Abigail Guerier, Jacqueline M. Bishop, Stuart M. Crawford, Anne Schmidt-Küntzel &
Ken J. Stratford

Parentage analysis in a managed free ranging population of southern white rhinoceros: Genetic diversity, pedigrees and management

Abigail Guerier, Ken Stratford
Ongava Research Centre,
PO Box 58, Okaukeujo, Namibia

Abigail Guerier
Mammal Research Institute, Department of Zoology and Entomology,
University of Pretoria, Pretoria 0002, South Africa

Jacqueline Bishop
Department of Zoology, University of Cape Town,
Private Bag X3, Rondebosch 8001, Cape Town, South Africa

Stuart Crawford
Ongava Game Reserve,
PO Box 58, Okaukeujo, Namibia

Anne Schmidt-Küntzel
Life Technologies Conservation Genetics Laboratory,
Cheetah Conservation Fund,
PO Box 1755, Otjiwarongo, Namibia

Corresponding Author:	Ken Stratford
Email:	kenjstratford@gmail.com
Telephone:	00264-67-687075 (Namibia),
	0027-76-225-9033 (South Africa)
Fax:	00264-88-616411 (Namibia)

Abigail Guerier, Jacqueline M. Bishop, Stuart M. Crawford, Anne Schmidt-Küntzel & Ken J. Stratford

Parentage analysis in a managed free-ranging population of southern white rhinoceros: Genetic diversity, pedigrees and management

Abstract Small populations are vulnerable to the consequences of breeding within closed groups as the loss of genetic variability can lead to inbreeding depression. Here, we use microsatellite genotypes to assess variability and parentage within a small, managed population of southern white rhinoceros in northern Namibia. Tissue samples gathered from either a modified biopsy darting technique or ear notches allowed us to obtain genotypic data for all individuals in the population. As expected for this species, genetic variability in the population was relatively low (overall H_{obs} 0.45). In combination with detailed management records for the period 1993-2009, we were able to assign both parents for all 23 offspring. Only one calf of seven in the F_2 generation arose from father-daughter inbreeding within the population. Our analysis revealed that paternity was initially dominated by a single founder bull siring 10 of 13 calves over nine years; paradoxically, the other founder bull was selected for removal based on observations suggesting he was behaviourally dominant and therefore the likely sire of most calves. We also found that young introduced bulls were breeding successfully within six months of their arrival, well before having established their home ranges. We argue that in order to optimally manage and conserve the southern African rhino meta-population it is essential to have accurate pedigree information and genetic data for all individuals in the numerous small populations that are key to the survival of the species. (379 words)

Keywords Microsatellites, tissue biopsy, biopsy dart method, inbreeding, *Ceratotherium simum simum*

Electronic supplementary material The online version of this article (doi: to come) contains supplementary material, which is available to authorized users.

Introduction

Having literally been brought back from the brink of extinction nearly a century ago, the recovery of the southern white rhinoceros, *Ceratotherium simum simum,* is one of the great success stories of modern wildlife conservation. In the late 1880's as few as 50-
5 100 animals survived in the iMfolozi River valley of the Kwazulu-Natal region of South Africa (Emslie and Brooks 1999; Skinner and Chimimba 2005). Today, thanks to intensive protection and conservation efforts, there are over 17,500 animals in southern and east Africa (International Rhino Foundation, 2009) and the species has been re-classified as Near Threatened in the IUCN Red List of Threatened Animals. Throughout
10 their current distribution, most populations of white rhinoceros are managed as small isolated groups and maintaining genetic diversity in these populations is a key concern of conservation biologists (Emslie and Brooks, 1999; Frankham et al. 2002; Heller et al. 2010). As genetic diversity is an essential component for adaptive evolution its loss via genetic drift and/or selection decreases the probability of any population persisting
15 through time (Frankham 2005). Small populations are particularly susceptible to the loss of diversity via enhanced genetic drift as well as potential inbreeding; in turn this increases the risk of further losses of both adaptive and neutral genetic diversity (Frankham et al. 2002; Spielman et al. 2004). At the individual level, loss of genetic diversity via inbreeding is conclusively linked to reduced reproductive fitness; mating
20 between related individuals or individuals with similar genetic background leads to short-term inbreeding depression in virtually all species studied thus far, both in captivity and in the wild (cf. Keller and Waller, 2002; Charlesworth and Willis, 2009). Minimizing founder effects, inbreeding and further loss of genetic diversity in isolated populations, and in particular those descended from small founder numbers, is a major
25 challenge for conservation managers. Moreover, a number of recent studies highlight the central role of drift in isolated large mammal populations (Haag et al. 2010; Heller et al. 2010); in species where dispersal is no longer possible and the effective population size is minimal, management informed by the breeding structure of each population is critical.

