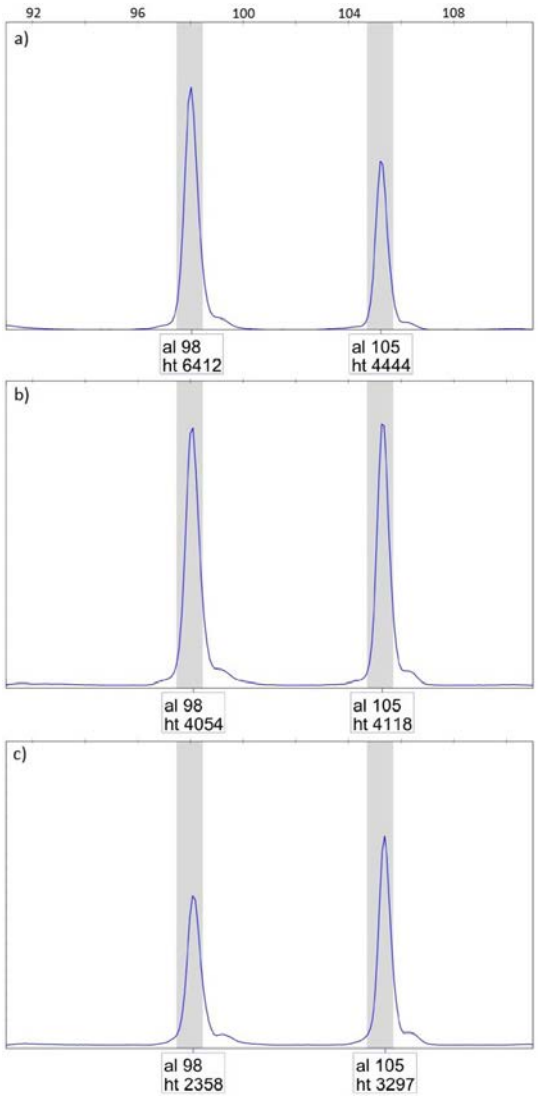


Figure 2.6. Electropherograms demonstrating variation in heterozygote balance for the sexing locus ZF1. Typically, heterozygotes display a lower peak height for the longer allele, for example (a), however, average heterozygote balance was 1.0, as (b), and could be imbalanced such that the longer allele has a higher peak (c). Size in bp is shown above the graphs. Grey bars indicate allele bins for *D. bicornis.*

Stutter peaks were observed for all STR markers at one to several repeats fewer than the true allele. Stutter peaks occurred as typical for dinucleotide STRs, ranging from one repeat shorter than the true allele (-1R) to several peaks shorter, with peak height decreasing with distance from the true allele. Mean stutter ratio ranged from 0.15 (DB23) to 0.49 (BlRh1C) (Table 2.5). The stutter ratio increased with the number of repeat units within the allele (Fig. 2.7). The sexing locus, ZF1, did not produce stutter.

A number of individual alleles displayed high stutter ratios (Fig. 2.8), and 13 alleles across six different loci had mean stutter ratios between 0.5 and 0.7. IR12 alleles 25 and 26 produced large stutter peaks (Fig. 2.8). Allele 25 had a mean stutter ratio of 0.65, however, this allele was only observed on three occasions in a single individual in this reproducibility study for which it was the longer of the two alleles. Allele 26 produced the highest stutter ratio of any locus and had a mean stutter ratio of 0.70. This allele was observed on three occasions in two individuals, once as a homozygote (mean stutter ratio = 0.69) and once as the longer heterozygous allele (mean stutter ratio = 0.70). All other stutter ratios observed for this locus were below 0.60. Typically, stutter was larger for the longer alleles (Table 2.6).

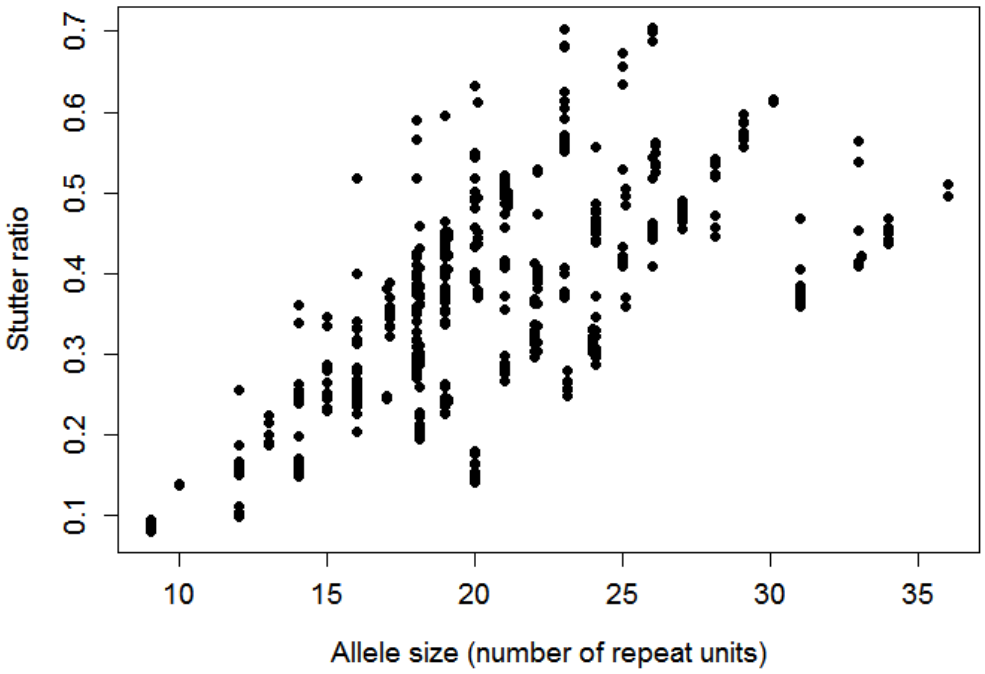
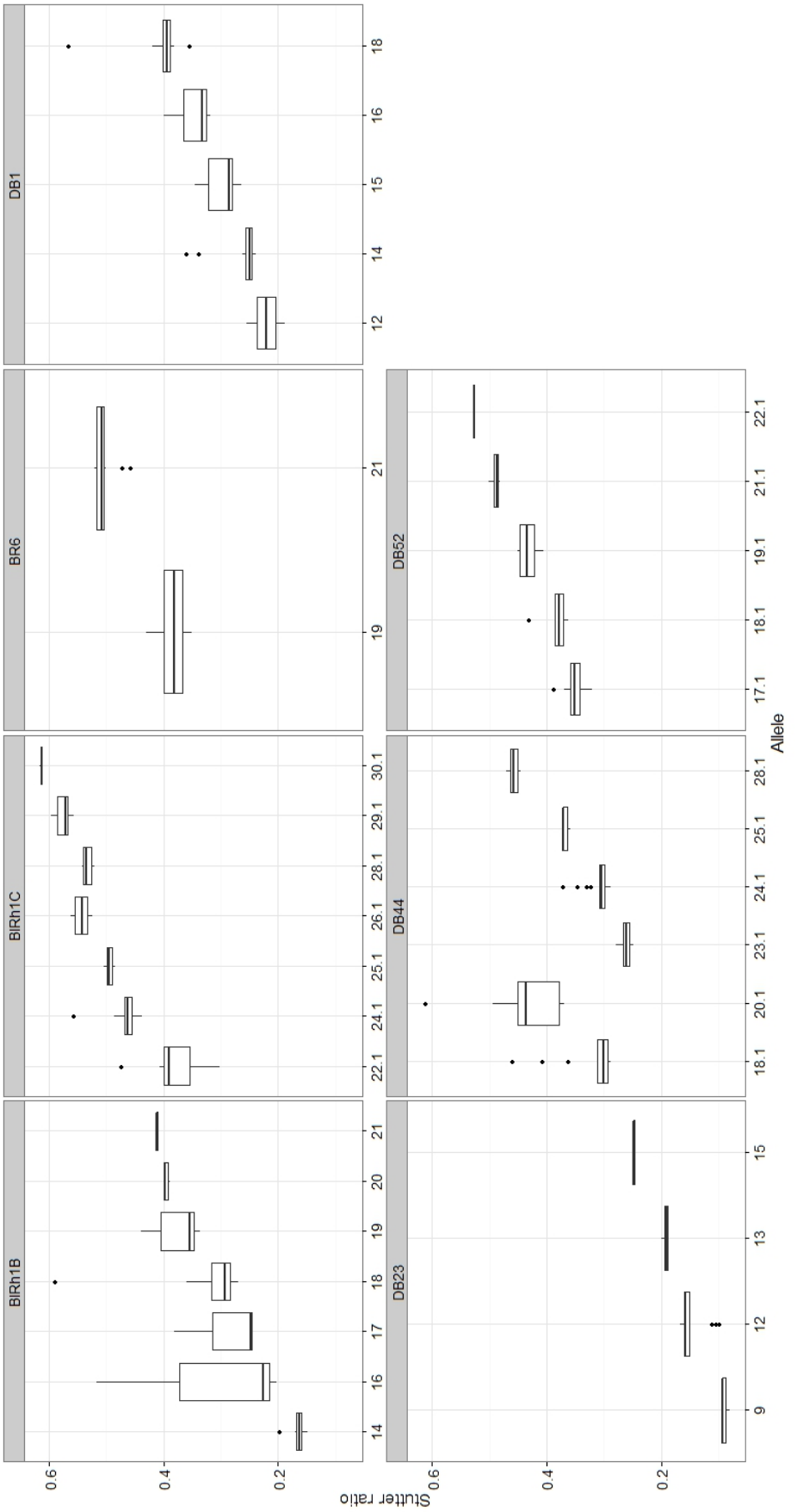


Figure 2.7. Stutter ratio increased with allele size across all loci, measured as the number of repeat units within each allele.


BIRh1B
BIRh1C
BR6
DB1
DB23
DB44
DB52
Stutter ratio
0.2
0.4
0.6
Allele

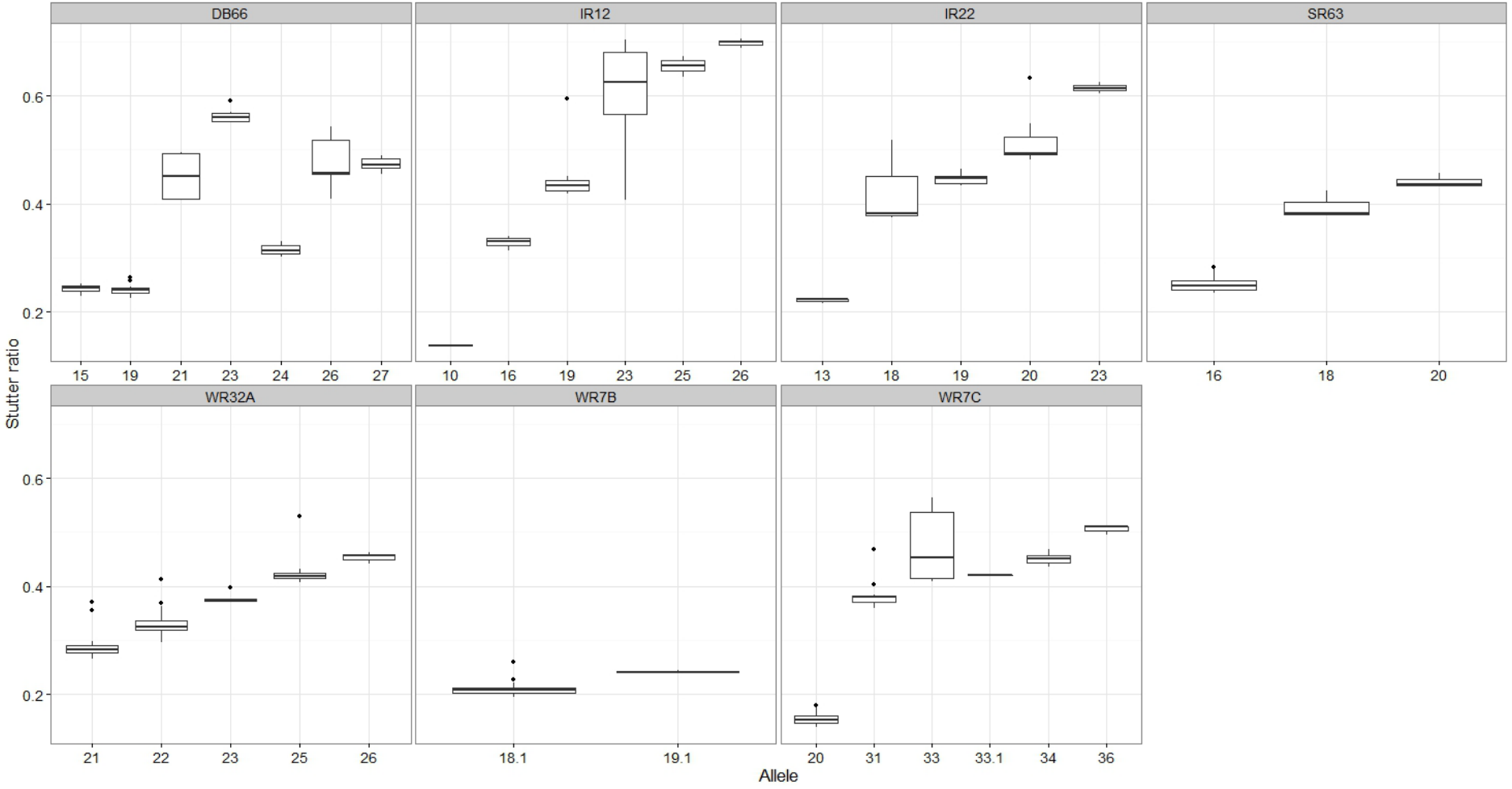


Figure 2.8**.** Box plots showing stutter ratios for all markers. Dark lines represent the mean for each locus and the boxes represent the first and third quartiles. Whiskers illustrate the limits of the data, to a limit of 1.5 times the interquartile range. Circles represent outliers.

Table 2.6. Mean stutter ratios for the short and long alleles individually for each marker. Stutter ratios were typically larger for the longer allele. Stutter could not be calculated for the long allele for WR7B as all heterozygotes carry consecutive alleles, masking the stutter peak of the long allele.

Locus	Mean stutter short allele	Mean stutter long allele
BlRh1B	0.401	0.379
BlRh1C	0.562	0.550
BlRh37D	0.206	0.312
BR6	0.401	0.513
DB1	0.300	0.395
DB23	0.139	0.184
DB44	0.314	0.375
DB52	0.371	0.487
DB66	0.300	0.477
IR12	0.433	0.676
IR22	0.556	0.529
SR63	0.333	0.418
WR32A	0.329	0.423
WR7B	0.212	N/A
WR7C	0.276	0.441

2.3.1.6 Sensitivity

The sensitivity of the method varied by locus (Table 2.7) and sample (Fig. 2.9). Three loci, BlRh1B, BlRh37D and ZF1, amplified all five samples at the lowest template concentration of 0.15 ng/µL. WR7C performed particularly poorly at low template concentrations, and dilutions below 5 ng/µL did not generate products that could be scored for one of the five samples used in the sensitivity study (see also Appendix III). Average minimum template concentration, however, was below 1 ng/µL for all loci except WR7C (Table 2.7).

Heterozygote imbalance (where the shorter allele has a lower peak height than the longer allele) occurred occasionally in nine of the 16 loci. Peak imbalance greater than 10% occurred in 6.5% of heterozygous amplifications, such that the shorter allele had a peak height at least 10% lower than the longer allele. On all occasions except one, this imbalance occurred during the sensitivity study with a stating template of 0.62 ng/µL or less. On no occasion, however, did heterozygote imbalance affect genotype scoring using the designated thresholds.

Table 2.7. Lowest template concentrations (ng/µL) with scored alleles are given for each locus for five individuals. Overall mean is shown, as well as mean excluding K14 as this sample showed increased sensitivity compared to other samples.

Sample	BlRh1B	BlRh1C	BlRh37D	BR6	DB1	DB23	DB44	DB52	DB66	IR12	IR22	SR63	WR32A	WR7B	WR7C	ZF1
KA	0.15	0.15	0.15	0.31	0.15	0.15	0.62	0.15	0.62	0.62	0.15	0.31	0.15	0.15	1.25	0.15
K06	0.15	0.15	0.15	0.31	0.15	0.15	0.31	0.15	0.31	0.62	0.15	0.31	0.15	0.15	2.5	0.15
K10	0.15	0.15	0.15	1.25	0.15	0.15	0.31	0.31	0.31	0.62	0.15	0.31	0.31	0.15	2.5	0.15
K14	0.15	0.62	0.15	1.25	0.15	0.31	1.25	0.62	2.5	1.25	0.15	1.25	0.31	0.31	5	0.31
K35	0.15	0.15	0.15	0.62	0.31	0.15	0.62	0.31	0.62	0.62	0.31	0.31	0.15	0.15	2.5	0.15
Mean	0.150	0.244	0.150	0.748	0.182	0.182	0.622	0.308	0.872	0.746	0.182	0.498	0.214	0.182	2.750	0.182
Mean (excl K14)	0.15	0.15	0.15	0.623	0.19	0.15	0.465	0.23	0.465	0.62	0.19	0.31	0.19	0.15	2.188	0.15

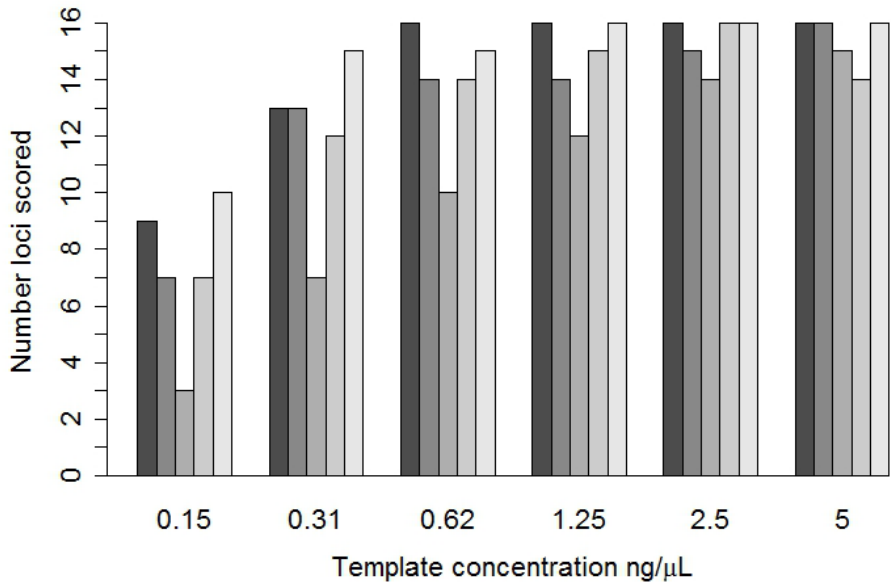


Figure 2.9. Number of loci that could be scored for five individuals at decreasing DNA template concentrations (ng/µL). Bar colours reflect sample in order of increasing lightness for KA, K06, K10, K14 and K35.

2.3.1.7 PCR-based procedures

Tests varying the thermocycling parameters demonstrated that the markers are robust to decreases in T_A and increased number of cycles; however, increasing T_A produced high failure rates (Table 2.8). At decreased T_A (–2 °C and –4 °C of standard) and increased cycle number (up to an addition 15 cycles above the standard), product yield was typically greater and thus required greater dilution before capillary electrophoresis, yet ease of allele scoring was not altered and no spurious peaks were observed. For loci with a standard T_A of 60 °C, two loci (WR32A and WR7C) demonstrated reduced amplification success at T_A +2 °C, while three loci (DB23, WR32A and WR7C) had reduced amplification at T_A +4 °C. Amplification success was greatly reduced above the normal T_A for loci where the standard T_A was 65 °C. All samples failed to amplify for IR12 and SR63 at T_A +2 °C, and the only locus to amplify at T_A +4 °C was IR22. Increased cycling, up to 43 cycles, had no effect on the scoring success of any locus except by increasing product yield, therefore warranting further dilution prior to electrophoresis.

2.3.1.8 Allele sequencing

DNA sequences from a minimum of two alleles of each of the 15 STR loci (48 alleles in total) confirmed that, with one exception, fragment length variation reflects variation in repeat unit number (Table 2.9). Nomenclature was assigned accordingly. A single intermediate allele was found at WR7C allele 33.2 (166 bp length). The repeat motif within this fragment contained an entire repeat unit (2 bp) fewer than the next consecutive allele (allele 34 – 167 bp length) although only a single base pair separates the alleles on capillary electrophoresis. Alleles were present at each 2 bp interval for this dinucleotide marker, including 165 bp and 167 bp, suggesting that allele 33.2 may have arisen due to microvariation in the flanking regions of the STR. Allele sequencing was unable to generate sequences of sufficiently high quality to detect such microvariation.

The variable region of the motif structures were the same as previously published for most loci. However, all loci were originally isolated from either a different species or subspecies of rhinoceros, and as such, some of motif structures differ from that originally described for the locus. BlRh1C was found to contain a variable GCAC insertion. DB44 was found to contain a C/G transition in one of three *D. b. michaeli* alleles sequenced. The repeat motif for SR63 was a compound dinucleotide repeat and included an additional TC repeat unit compared to that published for *Dicerorhinus sumatrensis* (Scott 2008). WR7B appears to contain a T/A transition and WR7C contains a T/C transition which interrupt the simple repeat motif identified in *C. simum* (Florescu *et al.* 2003).

Table 2.8. Amplification success for three individual samples (K10, K22 and K35) at each of the varied annealing temperatures. Samples were either successful (white cell), amplified but peak heights were below threshold RFU value (pale grey) or failed (dark grey).

Standard T_A	Locus	– 4°C			– 2°C			+ 2°C			+ 4°C		
		K10	K22	K35	K10	K22	K35	K10	K22	K35	K10	K22	K35
60°C	BlRh1C												
	DB23												
	DB52												
	DB66												
	WR32A												
	WR7B												
	WR7C												
	BlRh37D												
65°C	BlRh1B												
	BR6												
	DB1												
	DB44												
	IR12												
	IR22												
	SR63												
	ZFX/ZFY												

Table 2.9. Sequence data for all loci detailing repeat motif structure. A minimum of two alleles were sequenced per locus. Allele fragment length is given alongside the proposed nomenclature.

Locus	Repeat Motif	Fragment length (bp)	Nomenclature	Originally published motif	Species locus was derived from	Reference
BlRh1B	$(GT)_9$ GCA $(TG)_{3.1}$	236	14	$(GA)_{10}$	*D. b. minor*	Scott 2008
	$(GT)_{13}$ GCA $(TG)_{3.1}$	244	18			
	$(GT)_{16}$ GCA $(TG)_{3.1}$	250	21			
BlRh1C	$(AC)_{15}$ $(GCAC)_2$ $(AC)_{3.1}$	121	22.1	$(GT)_{13}$	*D. b. minor*	Scott 2008
	$(AC)_{19}$ $(GCAC)_1$ $(AC)_{3.1}$	125	24.1			
	$(AC)_{22}$ $(GCAC)_2$ $(AC)_{3.1}$	135	29.1			

Table 2.9. Continued.

Locus	Repeat Motif	Fragment length (bp)	Nomenclature	Originally published motif	Species locus was derived from	Reference
BlRh37D	$(TG)_6$ $(AG)_{11}$ GA $(AG)_{5.1}$	201	23.1	$(AG)_{17}$	*D. b. minor*	Scott 2008
	$(TG)_6$ $(AG)_{12}$ GA $(AG)_{5.1}$	203	24.1			
	$(TG)_6$ $(AG)_{14}$ GA $(AG)_{5.1}$	207	26.1			
BR6	$(CA)_{19}$	141	19	$(CA)_{15}$	*D. bicornis*	Cunningham *et al.*1999
	$(CA)_{20}$	143	20			
	$(CA)_{21}$	145	21			
DB1	$(CA)_{14}$	153	14	$(CA)_{14}$	*D. b. minor*	Brown and Houlden 1999
	$(CA)_{15}$	155	15			
	$(CA)_{18}$	161	18			
DB23	$(CA)_{12}$	247	12	$(CA)_{12}$	*D. b. minor*	Brown and Houlden 1999
	$(CA)_{13}$	249	13			
DB44	$(AC)_{6.1}$ C $(AC)_7$ G $(CA)_4$	206	18.1	$(CA)_4G(CA)_{16}$	*D. b. minor*	Brown and Houlden 1999
	$(AC)_{6.1}$ G $(AC)_{13}$ G $(CA)_4$	218	24.1			
	$(AC)_{6.1}$ G $(AC)_{14}$ G $(CA)_4$	220	25.1			
DB52	$(CA)_{17.1}$	211	17.1	$(CA)_{21}$	*D. b. minor*	Brown and Houlden 1999
	$(CA)_{18.1}$	213	18.1			
DB66	$(CA)_{15}$	182	15	$(CA)_7TA(CA)_{16}$	*D. b. minor*	Brown and Houlden 1999
	$(CA)_{15}$ T $(AC)_{8.1}$	200	24			
	$(CA)_{16}$ T $(AC)_{8.1}$	202	25			
	$(CA)_{18}$ T $(AC)_{8.1}$	206	27			
	$(CA)_{19}$ T $(AC)_{8.1}$	208	28			

Table 2.9. Continued.

Locus	Repeat Motif	Fragment length (bp)	Nomenclature	Originally published motif	Species locus was derived from	Reference
IR12	$(CA)_{10}$	157	10	$(CA)_{18}$	*R. unicornis*	Scott 2008
	$(CA)_{25}$	186	25			
	$(CA)_{26}$	188	26			
IR22	$(CA)_{13}$	210	13	$(CA)_{22}$	*R. unicornis*	Scott 2008
	$(CA)_{19}$	221	19			
	$(CA)_{20}$	223	20			
SR63	$(AC)_{13}(TC)_3$	193	16	$(AC)_{19}$	*Dicerorhinus sumatrensis*	Scott 2008
	$(AC)_{16}(TC)_2$	196	18			
	$(AC)_{18}(TC)_2$	200	20			
WR32A	$(AC)_6$ CCCCATACGCAA $(AC)_{15.1}$	197	21	$(CA)_{14}$	*C. s. simum*	Florescu *et al.* 2001
	$(AC)_6$ CCCCATACGCAA $(AC)_{16.1}$	199	22			
	$(AC)_6$ CCCCATACGCAA $(AC)_{19.1}$	205	25			
WR7B	$(TG)_{12}$ A $(GT)_6$	223	18.1	$(TG)_{16}$	*C. s. simum*	Florescu *et al.* 2001
	$(TG)_{13}$ A $(GT)_6$	225	19.1			
WR7C	$(TC)_7$ CC $(TC)_4$ $(TA)_8$	137	20	$(CT)_{14}(AT)_{11}$	*C. s. simum*	Florescu *et al.* 2001
	$(TC)_7$ CC $(TC)_{17}$ $(TA)_6$	161	31			
	$(TC)_7$ CC $(TC)_{17}$ $(TA)_7$	163	32			
	$(TC)_7$ CC $(TC)_{14}$ $(TA)_{11}$	166	33.2			
	$(TC)_7$ CC $(TC)_{16}$ $(TA)_{10}$	167	34			

2.3.2 Population studies

2.3.2.1 Multiplexes

Both multiplexes (Table 2.2) successfully amplified all 13 samples included in the reproducibility study (Fig. 2.10). DNA profiles generated in multiplex were identical to those generated under singleplex conditions. Full profiles were also successfully generated from a single horn sample of each of *D. bicornis* and *C. simum*. No genotyping errors were detected upon re-amplification and independent analysis of 10 samples.

2.3.2.2 Genetic profiles

Profiles were generated for 52 *D. b. michaeli* individuals (24 females, 32 males) from across Kenya (Appendix IV). Allele frequency data is given in Table 2.10. Across all individuals, 98.5% of loci amplified to generate genotypes that could be scored; however two loci, BlRh37D and WR7C, exhibited higher failure rates than other loci (Appendix IV). A full summary of allelic frequencies for loci as well as repeat structures are shown in Appendix V.

Following population genetic tests to determine the suitability of the markers for individual identification, two loci, BlRh37D and IR12 were deemed unsuitable. These loci both deviated from HWE after sequential Bonferroni correction, BlRh37D (p=0.0002) and IR12 (p<0.0001) (Table 2.11). An excess of homozygotes was detected by Micro-checker (Van Oosterhout *et al.* 2004) for the same two loci, although no scoring errors were evident. BlRh37D also had increased failure rate compared to other loci (Appendix IV), and five individuals failed to amplify despite repeated amplification attempts. IR12 was found to be in linkage disequilibrium with the sexing marker ZF1 (p=0.00002). Full details of linkage disequilibrium results are shown in Appendix VI. Further analysis of IR12 revealed that all male individuals were homozygous for this locus, whilst females contained a mixture of homozygous and heterozygous individuals, supporting the conclusion that IR12 is sex-linked in *D. b. michaeli*. STR loci BlRh37D and IR12 were therefore found to deviate from two assumptions of the RMP calculation, HWE and linkage equilibrium, and were thus excluded from further calculations.

The remaining 14 loci were found to be in HWE and all locus combinations were found to be in linkage equilibrium, and could thus be deemed suitable for inclusion in RMP calculations. For these final 14 loci, mean number of alleles was 5.93 (range 2-8), mean heterozygosity was 0.663 (±SE 0.051), mean observed heterozygosity was 0.638 (±0.048) and mean PIC was 0.623 (±SE 0.050) (Table 2.11).

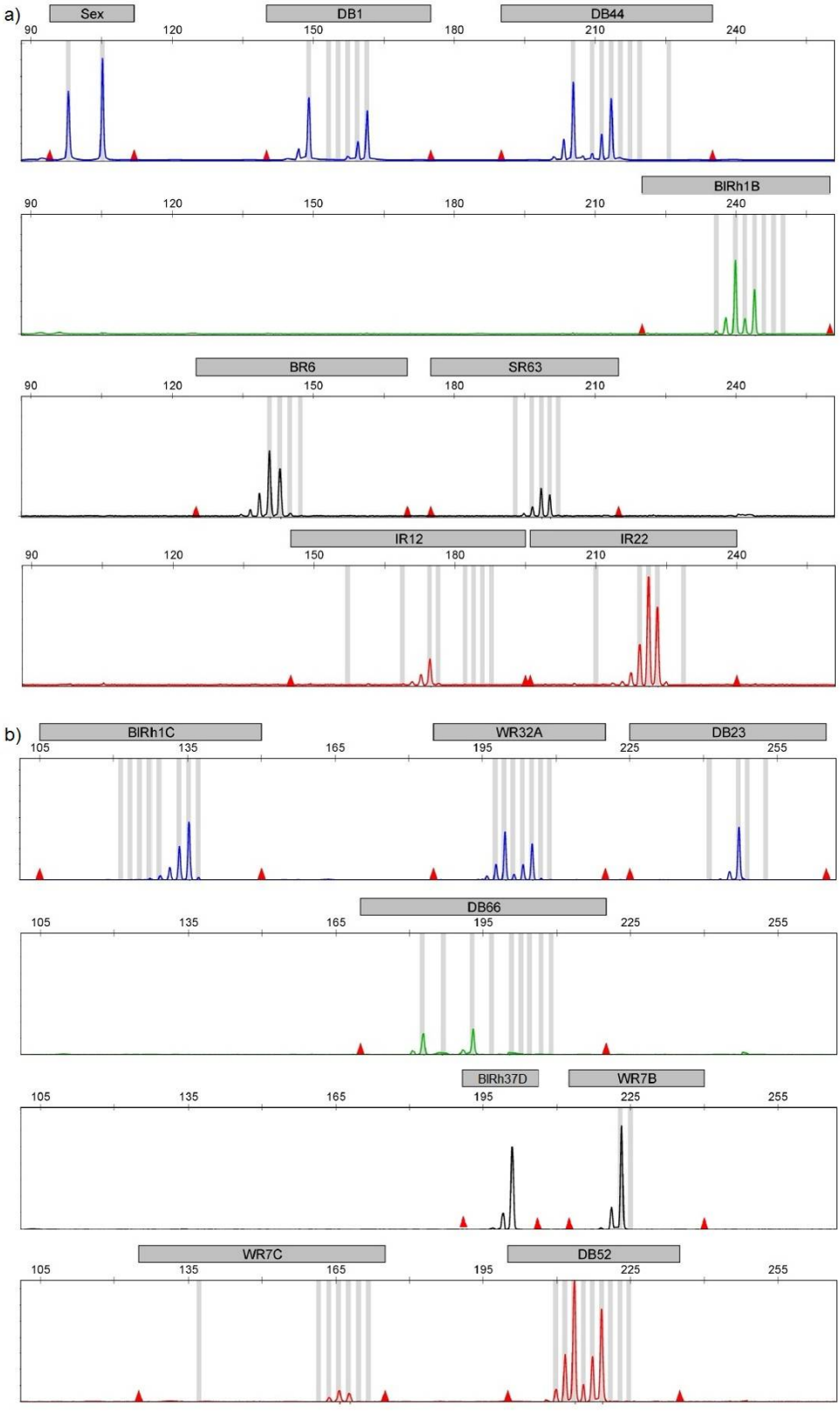


Figure 2.10. Electropherograms displaying typical genotypes produced using multiplexes at 65°C (a) and 60°C (b). Fluorescently labelled PCR products were run in the same lane on the 3130xL, and results were visually separated by GeneMapper into four panels by colour. 6FAM is blue, VIC is green, NED is black and PET is red. Size in bp is shown above the graphs. Locus name is indicated above peaks and vertical grey bars indicate allele bins for D. b. michaeli.

Table 2.10. Allele frequency data for 52 Kenyan *D. b. michaeli* individuals.

Allele	BlRh1B (N:55)	BlRh1C (N:56)	BlRh37D (N:51)	BR6 (N:56)	DB1 (N:56)	DB23 (N:56)	DB44 (N:55)	DB52 (N:56)	DB66 (N:54)	IR12 (N:55)	IR22 (N:56)	SR63 (N:54)	WR32A (N:56)	WR7B (N:56)	WR7C (N:49)
10	–	–	–	–	–	–	–	–	–	0.109	–	–	–	–	–
11	–	–	–	–	–	0.080	–	–	–	–	–	–	–	–	–
12	–	–	–	–	0.009	0.554	–	–	–	–	–	–	–	–	–
13	–	–	–	–	–	0.339	–	–	–	–	0.054	–	–	–	–
14	0.127	–	–	–	0.313	–	–	–	–	–	–	–	–	–	–
15	–	–	–	–	0.196	0.027	–	–	0.157	–	–	–	–	–	–
15.1	–	–	–	–	–	–	–	0.009	–	–	–	–	–	–	–
16	0.064	–	–	–	0.134	–	–	–	–	0.018	–	0.556	–	–	–
16	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
17	0.018	–	–	–	0.036	–	–	–	–	–	–	–	–	–	–
17.1	–	–	–	–	–	–	–	0.277	–	–	–	–	–	–	–
18	0.409	–	–	–	0.313	–	–	–	–	–	0.045	0.120	–	–	–
18.1	–	–	–	–	–	–	0.191	0.250	–	–	–	–	–	0.902	–
19	0.164	–	–	0.188	–	–	–	–	–	0.500	0.223	0.019	–	–	–
19.1	–	–	–	–	–	–	–	0.071	–	–	–	–	–	0.098	–
20	0.109	–	–	0.357	–	–	–	–	0.176	0.045	0.670	0.259	–	–	–
20.1	–	–	–	–	–	–	0.009	0.116	–	–	–	–	–	–	0.306
21	0.109	–	–	0.420	–	–	–	–	–	–	–	0.046	0.223	–	–
21.1	–	–	–	–	–	–	0.027	0.170	–	–	–	–	–	–	–
22	–	–	–	0.036	–	–	–	–	0.102	–	–	–	0.232	–	–
22.1	–	0.116	–	–	–	–	0.064	0.036	–	–	–	–	–	–	–
23	–	–	–	–	–	–	–	–	–	0.073	0.009	–	0.098	–	–
23.1	–	0.196	0.735	–	–	–	0.055	0.036	–	–	–	–	–	–	–
24	–	–	–	–	–	–	–	–	0.148	0.064	–	–	0.098	–	–
24.1	–	0.268	0.108	–	–	–	0.409	0.036	–	–	–	–	–	–	–
25	–	–	–	–	–	–	–	–	0.111	0.073	–	–	0.205	–	–
25.1	–	0.116	–	–	–	–	0.227	–	–	–	–	–	–	–	–

Table 2.10. Continue

Allele	BlRh1B (N:55)	BlRh1C (N:56)	BlRh37D (N:51)	BR6 (N:56)	DB1 (N:56)	DB23 (N:56)	DB44 (N:55)	DB52 (N:56)	DB66 (N:54)	IR12 (N:55)	IR22 (N:56)	SR63 (N:54)	WR32A (N:56)	WR7B (N:56)	WR7C (N:49)
26	–	–	–	–	–	–	–	–	0.028	0.118	–	–	0.063	–	–
26.1	–	0.098	0.157	–	–	–	–	–	–	–	–	–	–	–	–
27	–	–	–	–	–	–	–	–	0.102	–	–	–	0.080	–	–
27	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
28	–	–	–	–	–	–	–	–	0.176	–	–	–	–	–	–
29	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
29.1	–	0.161	–	–	–	–	–	–	–	–	–	–	–	–	–
30	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
30.1	–	0.027	–	–	–	–	–	–	–	–	–	–	–	–	–
31	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.214
32	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.082
33	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.204
33.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.031
34	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.133
35	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.020
36	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.010

Table 2.11. Descriptive statistics of all loci, including number of alleles in *D. b. michaeli* (Na), expected heterozygosity (H_E), observed heterozygosity (H_O), polymorphism information content (PIC) and probability of deviation from HWE (p).

Locus	Na	H_E	H_O	PIC	HWE (*p*)[a]
BlRh1B	7	0.761	0.709	0.734	0.1631
BlRh1C	8	0.826	0.768	0.804	0.0214
BlRh37D	3	0.423	0.216	0.383	**0.0001**
BR6	4	0.660	0.661	0.593	0.2220
DB1	6	0.747	0.696	0.704	0.1831
DB23	4	0.571	0.429	0.495	0.0217
DB44	8	0.735	0.768	0.697	0.4498
DB52	9	0.810	0.804	0.784	0.1501
DB66	8	0.858	0.815	0.841	0.0380
IR12	8	0.707	0.255	0.684	**0.0000**
IR22	5	0.497	0.429	0.447	0.0118
SR63	5	0.607	0.556	0.553	0.0125
WR32A	7	0.824	0.768	0.801	0.2366
WR7B	2	0.177	0.161	0.161	0.4189
WR7C	8	0.796	0.714	0.768	0.0988
ZFX/Y	2	0.408	0.571	0.325	

[a] Significant deviations from HWE following sequential Bonferroni correction are marked in bold.

2.3.2.3 Profile matching

PI_{ave} were modelled for the final 14 loci for the wild Kenyan population (Figure 2.11). Assuming an unrelated population (simulating no inbreeding), PI_{ave} ranged between 1.06 x 10^{-11} with minimal population structure (F_{ST} = 0.05) to 1.37 x 10^{-8} with reasonable population structure (F_{ST} = 0.2). With increased population relatedness to the level of cousins (simulating a higher level of inbreeding) PI_{ave} increased to between 3.13 x 10^{-10} (F_{ST} = 0.05) and 1.16 x 10^{-7} (F_{ST} = 0.2).

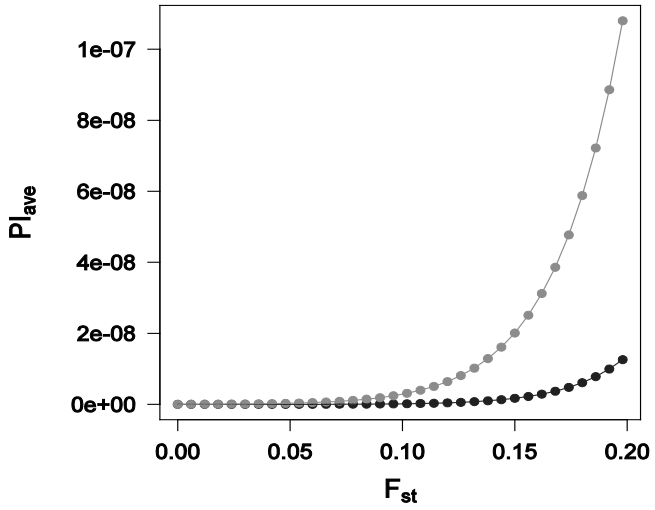


Figure 2.11. PI_{ave} values for an unrelated population (black) and a population related as cousins (grey) for varying levels of population structure (F_{st}).

Cumulative likelihood RMPs were calculated for the profiles generated from the four wild *D. b. michaeli* to test sensitivity of the loci (Table 2.12). The RMP 99% source attrition threshold was 7.46 x 10^{-4}. This threshold value was not met at DNA template concentrations below 0.62 ng/µL for individual K14 and at 0.15 ng/µL for K10 and K35. All other profiles, however, produced in the sensitivity study generated sufficient RMP to assign confidence that the profile is highly unlikely to occur more than once within the population.

Table 2.12. Match probabilities for four samples (θ=0.1) with varying numbers of scored loci as template DNA concentration increased. A maximum of 14 loci could be scored (number scored in parentheses). BlRh37D and IR12 were not included in probability calculations. Match probability for the fifth sample assessed in the sensitivity study was not included as it was a captive rhino and allele frequencies were unknown. Match probabilities highlighted in bold are below the 99% confidence threshold value.

Template concentration	K06	K10	K14	K35
0.15 ng/µL	8.68E+05 (9)	6.73E+04 (7)	1.61E+02 (4)	2.50E+04 (6)
0.31 ng/µL	4.69E+08 (12)	1.58E+09 (12)	5.46E+03 (7)	3.68E+09 (12)
0.62 ng/µL	6.66E+10 (14)		2.70E+06 (10)	1.95E+10 (13)
1.25 ng/µL			3.24E+07 (12)	
2.5 ng/µL		8.04E+09 (13)	4.57E+08 (13)	4.11E+11 (14)
5 ng/µL		5.35E+10 (14)	3.77E+09 (14)	

2.4 Discussion

With ever growing demand for rhinoceros horn, there is increasing need for forensic methods to individually identify them and trace their movements in trade. This study validates an individual profiling system for *D. bicornis* which could enable samples recovered in trade to be traced back to their source. The developmental validation carried out in this study followed guidelines as appropriate for wildlife (Budowle *et al.* 2005; Linacre *et al.* 2011) where possible for this critically endangered species. Thirteen STR markers and a sexing marker (zinc-finger protein) have passed this forensic validation and, provided allele frequencies from an appropriate population are available, can be used for testing in criminal casework.

2.4.1 Developmental Validation

2.4.1.1 Different starting materials

Previous studies have shown that both mitochondrial DNA (Hsieh *et al.* 2003; Peppin 2009) and nuclear DNA (Peppin 2009; Peppin *et al.* 2010; Harper *et al.* 2013) can be isolated from rhinoceros horns. This validation has demonstrated that this marker system produces consistent profiles from blood and horn and thus the profiles are comparable. The extraction protocol used here was, however, less than optimal. The method used in this validation used 100 mg of starting material and yielded both low quantity and low quality DNA. Obtaining horn material of this quantity may not always be feasible from historic horns which may be valuable scientifically or artistically and the physical alteration caused by drilling to recover large quantities of material for DNA analysis may be unacceptable. This is particularly likely to be true for mounted horns which remain attached to the skull. Furthermore, historic horns are more likely to contain degraded DNA which is typified by highly fragmented DNA (Pääbo *et al.* 2004). The extraction was therefore optimized to minimize horn material requirements and maximize DNA yield (Chapter 3).

2.4.1.2 Cross-amplification

Human and canine DNA are possibly the most likely to be encountered in a wildlife forensic laboratory, originating from both the analysts and their pets' hairs. Cross-amplification was observed from DNA of these two species at two loci, WR7B and ZF1. Amplification with WR7B generated strong PCR products for both human and canine DNA, which were larger and

smaller (respectively) that the observed *D. bicornis* amplicons. The sexing marker, ZF1, produced a strong ZFX amplicon for canine DNA and a weak amplicon for human DNA. Male DNA was not, however, amplified during this study, and therefore amplicon sizes are unknown.

ZF1 seems well suited to detect amplification of contaminant DNA. This marker is not an STR, but rather homologous zinc finger protein genes which show size polymorphism between males and females. Peppin *et al*. (2009) found that ZFX and ZFY fragment lengths generated for *C. simum* and *R. unicornis* were the same as found for *D. bicornis*. Given the phylogenetic distance between *R. unicornis* and the African rhinoceros species (Tougard *et al.* 2001; Scott 2008), it seems likely that these same may also be true of ZF1 fragment sizes for other extant rhinoceros species. ZF1 alleles can be assumed to be fixed across rhinoceros, although this should be empirically tested, and therefore it is likely that any observed amplicons with fragment sizes that differ from the internally validated ZFX and ZFY fragments for *D. bicornis* originate from DNA of a non-target species. Human and canine amplicons produced by WR7B were distinct from the two alleles that have been observed in *D. b. michaeli*. The human amplicon was three bp larger and the canine amplicon 11 bp shorter than the largest and shortest *D. b. michaeli* amplicons. At least one additional allele has been observed in other subspecies of *D. bicornis* and a second in *C. simum* (Harper *et al.* 2013), therefore it is entirely possible that both the human and canine amplicons could overlap with those from rhinoceros, and therefore detection of contamination at this locus may not be possible.

No other species were tested for cross-amplification with the marker panel through this validation. It is clear that two markers, WR7B and ZF1, are capable of cross-amplifying mammalian DNA as demonstrated by positive amplification of human and canine DNA. ZF1 in particular could be predicted to amplify a wide range of species, and indeed alternative primer pairs for this gene have been utilised successfully to amplify an 800 – 1000 bp region for a range of mammalian species (Shaw *et al.* 2003). Other markers are expected to be rhinoceros specific, although further testing would be required to confirm this.

The ability to detect contamination may be particularly valuable when working with historic rhinoceros horns which may contain low quantity and quality DNA as such samples are more likely to be contaminated by higher quantity and/or quality DNA (Butler 2005). Adherence to quality control and quality assurance measures should minimise the likelihood of contamination, either from external sources or between samples. If any species is routinely processed within the same laboratories as rhinoceros casework (ether pre– or post-PCR),

then it would be prudent to test this marker panel for cross-contamination during the internal validation.

It cannot be excluded that non-target DNA could be present within the sample being analysed, for example in TAM or weapons used for poaching may contain traces of DNA from multiple species. This situation can normally be identified by carrying out species identification tests prior to individual identification tests. This marker set, however, should remain useful in such situations, as 12 of the 14 markers in the final marker panel are likely to be rhinoceros specific. Novel alleles detected under such circumstances could be indicative of non-specific amplification, particularly if present at multiple loci and in conjunction with the presence of non-rhinoceros amplicons at WR7B and ZF1. Such novel alleles would warrant further investigation.

It is hoped that this marker panel will be valuable for forensic analysis of all rhinoceros species. All primer pairs except BlRh37D cross-amplified DNA from *C. simum* in this study. A study to characterise STR markers in black and white rhinoceros by Harper *et al.* (2013) found that, using the same primer pairs, BlRh37D did amplify in *C. simum* although it was monomorphic. This variation in amplification success of BlRh37D in *C. simum* is likely to be due to null alleles, as was found for *D. bicornis* in this validation, which has not been detected by Harper *et al.* (2013). Three loci were found to be monomorphic in the *C. simum* individuals assessed. DB23 and IR22 were also found to be monomorphic in *C. simum* by Harper *et al.* (2013), although BR6 was polymorphic. The *C. simum* samples analysed in this study originated from captive rhinoceros, and it is likely that with only four individuals assessed, variation was not fully captured.

With the exception of BlRh37D, however, inter-species amplification of the remaining markers suggests this panel will be a useful forensic profiling tool for other rhinoceros species. Indeed a further seven markers were shown to amplify DNA from both *D. bicornis* and *C. simum* by (Harper *et al.* 2013). Pending full validation of the markers for forensic validation and assessment of inter-species amplification of DNA from Asian rhinoceros species, there is potential for a 21-marker panel for application to any rhinoceros horn.

Whilst this study has only assessed the marker panel in a single *D. bicornis* subspecies, the Kenyan population of *D. b. michaeli*, it is expected that the validated markers would be suitable for application to other subspecies. Indeed, all loci were shown to amplify in *D. b. minor* (Harper *et al.* 2013). STR analyses by (Harley *et al.* 2005) of the four subspecies currently recognised by the IUCN (Emslie 2012) reveal genetic differentiation between the

subspecies at five of the same loci included here. This would suggest that the subspecies have been separate for long enough for genetic drift to operate but not long enough for private alleles to differentiate subspecies. Furthermore, the same markers have been used in numerous populations of all subspecies in many studies (Brown & Houlden 1999; Cunningham *et al.* 1999; Scott 2008; van Coeverden de Groot *et al.* 2011; Muya *et al.* 2011; Harper *et al.* 2013). Therefore it is expected that this developmental validation is applicable to *D. bicornis* across its range, providing that the necessary population genetic analyses are carried out for new populations.

2.4.1.3 Reproducibility

The marker set was shown to be reproducible both during the PCR and capillary electrophoresis phases. D-values did not exceed 0.55 bp for any individual allelic amplification (see Appendix I for schematic explaining D-values). In other words, the largest drift observed between the smallest and largest observed fragments for any allele was 0.275 bp above and below the median. Thus demonstrating that between-run variation was within acceptable limits (Dawnay *et al.* 2008). Bins for allele designation can therefore be set accordingly with confidence that there is no overlap between consecutive alleles, even for alleles which differ by only one bp as the worst observed drift seen in this study would still separate these by at least 0.45bp. This marker system was shown to be very precise; however development of an allelic ladder would not only provide confidence in between-run precision but also greatly facilitate inter-laboratory calibration.

Tri- or tetranucleotide markers are preferred for forensic testing (Linacre *et al.* 2011), however only dinucleotide STRs have thus far been published for any rhinoceros species. Dinucleotide STRs in particular suffer from increased stutter which can make allele designation ambiguous (Shinde *et al.* 2003). The use of dinucleotide STRs is not ideal; however, lack of funding for such wildlife research prevents the use of next-generation sequencing technologies to search for STRs with larger repeat units. Furthermore, ISFG guidelines indicate that "the use of dinucleotide repeats in forensic genetics is not recommended, except for those markers that are already used widely in animal genetic studies" (Linacre *et al.* 2011), and all dinucleotide STRs assessed here are currently used in South Africa for identification of rhinoceros (Harper *et al.* 2013).

Stutter was observed to a maximum ratio of 63% of the peak height of the true allele (IR22), though means ranged between 0.21 (WR7C) and 0.49 (IR22). Stutter ratio increases in STRs

with shorter repeat units (dinucleotide > trinucleotide > tetranucleotide > pentanucleotide), and longer alleles have greater stutter (Murray *et al.* 1993; Ellegren 2004; Butler 2014). Stutter was observed to increase with allele length (see figure 2.7), and was greater in most cases than the standard 15% threshold applied to tetranucleotides used in human profiling panels (Butler 2014). Thresholds for dinucleotide STRs have not previously been published, and dinucleotide STRs are not used in human profiling systems. Stutter ratios observed here for *D. bicornis* are less than those observed by Andreassen *et al.* (2012), the only study to report stutter ratios for a dinucleotide STR. As such, the maximum stutter ratios observed for each marker during this reproducibility study could be considered as thresholds.

Heterozygote balance, such that the shorter allele exhibited weaker amplification, was rare with the exception of DB44. However, even for this locus, heterozygote balance did not occur below 0.89. This means that the shorter allele was never lower than 89% of the height of the longer allele, and even with a maximum stutter ratio of 61% for this locus, there was no overlap between the largest stutter and minimum heterozygote balance. For all other STR loci, the difference between maximum stutter and minimum heterozygote balance was greater than for DB44 (Figure 2.4), and it is therefore unlikely that the shorter allele could be mistaken for a stutter peak for any locus (Andreassen *et al.* 2012). The sex marker ZF1 did show heterozygote imbalance as low as 0.72, however no stutter was generated at this locus.

It is important to be able to distinguish between heterozygotes and homozygotes to generate a true genotype, and clear guidelines should be defined to avoid genotyping errors generated through mistyping caused by heterozygote imbalance. Large heterozygote imbalance in weak samples can cause a heterozygote to mistakenly be typed as a homozygote if the peak height of the shorter allele is greater than the stochastic threshold, but the peak height of the longer allele is below the analytical threshold (Andreassen *et al.* 2012). With one exception, all longer alleles observed during the reproducibility study were at least 50% of the height of the shorter allele (heterozygote balance of 2.0). Thus, the stochastic threshold should be at least two times the analytical threshold. Heterozygote imbalance was more apparent for amplifications from low DNA template concentrations and alleles differing by a large number of repeats as found for WR7C. Thus it may be prudent to set a higher stochastic threshold when working with potentially low DNA quality and quantity samples such as historic horns.

2.4.1.4 Sensitivity

A number of loci in the marker system tested were very sensitive to low template DNA concentration. However, variation in sensitivity between samples was also evident (Table 2.6). While variation in DNA quantity and quality is not unusual between samples, sample K14 showed lower amplification success than other samples at concentrations below 5 ng/µl. Inaccuracies during the serial dilution of K14 may have resulted in weaker dilutions than expected. Alternatively, inhibitors could be present in the extract which could reduce amplification success. If this were true, variability between loci remains particularly evident as the dilution of K14 expected to be 0.15 ng/µL was amplified by some loci but not others. Sensitivity does not appear to be purely correlated with fragment length, even for sample K14, as might be expected as primers typically preferentially amplify shorter fragments. Previous studies have shown that inhibitors that bind to DNA may affect alleles of all sizes (Mccord *et al.* 2011). It seems unlikely that inhibitors are present within only sample K14 as extraction methods were identical across samples. Furthermore, although the minimum template concentration at which K14 produced genotypes that could be scored was greater than for other samples, it was only increased slightly (Appendix VI). Therefore, the mean minimum template concentration including K14 was rarely increased much above that without, and for no locus was it doubled. It is most likely therefore that the dilution series were not made equally because samples were not sufficiently homogenized at each stage and thus actual DNA concentrations varied.

Markers BR6, DB66 and WR7C showed greatest mean sensitivity to template concentration. Excluding sample K14, all samples could be scored at template concentrations of 0.62 ng/µL or less for all markers except sample K10 for BR6 and all samples for marker WR7C, which showed greatly increased sensitivity to template concentration. Marker BR6 was originally isolated from and primers designed for southern African *D. bicornis* (subspecies not stated but does not include *D. b. michaeli*) (Cunningham *et al.* 1999). WR7C was isolated from *C. simum* (Florescu *et al.* 2003) and primers were designed using sequence data from the same species (Nielsen *et al.* 2008). Whilst it is clear that cross-species amplification is successful, it may be that microvariation has reduced the binding affinity of the primers in non-target species or subspecies, which has generated greater sensitivity to template concentration. Nonetheless, many loci can be scored at very low template concentrations which are much lower than is typical from rhinoceros horn extractions (Peppin 2009; Harper *et al.* 2013; this study). Furthermore, this developmental validation has clearly demonstrated that an experienced analyst should have no difficulties in scoring genotypes and determining

whether or not a homozygote is truly homozygous even when using low template concentrations if applying the appropriate stochastic and analytical thresholds.