30

Founded in 1993, the population of southern white rhinoceros (hereafter referred to as white rhinoceros) described in this study is now producing second-generation calves, providing an ideal study group to assess the value of breeding structure data in the

management of small populations. Traditionally, populations of white rhinoceros have been managed by controlling the sex ratio of adult individuals (maintaining a limited number of males in the population) and employing an exchange (or sales) programme for dominant breeding males (Bothma 2002). Mating success and identification of breeding bulls is commonly determined from behavioural observations and monitoring of the population in the field. Molecular analysis of parentage in an increasing number of species has, however, revealed that the observed social mating system in natural populations is not always supported by genetic parentage assignment; across a wide range of species genetic data indicates that the apparent 'dominant' male is not always the father (cf. Avise et al. 2002; Westneat and Stewart 2003). In white rhinoceros, field-based studies suggest that sexually receptive females associate primarily with dominant territorial males (Rachlow et al. 1999). However females have overlapping home ranges that can include the territories of a number of mature males, potentially leading to numerous mating opportunities (Owen-Smith 1988; Rachlow et al. 1999). Observations of white rhinoceros matings are very rare (Owen-Smith 1977) and whilst the dominant male will attempt to keep and cover oestrous females in his range, he is not always successful (Owen-Smith 1988).

In this study we use microsatellite genotypes together with current and historic field observations collected at Ongava Game Reserve, Namibia to determine the genetic mating system and create an accurate pedigree for this population of white rhinoceros. Our initial pedigree analysis clearly showed that most parentage relationships could not be resolved using observational data alone. For paternity, only the period during which bulls might be responsible for siring offspring could be determined from behavioural observations. Furthermore, since calves typically disperse from their mothers before they are individually marked, maternity cannot always be assigned with confidence from observations alone. As a result, molecular data were generated to evaluate both maternal and paternal parentage. DNA microsatellites are the marker of choice for population studies such as this (Bruford and Wayne 1993; Frankham et al. 2002, Jones and Ardren 2003) and have been successfully employed to investigate a range of questions in natural populations including parentage queries, mating systems (social vs 'genetic'), inbreeding, relatedness and reproductive skew (Swart and Ferguson 1997; Alderson et al. 1999; Banks et al. 2003; Piggott and Taylor 2003; Bishop et al. 2004; Hampton et al. 2004; Lorenzen and Siegismund, 2004; but see Dakin and Avise 2004).

Using data from 11 microsatellite loci together with behavioural observations for 31 individuals, we have been able to construct an accurate pedigree for all the animals in the Ongava population. Using this pedigree we assess the value of this approach in understanding the current structure of the population, and its potential application in the
5 management of white rhinoceros.

Materials and Methods

Study site and population

The habitat on the Ongava Game Reserve is Karstveld, with vegetation dominated by *Colophospermum mopane* shrub and woodland. A ridge of dolomite hills with *Terminalia* and *Acacia* tree species provides some differentiation of habitat (Berry and Loutit 2002). The region receives an average annual rainfall of 380mm which feeds a number of natural dams on the reserve, although most of these only contain water during the rainy season (Jan-Mar). Most of the water sources are artificial waterholes, which are opened in rotation to encourage grazing and browsing across the reserve. The population of white rhinoceros in this study was founded in 1993 with six adult individuals (two males and four females); by 2009, the founders plus a further two young adult bulls (introduced in 2002) together with their offspring, had produced a total of 41 calves.

The reserve is 300km^2 in size and patrolled throughout the year by resident Anti-Poaching Units (APU) to track and monitor the rhinoceros. All sightings are recorded with associated GPS coordinates and images. Additional data from observations made by researchers and other members of staff on the reserve are also recorded. All other information regarding the population, such as births, deaths, disease, translocation, sales, etc. is documented by the reserve management. Management's historic records were used to create individual rhinoceros profiles.