Understanding the sensitivity of the marker set will be valuable in its application to historic horn samples. Horn samples were found to yield lower DNA quantities than either blood or tissue samples, and degradation of DNA in historic samples could reduce its quality. Amplification of low copy number DNA can result in sporadic contamination, allelic dropout and excessive stutter amplification (Gill 2001b). While none of these stochastic effects were observed during this validation study, it is important to be aware that these could be more likely when amplifying DNA from historic samples. For loci which are sensitive to template concentration and for which genotypes cannot be scored because PCR products are too weak, increased cycling during amplification may help to generate stronger products.

2.4.1.5 PCR conditions

Increased annealing temperatures resulted in greatly reduced amplification success, the marker set was robust to reduced annealing temperatures and increased cycle number. For loci with a standard T_A of 65 °C, an increase of 4 °C caused amplification failure in all but one locus (IR22), whilst an increase of 2°C caused amplification failure or weak product yield in a single individual for all but two loci (IR22 and ZF1) and for all samples at two loci only (IR12 and SR63). This suggests that above 65 °C, PCR conditions are too restrictive and a high failure rate can be expected. For loci with a standard T_A of 60 °C, DB66 WR7C showed increased amplification failure at increased T_A, but other loci were successfully able to amplify at increased of both 2 °C and 4 °C. Maintaining a high T_A is advisable in forensic systems, especially when low copy number samples are processed, to avoid primers binding to non-target DNA. Therefore maintaining annealing temperatures at 60 °C and 65 °C would balance the need to avoid non-specific binding whilst maximising amplification success.

The marker system was shown to be robust to reduced annealing temperatures and increased cycle numbers. At lower temperatures, binding specificity of primers is relaxed and primers are able to bind to DNA more easily. Therefore reduced annealing temperature can improve the product yield if a PCR is not working effectively, but it can also allow unwanted non-specific PCR products to be amplified which complicates analysis. This was not observed throughout this study, suggesting that reduced annealing temperatures could be used if required. This could benefit analysis of historic horn samples which fail to amplify under standard conditions or which yield low quantity PCR products. This study found that

increased cycling does not generate any spurious results and should not negatively affect the analysis. Although it is worth bearing in mind that increasing cycling parameters would increase the likelihood of amplifying non-target DNA; therefore if mixed or contaminated samples are suspected, analysis could become ambiguous or challenging. Increased cycling parameters should be used with caution.

2.4.1.6 Allelic sequencing

Allele sequencing showed that, with one exception (WR7C allele 33.2), fragment length reflected repeat unit number, and nomenclature is proposed accordingly (Table 2.8). Use of a standardised nomenclature means that allele designation does not rely upon observed fragment length, which can vary between laboratories due to a number of factors. Variation can be observed between individual genome analysers of the same model. This can cause difficulties when comparing profiles generated at two different laboratories. For example, a dinucleotide locus genotyped at one laboratory could produce alleles A and B at 156 and 158 bp and at another laboratory these may be measured as 157 and 159 bp and at a third 155 and 157 bp. The two alleles measured as 157 bp would be assumed to be the same but in fact they are actually two different alleles that have drifted possibly because of variation between genome analysers and electrophoresis protocols. This difficulty can be overcome by ensuring that all laboratories comparing data have analysed standardised samples and therefore calibrated their data. This processed is simplified if those alleles are described using a nomenclature that cannot be confused with the fragment size.

There are two options which are typically employed for allelic nomenclature: alphabetic or the number of repeat units within the allele. Alphabetic nomenclature is useful when the repeat motifs of alleles are unknown, but this system struggles to cope when new alleles are discovered after nomenclature is assigned (Van De Goor *et al.* 2010). Most forensic systems use nomenclature that reflects the number of repeat units within each allele (Bär *et al.* 1997; Gill *et al.* 1997; Hellmann *et al.* 2006), which is the system adopted here. This system allows the assignment of new alleles whenever they are discovered as well as allowing for description of partially complete repeat motifs. Allelic sequencing is, however, required in order to assign nomenclature.

Quality of the sequences generated here was sufficient to determine the number of repeat motifs for all loci, although variation in flanking regions could not be reliably assessed. Across all alleles within a locus, no discrepancies were observed between motif repeat number and

genotyping fragment length and thus the nomenclature system for each locus accurately reflects the variation within it. Producing high quality sequences of microsatellites is very challenging, in large part due to stutter products generated during the PCR amplification which typically present as mixed sequences at the 3' end of the repeat motif. Whilst cloning may improve the quality of the sequences produced, both the true allele and stutter fragments will be cloned thus making it difficult to determine the true repeat length. Whether or not repeat motif sizes can be accurately determined, a best estimate for all alleles can be used to assign allelic nomenclature. In this instance, the nomenclature could be inaccurate if complicated sequence traces have been interpreted incorrectly. However, ultimately its function does not change. The nomenclature essentially becomes an arbitrary name for each allele. Calibration between laboratories cannot rely solely on sequencing of alleles to determine motif size and therefore which fragments correspond to which allele. It must include cross-calibration of a subset of samples, and, due to the wide range of variables that can affect genotype scoring between laboratories, it is essential that inter-laboratory calibration of the marker systems includes genotyping of a range of samples at both laboratories in order to standardise genotype scoring. This would be further aided by the development of an allelic ladder.

Variation in two different repeat regions within an imperfect STR, as found in BlRh1C, SR63 and WR7C, is not ideal as multiple alleles could be represented by a single fragment size. Although some alleles are therefore cryptic to analysers, profiles generated under these circumstances are not inaccurate, but rather have a conservative effect on RMP calculations by not discriminating between these alleles. This reduces the discriminatory power of the locus. It seems typical, however that loci with such alleles are highly variable and therefore are likely to be sufficiently informative to remain valuable in the marker panel.

2.4.2 Population studies

Population data presented here were generated for the primary purpose of assessing the suitability of the markers rather than generating full allele frequency data for the wild Kenyan black rhinoceros population. A representative sample from a human population is ideally from a minimum of 200 individuals (ISFG guidelines - Linacre *et al.* 2011), however for wildlife this sort of number is not always possible. An effective allele frequency database should include representative samples (e.g. 30-50 unrelated individuals) from as many populations as possible. The samples utilised in this study were from reserves across Kenya but are not a

fully representative cross-section of the metapopulation (Ashley *et al.* 1990). More extensive allele frequency distribution data should be generated before application in criminal casework. The population analysis did, however, reveal possible null alleles at BlRh37D and linkage disequilibrium between IR12 and ZF1. An alternative primer pair for locus BlRh37D may improve primer binding and remove the problem of null alleles, however genotypes generated using the primer pair tested in this study were not included in further forensic or population genetic studies of *D. b. michaeli*. Deviation from HWE for IR12 was shown to be caused by linkage disequilibrium with the sexing marker ZF1, and as such all males were homozygous. Both BlRh37D and IR12 have been proposed by Harper *et al.* (2013) for use in forensic identification of *D. bicornis* and *C. simum*. Whilst it is possible that null alleles and sex-linkage of these markers do not occur in white rhinoceros or subspecies of black rhinoceros not covered our study, relevant population genetic data from these markers, should be analysed prior to use in forensic casework to ensure they do not deviate from HWE, or that such inheritance patterns are accounted for in the RMP calculation.

Except IR12, all loci pairs were found to be in linkage equilibrium despite a number of markers being apparently located on the same chromosome on the *E. caballus* genome. *E. caballus* has 64 chromosomes whilst *D. bicornis* has 84 chromosomes, thus it is possible that the loci are located on different chromosomes in *D. bicornis*. The data presented illustrates that the remaining loci are sufficiently separated to segregate independently in this species. This is a requirement for the calculation of random match probabilities (Butler 2005; Dawnay *et al.* 2008), which are invaluable to interpret the power of a match between profiles recovered from different locations (Weir 2012). The remaining 14 markers (13 STR and sexing marker) are therefore recommended for individual identification of *D. bicornis*.

The 13 STRs and the sexing marker can be amplified in two multiplexes of seven primers each. These multiplexes have been combined in such a way to avoid overlap of allele size ranges for *D. b. michaeli* even if new alleles are found. Preliminary tests (data not shown) demonstrate that these multiplexes are also effective for other subspecies of *D. bicornis*, although they may need to be re-combined for other species of rhinoceros. The multiplexes described here have not, however, been validated for forensic use and further testing of the multiplex system would be required for use in casework as the complexity of the mixture may affect the sensitivity and robustness of some markers. They may, however, prove valuable to reduce costs and time to generate population data to estimate allele frequency distributions as well as for use in conservation genetic research. No genotyping errors were detected upon re-amplification and analysis of 10 samples in multiplex. This is a good

indication that the multiplex is not only reproducible but that genotype scoring is unambiguous.

2.4.2.1 RMP calculations

The number of loci profiled is a key determinant of the RMP value (Foreman & Evett 2001). Thus, the discriminatory power could be reduced in low DNA quantity samples. Of the four samples in the sensitivity study for which RMPs were calculated, two samples produced profiles with enough loci to generate sufficient match probabilities to reach the RMP 99 % threshold when just 0.15 ng of DNA was added per amplification (Table 2.11). At the other extreme, one sample required 0.62 ng of DNA to be added per amplification to generate a sufficient DNA profile to meet the 99 % RMP threshold. In spite of the variation in sensitivity between samples, it is important to note that partial profiles remain valuable in reducing the pool of potential matches.

2.4.3 Conclusions

This developmental validation has shown the final marker set to be robust, reliable and reproducible in simplex, and PI_{ave} values and RMP calculations demonstrate its high discriminatory power. Generation of further allele frequency information is necessary for all *D. bicornis* subspecies for inclusion in RMP calculations. This STR profiling system has the potential to be a useful forensic tool for DNA analysis of *D. bicornis* samples in criminal casework.

Chapter 3 – Optimisation of DNA extraction from historic horns

3.1 Introduction

3.1.1 Historical collections of rhinoceros horns

Both Western and Eastern cultures have long traditions of collecting rhinoceros horns. They have been used artistically in Asian cultures for thousands of years; decoratively engraved cups, bowls and other ornaments have been produced from rhinoceros horns since the Tang dynasty (c. 618-907 AD) (Martin & Martin 1982). African rhinoceros have been primarily hunted for sport by Europeans, largely throughout the 18th to 20th centuries (Martin & Martin 1982; Emslie & Brooks 1999). Their horns made popular trophies and decorative carvings (Martin & Martin 1982). Many trophies remained in range states, displayed in both private homes and public venues such as country clubs, public houses and hotels, and other trophies were exported as their collectors returned home, many to Europe. Thousands of rhinoceros horns were brought back to the UK. Many can still be found in stalely homes and private collections, and many were donated to museums and scientific collections. A number of historic Asian cups and bowls can also be found in the UK in antique and artwork collections.

3.1.2 The changing state in horn sources

The primary source of rhinoceros horns for trade has always been live animals (Leader-Williams 1992). Demand for horn has been detrimental to all species of rhinoceros, and indeed caused the near collapse of both the white rhinoceros in the early 20th century (Rookmaaker 2000) and the black rhinoceros later that the same century (Emslie & Brooks 1999) (see section 1.1). In the mid to late 20th century, demand for rhinoceros horn stemmed from the Yemeni jambiyas and East Asian TAM. Rhinoceros were numerous across unprotected African savannahs at this time, although their numbers were reduced across Asia. Prior to the CITES agreement to restrict trade in rhinoceros products, exportation from Africa was cheap, easy and unrestricted (Martin *et al.* 1997). Even after the ban had been implemented, its enforcement was somewhat limited across many African range states, although its import into Yemen became increasingly restricted (Martin *et al.* 1997)**.** There was therefore no need for traders of rhinoceros horn to look for a horn source anywhere but live rhinoceros.

The recent resurgence in the use of rhinoceros horn in modern forms of TAM has increased demand dramatically and the retail value of horns. Supply of horns is somewhat impeded by the trade ban regulated through CITES (Cites 2011), under which imports and exports are illegal therefore making movement of illegal horns difficult. In contrast with the last peak in the horn trade of the mid-late 20th century, live rhinoceroses as sources of horns are now diminished, and demand remains greater than supply. Although live rhino remain the primary source of horns entering the illegal trade, the challenges faced in poaching, as well as the very high value of horns has enticed some to seek other sources, including private and state-owned collections.

The retail price of rhinoceros horns has increased from approximately US$4 700/KG in 1993 to around US$50 000/KG in 2009 (Graham-Rowe 2011), and the value per unit weight has since become greater than diamonds or cocaine (Biggs *et al.* 2013). This massive rise in the retail price of rhinoceros horns has changed the modern horn trade and it has become lucrative to organise crime (Milliken & Shaw 2012). Poaching has changed from traditional tracking by foot to large, expensive operations involving helicopters to track and dart the rhinoceros and ground crews to rapidly remove the horns. Furthermore, alternative sources of horns have been exploited, including pseudo-hunts by Vietnamese nationals (see section 1.3.3), siphoning official stockpiles and natural mortalities, and thefts from private collections. The illegal movement of these horns has been aided by the cash, personnel and trade-routes available to organised crime groups (Milliken & Shaw 2012).

Personal, state and museum collections of horns (referred to henceforth as “historic”, although may include legally obtained horns from recently deceased rhinoceros) cannot be excluded as sources during the last peak in the horn trade; however the lack of discussion about such horns in the literature throughout this period suggests that occurrences were minimal (especially in comparison with the numbers of live animals being poached). Government and other legal stockpiles of horns, which typically include both recent and historic horns, have been recognised as a potential target for thefts and illegal sales for many years (Emslie & Brooks 1999) and between 2000 and 2005 these were believed to have been the primary source of African horns for the illegal markets (Milledge 2008). More recently, however, the horn trade has begun impacting horn collections in both range and non-range states.

Reported thefts from private, state and museum collections have been reported in South Africa since 2002 (Table 3.1) (Milliken & Shaw 2012). Between 2009 and 2011, numbers of

horns stolen from sites across Europe and North America rose sharply (Table 3.2). Whilst the total numbers of horns stolen is low in comparison with numbers of live individuals which have been poached, the loss of the horns alone is not the sum of the value of the theft. Horns in such collections rarely hold any monetary value, although they may hold significant scientific, historic or artistic value. However, thefts of horns typically involve burglaries and there are therefore associated costs with damage to property, as well as harm and threat to humans at the scene and law enforcement time and costs.

Table 3.1. Reported thefts or attempted thefts of horns from state and private collections in South Africa (note Number of locations targeted includes unsuccessful attempts, and number of horns stolen includes theft of fake horns). Source: Milliken & Shaw (2012).

Year	Number of locations targeted	Number of horns stolen
2002	1	1
2006	1	1
2007	2	3
2008	2	3
2009	2	17
2010	2	25
Unknown (2008-2010)	7	15

Table 3.2. Details of reported thefts across Europe 2009 – 2011. Source Milliken & Shaw

Country	Number of horns stolen*	Type of locations
USA	2	Cheque casher
Germany	13	Zoo, Museum
UK	> 8	Auctioneer, Museum, Zoo
France	5	Museum
Portugal	2	Museum
Belgium	6 (2)	Museum
Italy	3	Museum
Czech Republic	3	Castle
Sweden	1	Museum
Netherlands	2	Museum
Austria	2	Auctioneer, Taxidermist

(2012)

* Brackets indicate number of attempted thefts resulting in no horns stolen

Legal sales of trophy and artistic worked horns have also been recognised as an increasingly popular source of horns. Sales and applications for intra-EU trade of "antiques" and "worked specimens", which was legal under CITES legislation, showed a dramatic rise (Ahvla 2012; European Commission 2013). Furthermore, the prices fetched for such pieces at auction were typically far greater that the valuation, which reflected its artistic value only (Viscardi 2012; European Commission 2013). Auction prices were actually correlated with an items weight and buyers were not typically interested in their artistic value (European Commission 2013). In all likelihood, many of these horns were bought for subsequent illegal trade. Indeed members of an organised crime group known to be involved in thefts of rhinoceros horns regularly visited auctioneers and antiques dealers (European Commission 2013). In 2012, the UK tightened regulations regarding the commercial sale of rhinoceros horn to alter the status of mounted horns to "unworked" which therefore cannot be traded, and horns can no longer be re-exported except as a genuine cultural or artistic exchange between reputable institutions, movement of a family heirloom upon relocation or bequest, or as part of a bona fide scientific research project (Ahvla 2012). These changes have been adopted in an attempt to inhibit movement of genuine and false artistic specimens into the illegal market.

As discussed previously in Chapter 1, law enforcement can be greatly improved by the ability to confidently matching a horn to a crime. The case of the attempted smuggling of horns from a deceased captive rhino from Colchester zoo was the first demonstration of the application of DNA profiling to rhinoceros horn for a UK crime (Bhattacharya 2010). Comparison of the profile obtained from the horn to that obtained from a stored blood sample from the rhino demonstrated a match between the two samples (Bhattacharya 2010). This evidence was valuable in building a case against the perpetrator. Furthermore, many of the individuals involved in the trade of horns from poached rhinoceros are also likely to be involved in moving and trading in historic horns. Therefore being able to target such individuals may be beneficial to reduce all forms of trade.

Whilst it has been demonstrated that DNA can be isolated and amplified from modern horns using different techniques (Chapter 1; Peppin 2009; Peppin *et al.* 2009; Harper *et al.* 2013), historic horns are more likely to contain degraded DNA and possibly be composed of toughened material, making extraction of good quality DNA more challenging. DNA obtained from modern horns under the extraction method utilised in Chapter 2 – extraction of 100 mg of powdered horn material with the Qiagen DNA Investigator Kit – produced sufficient DNA quantity and quality for profiling. PCR product strength was not, however, particularly strong, reflected in low RFU heights, suggesting that degraded samples may not always yield

sufficient DNA and allelic- or locus dropout may be more likely. Therefore a variety of different extraction methodologies were trialled using both modern and historic horns to determine an optimal technique. This optimum would provide a balance between high DNA yield (both quality and quantity), ease and speed of processing and maximal limitation of cross-contamination likelihood.

3.2 Methods

3.2.1 Samples and extraction kits used

Cross sections through the vertical axis of two modern horn samples from recently deceased captive *D. bicornis michaeli* rhinoceros were obtained from Port Lympne Wild Animal Park (Horns E and OU). The trials described here took place within two years of the death of the animal that produced horn E and it is therefore considered to be a fresh horn sample. This horn sample was also used in Chapter 2 to compare profiles obtained from different starting materials. Horn OU is of unknown age, but horn extraction trials were likely to have been carried out between 10 and 30 years since the individual's death, and the DNA may therefore have undergone some degradation. Horns were stored dry at room temperature.

A further 11 horn samples were obtained by Science and Advice for Scottish Agriculture (SASA) from museum collections. Sample information was anonymised before use in this study, and thus exact age and origin of samples was unknown. However, samples were expected to be between 50 and 130 years old. How DNA degrades as a rhinoceros horn ages is unknown, and is likely dependent on storage conditions (e.g. exposure to UV, dampness, temperature). The DNA within the museum samples is therefore expected to be more representative of truly historic horns than those of horns E and OU, although the exact details of the age continuum were unknown.

Material for extraction was obtained by drilling as described in Chapter 2 (section 2.2.1). From the two recent horn samples, material was drilled from across the base of the cross section, no more than 3 cm depth. Museum samples were drilled from the centre of the horn base.

To determine the optimal extraction method for rhinoceros horns, particular details of the protocols were adjusted as detailed is subsequent sections below. Extractions were carried out using one of two extraction kits: Qiagen® QIAamp DNA Investigator Kit (hereafter

referred to as the Qiagen method) or Thermo Scientific KingFisher™ Cell and Tissue DNA Kit (hereafter referred to as the Kingfisher method). For the Qiagen method, two different types of silica spin-columns were tested: QIAamp MinElute columns, which are included in the Qiagen DNA Investigator Kit, and QIAamp Mini columns, which are included in the Qiagen DNEasy Blood and Tissue Kit. The standard protocols for DNA extraction were followed for both the Qiagen and the Kingfisher method, except for the stages undergoing optimisation as detailed in section 3.2.2. Extraction controls were included for all extractions.

3.2.1.1 Qiagen method – standard methodology

Digested material was centrifuged for one minute at 10 000 rpm prior to transferring the supernatant to the spin column in order to prevent undigested material from blocking the filter. The spin-column was then centrifuged at 8 000 rpm for one minute during which the DNA becomes bound to the silica filter and the filtered lysate is discarded. A number of standard washes were then performed as detailed in the protocol (see Appendix VII for further details). Finally, 50 µL of warmed Buffer ATE was applied to the centre of the silica membrane and incubated for 10 minutes. A final centrifugation at 14 000 rpm for one minute eluted the DNA into a microcentrifuge tube for storage.

3.2.1.2 Kingfisher method – standard methodology

Digested material was added to the KingFisher mL tubes with 25 µL of KingFisher magnetic beads and 360 µL of binding buffer. Wash Buffers and 50 µL of elution buffer were added to four additional tubes (see Appendix VIII for further details). The KingFisher mL instrument was then started and DNA extraction proceeds automatically. Finally the purified DNA is transferred to a microcentrifuge tube for storage.

3.2.2 Extraction optimisation

3.2.2.1 Sample preparation and addition of DTT with the Qiagen method

Two methods of sample preparation prior to DNA extraction were tested. Whether DNA yield was greater by using shavings from the drilling process in an unaltered state or by powdering them using liquid nitrogen and Retsch mixer mill (2 mins at 25 revolutions/s) was tested. It was thought that surface area could be increased by powdering the horn, thus improving the

digestion stage of the extraction. Furthermore, dithiothreitol (DTT) is often added to digestion of hair and nail samples for DNA extraction to help break down keratinous structures. DTT is an irritant, however, and its use should be avoided where possible. Its necessity in the digestion stage of rhinoceros horn extraction was therefore evaluated by assessing the yield with and without DTT.

Use of unaltered shavings or powdered horn material and DTT in the digestion stage of extraction was tested on three separate occasions in a factorial method. On the first two occasions, the powdered and DTT excluded combination was not tested. Digestion was carried out using 100 mg of horn material from horn E. Material was then used either unaltered in the digestion (shavings) or powdered as described in table 3.3. Powdering of horn material was carried out by dropping a 2 mL safe-lock Eppendorf containing 100 mg horn material into liquid nitrogen to snap freeze it. The Roche mixer mill (two minutes at 25 revolutions/s), was then used to pulverize the material.

Table 3.3. Sample preparation and DTT were compared under these combinations. Trial was carried out on three separate occasions and sample numbers for each are indicated in brackets.

	DTT included	DTT excluded
Shavings	Shavings + DTT (4, 4, 2)	Shavings – DTT (2, 2, 2)
Powdered	Powdered + DTT (4, 4, 2)	Powdered – DTT (0, 0, 2)

Samples were then processed following the standard Qiagen method using MinElute columns with the following exceptions. To the digestion, 20 µL of 1 M DTT was added and digestion volume was doubled to cover the sample fully, and samples were digested for 16 to 20 hours at 56 °C. Washes and elution was carried out under standard protocol except that spin columns were incubated for 10 minutes with 75 µL of warm ATE before elution. The best combination was tested on eleven museum samples. Gel electrophoresis was not carried out for these extracts as the extract volume was minimal.

3.2.2.2 Reduction of starting material and varying DTT with the Qiagen Method

The mass of horn material used for extraction was reduced from 100 mg to 50 mg. Using material from horn OU, two samples were extracted from shavings only and two samples were powdered using liquid nitrogen and Retsch mixer mill (2 mins at 25 revolutions/s).

Standard extraction protocol was carried out as described above, and included 40 µL of 1 M DTT.

Three museum samples were extracted using 50 mg of starting material. Horn material was powdered using the Retsch mixer mill (2 mins at 25 revolutions/s) without freezing in liquid nitrogen (this was deemed unnecessary to powder the horn material), and 30 µL of 1 M DTT was added to aid digestion. Standard extraction protocol was carried out as described above. Extractions were genotyped using multiplexes A and B using the standard cycling parameters of 28 cycles as well as 33 and 38 cycles.

To comparatively assess the effect of the mass of starting material on the quantity and quality of DNA extract, material from horn OU was extracted using 100 mg and 20 mg of material (six replicates of each mass). To half of the replicates, 10 µL of DTT was added to the digest and 30 µL of DTT to the other half. A further three samples starting with 100 mg of material (and 10 µL DTT) were extracted as standard but digested material was divided equally between two spin columns. All material was pulverized using the Retsch mixer mill (2 mins at 25 revolutions/s). Standard extraction protocol was carried out as described above.

3.2.2.3 Comparison of different extraction methods

The Qiagen method was compared with the automated KingFisher method. Qiagen extraction kits use silica membrane-based purification spin columns, and both MinElute and Mini columns were tested. The KingFisher method uses a magnetic bead system to purify DNA, and extractions were carried out on the KingFisher™ mL machine. The three purification methods were expected to differ in the starting mass for optimal DNA yield. The silica-membranes of the Qiagen spin columns can become blocked if too much DNA and cellular material is added to the extraction. Magnetic beads, on the other hand, may become saturated but yield was not expected to decrease if overloaded.

Three extracts from 50 mg and three extracts of 20 mg of horn OU were extracted with the Qiagen method, using MinElute and Mini columns, and the KingFisher method. Three replicates from a single museum horn were also extracted using the Qiagen method with MinElute columns. Extractions followed standard protocols for each method with the addition of 10 µL 1 M DTT to the digestion. Following DNA quantity check using gel electrophoresis, a 1:1 dilution of Qiagen method extracts from 20 mg starting material using Mini columns and all KingFisher extracts was carried out prior to PCR with multiplex B.

3.2.2.4 KingFisher method sample preparation

Sample preparation for digestion, shavings only or powdered horn material, was compared using material from two museum horns. Three replicates of each horn were prepared for extraction using each of the sample preparation methods. The cross-amplification success of STR loci was not fully understood, and the species of the two museum horns was not known. Thus the quality of each multiplex was scored as good if five or more loci amplified, weak if one to four loci amplified and failed if no loci amplified.

3.2.3 DNA quantification and quality check

Where indicated, DNA quantity was measured from 1.5 µL of DNA extract on the ThermoScientific NanoDrop 1000 UV/Vis- spectrophotometer. DNA quantity and quality was further assessed using gel electrophoresis visualisation on a 1 % agarose gel, 5 µL of undiluted DNA was run at 80 V alongside a Sigma PCR sizing ladder.

3.2.4 Amplification and capillary electrophoresis

No template controls (NTC) were included in all PCRs. Where indicated, all amplification of STR products was carried out using the final panel of 14 markers described in Chapter 2 under either singleplex or multiplex conditions. All PCR, capillary electrophoresis and genotype was carried out as described previously in Chapter 2 section 2.2.

3.3 Results

3.3.1 Sample preparation and DTT

Amplification success varied across extraction attempts varying sample preparation and DTT inclusion (Table 3.4). On the first two occasions, amplification success was variable for all treatments. On the third attempt, all loci amplified under multiplex conditions for all replicates of each treatment, although best amplification was observed from one of two shavings only/+ DTT replicates and both shavings only/– DTT replicates. These samples also showed greatest DNA yield and quality on gel electrophoresis (lanes 1, 5 and 6 of Figure 3.1). No amplification was observed in the extraction controls or NTCs.