Tissue sample collection

Tissue samples were collected as ear notches from individual rhinoceros for the period 2005-2009. Ear notches are taken during annual ear-marking and micro-chipping procedures used for individual identification; a total of 14 ear notch samples were collected and stored in 90% ethanol at 4 ºC. A further 18 samples were collected using biopsy darts from 2007-2008. In four cases we were able to obtain a biopsy sample from a calf while it was still with its mother (i.e. before dispersal) and then also later obtain an ear tissue sample for this animal during the annual marking procedure. Biopsy darting was only used for animals where ear notch tissues were unavailable. We explored a number of methods to obtain biopsy samples in the least invasive manner, and developed a modified method for biopsy darting rhinos using a CO_2 – powered Dan-inject rifle (model JM Special). Biopsies were collected using ethanol-sterilized

(99% ethanol) 20mm biopsy needles attached to 3ml darts. A dental barb (Maillefer 21mm barbed nerve broach) was inserted into each biopsy needle to reduce the chance of biopsy darts dropping off the animal without removing a core of tissue. Rhinoceros were darted at night from a distance of 10-25m. Animals were darted either in the front
5 or hind legs only, as the flexible skin in these positions allows the dart to fall as soon as the limb is moved. The dart was only retrieved from the ground once the individual had moved off of its own accord. The biopsy needles were placed in sterile collection tubes containing 90% ethanol. In the laboratory biopsy samples were removed from the needle and barb and stored in 90% ethanol at 4 ºC.

10

One of the original founder males had been sold from the reserve in 2004; this individual was subsequently shot and prepared as a trophy. We were able to source a small piece of chemically preserved skin from the trophy which was stored at room temperature until use.

15

DNA extraction and microsatellite genotyping

DNA extractions from ear notch and biopsy tissue samples were carried out at the Ongava on-site laboratory while the trophy mount sample was processed at the Life Technologies Conservation Genetics Laboratory of the Cheetah Conservation Fund
20 (CCF), Namibia. DNA from ear notch and biopsy tissue samples was extracted using a DNeasy Tissue Kit (Qiagen) following the manufacturer's protocol for animal tissues and stored at 4ºC. DNA from the trophy mount sample was extracted using the DNeasy Tissue Kit (Qiagen) but with several modifications to the protocol including additional wash steps to eliminate excess salts from the trophy preparation process, concentration
25 of the eluate and a combination of multiple extractions to maximize the amount and concentration of DNA (see Online Resource S1 for details). DNA quality was checked on 1.5% agarose gels and visualised under UV with ethidium bromide. Samples were also checked for their concentration using a Nanodrop spectrophotometer (Thermo Scientific) and these concentrations informed the amounts of template used in PCR (10
30 to 50 ng). DNA concentrations from the trophy sample were tested using a Beckman DU 650 spectrophotometer (Beckman Coulter).

Individuals were genotyped at 11 microsatellite loci developed in a number of rhinoceros species (Table 1, see Online Resource S3 for full specification). Loci were

selected based on data from Coutts (2009) and included 7B and 7C (Florescu et al. 2003), WR1 and WR2 (reported in Nielsen et al. 2008, redesigned from loci reported in Florescu et al. 2003), BR6 (Cunningham et al. 1999), DB1, DB44 and DB49 (Brown and Houlden 1999) and Rh7, Rh8 and Rh9 (Zschokke et al. 2003). PCR conditions
5 followed Coutts (2009) and Menotti-Raymond et al. (2005) and cycle conditions were adapted where needed (see Online Resource S2 for details). Microsatellite amplification and genotyping was carried out at the University of Cape Town (UCT) and CCF. To standardize the genotyping calls and determine the intrinsic genotyping error rate across the data set PCR amplification and analysis was repeated for a random sub-sample of
10 individuals in both laboratories.

Analysis

The presence of null alleles, stutter errors or short allele dominance was determined using the program MICROCHECKER version 2.2.3 (Van Oosterhout et al. 2004).
15 Observed and expected heterozygosity, together with the polymorphic information content (PIC) and tests for deviation from Hardy-Weinberg expectations were calculated using Cervus 3.0 (Kalinowski et al. 2007).

We investigated individual parentage using a manual assignment approach based on
20 exclusion criteria together with computational assignments calculated within a likelihood framework. For the manual assignments, at each locus the genotype profile of an offspring was compared to the genotype profile of all possible adult males and females. We used strict exclusion criteria to eliminate non-matching parents. Results representing all possible parent-offspring combinations were recorded. We then refined
25 the process by constraining the numbers of candidate parents based on the information derived from our historic records database; this included date of birth for each animal, sex of individuals, tenure times on the reserve (for bulls) and groups of calves that were all born in the same year (See Online Resource S5 for details). The results from this analysis were recorded as the set of possible parent pairs for each offspring. Because of
30 the low levels of genetic variation observed in this species, we excluded possible parent pairs based on genotype mismatches at one or more loci. To explore the possibility that we erroneously excluded candidate parents due to incorrect genotypes at a particular locus, we analysed the distribution of mismatches across the loci used in this study to

assess whether any one particular locus accounts for a disproportionate number of single locus mismatches (See Online Resource S6 for a detailed analysis).

We also used two independent computational methods to determine parentage in our data set, Cervus 3.0 (Kalinowski et al. 2007) and Colony 2.0.1.3 (Jones and Wang 2009); both programs implement likelihood methods to determine parentage using multi-locus genotype data.

Through simulation, Cervus generates criteria for parentage assignment at a given level of confidence for all offspring analysed. As for the manual analysis, initially we allowed Cervus to select candidate parents from all males and all females. Subsequently the number of candidate parents was constrained using our historic records. When running simulations the following parameters were used: 100,000 offspring, 1% mistyped loci and 96.9% typed loci (determined by Cervus from our data set). In all simulations 100% of candidate mothers and fathers were represented in the data set. Confidence was calculated using LOD (Likelihood-of-Difference) scores at both 95% (strict) and 80% (relaxed) confidence levels. Each analysis required a minimum of six genotyped loci and possible parent pairs were ranked based on their joint LOD score. Because relatively small samples sizes together with a large number of highly related individuals can significantly bias parentage assignment, we tested whether the use of our samples as the reference population influenced parentage assignment in Cervus. To do this we repeated the analysis using genotype data from a sub-set (n=49) of individuals sampled from the Hluhluwe-Imfolozi Game Reserve, SA as the reference population (Coutts et al. unpublished data).

Colony 2.0.1.3 (Jones and Wang 2009) was also used to determine parentage for each calf in our data set. As in Cervus we first ran the analysis with unconstrained parents and then excluded mothers and fathers accordingly based on historic records and our pedigree data. We implemented the full-likelihood method and a polygamous mating system without inbreeding; in the final analysis we used medium precision and medium run lengths, as tests with higher precision and longer runs did not make a significant difference to the assignments. We also tested a number of genotyping error rates (0.001, 0.01) for type I and II errors and found that, as expected, with higher error rates assignment probabilities were reduced; nonetheless, we did not see a difference in

assignment order. In the final analysis we used an error rate of 0.001. Because we have a closed data set i.e. all individuals were sampled, we set the probability that both the father and the mother were amongst the candidate male and female parents to one. Unlike Cervus, Colony asks for an exclusion list of candidate parents. For the constrained runs, each offspring was therefore given a list of individuals that could be neither the father nor the mother. We also replicated the date-of-birth refinement process by sequentially removing mothers who had unambiguous assignments to calves with the same year of birth.

The ability to differentiate between candidate parents is strongly influenced by levels of genetic variation. Because we had a number of known mother/calf combinations (n=5) we also tested the intrinsic power of the data set to differentiate between these true mothers and other female candidates by calculating how often the three parentage assignment methods used identified the true mother as the most likely candidate. Finally, to explore the degree to which individual genetic similarity may influence the assignment of parentage and possibly male mating success among Ongava rhinoceros we calculated the shared allele distance *ds* (Chakraborty and Jin 1993) between all individuals in the population. Where relevant, results are reported as Mean ± Standard Error.

Results

Microsatellite genotyping

The combination of materials from ear notches, biopsy samples, and the trophy mount gave a total of 35 samples (including four duplicates). These represented all rhinoceros aged at least two years on Ongava at the end of 2009, plus three recently deceased animals, as well as animals sold during 2004 - 2009. The final genotype data set consisted of 376 genotypes (35 samples for 31 individuals across 11 microsatellite loci). For the trophy sample, two of the loci showed evidence of allele dropout, and four PCR repeats were required in order to obtain the complete genotype. For the total dataset, nine genotype gaps remained (2.3%). Of the 35 samples, 29 (83%) produced a genotype at every locus and with exception of the trophy sample no one sample had more than one missing genotype (See Online Resource S4 for further details of error rates).