Table 3.4. Amplification success for replicate extractions using two types of sample preparation and inclusion or exclusion of DTT during digestion. Extractions were trialled on three different occasions and amplification was carried out using either singleplex reactions (loci BlRh1B and DB23) or with both multiplexes A and B.

Sample preparation	Extraction attempt (amplification conditions)	DTT	DTT excluded
Shavings only	1 (Singleplex)	No amplification	1 of 2 amplified both loci
	2 (Multiplex)	No amplification	1 of 2 amplified all loci
	3 (Multiplex)	2 of 2 amplified all loci	2 of 2 amplified all loci
Powdered	1 (Singleplex)	No amplification	Not tested
	2 (Multiplex)	1 of 4 weak amplification	Not tested
	3 (Multiplex)	2 of 2 amplified all loci	2 of 2 amplified all loci

Of the 11 museum horns extracted using no sample preparation (shavings only, no DTT added), four samples amplified more than 10 loci, and six samples between one and eight loci. Two loci, DB44 and BlRh1B, failed to amplify in any samples. No amplification was observed in the extraction controls or NTCs.

3.3.2 Reduction of starting material and varying DTT with the Qiagen method

Reducing the mass of starting material from 100 mg to 50 mg improved DNA yield (DTT was added to all samples). Full genotype profiles were obtained from both samples from

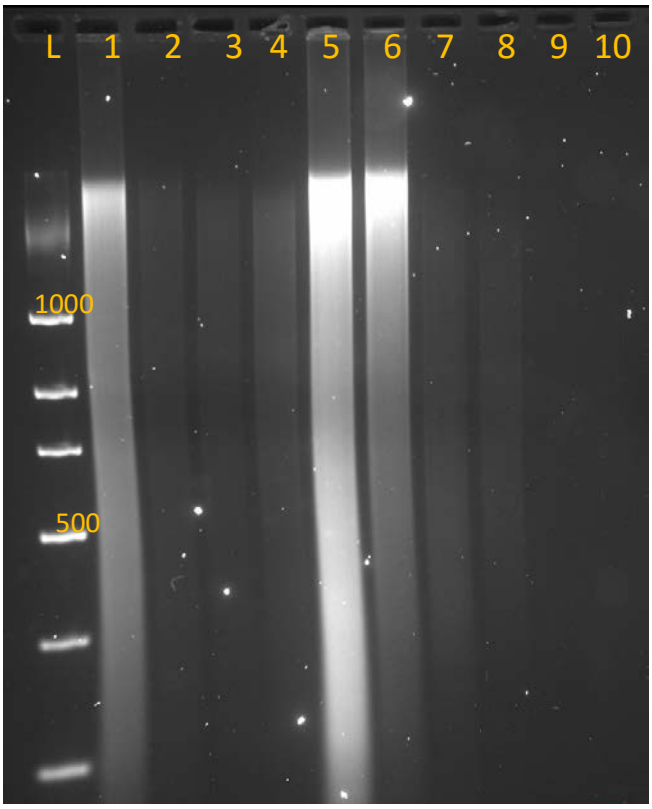


Figure 3.1. Gel electrophoresis visualisation of DNA extracts from the third attempt to compare sample preparation methods and addition of DTT to the digestion stage of extraction. Lanes are indicated by numbers. Ladder (L) is loaded in the first lane and 500 bp and 1000 bp bands are indicated. Lanes 1 – 4 are extracts which used DTT and 5 – 8 did not have DTT added. Extracts from unaltered shavings were run in lanes 1, 2, 5, and 6, and powdered horn material in lanes 3, 4, 7 and 8. Lanes 9 and 10 contain extraction controls

powdered horn and one of the two samples from shavings only. Weak amplification was observed from the second shavings only sample, but no loci could be scored because RFU values of peaks were below set thresholds. All three museum samples extracted from 50 mg of horn material failed to amplify either multiplex under the standard 28 cycles. Two samples amplified only multiplex A at 33 cycles, and all loci at 38 cycles. The third sample failed to amplify under any conditions. No amplification was observed in the extraction controls or NTCs.

Samples extracted using 20 mg of starting material yielded superior DNA quantity and quality than 100 mg of starting material upon visual inspection on gel electrophoresis (Figure 3.2), although this was not evident from DNA quantification using the Nanodrop (Table 3.5). Furthermore, samples with only 10 µL of DTT added to the digestion yielded more DNA than those with 30 µL of DTT, evident from both gel electrophoresis and DNA quantification. All multiplex A loci amplified from all samples. Multiplex B showed more variable amplification success (Table 3.6). Sample D3 (20 mg horn and 30 µL DTT) failed to amplify any multiplex B

loci. Given that sample D3 did amplify all loci for multiplex A, it seems likely that PCR failure accounts for lack of amplification of multiplex B. Considering this, DNA extractions from 20 mg of material consistently amplified strong PCR products across all loci. Furthermore, adding 10 µL of DTT in digestion yielded greater DNA quantity than 30 µL, amplification success does not appear to differ. No amplification was observed in the extraction controls or NTCs.

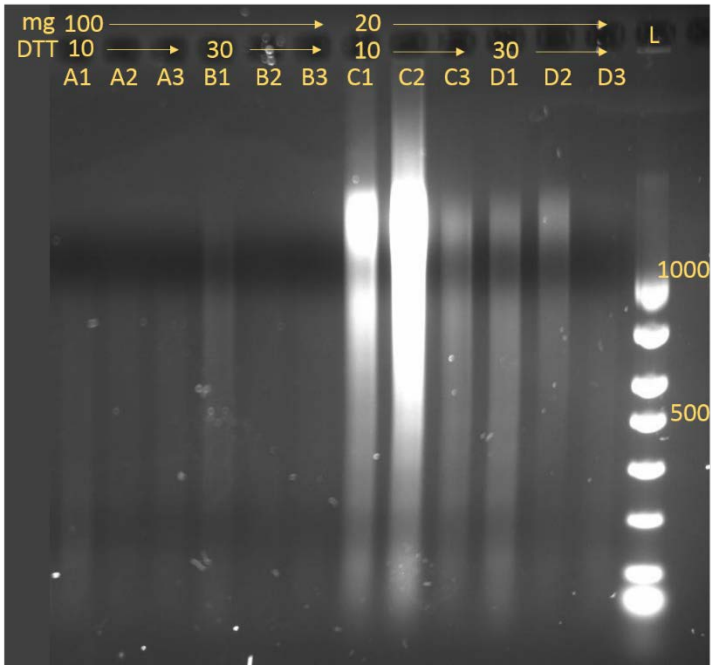


Figure 3.2. Gel electrophoresis visualisation of DNA extracts comparing mass of starting material (first line of notation) and volume of DTT (µL) added (second line of notation). Lanes are labelled using extraction IDs in Table 3.5.

Table 3.5. DNA quantities (ng/µL) determined using Nanodrop 1000 UV-Vis spectrophotometer for extractions comparing mass of starting material and volume of DTT added.

Mass starting material (mg)	Volume DTT added (µL)	Extraction ID	DNA quantity (ng/µL)	Mean DNA quantity (ng/µL) per treatment
100	10	A1	13	19.33
		A2	25	
		A3	20	
	30	B1	35	22.67
		B2	13	
		B3	20	
20	10	C1	15	22.00
		C2	47	
		C3	4	
	30	D1	5	4.33
		D2	3	
		D3	5	

Table 3.6. Genotyping success of multiplex B of extracts comparing 100 mg with 20 mg of starting material and 10 µL DTT with 30 µL DTT. + indicates scored genotype. Weak indicates that amplification was observed but peaks were below threshold RFU and could therefore not be scored. Failures are indicated as Fail.

Starting material (mg)	Volume DTT added (µL)	Sample ID	BlRh1C	WR32A	DB23	DB66	WR7B	WR7C	DB52
100	10	A1	+	+	+	+	+	Fail	+
		A2	+	+	+	+	+	Fail	+
		A3	+	+	+	+	+	Fail	+
	30	B1	+	+	+	+	+	+	+
		B2	+	+	+	+	+	Fail	+
		B3	+	+	+	Weak	+	Fail	+
20	10	C1	+	+	+	+	+	+	+
		C2	+	+	+	+	+	+	+
		C3	+	+	+	+	+	+	+
	30	D1	+	+	+	+	+	+	+
		D2	+	+	+	+	+	+	+
		D3	Fail	Fail	Fail	Fail	Fail	Fail	Fail

3.3.3 Comparison of different extraction methods

Visual inspection on gel electrophoresis showed variation in DNA quality and quantity between extraction methods from different starting masses (Figure 3.3). Extractions using the Qiagen method from 20 mg starting material showed brighter bands on gel electrophoresis, suggesting higher DNA yield, than from 50 mg starting material for both MinElute and Mini columns. All bands were equally bright for extractions using the KingFisher method from both starting masses. The replicates from the museum sample were not run on gel electrophoresis.

For all replicates with visible DNA on gel electrophoresis (Figure 3.3), genotyping of multiplex B was successful for all seven loci (Table 3.7). Qiagen method extracts using the MinElute columns from 50 mg starting material varied in genotyping success between two and six loci scored. Replicates of the museum sample extracted from 20 mg of horn with the Qiagen method using the MinElute column showed consistently poor genotyping success. No amplification was observed in the extraction controls or the NTCs.

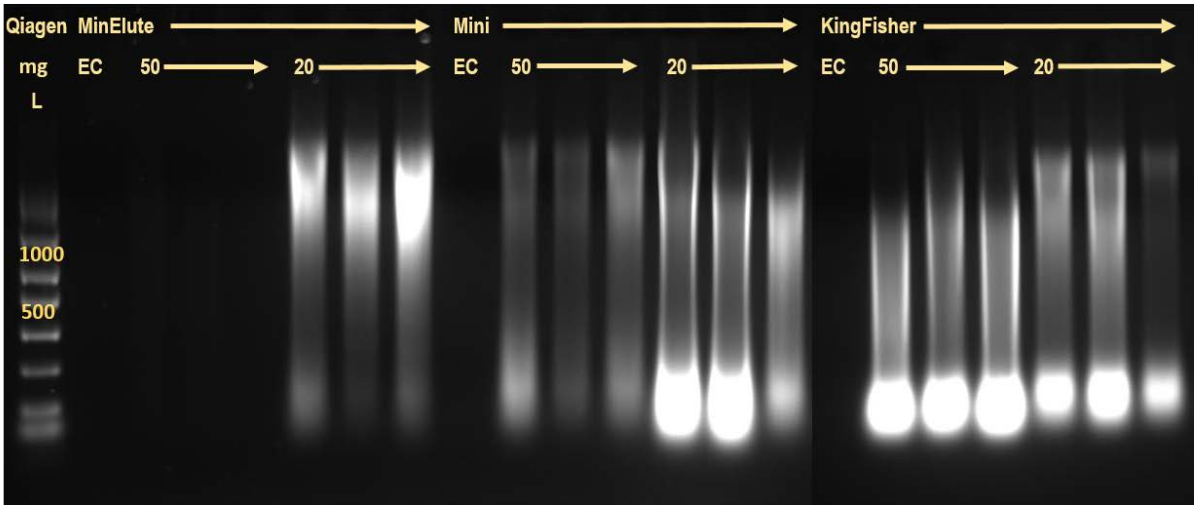


Figure 3.3. Gel electrophoresis visualisation of extracts comparing the two Qiagen column and KingFisher methods (first line of notation) from 50 mg and 20 mg of starting material (second line of notation). Lane labelled L contains the ladder (500 and 1000 bp markers are indicated).

Table 3.7. Genotyping success of multiplex B for extracts comparing different DNA purification methods. + denotes scored genotype. Weak indicates that amplification was observed but peaks were below threshold RFU (100 RFU for heterozygotes and 200 RFU for homozygotes) and were not be scored. Fail denotes complete absence of amplification.

Extraction method	Mass starting material (mg)	Replicate	BlRh1C	DB23	DB52	DB66	WR7B	WR7C	WR32A
Qiagen MinElute Column	50	a	+	+	+	+	+	Fail	+
		b	+	+	Weak	Fail	+	Fail	+
		c	Weak	Weak	Fail	Fail	+	Fail	+
	20	a	+	+	+	+	+	+	+
		b	+	+	+	+	+	+	+
		c	+	+	+	+	+	+	+
		Museum a	+	Weak	Weak	Fail	+	Fail	+
		Museum b	+	Fail	Fail	Fail	+	Fail	+
		Museum c	+	Weak	Weak	Fail	+	Fail	+
Qiagen Mini Column	50	a	+	+	+	+	+	+	+
		b	+	+	+	+	+	+	+
		c		+	+	+	+	+	+
	20	a	+	+	+	+	+	+	+
		b	+	+	+	+	+	+	+
		c	+	+	+	+	+	+	+
KingFisher Method	50	a	+	+	+	+	+	+	+
		b	+	+	+	+	+	+	+
		c	+	+	+	+	+	+	+
	20	a	+	+	+	+	+	+	+
		b	+	+	+	+	+	+	+
		c	+	+	+	+	+	+	+

3.3.4 KingFisher method sample preparation

Variation in DNA yield between two museum samples was evident between samples on gel electrophoresis (Figure 3.4). There appeared to be no difference within a sample whether shavings or powdered horn material was used (Figure 3.4). Both sample preparation methods showed good amplification for multiplex A, but sample 2 showed weaker amplification of multiplex B using the shavings only preparation (Table 3.8). No amplification was observed in the extraction controls or the NTCs.

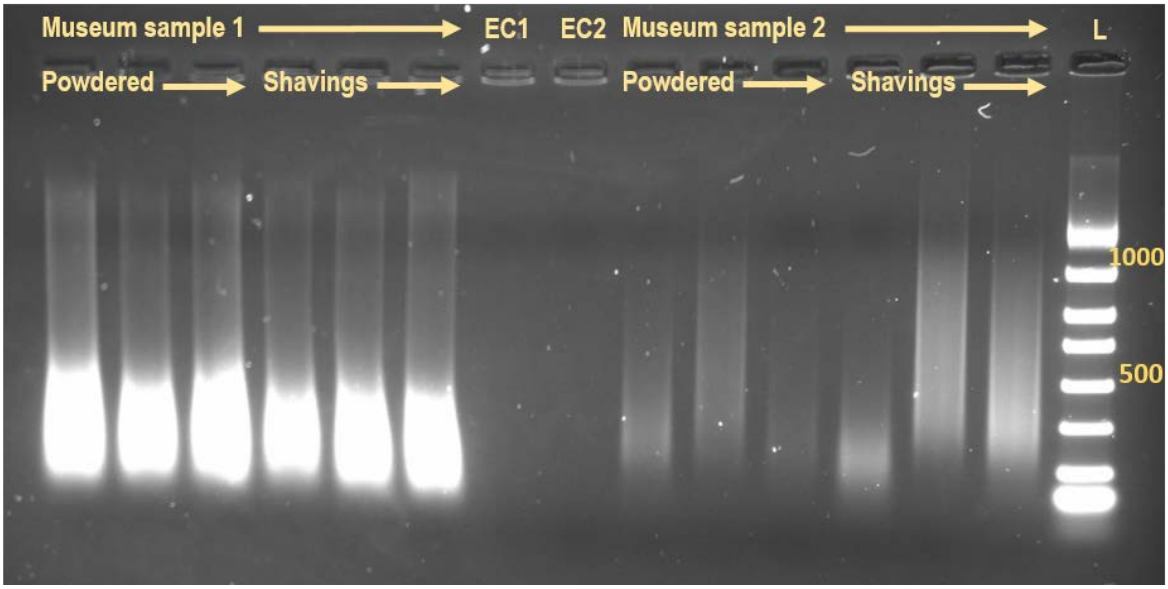


Figure 3.4. Gel electrophoresis of extracts from two museum samples (first line of notation) using the KingFisher method from powdered horn or shavings only (second line of notation). Lane labelled L contains the ladder (500 and 1000 bp markers are indicated). Extraction controls are labelled EC.

Table 3.8 Genotyping success of multiplexes A and B for two museum samples extracted using the KingFisher method under different sample preparation methods. Genotyping success of each multiplex was scored as good (5 – 8 loci scored) or weak (1 – 4 loci scored).

Sample	Preparation method	Replicate	Multiplex A score	Multiplex B score
Museum 1	Powdered	A	Good	Good
		B	Good	Good
		C	Good	Good
	Shavings	A	Good	Good
		B	Good	Good
		C	Good	Good
Museum 2	Powdered	A	Good	Weak
		B	Good	Weak
		C	Good	Weak
	Shavings	A	Good	Good
		B	Good	Good
		C	Good	Good

3.4 Discussion

3.4.1 Sample preparation and DTT addition with the Qiagen method (MinElute columns)

Extraction quality was highly variable when extracting from 100 mg of a modern horn using the Qiagen method. Amplification success was typically minimal. Across extraction attempts comparing DTT and horn preparation methods, using the horn shavings in their unaltered state tended to be most successful, although DNA yield clearly remained variable as did the effect of DTT addition.

Rhinoceros horn is often considered to be similar to horns of other mammals, but in fact its structure is keratinised rather than containing a bony core (Modell 1969; Geerinckx *et al.* 2007). The preparation of horn material by freezing with liquid nitrogen and powdering with the Retsch mixer mill was adapted from a method for extraction of DNA from ivory (Comstock *et al.* 2003) under the assumption that whilst not as hard as ivory, rhinoceros horn is tough enough to require mechanical breakdown prior to chemical digestion. Whether this process was necessary could not be determined during this optimisation trial. Furthermore, no strong conclusions could be drawn about the effect of DTT on the extraction process using either starting material. This could suggest that DTT has little effect on the digestion success of modern horn, and proteinase K in combination with the digestion buffers is sufficient to break down the keratinous material.

When trialling the Qiagen method on museum samples (using unaltered shavings without DTT), amplification success was again very variable. Firstly, it is important to note that the species of origin of the eleven horn samples originated was not known, nor the ability of the loci to cross-amplify DNA from species other than *D. bicornis*. A study by Harper *et al.* (2013) assessed the 14 loci used in this study using the same primer pairs, and found that they were all polymorphic in *C. simum* and *D. bicornis*. Scott (2008) found that ten of the loci cross-amplified all rhinoceros species but did not test against *R. sondaicus*; however, five of the primer pairs differed from those used here and by Harper *et al.* (2013) and thus their cross-amplification success may differ. Four STR loci (BlRh1B, BlRh1C, DB23 and DB66) have not yet been tested for cross-amplification in Asian rhinoceros species. Therefore, reduced amplification success cannot solely be considered a failure of the extraction process, but could in fact be a result of failure to cross-amplify. Failure of the largest fragments to amplify suggests that the extracts may be primarily composed of fragmented DNA. This fragmentation could be a consequence of the age of the horn sample as DNA is known to fragment with age (Pääbo *et al.* 2004), a result of a poor extraction process or that PCR

conditions are too stringent for historic samples. Amplification success using the Qiagen method with MinElute columns could be considered variable from historic samples, which is not conducive to forensic testing of such material.

3.4.2 Reduction in starting material and varying DTT with the Qiagen method (MinElute columns)

Previous studies suggest that extraction from modern horns should provide sufficient material for DNA profiling, even when using less starting material (Peppin 2009; Harper *et al.* 2013); however, extraction attempts from 100 mg of material had shown very variable DNA yield. It could be possible that when using the MinElute columns, which are specifically designed to maximise DNA yield from samples containing very little DNA, the columns were overloaded when starting with 100 mg of horn material. Indeed, halving the mass of starting material to 50 mg did result in improved amplification success when using a modern horn sample, although this effect was lost when amplifying from historic samples. For these, the addition of a further ten cycles during the PCR improved the amplification success; however this is not desirable for forensic analysis of horns as the likelihood of amplifying artefacts and contaminants is increased with increasing numbers of PCR cycles.

Reduction in the mass of starting material to 20 mg further improved DNA yield from extractions with the MinElute columns. Increasing the volume of DTT added to the digestion from 10 µL to 30 µL did not appear to be beneficial. The increase in yield following reduction in starting material suggests that DNA quantity within rhinoceros horn is sufficiently high, and that by overloading the spin columns, they may become blocked and thereby reducing its ability to recover the DNA during the elution stage. This may be exaggerated when using the MinElute columns due to their enhanced capability of capturing and eluting DNA from low quantity samples rather than from those with abundant DNA.

3.4.3 Comparison of different extraction methods

The Qiagen method using MinElute columns was the most sensitive to starting material mass, showing complete failure to extract from 50 mg of horn material in this instance. At 20 mg of horn material, however, this method was the only one to elute a high concentration of high molecular weight DNA. The Qiagen method with Mini columns and the KingFisher method generated sufficient quality and quantity DNA at both starting masses for successful genotyping, although more fragmented DNA was present on gel electrophoresis compared with the successful MinElute extractions. For use in DNA profiling, which utilises STRs less

than 300 bp in length, this fragmentation does not appear to be a significant issue as all loci could be genotyped for extractions using the Mini columns and the KingFisher method. The museum sample showed no visible DNA on gel electrophoresis and genotyping success was low. This, however, was likely due to use of the MinElute column which seems to be the poorest extraction method.

In choosing the best method for extraction from historic horns, there are a number of advantages and disadvantages to the different methods. The Qiagen method using MinElute columns would be beneficial if only very small amounts of starting material containing trace amounts of DNA are available, but if DNA quantity and quality is high, extraction results appear highly variable. The Qiagen method with Mini columns was less sensitive to starting material, and therefore likely to be most suitable to horn material of varying ages. The amount of DNA within a given quantity of rhinoceros horn is likely to differ between horns of different ages, storage conditions over the lifetime of the horns and individual variation in cellular density within the horn. Therefore an optimal method cannot afford to be overly sensitive to the starting material. In this respect, the KingFisher method is the most robust. Furthermore, this method is semi-automated, reducing pressure on person time.

3.4.4 KingFisher method sample preparation

Comparison of extractions from unaltered shavings with powdered horn material showed no apparent difference between the two treatments for museum horns. Information about the history of the horn sample was anonymised, thus by chance, both horns could be recent. Material from older horns were presumed more likely to be tougher than that from more recent horns and therefore more difficult to digest. One horn showed much greater DNA yield that the other on gel electrophoresis (Figure 5), thus it is clear that the quality of the starting material differed. Nonetheless, genotyping success was high for all extractions which did not involve powdering the horn material prior to digestion. The second sample showed reduced genotyping success for extractions from powdered material, which may be suggestive of DNA damage due to the pulverisation. Therefore, chemical digestion using DTT in addition to the standard proteinase K and digestion buffers was considered sufficient for digestion of rhinoceros horn.

3.4.5 Conclusions

Due to limited time and amount of rhinoceros horn available for this optimisation, particularly from museum samples, the completion of a fully factorial optimisation test could not be carried out. The results from the KingFisher trial indicated the method was capable of generating sufficient quality extracts from the two museum samples tested. The Qiagen method with Mini columns also yielded sufficient quantities of DNA for profiling from small quantities of starting material. It is therefore likely to be suitable for extraction from historic horns of variable quality and quantity, whilst minimising the possibility of contamination between samples. Both methods, however, require further testing on older horns of known ages.

DNA purification using the Qiagen method with MinElute columns was highly variable throughout these studies. These columns are designed to enable extraction of DNA from a range of forensic sample types, and to be particularly effective at recovering DNA from minute samples. It seems likely that the risk of blocking or overloading these columns is quite high, and may therefore not be well suited to horn samples due to both variation in DNA content and tough keratinous material which may be transferred to the spin columns.

Where tested, the difference in results between mechanically powdering the horn sample and simply using shavings was minimal, although shavings has a slight tendency for greater extraction success. Exclusion of the liquid nitrogen from pulverisation of starting material reduces the burning hazards associated with this process, and excluding the powdering of the horn material reduces the potential of additional source of contamination from the beads.

It is important to note that rhinoceros horn does not always appear to digest fully as is frequently the case with other source materials such as tissue or blood. Typically, a honeycomb-like structure appears to remain after 16 hours of digestion. However, DNA yield can be very high despite this apparent lack of digestion, suggesting that the cellular material has been digested and DNA released into the supernatant. As stated in the methods, it is very important when using silica spin columns to purify DNA from rhinoceros horns, to carry out an additional centrifuge step prior to transferring the supernatant to the spin columns. This ensures that the largest cellular components are not transferred and therefore will not block the spin columns. The addition of DTT may aid digestion of the material, although definitive conclusions cannot be drawn from the tests carried out here. These test are indicative,

however, that addition of a small amount of DTT (10µL of 1M DTT) does not reduce extract quality or quantity.

Chapter 4 – General Discussion

4.1 Summary

4.1.1 Developmental validation

A panel of 15 STR markers and one sexing locus was rigorously assessed to determine its suitability for use in forensic individualisation of *D. bicornis* horns. Two loci, BlRh37D and IR12, were found to deviate from Hardy-Weinberg equilibrium in a wild Kenyan population of *D. bicornis michaeli* and were thus rejected. The remaining 13 STR markers and sexing locus were determined suitable for forensic use.

This developmental validation attempted to follow guidelines established for forensic analysis of non-human DNA by Linacre *et al*. (2011) with consideration of the recommendations by Budowle *et al*. (2005) and those established for human systems (SWGDAM 2012). The ability of this developmental validation to meet the recommendations of the International Society for Forensic Genetics (ISFG) (Linacre *et al.* 2011) that relate to individual identification are as below. Recommendation numbers 1, 12 and 13 relate to casework procedures, and number 3 refers to species specificity testing only; these recommendations are therefore not discussed.

Recommendation # 2 – Use of voucher specimens

Samples of captive *D. bicornis* were obtained from Port Lympne Wild Animal Park. Species and sex were therefore not ambiguous. Their use lends confidence to the method and ensures there is no uncertainty over identification of species or subspecies of samples.

Recommendation # 4 – Nucleotide sequence and mapped location of primers

Genomes sequences were not available for any rhinoceros species, and thus loci synteny was established with the genome of a closely related species, *E. caballus*, for 12 of 15 STR loci. This prohibited determination of the exact positions of loci as *D. bicornis* have 84 chromosomes and *E. caballus* have 64. By using a combination of locus mapping to the *E. caballus* genome (section 3.1) and estimating linkage disequilibrium between loci for a wild population of *D. bicornis* (section 3.3), it was possible to predict with some degree of confidence whether two loci are located on the same chromosome and therefore not independently segregating during meiosis.

Three loci (BlRh1C, DB66 and WR7C) mapped to a *E. caballus* chromosome 5 and two (DB1 and SR63) to chromosome 14; however, none of these loci were found to be in linkage disequilibrium. This suggests that *E. caballus* chromosomes 5 and 14 may not have single homologs in *D. bicornis* or that the distance between loci on a single chromosome is enough to enable high enough recombination rates to emulate independent segregation.

The position of IR12 could not be determined on the *E. caballus* genome as the flanking regions have not previously been published and primers are located too close to the repeat region, preventing the generation of high quality sequence here (see section 3.2). However, homozygosity across all males and significant linkage disequilibrium with ZF1 provides compelling evidence to suggest that it is linked to ZFX and therefore likely to be located on the X chromosome.

Recommendation # 6 - Specificity and reproducibility

All primer pairs except WR7C and ZF1 showed no amplification of human or domestic dog DNA, suggesting that they are highly specific to rhinoceros species. Although WR7C and ZF1 primers did amplify fragments from both human and domestic dog DNA, these peaks were easily distinguishable from *D. bicornis* alleles. DNA from these two species is likely to be the most commonly encountered in the wildlife forensics laboratory in which this study was undertaken. However, if samples from other species are regularly process in the same laboratory, especially species more closely related to rhinoceros including *Equus caballus*, then it would be prudent to include such specificity tests within the internal validation.