The data set contained genotypes from five calves that were sampled via biopsy dart when with their mothers in 2006 and 2007 - hence the identity of the mother of these calves is known. In 2007 a total of eight calves dispersed; ear notch tissue samples were obtained from all of these. One of the five calves (Zac) with known mother (Sharon) did not disperse. The genotype profiles of the other four darted individuals could be perfectly matched within the set of eight calves.

The summary statistics for the microsatellite loci used in this study are shown in Table 1. For three samples we found that none of the genotypes for candidate fathers matched the genotype of the offspring. In particular these mismatches occurred at locus Rh8. None of these genotypes changed after repeated amplification and analysis. This is compatible with the presence of a null allele at this locus, which was confirmed by MICROCHECKER and Cervus.

Parentage assignment

In order to explore the range of solutions for different assumptions, we performed parentage analysis for the 23 offspring using both manual and software techniques under unconstrained (all possible parents) as well as constrained (sets of candidate parents specified according to historic records, including calving dates) conditions. We also repeated the analysis excluding locus Rh8 due to a possible null allele. For all constrained manual assignments the trophy sample was not excluded from any potential

assignments based on a homozygous genotype, because of the increased possibility of allelic drop out for this sample.

Despite the generally low levels of allelic diversity at all loci genotyped (Table 1), in 16 of 23 cases it was possible to uniquely assign parentage to one set of parents using manual assignment methods. For five of the 23 cases there were two possible candidate parent pairs, and for the remaining two cases there were three possible candidate parent pairs. Adding pedigree constraints significantly reduced the number of possible parent pair combinations. Unconstrained assignments based on the likelihood approach used in Cervus resulted in only five parent pair combinations that were possible given the pedigree information for the population. Once the number of candidate parents had been constrained according to the pedigree, 18 of the 23 'most-likely' Cervus parent pairs were assigned at the 95% confidence level. For the remaining offspring, four individuals were assigned parent pairs at the 80% confidence level and in only one instance was the parent pair assignment not statistically significant. In the latter case, the mother could be assigned with high confidence, but Cervus could not discriminate between two possible candidate fathers. When the Cervus analysis was run using a different reference population from Hluhluwe-Imfolozi, the top ranked pairs assigned were identical to those assigned when using the Ongava sample set as the reference population.

Similarly to Cervus, in the unconstrained case, Colony assigned parents that were not possible given the pedigree information for the population. Once the candidate parent pairs were constrained by pedigree, Colony assigned exactly the same fathers as Cervus for all offspring (i.e. the statistically most-likely father was the same for both software analyses). In 3 of 23 cases, the top-ranked mother assigned by Colony did not agree with that assigned by Cervus, but in all three cases that assignment was included in the possible parent pair combination derived from the manual analysis.

A comparison between the different modes of parentage assignments showed that in all instances where there was an unambiguous manual assignment (i.e. only one parent pair possible), both the Cervus and Colony 'most-likely' assignments were the same parent pair. For the remaining seven manual assignments, both the Cervus and the Colony'most-likely' assignments based on a pedigree-constrained dataset were always

one of the manual assignment possibilities (see Table 2 for a summary of parentage assignment for each of the analysis methods).

We analysed further the five offspring with known mothers. When allowed to consider
5 all candidate mothers, the manual assignment always included the known mother in the set of candidates. Under similar assumptions, Cervus chose the true mother as the most-likely mother in four of five cases, with the true mother ranked second in the other case. Colony also ranked the true mother as the most likely mother in four of five cases, also ranking the true mother second in a fifth case (different from the case in Cervus).
10

We also explored the potential problem of 'false' parent pair exclusions due to genotype mismatches arising from genotyping errors. In summary, relatively few parent-pair exclusions occurred as the result of single locus mismatches (13%) while approximately three-quarters (77.4%) of candidate parent pairs were excluded based on genotype
15 mismatches at more than one locus. For those parent pairs excluded based on single locus mismatches a difference in the distribution of mismatches over the 11 loci was observed, however in the context of the low error rate in the full study this is unlikely to have led to a significant number of erroneous exclusions (see Online Resource S6 for further details).
20

The complete set of parentage assignments allowed us to resolve the ambiguities in the pedigree data and fully define the four founder matrilines (see Figure 1).

Pairwise genetic distance estimates and parentage assignment
25 Adult breeding males were, on average, as genetically similar to each other as they were to the females with which they sired offspring (mean pairwise male ds=0.40 ±S.E=0.01; mean parental pair ds=0.43±0.02). The most successful male, Derek (13 successful sires in the study period), was on average as genetically similar to females with which he successfully sired calves (ds=0.42±0.02; range 0.22-0.50) as compared to Tony
30 (ds=0.40±0.03; range 0.31-0.50) or Jeff (ds=0.48±0.06; range 0.36-0.50), suggesting that male mating success at Ongava is not influenced by genetic similarity among mated pairs. The mean shared allele distance among founder breeding females was ds= 0.42±0.02. These results together suggest that assignments to potential sires are not biased to a particular candidate father by chance genetic similarity among individuals.

Indeed, the mean overall pairwise *ds* (0.37±0.004) at Ongava was similar to that for our alternative reference population sample (Hluhluwe-Imfolozi *ds*= 0.401±0.01; n=49) derived from the original source and largest population of southern white rhinoceros in southern Africa.

5

Levels of inbreeding

Reviewing the composition of the population by generation shows an F_0:F_1:F_2 structure of 4:16:7 (excluding the four breeding males). The descriptive statistics for all loci grouped by generation are given in Table 3. The mean heterozygosity across all 31

10 individuals, H_{pop}, is 0.447, and the polymorphic information content (PIC) is 0.357. To assess whether the population shows signs of progressive inbreeding at the genetic level, we calculate heterozygosity H, across all loci for each generation in Cervus (See Table 3). We then estimated the effective inbreeding coefficient, F_e, for the F_2 generation for expected values of H, giving $F_{e,F2} = 1 - (H_2 / H_0) = 0.157$ (Frankham et

15 al. 2002).

All offspring males are removed from the population before they have the opportunity to become candidate fathers, and no F_2 female produced calves within the time period of this study. Instances of inbreeding in this white rhinoceros population are therefore

20 restricted to F_1 females producing an F_2 offspring with their father. Examination of Figure 1 shows that father-daughter inbreeding has occurred only once - Anne's father is founder bull Derek, and he is also the father of her calf, #22. All other six F_2 offspring have different fathers and grandfathers. Thus from the parentage assignment data this population has one of seven possible calves that is the result of inbreeding.

25

Discussion

To our knowledge this is the first study in which it has been possible to assign with statistical confidence a complete set of parents to all offspring in a managed free-ranging population of white rhinoceros. The two major factors contributing to the success of this study were that we were able to obtain DNA via a tissue sample from every individual in the population, and that we had comprehensive historic information for each rhinoceros allowing us to constrain the number of candidate parents for any one offspring.

We initially tried faecal samples as a DNA source, but found that extraction quality and quantity issues repeatedly limited the chances of successful and reliable PCR amplification of the target loci. This is consistent with Nielsen et al.'s (2008) conclusion that faecal DNA from either white or black rhinoceros is not an adequate source of DNA for genetic analysis (however black rhinoceros dung did provide adequate DNA in other studies: Cunningham et al. 1999; Garnier et al. 2001; Harley et al. 2005, Guerier et al. in preparation). In the white rhinoceros several reasons may exist for this, including the possibility that the DNA obtained was particularly degraded, the presence of inhibitors , or perhaps more simply that there is insufficient rhinoceros DNA present in the samples. Other collections and storage methods (e.g. combining a short period of storage of the sample in ethanol with subsequent desiccation using silica, per Nsubuga et al. 2004) might yield higher concentrations of DNA.

DNA from the sample obtained from the trophy mount specimen required significant effort in both logistics and laboratory time. This included an environment free of any rhinoceros DNA, with repeated attempts of amplification under different conditions. This effort did however prove to be worthwhile. For example, there were a number of instances where it was not possible to exclude the other founder bull from being the father, yet the trophy sample genotype, even at only six of the 11 loci, was sufficient to exclude that individual. This emphasizes the importance of getting samples from all candidate parents.