All loci were determined to be highly reproducible, and d-values, heterozygote balance and stutter ratio were documented. D-values, which describe the observed drift of alleles between capillary electrophoretic runs, did not exceed 0.55 bp; thus demonstrating that variation does not affect allele calling and the system is precise. Mean stutter ratios were all below 0.5 for STR loci (the sexing locus, ZF1, did not produce stutter), although the greatest observed stutter ratio was 0.70. Heterozygote balance ranged between 0.72 and 2.74, but 90.8 % of amplifications were greater than 1.0. Stutter and heterozygote imbalance are known to be more pronounced for dinucleotide repeats than STRs with larger repeat units; however for this panel, neither stutter or heterozygote balance caused allele scoring difficulties.

Recommendation # 7 – Use of tetranucleotide STRs

This profiling system includes 15 dinucleotide STRs, all of which have been used in forensic profiling in South Africa (Harper *et al.* 2013). Dinucleotide STRs are more challenging than those with larger repeat units as increased stutter and heterozygote imbalance can cause scoring ambiguities, and thus ISFG guidelines promote the preferential use of tetranucleotide STRs, "except for those markers that are already used widely in animal genetic studies" (Linacre *et al.* 2011). The only STR markers discovered thus far for any rhinoceros species have been dinucleotides, despite extensive searches (Brown & Houlden 1999; Cunningham *et al.* 1999; Zschokke *et al.* 2003; Florescu *et al.* 2003; Scott *et al.* 2004; Scott 2008) and whole genome sequencing to identify tetranucleotide STRs remains prohibitively costly. Given the extensive use of these markers for profiling of rhinoceros, including the black rhinoceros subspecies *D. b. minor* (Harper *et al.* 2013), the markers can be considered "widely used in animal genetic studies" and this work validates their use in forensic casework. Stutter ratios and heterozygote balance were reported.

Recommendation # 8 – Allele ladder and allelic sequencing

Due to time constraints, an allelic ladder has not yet been developed for this STR panel for *D. bicornis*. An allelic ladder would greatly improve confidence that electrophoretic variation does not affect allele calling, as well as inter-laboratory calibration. In the absence of an allelic ladder, a few samples of known genotypes should be included as positive controls with every run (Andreassen *et al.* 2012).

Allele sequencing showed that, with one exception (WR7C allele 33.2), fragment length reflected repeat unit number, and nomenclature is proposed accordingly. This nomenclature standardizes allele designation and does not rely upon fragment length which can vary between electrophoretic conditions and should benefit inter-laboratory standardisation (Budowle *et al.* 2005). Producing high quality sequences of microsatellites is very challenging, in large part due to stutter products generated during the PCR amplification which typically present as mixed sequences at the 3' end of the repeat motif. Whilst cloning may improve the quality of the sequences produced, both the true allele and stutter fragments will be cloned thus making it difficult to determine the true repeat length. Quality of the sequences generated here was sufficient to determine the number of repeat motifs for all loci, although variation in flanking regions could not be reliably assessed. Across all alleles within a locus, no discrepancies were observed between motif repeat number and

genotyping fragment length and thus the nomenclature system for each locus accurately reflects the variation within it. Due to the wide range of variables that can affect genotype scoring between laboratories, however, it is essential that inter-laboratory calibration of the marker systems includes genotyping of a range of samples at both laboratories in order to standardise genotype scoring.

Recommendation # 10 – Allelic frequencies

Allelic frequencies were calculated for 52 wild Kenyan *D. bicornis* to assess the discriminatory power of the loci and to identify any deviations from Hardy-Weinberg equilibrium or other assumptions of individualisation tests. Two loci, BlRh37D and IR12, were found to deviate from HWE, the former likely due to the presence of null alleles and the latter due to linkage with the X chromosome. After the exclusion of these loci, PI_{ave} values were found to range between 1.06×10^{-11} (unrelated population, Fst=0-.05) and 1.16×10^{-7} (relatedness of cousins, Fst=0.2).

ISFG guidelines suggest that the sample size required for an allele frequency database must be sufficient to account for sampling errors, and depends upon levels of diversity across loci and how many individuals are potential contributors to that population. A sample size of 200 individuals is suggested as typical (Linacre *et al.* 2011). Obtaining samples from rhinoceros is particularly challenging as they must be anaesthetised, which can be very risky for the rhinoceros. Furthermore, 200 individuals represents just under a third of the entire Kenyan population of 631 *D. bicornis michaeli* population, or a quarter of the entire subspecies (799 individuals) (Emslie & Knight 2014).

The Kenyan *D. bicornis michaeli* are now primarily restricted to national parks, reserves and wildlife conservancies and are considered to be a single meta-population (Emslie, R.H., Amin, R., Kock 2009; Muya *et al.* 2011). This follows a decline to around 400 individuals by the early 1990s (Emslie & Brooks 1999) within four sanctuaries, and subsequent population growth and translocations among approximately 16 managed populations (Okita-Ouma *et al.* 2007; Emslie, R.H., Amin, R., Kock 2009; Muya *et al.* 2011). Gene-flow was found to be high amongst populations by Muya *et al*. (2011) except with the Masai Mara, thus allelic frequencies can be expected to be similar across most of the meta-population. The Masai Mara population was highlighted by Muya *et al*. (2011) as distinct from other

populations, due to either demographic isolation or genetic interaction with the Tanzanian population which likely contains *D. bicornis minor*. Whilst a sample size of 52 individuals, as used for population genetic studies here, could be considered too small, these individuals were composed of a subset of most of these populations and the allele frequencies are likely to be a good representation of this meta-population. Further genetic datasets would be needed, however, to extend this method for forensic use beyond the Kenyan population.

Recommendation # 11 – Kinship factor

Calculations of probability of identity and RMP should include a kinship factor appropriate for the population under consideration (Linacre *et al.* 2011). The Kenyan rhinoceros, with the exception of the Masai Mara population, is considered a single meta-population (Emslie, R.H., Amin, R., Kock 2009; Muya *et al.* 2011) and therefore it is appropriate to apply a single kinship factor. Using data from (Muya *et al.* 2011) which considered 145 individuals across the meta-population, a θ value of 0.1 could be considered conservative and was therefore applied.

For a wildlife species such as *D. bicornis*, perhaps the greatest limitations to a developmental validation are the availability of samples, funds and available genetic markers. These limitations make it challenging to fully evaluate such a marker system following all guidelines, as has become established for domestic species (Menotti-Raymond *et al.* 2005; Coomber *et al.* 2007; Kun *et al.* 2013; Wictum *et al.* 2013). This validation has instead attempted to assess the marker system as fully as possible, given these limitations, and is in line with other developmental validations for wild animal systems (Dawnay *et al.* 2008; Andreassen *et al.* 2012).

4.1.2 Horn extraction optimisation

The optimisation of DNA extraction from rhinoceros horns was carried out to ensure that when extracting from historic horn samples, maximal DNA yield could be obtained from minimal starting material. This would ensure that minimal damage is caused to historic horns with scientific, artistic and social value. This demonstrated that mechanical breakdown of

horn material has relatively minor impact on DNA yield. The mass of starting material and purification method, on the other hand, greatly affected DNA yield.

Due to the solidity of rhinoceros horn in comparison with other tissues typically used in extractions, for example tissue or blood, mechanical breakdown of the material prior to chemical digestion was expected to improve DNA yield. The use of liquid nitrogen to snap freeze the horn was excluded from the protocol as unnecessary and hazardous. The effect of powdering the horn prior to digestion was typically variable. The introduction of milling balls when using the Retsch mixer mill to powder the horn, however, is an additional potential source of contamination. Therefore, as the mechanical breakdown did not seem to improve the DNA yield, it was deemed unnecessary.

The comparison of different extraction methods clearly demonstrated that the Qiagen method using MinElute columns (the standard DNA Investigator Kit) is very sensitive to overloading of starting material. The MinElute column is optimised for the extraction of DNA from trace samples. Rhinoceros horns, including museum samples, typically yield high quantities of DNA and thus alternative methods for standard extractions were superior. Both the Qiagen DNEasy Kit and the Thermo Scientific KingFisher method yielded more than sufficient DNA quantity and quality for STR analysis. When extracting from large numbers of samples, the Thermo Scientific KingFisher method has the advantage of being automated. However, contamination was found to have occurred between wells within a single run and thus the machine was not suitable for forensic use. Given the wide usage of this machine amongst laboratories, it is likely that this contamination problem is specific to the machine used, although with such automated machines it is always prudent to use a number of extraction controls. The Qiagen DNEasy kit, on the other hand, uses individually sealed tubes which greatly reduces the possibility of contamination but does increase person time.

Overloading of the Qiagen spin columns is known to reduce DNA yield and is warned against in the kit protocols. These protocols describe appropriate quantities of samples for standard extractions, such as blood or tissue. However, it was important to determine what the optimal mass of rhinoceros horn was. Using only 20 mg of rhinoceros horn generated high DNA yield when using either the Qiagen DNeasy Kit or Thermo Scientific KingFisher Kit, although yield did vary amongst samples as can be expected, particularly for historic samples.

This optimisation suggests that the Qiagen DNEasy Kit offers the most robust method for extraction of DNA from historic rhinoceros horns. Optimal yield can be obtained from 20 mg of horn material digested following standard protocol with the addition of 10 µL of 1M DTT.

4.2 SNPs as an alternative to STRs

Advances in DNA analysis technologies in recent years has made it possible to identify SNPs and genotype them in a reliable high-throughput manor. SNP analysis utilises profiles generated from many single base pair sites, using only short flanking regions on either side of the SNP where primers bind, and analysis is therefore reliant on much shorter DNA fragments that STRs. They are therefore perhaps more likely to amplify from degraded DNA sources (Butler *et al.* 2007), although development of miniSTRs (short length STR fragments) can be very successful with degraded DNA (Dixon *et al.* 2006). SNPs have a much lower mutation rate than STRs (Amorim & Pereira 2005) and are thus very useful for familial testing. Furthermore, the very nature of high throughput technologies used in SNP analysis allows large volumes of data to be obtained rapidly (useful for generating databases of allele frequencies), as well as rapid analysis of samples. The use of SNPs for forensic testing in replacement of STRs has therefore been widely discussed in recent years (Gill 2001a; Amorim & Pereira 2005; Sobrino *et al.* 2005; Dixon *et al.* 2006; Butler *et al.* 2007; Ogden 2011).

For wildlife DNA forensics, STRs are widely available for many species due to their widespread use in population and conservation genetic research (Cassidy & Gonzales 2005; Linacre *et al.* 2011). If STR loci have not previously been identified for a particular species, it is often possible to cross-amplify loci from related species; although this is often associated with an increased rate of null alleles and reduced reproducibility (Cassidy & Gonzales 2005). To obtain similar discriminatory power, approximately three times as many SNP loci are required as STR loci. Typically 40 – 60 SNPs have a similar discriminatory power to 13 – 15 STRs (Chakraborty *et al.* 1999; Gill 2001a). SNPs have not yet, however, become widely used in DNA forensics, even for human systems.

The use of SNPs in forensic analysis is made challenging by both their lack of widespread usage in wild populations and challenges associated with their analysis. Match probability calculations are dependent upon databases of allele frequencies for populations of individuals (Butler 2005). Whilst this is currently restrictive for human DNA forensics as large databases already exist for STRs but would need to be collected for SNPs (Butler *et al.* 2007), it is perhaps somewhat less problematic for wildlife DNA forensic systems as population data is not always available prior to the establishment of the method. However, where STR markers have already been applied in population genetic research, allele frequency data may already exist. Development of a SNP based panel for wildlife species will often be reliant of first identifying a panel of SNPs, whereas STR markers are widely available from population genetic research or cross-amplify another closely related species. This situation may be

changing, however, as SNP chips (microarrays containing thousands probes or primers targeting SNPs) and entire genome sequences become available for a wider range of species.

There are also a number of methodological and analytical challenges to the application of SNPs in forensic systems. Amplification failures of SNPs impact analysis more significantly and are more challenging to deal with than STRs (Butler *et al.* 2007). Furthermore, SNP analysis methods are not yet standardized and currently a wide variety of platforms and therefore chemistries are used, which result in varying raw data and data interpretation methods and software for which standards have not yet been established (Butler *et al.* 2007). The increased number of loci required in SNP panels increase the potential for artefacts, and whilst certain artefacts of STR systems such as stutter would no longer arise, stutter does provide confidence that a true allele is present and is not an artefact in its own right, which would be lost in SNP analysis (Butler *et al.* 2007). Greater expert analysis would be required to deal with the complexities of the software and data interpretation (Butler *et al.* 2007), which may be problematic for wildlife DNA forensic systems as much wildlife crime, and subsequent DNA analysis, occurs in developing countries in which these experts are not always readily available. Perhaps of greatest concern in developing human SNP systems has been the difficulties it presents to mixture analysis (Butler *et al.* 2007). This is caused by the limited number of alleles available at a locus. STR panels use loci with a wide range of loci, thus it is possible to detect the presence of three or four alleles at a locus if a mixture of more than one individual's DNA is present within the sample. For SNPs, however, there is a maximum of four possible alleles (one allele for each of the four bases A, G, T or C), but frequently SNPs are bi-allelic. This greatly reduces the likelihood of non-overlapping alleles and therefore the ability to detect mixtures (Butler *et al.* 2007).

SNPs are, at present, extensively used in both human and wildlife research. Their application to human forensic analyses has been widely debated for a number of years for both humans (Gill 2001a; Amorim & Pereira 2005; Sobrino *et al.* 2005; Butler *et al.* 2007; Budowle & Van Daal 2008) and recently wildlife (Ogden 2011), and indeed are now developed and available for use in a number of specific circumstances (Schwenke *et al.* 2006; Webb & Allard 2009; Walsh *et al.* 2011; Kanthaswamy *et al.* 2012)**.** STRs, however, remain the marker of choice at present (Wictum *et al.* 2013; Iyengar 2014).

4.3 Population genetic potential of this marker panel

All of the markers used in this panel have been previously published and are therefore available for population genetic research. The combination of these 14 markers into two

D. bicornis specific multiplexes which can generate data rapidly and cost effectively has not previously been published. These multiplexes therefore could be useful in generating population genetic data across the species range and in understanding how contemporary genetic variation relates to historical variation.

The population structure of *D. bicornis* across subspecies is still poorly understood. Many studies have assessed genetic diversity within a country (e.g. Muya *et al.* 2011) or within specific populations (e.g. (Garnier *et al.* 2001; Scott 2008; Kim 2009; van Coeverden de Groot *et al.* 2011; Karsten *et al.* 2011). Two studies have begun to address the limited inter-population comparisons needed to understand the wider population structure of the species (Harley *et al.* 2005; Harper *et al.* 2013; Kotzé *et al.* 2014), although either resolution, geographic variation or population analyses were limited. The greatest challenge to genotyping rhinoceros most likely remains the ability to obtain high quality DNA samples; however, the studies noted above demonstrate that many samples have been collected over recent years. Were a connected approach undertaken, using the same markers, it would be possible to begin gaining a clearer picture of the population structure of *D. bicornis* across its range.

Whilst numerous studies have found that high levels of diversity remain within *D. bicornis* despite 95 % population loss (Scott 2008; Kim 2009; van Coeverden de Groot *et al.* 2011; Karsten *et al.* 2011; Muya *et al.* 2011; Kotzé *et al.* 2014), the comparison with historic variation is yet to be made. This study found that horns were a good source of DNA, as did Harper *et al*. (2013), and therefore horns in private and museum collections should be able to act as a DNA record of pre-population decline genetic variation. This could help to identify contemporary populations that have maintained historic variation, allowing management programmes to better maintain that variation. Conversely, it could reveal that populations managed for such purposes need not be. Furthermore, *D. bicornis* occupy a range of habitat types, and these neutral markers may be able to act as a proxy for adaptive variation within those environments, either historically or remaining within contemporary populations. Finally, understanding how genetic variation has changed within a species that has undergone population decline and subsequent expansion may improve our understanding of how genetic diversity changes under such circumstances.

4.4 Future forensic applications of this marker panel

A panel of 13 STR markers and one sexing marker have been selected by this developmental validation for forensic individualisation of *D. bicornis*. Within the UK, it is hoped that this will

be useful in profiling rhinoceros horn in museum collections and of captive rhinoceros. These profiles can then act as a unique record for each individual. Should a crime occur and a horn be recovered in trade, DNA profiling should be able to connect the stolen horn to the crime.

Allele frequency estimates are essential to calculate RMP calculations. During the second half of the 20th Century, 93 % of the *D. bicornis* population was killed (Emslie & Brooks 1999). Such a dramatic decline of a population is likely to result in different allele frequency distributions of populations before and after, and many populations of *D. bicornis* were wiped out completely (Emslie 2012). Profiles generated from historic horns should not be compared to those from contemporary *D. bicornis*, as this is not an appropriate population of potential matches. Instead, profiles should be compared to population data collected from a large number of historic horns.

At present these markers have only been assessed as fit for purpose for *D. bicornis*. Historic collections are likely to include horns from all species of rhinoceros, and captive populations also include *C. simum* and *R. unicornis*. Therefore, it would be beneficial to extend this developmental validation to determine marker suitability for all rhinoceros species. The discriminatory power of a marker set is dependent upon the number of loci and their variability. Only 11 of the 13 STR loci are expected to be polymorphic in *C. simum* (Harper *et al*. 2013; Chapter 2) and locus variability is unknown for Asian species of rhinoceros. In validating a marker panel for other rhinoceros, it may therefore be useful to include a greater number of markers, possibly including the additional seven markers used by Harper *et al*. (2013). A profiling system validated for all rhinoceros species would be a powerful tool for law enforcement, providing a means to trace poached and museum specimens through the black market around the world.

References

Ahvla (2012) *Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) - Additional guidance notes for traders in worked specimens (AHVLA GN7)*.

Alacs EA, Georges A, FitzSimmons NN, Robertson J (2010) DNA detective: A review of molecular approaches to wildlife forensics. *Forensic Science, Medicine, and Pathology*, **6**, 180–194.

Alroy J (2001) A multispecies overkill simulation of the end-Pleistocene megafaunal mass extinction. *Science (New York, N.Y.)*, **292**, 1893–1896.

Amin R, Adcock K, Emslie RH, Mulama M (2006) An integrated management strategy for the conservation of Eastern black rhinoceros in Kenya. *International Zoo Yearbook*, **40**, 118–129.

Amorim A, Pereira L (2005) Pros and cons in the use of SNPs in forensic kinship investigation: A comparative analysis with STRs. *Forensic Science International*, **150**, 17–21.

Anderson JD, Honeycutt RL, Gonzales R a *et al.* (2002) Development of microsatellite DNA markers for the automated genetic characterization of white-tailed deer populations. *Journal of Wildlife Management*, **66**, 67–74.

Andreassen R, Schregel J, Kopatz A *et al.* (2012) A forensic DNA profiling system for Northern European brown bears (Ursus arctos). *Forensic Science International: Genetics*, **6**, 798–809.

Anonymous (1983) Yemen acts to halt rhino horn daggers; scientific tests fail to show rhino horn effective as medicine. *The Environmentalist*, **3**, 153.

Ansell R (2013) Internal quality control in forensic DNA analysis. *Accreditation and Quality Assurance*, **18**, 279–289.

Ashley M V, Melnick DJ, Western D (1990) Conservation genetics of the black rhinoceros (Dicornis bicornis), I: Evidence from the mitochondrial DNA of three populations. *Conservation Biology*, **4**, 71–77.

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

Baillie JEM, Hilton-Taylor C, Stuart SN (2004) *A Global Species Assessment: 2004 IUCN Red List of Threatened Species*. IUCN, Gland, Switzerland and Cambridge, UK.

Bär W, Brinkmann B, Budowle B *et al.* (1997) DNA recommendations. Further report of the DNA Commission of the ISFG regarding the use of short tandem repeat systems. *Forensic Science International*, **87**, 179–184.

References

Berger B, Berger C, Hecht W *et al.* (2014) Validation of two canine STR multiplex-assays following the ISFG recommendations for non-human DNA analysis. *Forensic Science International: Genetics*, **8**, 90–100.

Bhattacharya S (2010) Tracking the rhino killers. *New Scientist*, **208**, 8–9.

Biggs D, Courchamp F, Martin R, Possingham HP (2013) Legal Trade of Africa ' s Rhino Horns. *Science*, **339**, 1038–1039.

Blackburn TM, Cassey P, Duncan RP, Evans KL, Gaston KJ (2004) Avian extinction and mammalian introductions on oceanic islands. *Science (New York, N.Y.)*, **305**, 1955–1958.

Broad S, Mulliken T, Roe D (2003) The Nature and Extent of Legal and Illegal Trade in Wildlife. In: *The Trade in Wildlife: Regulation for Conservation* (ed Oldfield S), pp. 3–22. Earthscan Publications Ltd, London.

Brook S, Van Coeverden De Groot PJ, Mahood S, Long B (2011) *Extinction of the Javan Rhinoceros (Rhinoceros sondaicus) from Vietnam*.

Brook SM, van Coeverden de Groot P, Scott C *et al.* (2012) Integrated and novel survey methods for rhinoceros populations confirm the extinction of Rhinoceros sondaicus annamiticus from Vietnam. *Biological Conservation*, **155**, 59–67.

Brown SM, Houlden BA (1999) Isolation and characterization of microsatellite markers in the black rhinoceros (Diceros bicornis). *Molecular ecology*, **8**, 1559–1561.

Budowle B, Chakraborty R, Carmody G, Monson KL (2000) Source Attribution of a Forensic DNA Profile. *Forensic Science Communications*, **2**, 1–8.

Budowle B, Garofano P, Hellman A *et al.* (2005) Recommendations for animal DNA forensic and identity testing. *International Journal of Legal Medicine*, **119**, 295–302.

Budowle B, Van Daal A (2008) Forensically relevant SNP classes. *BioTechniques*, **44**, 603–610.

Burke T, Bruford M (1987) DNA fingerprinting in birds. *Nature*, **327**, 149–152.

Burney DA, Flannery TF (2005) Fifty millennia of catastrophic rextinctions after human contact. *Trends in Ecology and Evolution*, **20**, 395–401.

Butler JM (2005) *Forensic DNA typing: biology, technology, and genetics of STR markers*. Elsevier, Burlington.

Butler JM (2014) *Advanced Topics in Forensic DNA Typing: Interpretation*. Academic Press, Oxford.

Butler JM, Coble MD, Vallone PM (2007) STRs vs. SNPs: Thoughts on the future of forensic DNA testing. *Forensic Science, Medicine, and Pathology*, **3**, 200–205.

Caratti S, Rossi L, Sona B *et al.* (2010) Analysis of 11 tetrameric STRs in wild boars for forensic purposes. *Forensic Science International: Genetics*, **4**, 339–342.

References

Cassidy BG, Gonzales R a (2005) DNA testing in animal foresics. *Journal of Wildlife Management*, **69**, 1454–1462.

Cerdeno E (1995) Cladistic Analysis of the Family Rhinocerotidae (Perissodactyla). *American Museum Novitates*, **3143**, 1–25.

Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B (1999) The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. *Electrophoresis*, **20**, 1682–1696.

Chambers GK, Curtis C, Millar CD, Huynen L, Lambert DM (2014) DNA fingerprinting in zoology: past, present, future. *Investigative Genetics*, **5**, 3–13.

Cites (2011) *Appendices I, II, and III*. UNEP, Geneva.

CITES Secretariat (2013) *Species trade and conservation: Rhinoceroses. CoP16 Doc. 54.2 (Rev. 1)*.

Comstock KE, Ostrander EA, Wasser SK (2003) Amplifying Nuclear and Mitochondrial DNA from African Elephant Ivory: A Tool for Monitoring the Ivory Trade. *Conservation Biology*, **17**, 1840–1843.

Convention on Biological Diversity (1992) Text of the Convention on Biological Diversity. *Convention on International Trade in Endangered Species of Wild Fauna and Flora*, 28.

Coomber N, David VA, O'Brien SJ, Menotti-Raymond M (2007) Validation of a short tandem repeat multiplex typing system for genetic individualization of domestic cat samples. *Croatian medical journal*, **48**, 547–555.

Corlett RT (2007) The impact of hunting on the mammalian fauna of tropical Asian forests. *Biotropica*, **39**, 292–303.

Cumming DHM, Toit RF Du, Stuart SN (1987) *Africa's elephants and rhinos: status survey and conservation action plan*. IUCN, Gland, Switzerland.

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

Dawnay N, Ogden R, Thorpe RS *et al.* (2008) A forensic STR profiling system for the Eurasian badger: A framework for developing profiling systems for wildlife species. *Forensic Science International: Genetics*, **2**, 47–53.

Dawnay N, Ogden R, Wetton JH, Thorpe RS, McEwing R (2009) Genetic data from 28 STR loci for forensic individual identification and parentage analyses in 6 bird of prey species. *Forensic Science International: Genetics*, **3**, e63–9.

Diamond JM (1989) The present, past and future of human-caused extinctions. *Philosophical Transactions: Biological Sciences*, **325**, 469–477.

Diamond JM, Diamond JM (1989) Overview of recent extinctions. In: *Conservation for the twenty-first century* (eds Western D, Pearl MC), pp. 37–41. Oxford Unversity Press, New York.

Dixon LA, Dobbins AE, Pulker HK *et al.* (2006) Analysis of artificially degraded DNA using STRs and SNPs-results of a collaborative European (EDNAP) exercise. *Forensic Science International*, **164**, 33–44.

DNA Advisory Board (2000) Quality Assurance Standards for Forensic DNA Testing Laboratories. *Forensic Science Communications*, **2**, 1–14.

Eisen JA (2001) Mechanistic basis for microsatellite instability. In: *Microsatellites, Evolution and applications* (eds Goldstein DB, Schlötterer C), pp. 34–48. Oxford University Press, Oxford.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature reviews. Genetics*, **5**, 435–445.

Ellis R (2005) *Tiger Bone & Rhino Horn*. Island Press, Washington, DC.

Emslie R (2012) Diceros bicornis. *IUCN 2012. IUCN Red List of Threatened Species. version 2012.2*. <http://www.iucnredlist.org/>. Downloaded on 11 January 2013.

Emslie R, Brooks M (1999) *Status survey and conservation action plan: African rhino*. IUCN/SSC African Rhino Specialist Group, Gland, Switzerland and Cambridge, UK.

Emslie RH, Knight MH (2014) *Update on African rhino status and trends: from IUCN SSC African Rhino Specialist Group (AfRSG). Report to CITES Standing Committee 65th meeting*.

Emslie RH, Milliken T, Talukdar BK (2013) *African and Asian Rhinoceroses – Status, Conservation and Trade. A report from the IUCN Species Survival Commission (IUCN/SSC) African and Asian Rhino Specialist Groups and TRAFFIC to the CITES Secretariat pursuant to Resolution Conf. 9.14 (Rev. CoP15)*.

Emslie, R.H., Amin, R., Kock R (2009) *Guidelines for the in situ Re-introduction and Translocation of African and Asian Rhinoceros.* (RH Emslie, R Amin, R Kock, Eds,). Iucn.

European Commission (2013) *Guidance document: Export, re-export and intra-Union trade of rhinoceros horns*.

Florescu A, Davila JA, Scott C *et al.* (2003) Polymorphic microsatellites in white rhinoceros. *Molecular Ecology Notes*, **3**, 344–345.

Foose TJ, van Strien NJ (1997) *Asian rhinos - status survey and conservation action plan*. IUCN, Gland, Switzerland and Cambridge, UK.

Foose TJ, Wiese RJ (2006) Population management of rhinoceros in captivity. *International Zoo Yearbook*, **40**, 174–196.

References

Foreman LA, Evett IW (2001) Statistical analyses to support forensic interpretation for a new ten-locus STR profiling system. *International Journal of Legal Medicine*, **114**, 147–155.