Although in this study the loci selected were known to have some variation in white rhinoceros (Coutts, 2009), several of the microsatellite loci used were developed in other rhinoceros species (see Table 1). Hence the possibility exists that they might not

provide the same levels of individual variation when used in cross-species studies. Locus Rh8, a marker developed for Indian rhinoceros (Zschokke et al. 2003), appears to have a null allele in white rhinoceros (see also Coutts, 2009). Locus Rh7 revealed low PIC values in white rhinoceros and was also isolated from Indian rhinoceros. Dakin and Avise (2004) show that null alleles can introduce substantial errors when they lead to false exclusion of parents during parentage assignments. We did exclude locus Rh8 during the final parentage analysis in order to avoid the possibility of false exclusions, but found no significant differences in assignments.

In this analysis we used a sequential approach to integrate pedigree information with genetic data. We collected both sets of data independently, first determined the genotype for each individual at each locus, and then refined the parentage assignment by restricting the number and identity of candidate parents based on pedigree information. We performed this in an iterative manner, first using parental presence (i.e. the only candidates were those breeding males and females on the reserve at the time) and then calf date of birth (mothers only produce a single offspring). Online Resource S5 gives further details. This approach has been used in the investigation of kinship and parentage across a number of species (Jones and Ardren 2003) and options for restricting candidate parents exist in parentage analysis software (e.g.: Cervus 3.0*,* Kalinowski et al. 2007 and Colony 2.0, Jones and Wang 2009)). According to the criteria suggested in Jones et al. (2010) – we have samples for all candidate parents, and our family size is small – the exclusion (our manual assignment) and categorical (as recommended in Cervus) methods we have used are most appropriate for our data set. There is however an alternative approach to the issue of integrating pedigree and genetic data, which is to simultaneously estimate parentage and the population-level parameters in which we are interested, termed full probability modelling. Hadfield et al. (2006), for example, present a Bayesian framework for this process, which they state provides a parentage assignment that may be less biased. The primary basis for implementing this rigorous statistical methodology is to compensate for genetic data sets that may have high rates of genotyping errors, or pedigrees with limited information (see for example the software suggested by Walling et al. 2010). While our final set of genotypes does contain a negligible error rate (as shown by the genotype repeatability of 26 random samples), and our baseline pedigree information is both accurate and comprehensive, we chose to repeat our analysis using the alternative

modeling framework provided by Colony 2.0 (Jones and Wang 2009). The computational methods both gave very similar results, suggesting that our assignments were not biased by any one particular statistical methodology.

5 Very few parentage studies in white rhinoceros populations have been undertaken. Indeed this is true across all rhinoceros taxa, in particular for African rhinoceros, and success in those attempts at parentage assignment has been limited. Coutts (2009) attempted to confirm maternity and assign paternity in two populations of white rhinoceros. Both populations had 10 calves with 12 candidate fathers and the 10 analyses revealed that very few paternities could be assigned with statistical confidence, even after the removal of non-informative loci. In other rhinoceros taxa, to date only Garnier et al. (2001) have been able to unequivocally assign parentage to all 19 offspring in a small (n=33) wild population of black rhinoceros (*Diceros bicornis*) in Zimbabwe. The authors used DNA extracted from faecal samples to derive genotypes 15 across 10 microsatellite loci (on average 3 alleles per locus, range 2-7) and were successful in determining 100% of genotypes (here we determined 97.7%). In their analysis, most mothers were first assigned based on behavioural observations and paternity was then tested using Cervus. Up to 10 males were candidate fathers (in comparison with our study, where either 2 or 3 males were candidates), however they 20 had DNA samples from 9 of these 10. Kim (2009) assessed populations of black rhinoceros (*Diceros bicornis bicornis*) in Namibia using microsatellite markers and also conducted parentage studies in two small game reserves with limited or no pedigree information. His investigations revealed that even though the number of candidate fathers was small (five and two respectively in each of the reserves/farms) not all calves 25 could be assigned at the strict 95% confidence level.

The presence of duplicate samples in the data set allowed us to verify of the accuracy of our assignment results. We recommend including a subset of anonymous samples within the data set for any similar future studies to act as internal controls along with the 30 more typical methods such as multiple DNA extractions from each sample, repeat PCRs per sample, calculation of error rates, etc. (see reviews in Bonin et al. 2004; Selkoe and Toonen 2006; Scott 2008).