Garnier JN, Garnier JN, Bruford MW *et al.* (2001) Matings system and reproductive skew in the black rhinoceros. *Molecular Ecology*, **10**, 2031–2041.

Geerinckx T, Brunain M, Adriaens D (2007) Development of the Osteocranium in the Suckermouth Armored Catfish Ancistrus cf . triradiatus. *Journal of Morphology*, **268**, 254–274.

Gill P (2001a) An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. *International Journal of Legal Medicine*, **114**, 204–210.

Gill P (2001b) Application of low copy number DNA profiling. *Croatian medical journal*, **42**, 229–232.

Gill P, Brinkmann B, D'Aloja E *et al.* (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Forensic Science International*, **87**, 185–192.

Gill P, Buckleon J (2005) Biological basis for DNA evidence. In: *Forensic DNA Evidence Interpretation* (eds Triggs CM, Bckleton JS, Walsh SJ), pp. 16–20. CRC Press, Boca Raton.

Gill P, Jeffreys AJ, Werrett DJ (1985) Forensic application of DNA "fingerprints." *Nature*, **318**, 577–579.

Gill P, Kimpton CP, D'Aloja E *et al.* (1994) Report of the European DNA profiling group (EDNAP) — towards standardisation of short tandem repeat (STR) loci. *Forensic Science International*, **65**, 51–59.

Goor LHPV De (2011) Bovine and equine forensic DNA analysis. University of Utrecht.

Graham-Rowe D (2011) Biodiversity: Endangered and in demand. *Nature*, **480**, S101–S103.

Grogan J, Blundell AG, Landis RM *et al.* (2010) Over-harvesting driven by consumer demand leads to population decline: Big-leaf mahogany in South America. *Conservation Letters*, **3**, 12–20.

Groves CP (1972) Ceratotherium simum. *Mammalian Species*, **8**, 1–6.

Harley EH, Baumgarten I, Cunningham J, O'Ryan C (2005) Genetic variation and population structure in remnant populations of black rhinoceros, Diceros bicornis, in Africa. *Molecular Ecology*, **14**, 2981–2990.

Harper CK, Vermeulen GJ, Clarke AB, De Wet JI, Guthrie AJ (2013) Extraction of nuclear DNA from rhinoceros horn and characterization of DNA profiling systems for white (Ceratotherium simum) and black (Diceros bicornis) rhinoceros. *Forensic Science International: Genetics*, **7**, 428–433.

Heller N (2011) Why Has the Rhinoceros Come from the West? An Excursus into the Religious, Literary, and Environmental History of the Tang Dynasty. *Journal of the American Oriental Society*, **131**, 353–370.

Hellmann AP, Rohleder U, Eichmann C *et al.* (2006) A proposal for standardization in forensic canine DNA typing: Allele nomenclature of six canine-specific STR loci. *Journal of Forensic Sciences*, **51**, 274–281.

Hillman Smith K, Oyisenzoo M, Smith E, Hillman-smith K (1986) A last chance to save the. *Oryx*, **20**, 20–26.

Howard C, Gilmore S, Robertson J, Peakall R (2008) *Application of new DNA markers for forensic examination of Cannabis sativa seizures – Developmental validation of protocols and a genetic database*. National Drugy Law Enforcement Research Fund, Hobart.

Hsieh HM, Huang LH, Tsai LC *et al.* (2003) Species identification of rhinoceros horns using the cytochrome b gene. *Forensic Science International*, **136**, 1–11.

IUCN SSC African Rhino Specialist Group (2008) Ceratotherium simum ssp. simum. *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1*. <http://www.iucnredlist.org/>. Downloaded on 11 January 2013.

Iyengar A (2014) Forensic DNA analysis for animal protection and biodiversity conservation: A review. *Journal for Nature Conservation*, **22**, 195–205.

Jackson JB (2001) What was natural in the coastal oceans? *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5411–5418.

Jacquet J (2009) Silent water: A brief examination of the marine fisheries crisis. *Environment, Development and Sustainability*, **11**, 255–263.

Jeffreys AJ, Brookfield JF, Semeonoff R (1985a) Positive identification of an immigration test-case using human DNA fingerprints. *Nature*, **317**, 818–819.

Jeffreys AJ, Wilson V, Thein SL (1985b) Individual-specific "fingerprints" of human DNA. *Nature*, **316**, 76 – 79.

Jobin RM, Patterson D, Zhang Y (2008) DNA typing in populations of mule deer for forensic use in the Province of Alberta. *Forensic Science International: Genetics*, **2**, 190–197.

Kanthaswamy S, Premasuthan A, Ng J, Satkoski J, Goyal V (2012) Quantitative real-time PCR (qPCR) assay for human-dog-cat species identification and nuclear DNA quantification. *Forensic Science International: Genetics*, **6**, 290–295.

Karsten M, Van Vuuren BJ, Goodman P, Barnaud a. (2011) The history and management of black rhino in KwaZulu-Natal: A population genetic approach to assess the past and guide the future. *Animal Conservation*, **14**, 363–370.

Kim MH (2009) The utility of DNA microsatellite markers in conservation of a Namibian population of the Black rhinoceros (Diceros bicornis). Queen's University.

Kingdon J, Hoffmann M (2013) Family Rhinocerotidae. In: *Mammals of Africa, Volume V* (eds Kingdon J, Happold D, Hoffman M, et al.), p. 444. Bloomsbury, London.

Kotzé A, Dalton DL, du Toit R, Anderson N, Moodley Y (2014) Genetic structure of the black rhinoceros (Diceros bicornis) in south-eastern Africa. *Conservation Genetics*, **15**, 1479–1489.

Kun T, Lyons L a., Sacks BN *et al.* (2013) Developmental validation of Mini-DogFiler for degraded canine DNA. *Forensic Science International: Genetics*, **7**, 151–158.

Kuzmin Y V (2010) Extinction of the woolly mammoth (Mammuthus primigenius) and woolly rhinoceros (Coelodonta antiquitatis) in Eurasia: Review of chronological and environmental issues. *Boreas*, **39**, 247–261.

Lagrot I, Lagrot J-F, Bour P (2008) Probable extinction of the western black rhino, Diceros bicornis longipes: 2006 survey in northern Cameroon. *Pachyderm*, 19–28.

Leader-Williams N (1992) *World trade in rhino horn TRAFFIC review*. TRAFFIC International and the People's Trust for Endangered Species, Cambridge.

Levitus S, Antonov J, Boyer T (2005) Warming of the world ocean, 1955-2003. *Geophysical Research Letters*, **32**, 1–4.

Lin YC, Hsieh HM, Lee JCI *et al.* (2014) Establishing a DNA identification system for pigs (Sus scrofa) using a multiplex STR amplification. *Forensic Science International: Genetics*, **9**, 12–19.

Linacre A, Gusmão L, Hecht W *et al.* (2011) ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Science International: Genetics*, **5**, 501–505.

Linder JM (2008) The Impact of Hunting on Primates in Korup National Park, Cameroon: Implications for Primate Conservation. City University of New York.

Lorenzini R, Cabras P, Fanelli R, Carboni GL (2011) Wildlife molecular forensics: Identification of the Sardinian mouflon using STR profiling and the Bayesian assignment test. *Forensic Science International: Genetics*, **5**, 345–349.

Lyons SK, Smith F a, Brown JH (2004) Of mice, matodons and men: human-mediated extinctions on four cotinents. *Evolutionary Ecology Research*, **6**, 339–358.

Markova AK, Puzachenko AY, Van Kolfschoten T, Van der Plicht J, Ponomarev D V. (2013) New data on changes in the European distribution of the mammoth and the woolly rhinoceros during the second half of the Late Pleistocene and the early Holocene. *Quaternary International*, **292**, 4–14.

Martin EB (1979) *The international trade in rhinoceros products*. IUCN/WWF, Gland, Switzerland.

Martin EB (1981) The conspicuous consumption of rhinos [Part 2]. *Animal Kingdom*, **82**, 20–26.

References

Martin EB, Martin CB (1982) *Run rhino run*. Chatto & Windus, London.

Martin EB, Vigne L, Allan C (1997) *On a Knife's Edge: The Rhinoceros Horn Trade in Yemen*. TRAFFIC.

McCarthy M (2012) British zoos put on alert over rising threat of rhino rustlers. *The Independent*.

Mccord BR, Opel K, Funes M, Jantz LM (2011) An Investigation of the Effect of DNA Degradation and Inhibition on PCR Amplification of Single Source and Mixed Forensic Samples. *Project 2006-DN-BX-K006. Final Report to the US Department of Justice*, 1–66.

McGraw SN, Keeler SP, Huffman JE (2012) Forensic DNA Analysis of Wildlife Evidence. In: *Wildlife Forensics: Methods and Applications* (eds Huffman JE, Wallace JR), pp. 253–269. Wiley-Blackwell, Chichester.

Menotti-Raymond M, David V, Stephens J, Lyons L, O'Brien S (1997) Genetic individualization of domestic cats using feline STR loci for forensic applications. *Journal of forensic sciences*, **42**, 1039–1051.

Menotti-Raymond M, David V, Wachter L, Butler J, O'Brien S (2005) An STR forensic typing system for genetic individualization of domestic cat (Felis catus) samples. *Journal of forensic sciences*, **50**, 1061–1070.

Methods A (2004) Scientific Working Group on DNA Analysis Methods Validation Guidelines for DNA Analysis Methods. , 1–13.

Milledge S (2008) Illegal killing of African rhinos and horn trade, 2000-2005: the era of resurgent markets and emerging organized crime. *Pachyderm*, **43**, 2000–2005.

Milliken T, Shaw J (2012) *The South Africa - Viet Nam Rhino Horn Trade Nexus: A deadly combination of institutional lapses, corrupt wildlife industry professionals and Asian crime syndicates.* Johannesburg, South Africa.

Modell W (1969) Horns and Antlers. *Scientific American*, **220**, 114–122.

Montesh M (2012) *Rhino Poaching: a New Form of Organised Crime*. College of Law Research and Innovation Committee of the University of South Africa.

Moore MK, Kornfield IL (2012) Best Practices in Wildlife Forensic DNA. In: *Wildlife Forensics: Methods and Applications* (eds Huffman JE, Wallace JR), pp. 201–236. Wiley-Blackwell, Chichester.

Murray V, Monchawin C, England PR (1993) The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. *Nucleic acids research*, **21**, 2395–2398.

Muya SM, Bruford MW, Muigai AWT *et al.* (2011) Substantial molecular variation and low genetic structure in Kenya's black rhinoceros: Implications for conservation. *Conservation Genetics*, **12**, 1575–1588.

Nielsen L, Meehan-Meola D, Kilbourn A, Alcivar-Warren A (2008) Characterization of microsatellite loci in the black rhinoceros (Diceros bicornis) and white rhinoceros (Ceratotherium simum): Their use for cross-species amplification and differentiation between the two species. *Conservation Genetics*, **9**, 239–242.

Nowell K (2012) *Assessment of rhino horn as a traditional medicine.* CITES secretariat, Geneva.

Nowell K, Chyi WWL, Pei KCJ (1992) *The Horns of a Dilemma: The market for rhino horn in Taiwan*. Cambridge.

Nybom H, Weising K, Rotter B (2014) DNA fingerprinting in botany: past, present, future. *Investigative genetics*, **5**, 1.

Ogden R (2010) Forensic science, genetics and wildlife biology: Getting the right mix for a wildlife DNA forensics lab. *Forensic Science, Medicine, and Pathology*, **6**, 172–179.

Ogden R (2011) Unlocking the potential of genomic technologies for wildlife forensics. *Molecular Ecology Resources*, **11**, 109–116.

Ogden R, Dawnay N, McEwing R (2009) Wildlife DNA forensics - Bridging the gap between conservation genetics and law enforcement. *Endangered Species Research*, **9**, 179–195.

Okita-Ouma B, Amin B, Kock R (2007) *Conservation and management strategy for the black rhino (Diceros bicornis michaeli) and management guidelines for the white rhino (Ceratotherium simum simum) in Kenya (2007–2011)*. Nairobi.

Pääbo S, Poinar H, Serre D *et al.* (2004) Genetic analyses from ancient DNA. *Annual review of genetics*, **38**, 645–679.

Park SDE (2001) Trypanotolerance in West African cattle and the population genetic effects of selection. University of Dublin.

Parker PG, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: Choosing and using a molecular marker. *Ecology*, **79**, 361–382.

Peakall R, Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peppin L (2009) Development and application of molecular tools for the analysis of animal parts and derivatives. University of Wales, Bangor.

Peppin L, McEwing R, Ogden R *et al.* (2010) Molecular sexing of African rhinoceros. *Conservation Genetics*, **11**, 1181–1184.

Prothero DR, Guérin C, Manning E (1989) The history of the Rhinocerotoidea. In: *The Evolution of Perissodactyls* (eds Prothero DR, Schoch R), pp. 321–340. Clarendon Press & Oxford University Press, New York & London.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rhino A, Groups S, Milliken T, Emslie RH, Talukdar B (2009) *African and Asian Rhinoceroses – Status , Conservation and Trade*. Geneva, Switzerland.

Roewer L (2013) DNA fingerprinting in forensics: past, present, future. *Investigative genetics*, **4**, 22.

Rookmaaker K (2000) The alleged population reduction of the Southern White Rhinoceros (Ceratotherium simum simum) and the successful recovery. *Säugetierkundliche Mitteilungen*, **45**, 55–70.

Rovero F, Mtui AS, Kitegile AS, Nielsen MR (2012) Hunting or habitat degradation? Decline of primate populations in Udzungwa Mountains, Tanzania: An analysis of threats. *Biological Conservation*, **146**, 89–96.

Rütten KB, Schulz I, Olek K, Uhl G (2001) Polymorphic microsatellite markers in the spider Pholcus phalangioides isolated from a library enriched for CA repeats. *Molecular Ecology Notes*, **1**, 255–257.

Schneider P (2004) STR analysis of artificially degraded DNA—results of a collaborative European exercise. *Forensic Science International*, **139**, 123–134.

Schwenke PL, Rhydderch JG, Ford MJ, Marshall AR, Park LK (2006) Forensic identification of endangered Chinook Salmon (Oncorhynchus tshawytscha) using a multilocus SNP assay. *Conservation Genetics*, **7**, 983–989.

Scott C (2008) Microsatellite Variability in Four Contemporary Rhinocerous Species: Implications for Conservation. Queen's University.

Scott C, Foose T, Morales JC *et al.* (2004) Optimization of novel polymorphic microsatellites in the endangered Sumatran rhinoceros (Dicerorhinus sumatrensis). *Molecular Ecology Notes*, **4**, 194–196.

Shaw CN, Wilson PJ, White BN (2003) a Reliable Molecular Method of Gender Determination for Mammals. *Journal of Mammalogy*, **84**, 123–128.

Shinde D, Lai Y, Sun F, Arnheim N (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Research*, **31**, 974–980.

Slikas B (2003) Hawaiian Birds: Lessons From a Rediscovered Avifauna. *The Auk*, **120**, 953.

Sobrino B, Brión M, Carracedo A (2005) SNPs in forensic genetics: A review on SNP typing methodologies. *Forensic Science International*, **154**, 181–194.

Sparkes R, Kimpton C, Gilbard S *et al.* (1996) The validation of a 7-locus multiplex STR test for use in forensic casework. (II) Artefacts, casework studies and success rates. *International Journal of Legal Medicine*, **109**, 195–204.

Steadman DW (1995) Prehistoric extinctions of pacific island birds: biodiversity meets zooarchaeology. *Science (New York, N.Y.)*, **267**, 1123–1131.

Stuart AJ, Lister AM (2012) Extinction chronology of the woolly rhinoceros Coelodonta antiquitatis in the context of late Quaternary megafaunal extinctions in northern Eurasia. *Quaternary Science Reviews*, **51**, 1–17.

SWGWILD (2012) *SWGWILD Standards and Guidelines (Version 2.0)*.

Takahashi M, Kato Y, Mukoyama H, Kanaya H, Kamiyama S (1997) Evaluation of five polymorphic microsatellite markers for typing DNA from decomposed human tissues - Correlation between the size of the alleles and that of the template DNA. *Forensic Science International*, **90**, 1–9.

Talukdar BK, Emslie R, Bist SS *et al.* (2008) Rhinoceros unicornis. *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1*.

Thapa K, Nepal S, Thapa G, Bhatta SR, Wikramanayake E (2013) Past, present and future conservation of the greater one-horned rhinoceros Rhinoceros unicornis in Nepal. *Oryx*, **47**, 345–351.

The Aspinall Foundation (2013) Howletts & Port Lympne Wild Animal Parks Under Threat From Poachers.

Thomas R (2010) Surge in rhinoceros poaching in South Africa. *Traffic Bulletin*, **23**, 3.

Thomas CD, Cameron A, Green RE *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.

Tobe SS, Linacre A (2008) A multiplex assay to identify 18 European mammal species from mixtures using the mitochondrial cytochrome b gene. *Electrophoresis*, **29**, 340–347.

Tougard C, Delefosse T, Hänni C, Montgelard C (2001) Phylogenetic relationships of the five extant Rhinoceros species (Rhinocerotidae, Perissodactyla) based on mitochondrial cytochrome b and 12S rRNA genes. *Molecular phylogenetics and evolution*, **19**, 34–44.

UNODC (2012) *Bulletin on organised crime in Southern Africa*. United Nations Office on Drugs and Crime Southern Africa, Pretoria.

Urquhart A, Kimpton CP, Downes TJ, Gill P (1994) Variation in short tandem repeat sequences--a survey of twelve microsatellite loci for use as forensic identification markers. *International journal of legal medicine*, **107**, 13–20.

van Coeverden de Groot PJ, Putnam AS, Erb P *et al.* (2011) Conservation genetics of the black rhinoceros, Diceros bicornis bicornis, in Namibia. *Conservation Genetics*, **12**, 783–792.

Van De Goor LHP, Panneman H, Van Haeringen W a. (2010) A proposal for standardization in forensic equine DNA typing: Allele nomenclature for 17 equine-specific STR loci. *Animal Genetics*, **41**, 122–127.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

van Strien NJ, Manullang, Sectionov B *et al.* (2008) Dicerorhinus sumatrensis. *IUCN Red List of Threatened Species. Version 2012.2*. <http://www.iucnredlist.org/>. Downloaded on 28 September 2013.

van Strien NJ, Steinmetz R, Manullang B *et al.* (2013) Rhinoceros sondaicus. *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1*. <http://www.iucnredlist.org/>. Downloaded on 28 September 2013

Van NDN, Tap N (2008) *An Overview of the Use of Plants and Animals in Traditional Medicine*. TRAFFIC Southeast Asia, Viet Nam.

Vanleeuwe H, Gautier-Hion A, Cajani S (1997) Forest clearings and the conservation of elephants (Loxodonta africana cyclotis) north-east Congo Republic. *Pachyderm*, **24**, 46–52.

Vigne L, Martin E (2008) Yemen's attitudes towards rhino horn and jambiyas. *Pachyderm*, **44**, 45–53.

Viscardi P (2012) The Horns of a Dilemma: The Impact of the Illicit Trade in Rhino Horn. *The Natural Science Collections Association News*, 8–13.

Walsh S, Lindenbergh A, Zuniga SB *et al.* (2011) Developmental validation of the IrisPlex system: Determination of blue and brown iris colour for forensic intelligence. *Forensic Science International: Genetics*, **5**, 464–471.

Warchol GL (2004) The Transnational Illegal Wildlife Trade. *Criminal Justice Studies*, **17**, 57–73.

Webb KM, Allard MW (2009) Identification of forensically informative SNPs in the domestic dog mitochondrial control region. *Journal of Forensic Sciences*, **54**, 289–304.

Weir BS (2012) Statistics for Wildlife Forensic DNA. In: *Wildlife Forensics: Methods and Applications* (eds Huffman JE, Wallace JR), pp. 237–252. Wiley-Blackwell, Chichester.

Wetton JH, Carter RE, Parkin DT, Walters D (1987) Demographic study of a wild house sparrow populationby DNA fingerprinting. *Nature*, **327**, 147–149.

Whitaker JP, Clayton TM, Urquhart AJ *et al.* (1995) Short tandem repeat typing of bodies from a mass disaster: High success rate and characteristic amplification patterns in highly degraded samples. *BioTechniques*, **18**, 670–676.

White NE, Dawson R, Coghlan ML *et al.* (2012) Application of STR markers in wildlife forensic casework involving Australian black-cockatoos (Calyptorhynchus spp.). *Forensic Science International: Genetics*, **6**, 664–670.

Wictum E, Kun T, Lindquist C *et al.* (2013) Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework. *Forensic Science International: Genetics*, **7**, 82–91.

Wiliam Andrew Laurie (1978) The Ecology and Behaviour of the Greater One-Horned Rhino: Ph.D. Thesis. Cambridge University.

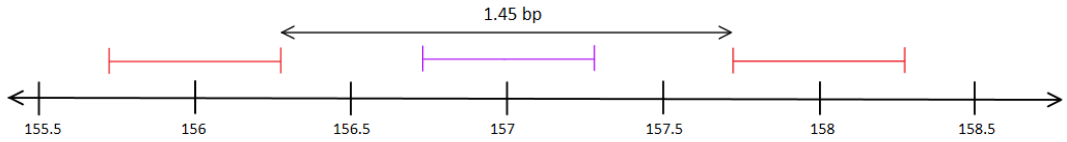
References

Wilson-Wilde L (2009) *Animal Investigators – How the World's First Wildlife Forensics Lab is Solving Crimes and Saving Endangered Species*. Scribner, New York.

Zschokke S, Gautschi B, Baur B (2003) Polymorphic microsatellite loci in the endangered Indian rhinoceros, Rhinoceros unicornis. *Molecular Ecology Notes*, **3**, 233–235.

Appendix I – D-distance schematic

Fragment size (bp) is shown along the black arrow at 0.5 bp intervals. The largest d-value observed throughout this validation was 0.55 bp. This is the distance between the largest and smallest of any measured fragment for an allele and thus fragments have been observed at 0.275 bp above the median and 0.275 bp below the median. For example an allele with a median of 156 bp may be observed anywhere between 155.725 bp and 156.275 (as shown by the red bar above 156 on the diagram). All markers used in this marker panel are dinucleotides (as shown by red bars), and therefore a minimum of 1.45 bp should be expected between the largest observed fragment for the shorter of two alleles and the smallest observed fragment of the larger allele (red bar above 158 bp). If two alleles were only a single bp apart, e.g. allele 156 and 157, then the smallest distance between the largest and smallest fragments of the two alleles could be expected to be 0.45 bp (the distance between red and purple bars).



Appendix II - Genotypes observed from two different starting materials

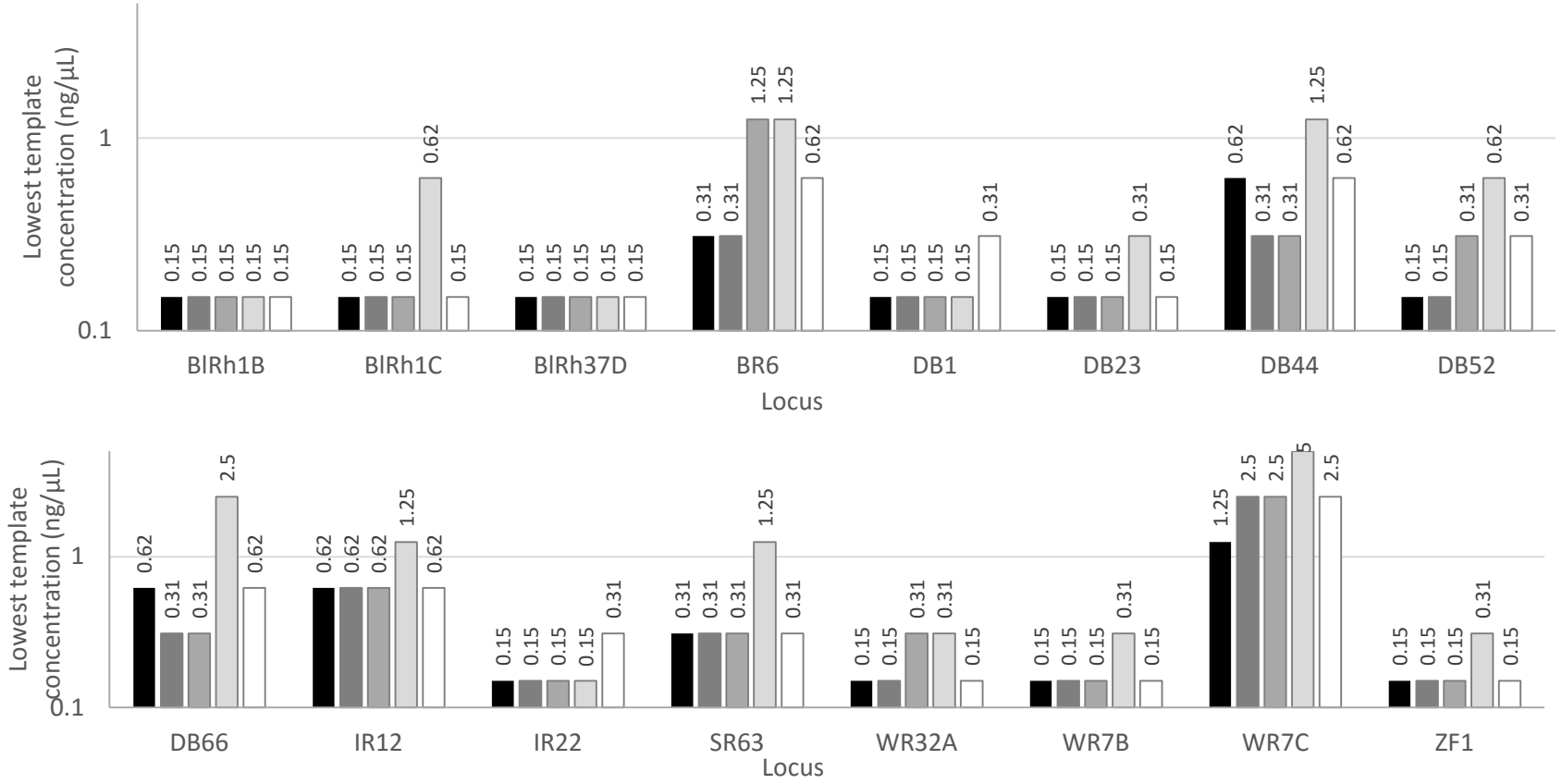
Genotypes observed from horn and blood for one *D. b. michaeli* (BR) and one *C. s. simum* (WR). WR samples were amplified in both singleplex and multiplex. *C. s. simum* alleles have not yet been characterized and assigned nomenclature, thus bp size is given. BlRh37D were only amplified in WR for the blood sample in singleplex. All other gaps in the table represent amplification failures. All genotypes were concordant. The heterozygous alleles for DB52 for WR horn in singleplex displayed heterozygote imbalance (0.82), such that the genotype could be mistakenly identified as homozygous.

	BlRh1B		BlRh1C		BlRh37D		BR6		DB1		DB23		DB44		DB52	
Allele	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
BR blood (singleplex)	18	20	25.1	28.1			20	20	14	16	12	12	24.1	24.1	18.1	23.1
BR horn (multiplex)	18	20	25.1	28.1	23.1	23.1	20	20	14	16	12	12	24.1	24.1	18.1	23.1
WR blood (singleplex)	242	244	133	135	243	243	133	133	157	159	236	236	217	217	217	219
WR blood (multiplex)	242	244	133	135	-	-	133	133	157	159	236	236	217	217	217	219
WR horn (singleplex)			133	135	-	-	133	133	157	159	236	236			217	219
WR horn (multiplex)	242	244	133	135	-	-	133	133	157	159	236	236	217	217	217	219

	DB66		IR12		IR22		Sex		SR63		WR32A		WR7B		WR7C	
Allele	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
BR blood (singleplex)	28	28			20	20	98	98	19	20	22	22	18.1	18.1		
BR horn (singleplex)	28	28	25	26	20	20	98	98	19	20	22	22	18.1	18.1	20	33
WR blood (singleplex)	204	204	180	180	208	208	98	105	189	189	197	199	225	227	152	152
WR blood (multiplex)	204	204	180	180	208	208	98	105	189	189	197	199	225	227		
WR horn (singleplex)	204	204	180	180	208	208	98	105	189	189	197	199	225	227		
WR horn (multiplex)	204	204	180	180	208	208	98	105	189	189	197	199	225	227		

Appendix III - Lowest concentration of template DNA from which each locus could be scored for each individual

Bar colours reflect sample in order of increasing lightness for KA, K06, K10, K14 and K35. Actual concentration (ng/µL) of template DNA is shown above the bar.



Appendix IV – Genotypes from population study

Population study genotypes in bp denoted as allele 1 / allele 2. Amplification failures following multiple attempts are denoted as 0/0.

Sample	BlRh1B	BR6	DB1	DB44	IR12	IR22	SR63	ZF1
K02	236/250	145/145	155/161	205/217	182/182	210/223	193/193	98/105
K05	240/244	143/145	157/161	205/217	174/174	223/223	197/201	98/105
K06	244/244	141/143	161/161	211/217	174/188	221/223	193/201	98/98
K07	244/246	141/141	153/161	205/217	182/182	210/221	193/193	98/105
K08	240/250	143/143	153/153	217/219	182/182	223/223	193/193	98/105
K10	244/244	143/145	153/153	209/217	184/186	221/223	193/193	98/98
K12	244/244	143/145	153/155	217/219	157/157	221/223	193/193	98/105
K13	236/244	143/145	153/155	217/219	174/174	221/223	193/193	98/105
K14	246/246	141/143	153/161	217/217	157/157	223/223	193/201	98/105
K15	236/248	143/145	155/155	215/225	174/188	223/223	193/201	98/98
K17	244/246	141/145	161/161	217/217	174/174	219/223	193/197	98/98
K18	236/244	143/145	153/155	217/219	174/176	223/223	193/193	98/98
K20	236/246	141/145	155/155	215/217	168/168	223/223	193/193	98/105
K22	244/250	143/145	155/161	219/219	174/188	210/223	193/193	98/98
K23	236/236	145/145	155/161	215/219	174/174	221/223	193/201	98/105
K29	240/244	143/145	153/157	205/217	184/184	223/223	193/201	98/98
K33	244/248	141/141	153/153	217/217	174/188	221/229	193/193	98/98
K35	242/246	141/145	153/155	205/217	174/174	219/221	193/193	98/105
K38	244/250	141/143	153/155	215/225	188/188	223/223	193/193	98/105
K47	0/0	143/145	153/159	205/217	174/174	223/223	197/203	98/105
K50	240/248	143/145	153/161	213/219	174/174	221/223	193/193	98/105
K51	240/244	143/143	157/157	205/213	184/184	223/223	201/201	98/105
K52	244/244	145/145	153/157	205/219	184/184	223/223	193/201	98/105
K53	244/248	143/143	157/159	213/217	174/174	223/223	201/203	98/105
K55	236/244	141/145	155/161	205/215	157/157	223/223	193/201	98/105
K56	244/244	141/145	153/161	217/219	174/174	223/223	193/201	98/98
K57	244/248	141/141	153/155	205/211	157/157	221/223	193/201	98/105
K58	244/248	141/145	153/161	215/219	157/157	223/223	199/201	98/105
K59	244/244	143/145	155/155	217/217	174/186	219/221	193/201	98/98
K60	236/246	143/145	153/157	217/217	174/182	221/223	193/197	98/98
K61	244/246	145/145	161/161	217/219	174/176	223/223	197/201	98/98
K62	244/244	143/145	155/155	217/219	186/186	223/223	193/193	98/105
K63	236/248	141/145	155/161	211/217	188/188	223/223	193/201	98/105
K65	248/250	143/145	153/155	205/217	174/182	210/210	193/193	98/98
K67	236/244	143/145	155/161	217/219	186/186	221/221	193/201	98/105
K68	236/244	145/145	153/155	217/219	174/186	223/223	193/193	98/98
K70	242/246	141/145	161/161	217/217	174/174	221/223	193/201	98/98

Appendix IV

Sample	BlRh1B	BR6	DB1	DB44	IR12	IR22	SR63	ZF1
K71	244/246	141/147	161/161	217/219	174/174	221/223	193/201	98/105
K72	236/246	145/147	157/157	213/219	174/174	219/221	201/201	98/98
K73	244/246	143/145	157/161	219/219	174/174	221/223	193/193	98/105
K74	244/246	145/145	157/161	217/219	174/176	219/221	201/201	98/98
K75	244/246	143/143	153/153	217/219	174/174	210/221	193/193	98/105
K76	240/248	145/147	161/161	217/217	188/188	223/223	201/201	98/105
K77	246/248	143/145	153/161	205/217	174/174	221/223	193/197	98/98
K78	236/248	141/147	153/161	217/219	188/188	221/223	193/197	98/105
K80	244/244	145/145	153/159	213/219	174/174	223/223	193/193	98/98
K82	244/244	141/145	153/161	205/219	174/188	221/221	193/201	98/98
K83	244/244	145/145	149/161	205/219	174/186	223/223	193/197	98/98
K85	244/246	143/145	161/161	213/217	176/176	221/223	193/197	98/105
K86	244/248	143/145	153/159	213/217	174/174	223/223	199/203	98/105
K87	244/250	143/143	153/161	205/217	174/174	223/223	197/197	98/105
K89	250/250	143/143	153/157	205/205	174/174	223/223	197/203	98/98

Sample	BlRh1C	BlRh37D	DB23	DB52	DB66	WR32A	Locus7B	Locus7C
K02	123/129	0/0	247/247	215/217	206/206	197/201	223/223	165/165
K05	121/123	201/201	249/249	217/219	196/200	197/205	223/223	137/137
K06	121/123	203/203	241/249	211/221	202/208	197/199	223/223	163/165
K07	125/129	201/203	247/247	223/225	206/206	197/201	223/223	163/165
K08	123/129	201/203	247/247	211/223	200/206	201/209	223/223	137/137
K10	121/129	201/201	241/247	213/219	196/200	199/207	223/225	137/137
K12	125/127	201/201	247/249	211/213	192/192	205/205	225/225	137/166
K13	127/135	0/0	247/247	211/219	182/206	203/205	223/223	161/165
K14	127/135	201/201	247/247	211/213	196/208	199/203	223/223	161/165
K15	129/135	201/207	247/249	213/215	182/208	203/203	223/223	163/165
K17	125/133	201/201	247/247	211/211	192/192	203/205	223/223	167/167
K18	135/135	201/201	247/247	211/219	0/0	199/205	223/223	161/161
K20	121/137	0/0	247/249	211/213	192/200	197/201	223/223	137/161
K22	125/127	207/207	247/247	211/217	200/206	203/207	223/225	161/165
K23	125/135	201/201	247/249	211/215	192/196	199/199	223/223	161/167
K29	121/123	201/201	247/249	219/219	200/204	199/205	223/223	137/166
K33	121/123	203/207	247/249	217/219	202/202	199/207	223/223	161/163
K35	125/125	201/201	241/249	211/221	202/208	197/199	223/223	161/171
K38	125/135	207/207	247/247	215/215	182/206	199/207	223/223	137/165
K47	135/135	201/201	247/249	213/213	182/208	205/209	223/223	0/0
K50	121/129	201/201	241/247	213/219	182/200	205/209	223/223	137/137
K51	123/123	201/201	247/247	213/225	182/208	197/209	223/223	137/165
K52	123/123	0/0	241/249	217/219	200/200	197/205	223/223	137/137
K53	123/123	201/201	247/253	213/219	192/208	197/209	223/223	137/163
K55	125/137	201/201	247/249	211/217	200/202	201/201	223/223	137/165
K56	129/133	201/207	247/249	217/219	202/206	197/203	223/223	0/0
K57	121/125	201/201	249/249	211/213	200/202	197/209	223/225	0/0

Appendix IV

Sample	BlRh1C	BlRh37D	DB23	DB52	DB66	WR32A	Locus7B	Locus7C
K58	123/125	201/201	247/249	211/211	192/202	201/209	223/225	161/165
K59	127/129	201/203	247/247	211/211	182/208	201/205	223/223	0/0
K60	125/135	201/201	247/247	215/219	196/204	199/199	223/223	137/161
K61	125/127	201/201	247/247	211/213	182/192	199/205	223/225	167/167
K62	125/125	201/201	241/249	211/211	182/192	199/205	223/225	0/0
K63	125/137	203/207	247/247	213/217	192/196	197/197	223/223	161/165
K65	123/129	0/0	247/249	217/219	202/206	197/201	223/223	161/165
K67	127/135	201/203	247/247	211/211	196/208	199/201	223/223	161/165
K68	125/125	201/201	249/249	211/213	192/208	199/205	223/223	137/165
K70	125/125	201/201	241/253	211/217	192/192	201/203	223/223	137/167
K71	125/127	201/201	247/247	211/213	182/192	199/199	223/223	137/167
K72	125/127	201/201	247/253	215/215	182/192	199/203	223/223	165/167
K73	127/129	201/201	247/247	213/213	182/192	199/199	223/225	165/169
K74	125/125	203/203	247/247	211/213	182/182	199/205	223/225	161/167
K75	121/123	207/207	247/247	213/217	206/208	197/207	223/223	163/165
K76	135/135	201/207	247/249	213/225	192/208	197/197	223/225	161/166
K77	121/125	201/207	241/247	213/221	202/208	207/207	223/223	163/163
K78	125/129	207/207	249/249	221/225	196/208	197/199	223/223	161/161
K80	121/127	207/207	249/249	219/223	202/208	197/197	223/223	161/166
K82	125/127	201/203	241/249	213/219	202/204	199/205	223/223	137/161
K83	121/125	201/201	247/247	207/211	196/208	205/205	223/223	167/169
K85	125/125	201/201	247/249	211/211	192/196	199/203	223/223	161/167
K86	123/135	201/201	247/247	213/223	208/208	209/209	223/223	137/137
K87	123/123	201/201	249/249	213/217	200/200	197/205	223/223	137/137
K89	123/135	201/201	249/249	213/219	182/200	197/205	223/223	137/167

Appendix V – Summarized locus details

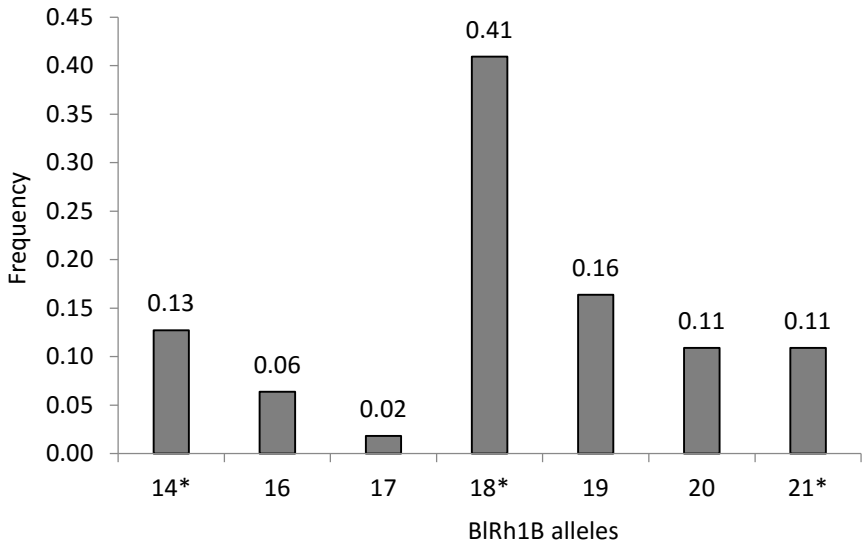


Figure S1
BlRh1B. Three alleles were sequenced (14, 18, 21). All alleles displayed the compound dinucleotide repeat structure $(GT)_nGCA(TG)_{3.1}$. Variation in repeat unit number was only observed for the $(GT)_n$ dinucleotide unit. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with GenBank accession AY606078. Sequence on GenBank:
CTAGGAAATTATTTGGGGAAACAAACCCATATAACAATGTCAAGAAGGGAAAAAAAGAGCAATCTGTATC
TGATTCTGCTCCCCAGTGATCAGTAACACCAAAGTCC$(GT)_{13}$GCA$(TG)_{3.1}$AGGCAGGGGGACATTGGTCA
CAACTAACACATGAAGGAATGAGAAAATCCCATGTTGAGTTGAACAGTAATAGGGTATACAACAATTTCT
AATATTGCATTTGGATAATGCTAGTCTTGATTTCTTTTTGTTGTATATTCTCCATTTTTCTCATATTATGTCC
TTGGTGATCCTTCTGTCTTCACTACATGGTTTGTGTAAAGTGTTCTCATCATCTTCCATGTCTTTACGTGT
TGTGGTTGTGGCATGGTAGTGATTTGCAAACCTTTATCTCCACATCAAAATCTCTCTCTTGGGTTCCAGAT
ATATGTTTGAATATCTCTTAGATAGAATCCTCTGTCTGTACATTATATTGTAAGGATTAAATGAGACAAGTT
GTGTGA

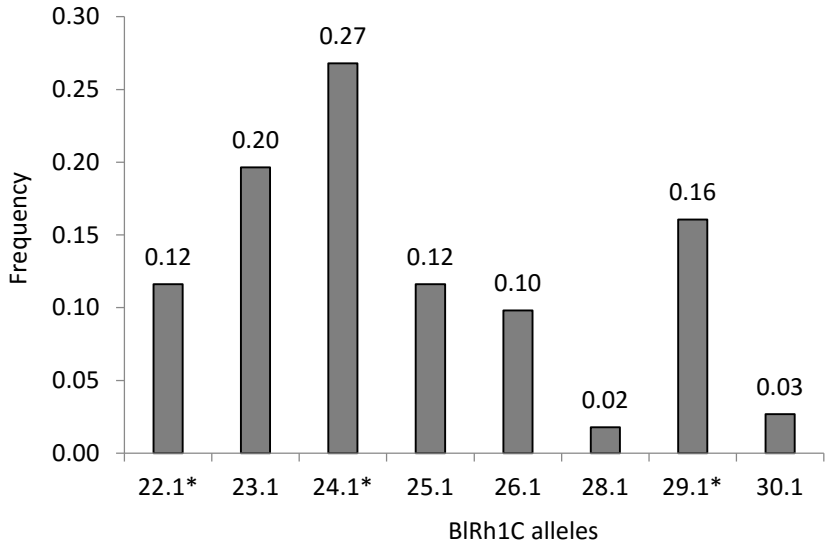


Figure S2
BlRh1C. Three alleles were sequenced (22.1, 24.1, 29.1). All alleles displayed the compound repeat structure $(AC)_n(GCAC)_m(AC)_{3.1}$. Variation in repeat unit number was observed at both the dinucleotide and tetranucleotide repeat units. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AY606079. Sequence on GenBank:
CCTGAAAATCCAAACGCTGCGTGAGATTCTTGGAAAGGTCACTCAAGGCACACAGAAAATGCCTCCT(AC

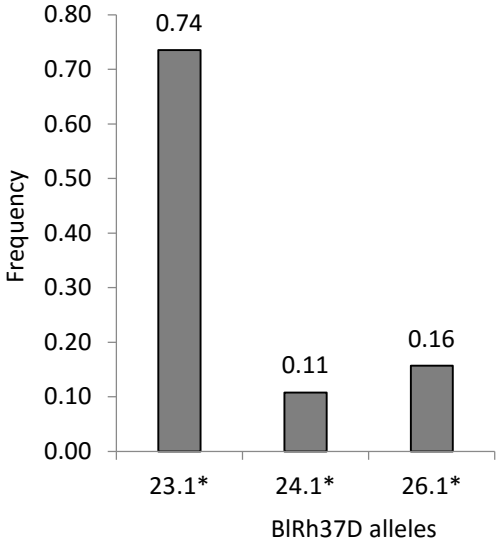


Figure S3
BlRh37D. Three alleles were sequenced (23.1, 24.1, 26.1). All alleles displayed the compound dinucleotide repeat structure $(TG)_6(AG)_nGA(AG)_{5.1}$. Variation in repeat unit number was only observed for the first $(AG)_n$ repeat unit. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AY606083. Sequence on GenBank:
CTTTTAGGTAGAGAATGTATACGGTAGTCATCCCTTATCTGTGGTTTTGCTTTCTGCTGTTTCAGTTACCT
GCGGTCAACTGCAGTCTGAAAATATTAAATGGAAAATTCCAGAAATACACAATTCATAAGTTTTAAGTTGC
ACACCATTTTGAGTAGCGTGATGAAATCTCTCATAGTTCCCCCTGGACGTGAATCATCTCTTTGGATCCA
CACTGTATACAATACCTGACTGTTAGTCACTTAGCAGCCAATCAGAGCCCTTCCCTAGGATTTTGCAAAT
GGAACCAAGGTTGGTAGAAACTGGACATGTGTAAACTTGGGAACTGTTATTAGTCATGTTTCTCCACTCA

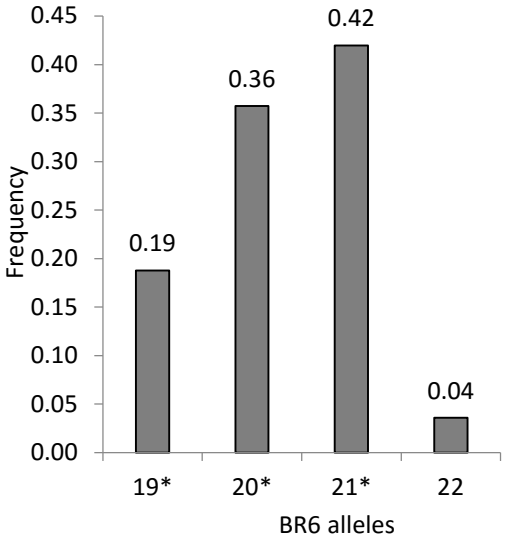


Figure S4
BR6. Three alleles were sequenced (19, 20, 21). All alleles displayed the simple dinucleotide repeat structure $(CA)_n$. No intermediate alleles were observed. No sequence data is available on GenBank for this locus, however the repeat motif is concordant with that originally described by Cunningham, Harley, and O'Ryan (1999).

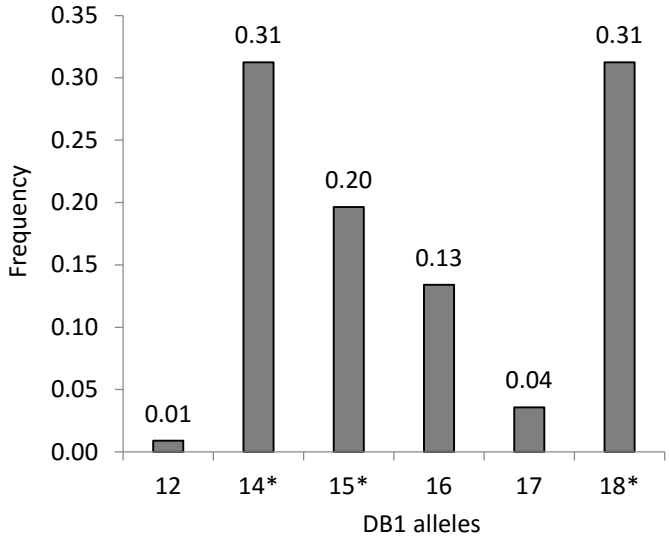


Figure S5
DB1. Three alleles were sequenced (14, 15, 18). All alleles displayed the simple dinucleotide repeat structure $(CA)_n$. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AF129724. Sequence on GenBank:
GGTTTGATAACAATTCTTTGAAGATTAGGTGTAAGTCATAGAGTGCTACACACACCATTCGTTGCCTCACT
AAATAGAAGAGGAGACAGACAAACGAAGAGTTCAATACACAGGAAGGGAAACCTACCTTAAGTAGACGA
CTAGAAGAGTGTCACACAGCGCTCGTAAACTTTGAGCTTCACTTGTCTCACCTTTGTGAAAAAAGTGACA
AAGAGAAAAGGAACATAAAAAGACGGTTAAAAACAGGTCTATACTCCAAATTTCATAGTTTGTAGTCATTA
AGTCCCCTTTTTTCAACTTTTGTTACCTTCGTTTGTAATAACTCCCAAATAACACTTACTCCGTTTCATAAA
ATGTATATAACAGAGTAGACCACCAGTATTGTTAGGTGACTT(CA)$_{15}$ACCCTCGTCCCAGGATAATAATAG
AGTTAAAATGTCTAATCAGGGACACTGAATCTCTACATTTCATTGGAAGTTCCAGTGTGTCGANCACTAAA
CGTCGCGATCCTGACGTCGGTCCGTCAGAATTAGTGATGCGATGTGATACAACGACTTTACTGAAATTTA
ATGAATCAAGTCACTTTCCATTCAACGGATGTAACAATTGCGTACGGTATATTACGAGTCTATAAGTTTAG
TTTTATTGAGTTAACTAGAGAACCTTTCTACGAGTCTTATTTTCTCGGAGACGGACCACTTAGTGTGCTCT
TACACTAAGTCCTTCTTCACTTCTTGTGGCGGTGAACCGTTATCGTGTTGTAATGAAAGACCGTCTTAAG
ACCCCCCCA

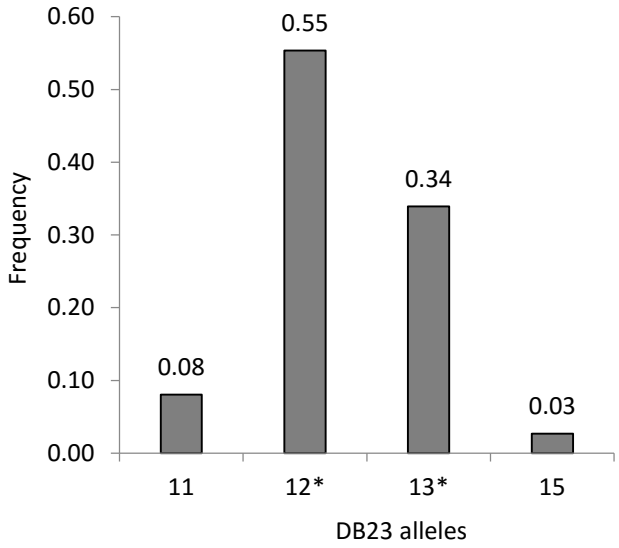


Figure S6
DB23. Two alleles were sequenced (12, 13). Both alleles displayed the simple dinucleotide repeat structure $(CA)_n$. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AF129734. Sequence on GenBank:
CCAATCTTCCTCAGCAATAAGGGGAGGATTAGCAACGGATGTTAGCTCAGGGCTAATCTTCCT$(CA)_{12}$AA
TTCATTATAAATTTAACTCAGAGTAAACAGATGGCTTCATTCTCCTTCAATCCGTAGAGGAATTGAGAAAT
AGNCCCAAACTCAGGGGCAGAGAATCAAGTTTAGAATATGAAAAATGTCTAGCAAAATNNCTCTTTTACA

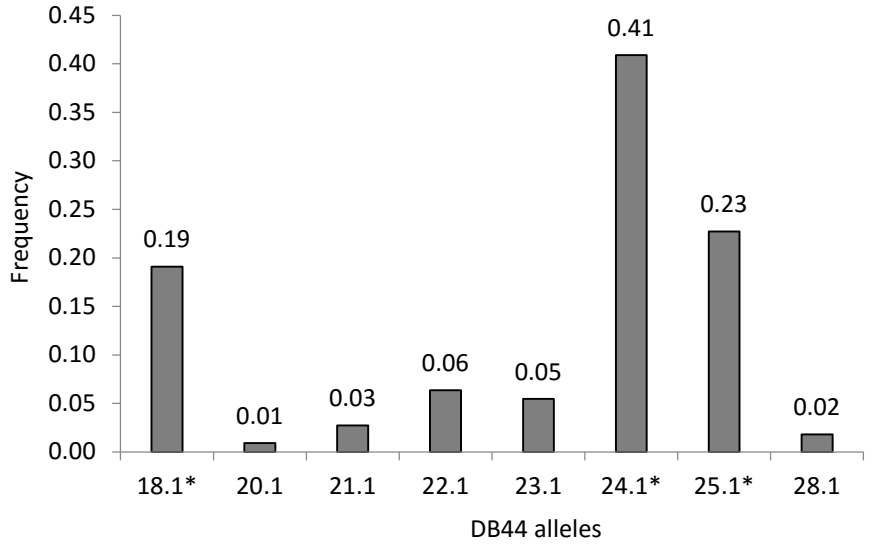


Figure S7
DB44. Three alleles were sequenced (18.1, 24.1, 25.1). All alleles displayed the compound dinucleotide repeat structure $(AC)_6AN^*(AC)_nG(CA)_4$. Variation in repeat unit number was observed only for the second (AC) dinucleotide unit, however G/C polymorphism was also observed at N*, although this does not affect fragment length nor, therefore, allele designation. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AF129730 except for the polymorphism observed at N*. Sequence on GenBank:
CCCTTCTAGAGGGAGACTAGGAGTCCCTGCCCCCAGCTTGTTGCCCCATCCCTGCCTCTGCCACCTTC
CA$(AC)_{16}G(CA)_4$GAATGGAGGTGGGGCTCCCAGCGCCAGCTCAGCCACAGGCAAATTAAAGAAACTGGA
AACGTGGAAAATATCAGCCCCGCTACTTGACATTCCACCCTTATGCTTCTAATCCCGAAGCCATGCTTTC
TGCTTCGGTCTTTCAGATCCCTTTGGAATCACGCCACCTTCCATTCACATCGAATGACTTGTCACCTCCC
CAGCACACCCGGGAGCCTGCCGGTGTCCGCACACCTCTCTCCGTGCAATTCTCTGCCCCTCCTCACCC
CGG

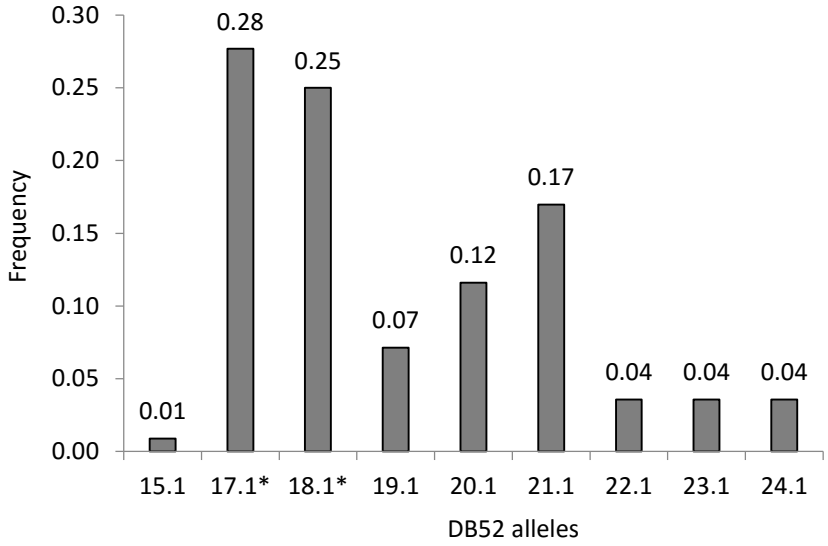


Figure S8

DB52. Two alleles were sequenced (17.1, 18.1). Both alleles displayed the simple dinucleotide repeat structure $(CA)_n$. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AF129732. Sequence on GenBank:

CTCCGAGTGAAAACAGACGTCACATGAAAATTCATCCTCAAGGGGTGGGAGGTGCTTGTGCAGGCGCG
CAGGCTGGCGGGGCGAGCGGGCGGAGGGACCAGTGTGTCTCCCTTGCTTAGACACCCAGTGGCACCA
GCAGACATTTCCTATCCCCATCTCTCGGTTTACAGGAAATGCCTGCTTGAGACAGACATGTCCTTTAACC
CCTCATGTGAAATGGACCGTCAGGCATTGGCAGGAAGCAGCTTGGAGGTTTCCACAGGACGCGTAGCT
CCTGGCTTGGCAGGTGTGGTGGCTTGTGTATGTGATAGCAG$(CA)_{21.1}$GAGCCAGGGGCACCCCCACTTA
CCCTGGGCTCTTTTGGAAATATCCATGCCCCCTGCCCCTTCCCAGAAATTCCATTCAGTGGCTATGGAG
GAACCCCATCCCTGTGTATTATGGATTACCATCCCGTTTTACAGATGAGGAAACTCACGG

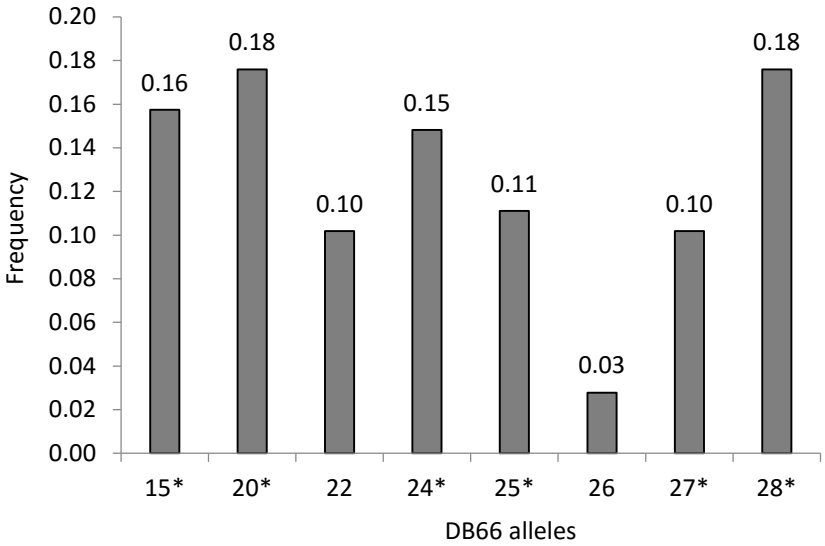


Figure S9

DB66. Six alleles were sequenced (15, 20, 24, 25, 27, 28). The shortest allele observed, allele 15, displayed a simple $(CA)_{15}$ repeat. All other alleles, however, displayed the compound dinucleotide repeat structure $(CA)_nTA(CA)_8$, varying at the first $(CA)_n$ repeat unit. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AF129733. Sequence on GenBank:

CATCTTCCTTAGCAAGAAGAGAAGGATTGGCATGGATGTTACCTCAGGGCTGATCTTCCT$(CA)_{17}TA(CA)_8$
AATATNGCGCANNGGGATAATGTGCCATATNTGTCAACACAATAAAAAGGGAAATTTCCCTCCACATATC
CTAATATANATNKCTTGCTAWATATAAGACCCTTCACCTGGNWKTGTAYTCAGAAAGATTTCACTCATTTC
CTACATATAATAWMTCCTCAACACAGCAGAAATGAGTTTGTTTTTAACTCAGAATAGTCTCAGAAAACTG
GAGGGCGGGAGACTTGTACTGGCTTCTCTGCCAGAGTTGCAGCGTAAGGGTAATCGCTGCTTCCTTTGC
AGACGATAGTTTCGTCTGGTGCTGCTGTCATTGGTTCTTAAGAAGAATAGTTGCTTCGACTGCAGTTCTT
TCCATTAGACTTAGAGAGTCTTCTACCAGCCCCAGTATTTCTCAGTCACAGCTGTCTATATAAATGACACT
GAGCTCCACTGAGCACACTGATGCTGGCTTCTCATTGCCAAGAGAAGAGGACTCTGGGAGAAAATACCT
CCAAGTGGAGAAACCAGTAGCCTCTGTGAGCTAATTCTGGGGCTGTGGAGTGAGTCTATCCAGGAAAAG
TCACAAATAAAGTGGAGGGCACAGATGATGGGGGATCCTCCTA

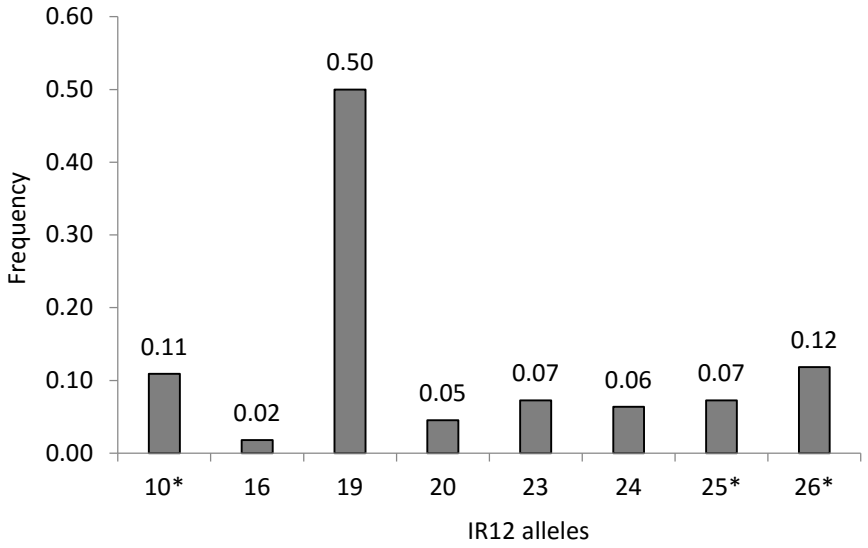


Figure S10
IR12. Three alleles were sequenced (10, 25, 26). All alleles displayed the simple dinucleotide repeat structure $(CA)_n$. No intermediate alleles were observed. No sequence data is available on GenBank for this locus, however the repeat motif is concordant with that originally described by Scott (2008).

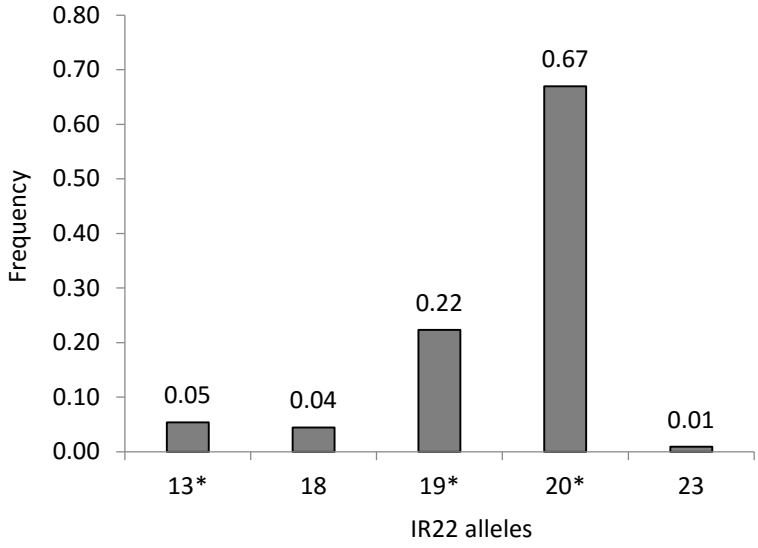


Figure S11
IR22. Three alleles were sequenced (13, 19, 20). All alleles displayed the simple dinucleotide repeat structure $(CA)_n$. No intermediate alleles were observed. No sequence data is available on GenBank for this locus, however the repeat motif is concordant with that originally described by Scott (2008).

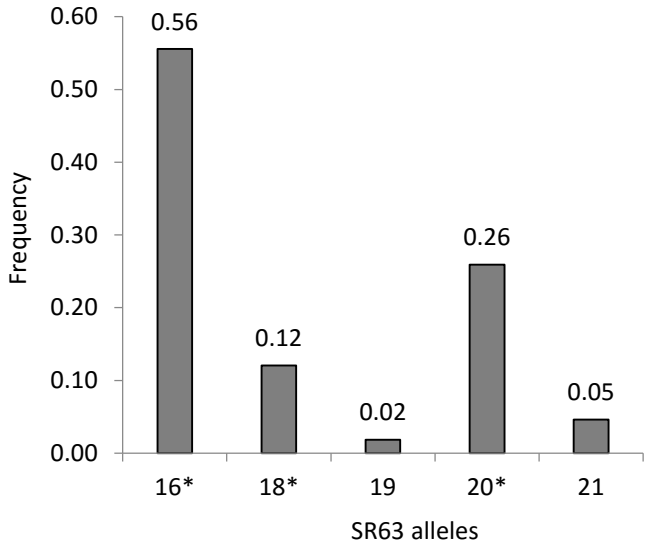


Figure S12
SR63. Three alleles were sequenced (16, 18, 20). All alleles displayed the compound dinucleotide repeat structure $(AC)_n(TC)_m$. No intermediate alleles were observed. The repeat motif observed here differs from the simple $(CA)_n$ repeat described by Scott (2008) and available on GenBank accession AY427965. This locus was, however, originally described for *Dicerohinus sumatrensis*. Sequence on GenBank:
CAGAAGCATTTCCTTGAGCAGAGTAGAATTTGGCATATTCAAGAAACAGATAGAATTCCATCACTGCTGG
AATATATTGGAGAGAGGAGTACCTCTCTCTCTCTTTCTCTGTCTCTCTCTT(AC)$_{19}$TCCAATTTCCCTGATA
CACACAGCATACAAAATAGGAATGAGGTGGATACAGAGGCTTCTG

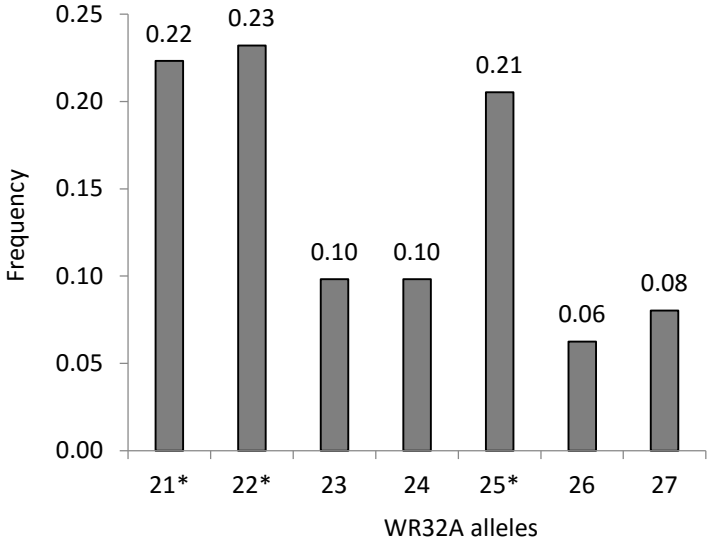


Figure S13
WR32A. Three alleles were sequenced (21, 22, 25). All alleles displayed the complex dinucleotide repeat structure $(AC)_6$CCCCATACGCAA$(AC)_n$. Variation was observed for the final $(AC)_n$ repeat unit only. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AY138541. Sequence on GenBank:
CTAGCAAAATCTCAAAGAGGTTTGATCCAACCATTCAAATTGTTTAGTATATACCCTGGTGGTTGAGCACT
GCTAGA(AC)$_6$CCCCATACGCAA(AC)$_{15.1}$GATTGCTGCCACCAAATAAATGAGTGGTCTCCAACTATCCCTG
GGCCCTTGGTGATTCCCTTAGTAATTGGCAGCAAGTTGTGTTCTTCACAAGCCAAGCCAGCACCTTCTGT
CATTCCCTCAGCAATATTCAAATCTCTACCCCAAATTTCTAGCCTCACTCTCAGCAAATGATCTCACTAAG
AAAATAGGG

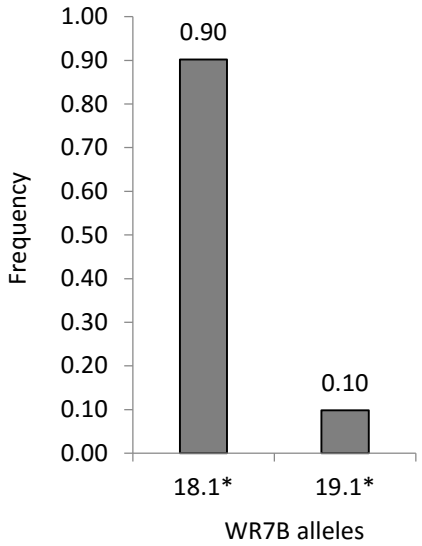


Figure S14
WR7B. Both alleles were sequenced (18.1, 19.1). Both alleles displayed the compound dinucleotide repeat structure $(TG)_nA(GT)_6$. Variation was observed for the first $(TG)_n$ repeat unit only. No intermediate alleles were observed. Repeat motif sequences generated here were concordant with the sequence available on GenBank accession AY138544. Sequence on GenBank:
CCCTCTGTGATTAAGCAAGGCAGAAGTGTGTTTACATTATGCCCAACCAACTTGTAATGAGAGGGCTACC
AAACCTTTC$(TG)_{16}A(GT)_5$TTTCCTTTCCACCATTCCAATATTTTTCCCCCTCGTTTCCCTTCTTTCTCTCTTT
ATGCGAATGTAAGAGAATGTGGCTTAACTTCAACTCTAGTGACTGTTCTTTCCCGCAAGTGTGTGTGTGC
GCGTCTTCCTTCCTGTTCATTCACAAAAAAGAAAAGCGACTCACTTCTCTGCTGTGGTCACCCCCTGG

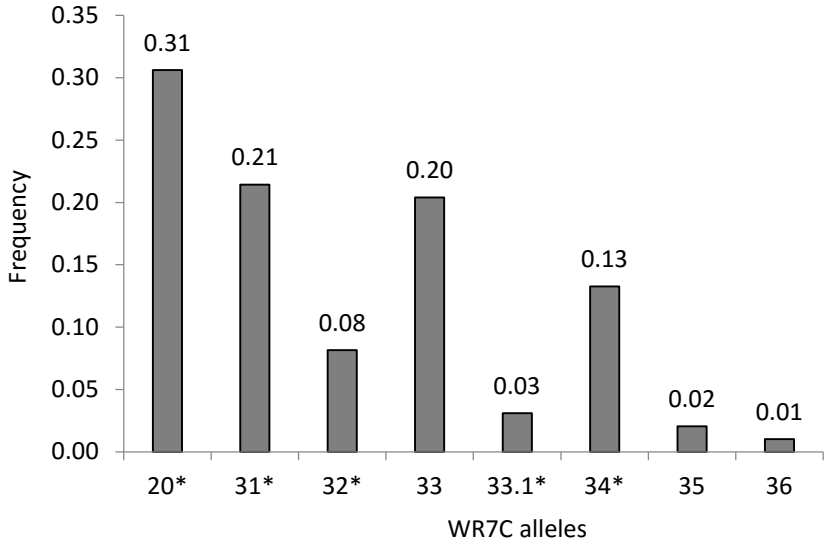


Figure S15
WR7C. Five alleles were sequenced (20, 31, 32, 33.1, 34). All alleles displayed the compound dinucleotide repeat structure $(TC)_7CC(TC)_m(TA)_n$. Variation was observed for the two latter repeat units, $(TC)_m$ and $(TA)_n$. A single intermediate allele was found between alleles 33 and 34, and was designated 33.1. This intermediate allele consists of an entire repeat unit fewer than allele 34. Unfortunately, no individuals were homozygous for allele 33 and thus it was not sequenced. Allele 33.1 most likely has the same number of repeat units as allele 33, but contains a microvariant outside of the repeat motif. Sequence data upstream and downstream of the repeat motif was of poor quality and therefore microvariation could not be determined.
Repeat motif sequences generated here were similar to the sequence available on GenBank accession AY138544 for *C. simum* which consist of $(TC)_{13}(TA)_{12}$. Sequence on GenBank:
CCAGAGCCTGATAAAAGCCCCAAAGCAGCCCCCAAGGTGAGAAGACAACGAGGAACTTCATCCAGATA
GAAGAGGAAGCAAATGTTTCCCAAAGCAGTGAAAACAGGTCTTGATTAGTGCTCAGGGTGAGTGGTCAG
TTCAAGTTTTTGCTCTGAGTATTATGTTGCTTGGAAGAAAGGGAATAGGATTCTCC$(TC)_{13}(TA)_{12}$AAGATTC
AAGTAAGAATCTTTGTAGAAGAAGCATGGATGAGTTCATCACCTTCTACATTTCTGTTTCATGTTTCTTCT
TAAAATCCATCATTCTCATTTCCATCAGAGTTCATAGGAAAATAACTGGAGTTTTAGTTTATTTCAAG

Appendix VI - Results of linkage disequilibrium tests

Pairwise significance values (p-value) and standard errors (S.E.) are listed along with alpha significance threshold values generated by applying sequential Bonferroni correction. *P*-values highlighted in bold were significant.

Locus 1	Locus 2	p-value	S.E.	Alpha threshold
IR12	Sex	**0.000006**	0.000002	0.00042
DB1	SR63	**0.00003**	0.000014	0.00042
BlRh1B	DB44	**0.00010**	0.0001	0.00042
BlRh1C	WR32A	**0.00020**	0.0001	0.00043
DB66	WR32A	0.00074	0.0003	0.00043
BlRh1C	DB52	0.0014	0.0004	0.00043
BlRh1C	DB66	0.0019	0.0005	0.00044
IR22	WR32A	0.0025	0.0003	0.00044
DB44	WR32A	0.0025	0.0005	0.00045
IR12	BlRh37D	0.0041	0.0003	0.00045
BlRh37D	WR32A	0.0041	0.0003	0.00045
BlRh1B	BlRh1C	0.0076	0.0008	0.00046
DB52	DB66	0.0082	0.0014	0.00046
IR22	BlRh37D	0.0095	0.0004	0.00047
IR12	WR32A	0.0104	0.0010	0.00047
BlRh1B	WR32A	0.0111	0.0012	0.00048
DB44	DB52	0.0122	0.0015	0.00048
DB52	WR32A	0.0125	0.0017	0.00049
BlRh1C	WR7C	0.0126	0.0013	0.00049
DB1	DB66	0.0138	0.0011	0.00050
DB23	DB66	0.0148	0.0007	0.00050
DB44	DB66	0.0153	0.0015	0.00051
BlRh1B	DB52	0.0161	0.0016	0.00051
DB1	WR7C	0.0168	0.0011	0.00052
DB44	BlRh1C	0.0192	0.0014	0.00052
DB44	DB23	0.0254	0.0008	0.00053
BlRh1B	DB1	0.0261	0.0011	0.00053
BlRh1B	WR7C	0.0266	0.0019	0.00054
BR6	DB52	0.0288	0.0013	0.00054
IR12	DB66	0.0288	0.0020	0.00055
DB1	DB52	0.0324	0.0017	0.00056
DB44	WR7B	0.0430	0.0006	0.00056
BR6	WR7C	0.0462	0.0015	0.00057
DB66	WR7C	0.0504	0.0034	0.00057

Appendix VI

Locus 1	Locus 2	p-value	S.E.	Alpha threshold
BR6	BlRh1C	0.0568	0.0015	0.00058
BlRh1B	DB23	0.0611	0.0014	0.00059
BlRh1B	DB66	0.0613	0.0035	0.00060
BR6	DB44	0.0699	0.0016	0.00060
DB1	DB44	0.0757	0.0021	0.00061
BlRh37D	DB66	0.0786	0.0018	0.00062
DB1	BlRh1C	0.0811	0.0024	0.00063
DB23	WR32A	0.0850	0.0020	0.00063
DB44	IR12	0.0962	0.0032	0.00064
WR32A	WR7C	0.108	0.0047	0.00065
BlRh37D	DB52	0.115	0.0025	0.00066
DB23	DB52	0.118	0.0027	0.00067
BR6	DB23	0.127	0.0012	0.00068
IR12	WR7B	0.133	0.0011	0.00068
DB44	SR63	0.136	0.0032	0.00069
IR12	DB52	0.138	0.0050	0.00070
IR22	WR7C	0.140	0.0032	0.00071
BlRh1B	BR6	0.141	0.0023	0.00072
DB23	WR7C	0.142	0.0024	0.00074
SR63	DB66	0.146	0.0035	0.00075
BR6	SR63	0.152	0.0019	0.00076
BR6	IR22	0.152	0.0020	0.00077
DB66	WR7B	0.161	0.0013	0.00078
IR12	BlRh1C	0.168	0.0042	0.00079
BlRh1B	IR12	0.168	0.0040	0.00081
DB52	WR7C	0.191	0.0065	0.00082
BlRh37D	DB23	0.212	0.0015	0.00083
BlRh1B	SR63	0.226	0.0039	0.00085
BR6	IR12	0.238	0.0028	0.00086
SR63	WR7B	0.256	0.0013	0.00088
BlRh1C	BlRh37D	0.258	0.0028	0.00089
SR63	BlRh1C	0.259	0.0040	0.00091
BR6	WR32A	0.279	0.0037	0.00093
IR22	BlRh1C	0.286	0.0043	0.00094
IR22	DB23	0.291	0.0024	0.00096
Sex	DB66	0.300	0.0010	0.00098
DB1	WR32A	0.300	0.0050	0.0010
IR12	IR22	0.311	0.0043	0.0010
IR22	DB66	0.321	0.0049	0.0010
Sex	BlRh1C	0.365	0.0010	0.0011

Appendix VI

Locus 1	Locus 2	p-value	S.E.	Alpha threshold
BR6	DB1	0.379	0.0024	0.0011
BlRh1B	Sex	0.387	0.0009	0.0011
BlRh1C	DB23	0.389	0.0031	0.0011
BlRh1B	BlRh37D	0.392	0.0029	0.0012
BlRh37D	WR7C	0.393	0.0034	0.0012
Sex	WR32A	0.397	0.0011	0.0012
IR12	WR7C	0.407	0.0061	0.0013
BlRh1B	IR22	0.423	0.0043	0.0013
DB44	Sex	0.433	0.0009	0.0013
SR63	WR32A	0.437	0.0050	0.0014
SR63	DB23	0.446	0.0026	0.0014
SR63	DB52	0.446	0.0056	0.0014
BlRh1C	WR7B	0.487	0.0017	0.0015
IR22	SR63	0.492	0.0038	0.0015
WR7B	WR7C	0.520	0.0018	0.0016
BR6	DB66	0.522	0.0041	0.0016
Sex	DB52	0.559	0.0010	0.0017
WR32A	WR7B	0.567	0.0019	0.0017
Sex	BlRh37D	0.586	0.0005	0.0018
IR22	Sex	0.599	0.0007	0.0019
DB1	IR22	0.615	0.0032	0.0019
Sex	WR7C	0.619	0.0009	0.0020
DB1	IR12	0.652	0.0038	0.0021
DB1	Sex	0.652	0.0007	0.0022
BR6	BlRh37D	0.654	0.0017	0.0023
BR6	WR7B	0.684	0.0010	0.0024
IR12	DB23	0.690	0.0027	0.0025
SR63	WR7C	0.704	0.0043	0.0026
DB1	DB23	0.765	0.0019	0.0028
BlRh1B	WR7B	0.802	0.0012	0.0029
DB44	WR7C	0.804	0.0046	0.0031
IR22	DB52	0.804	0.0041	0.0033
DB44	BlRh37D	0.816	0.0022	0.0036
DB1	BlRh37D	0.823	0.0016	0.0038
IR22	WR7B	0.827	0.0010	0.0042
SR63	BlRh37D	0.837	0.0016	0.0045
BR6	Sex	0.859	0.0004	0.0050
BlRh37D	WR7B	0.873	0.0005	0.0056
DB1	WR7B	0.890	0.0007	0.0063
IR12	SR63	0.905	0.0021	0.0071

Appendix VI

Locus 1	Locus 2	p-value	S.E.	Alpha threshold
Sex	DB23	0.914	0.0003	0.0083
SR63	Sex	0.929	0.0003	0.0100
DB23	WR7B	0.957	0.0003	0.0125
DB52	WR7B	0.966	0.0006	0.0167
DB44	IR22	0.993	0.0005	0.025
Sex	WR7B	1.000	0.0000	0.050

Appendix VII - Standard Qiagen Method

Adapted from the hair and nails protocol included with the Qiagen QIAamp DNA Investigator Kit (cat no. 56504). Italicised instructions are additional to the standard protocol.

1. Following digestion of horn material with 20 µL proteinase K, 300 µL Buffer ATL and 10 - 30 µL 1 M DTT (if included as detailed in Chapter 3), add 300 µL Buffer AL and mix by pulse-vortexing for 10 s.
2. Incubate at 70 °C with shaking at 900 rpm for 10 min.
3. Add 150 µL ethanol (96-100 %) and mix by pulse-vortexing for 10 s.
4. *Centrifuge at 10 000 rpm for 1 minute.*
5. Transfer the supernatant to the QIAmp MinElute column (in a 2 mL collection tube).
6. Centrifuge at 8 000 rpm for 1 minute. Place the QIAamp MinElute column in a clean 2 mL collection tube and discard the flow-through.
7. Add 500 µL Buffer AW1. Centrifuge at 8 000 rpm for 1 minute. Place the QIAamp MinElute column in a clean 2 mL collection tube and discard the flow-through.
8. Add 700 µL Buffer AW2. Centrifuge at 8 000 rpm for 1 minute. Place the QIAamp MinElute column in a clean 2 mL collection tube and discard the flow-through.
9. Add 700 µL ethanol (96-100 %). Centrifuge at 8 000 rpm for 1 minute. Place the QIAamp MinElute column in a clean 2 mL collection tube and discard the flow-through.
10. Centrifuge at 14 000 rpm for 3 minutes to dry the membrane completely. Discard the flow-through.
11. Place the QIAamp MinElute column in a clean 2 mL microcentrifuge tube. Apply 50 µL warm Buffer ATE to the centre of the membrane.
12. Incubate at room temperature for *10 minutes*. Centrifuge at 14 000 rpm for 1 minute.

Appendix VIII - Standard KingFisher Method

Adapted from the KingFisher Cell and Tissue DNA kit Kit (cat no. 97030196) for use with the KingFisher mL. Italicised instructions are additional to the standard protocol.

1. Digest of horn material with 200 µL of lysis buffer, 25 µL of proteinase K *and 10 - 30 µL 1 M DTT (if included as detailed in Chapter 3)*.
2. Place KingFisher mL tube strips on tube strip tray and prepare tubes as follows:

Tube	Tube name	Content	Reagent volume
A	Sample	Lysed sample	225 µL
		KingFisher Magnetic Beads	25 µL
		Binding buffer	360 µL
B	Wash 1	Wash buffer 1	600 µL
C	Wash 2	Wash buffer 2	600 µL
D	Wash 3	Wash buffer 3	800 µL
E	Elution	Elution buffer	*50 µL*

3. Switch on KingFisher mL and insert the tray into the instrument. Insert the tip combs into their slots and close the front lid.
4. Start the KF_TissueDNA_mL protocol.
5. After the run is completed, remove the tube strips and *transfer the purified DNA to 2 mL microcentrifuge tubes for storage*.