Given the additional concern that parentage assignment in species with low genetic diversity might be compromised, we undertook several additional analyses. We found that our manual assignments were not biased by an excessive number of exclusions of possible parent pairs due to a mismatch at just one locus. We also calculated the genetic distance between individuals in our population and found that distances were similar across the breeding elements of the population, suggesting that assignments to potential sires were not biased to a particular candidate father based on levels of individual genetic similarity. Finally, we used the presence of known mother-calf samples within the data set to confirm that the correct mother was selected by each of the three different assignment techniques. We therefore are confident that the pedigree results presented in this study are accurate and robust.

It is apparent that low levels of genetic diversity characterize southern white rhinoceros populations, and the results from our study are consistent with other microsatellite studies in this species (Seror et al. 2002; Florescu et al. 2003; Nielsen et al. 2008; Scott 2008; Coutts 2009). There does appear to be greater genetic diversity in the white rhinoceros population studied by Florescu et al (2003) (H_{Obs} = 0.597), however the authors selected for polymorphic loci, and this may have elevated the measures of diversity in their data. It is likely that the higher levels of variability found in microsatellite studies in black rhinoceros allows for more successful parentage analysis (Cunningham et al. 1999; Garnier et al. 2001; Harley et al. 2005). Research on other species supports this theory. For example, Tokarska et al. (2009) found they were unable to use microsatellites for parentage assignment in the severely bottlenecked European Bison where all surviving members of this species are thought to have descended from just two individuals; this is reflected in very low heterozygosity (H_{Exp}<0.3 across 17 microsatellites) and high levels of inbreeding (F_{bar}>0.5). Conversely, Archie et al. (2007) were able to assign parentage in a high proportion of elephant calves (152 of 183) using faecal samples typed across 11 microsatellites – this species shows relatively high heterozygosity (mean H_{Obs}=0.77 across 20 populations in 16 microsatellites, Comstock et al. 2002).

To date, our study population has produced just one inbred F_2 individual from seven offspring in the F_2 cohort. While many species demonstrate deleterious effects from inbreeding depression (Frankham et al. 2002), in both managed and natural populations

of free-ranging, long-lived species these effects are very difficult to measure. Krummenacher and Zschokke (2007) measured mortality rates of captive white rhinoceros and found a small but not significant increase in mortality of inbred individuals (non-inbred 14.7%, inbred 18.7%). In the one analysis to date, free-ranging
5 white rhinoceros populations did not display significantly reduced genetic diversity in seeded populations and also showed no evidence of inbreeding depression (Coutts 2009). In addition, given that all extant white rhinoceros are descended from just a few individuals (Emslie and Brooks 1999, Skinner and Chimimba 2005), it could be argued that inbreeding does not play a significant factor in reducing fitness in this species,
10 since genetic diversity is already very low. Nevertheless, unchecked inbred matings are likely to lead to further loss of genetic variation and some form of inbreeding depression in the future, and should be avoided (Edmands 2007). It is worth noting here that a lack of evidence for inbreeding depression (e.g. juvenile survival) does not support the conclusion that there is no inbreeding depression for other factors (e.g.
15 fecundity) or indeed that diverse fitness components will not interact to reduce overall fitness in the future (Hedrick and Kalinowski 2000; Edmands 2007).

Management implications

A complete genetic parentage analysis allows us now to review population management
20 decisions taken to date, as well as make recommendations for future management strategies that will both satisfy the micro-management objectives for Ongava's population and consider implications for the continued conservation of the southern white rhinoceros.

25 Since the population was founded in 1993 by six animals from Hluhluwe-Imfolozi Park, management decisions have been based on the traditional bull translocation strategies used to manage most stock herds. In theory, the dominant territorial bull is the most successful breeder. Among polygynous ungulates, territorial behaviour is believed to function primarily as a reproductive strategy to secure mates (Owen-Smith 1977).
30 Owen-Smith hypothesized that territorial white rhinoceros bulls achieved greater reproductive success than non-territorial bulls. Among white rhinoceros, dominant males occupy clearly demarcated breeding territories and defend them from other dominant males. These bulls also engage in mate-guarding behaviour (Hutchins and Kreger 2006). Rachlow et al. (1998) suggested that while mechanisms for female mate

choice among white rhinoceros were not clear and may be influenced by choice for high-quality resources within territories, it appears that territorial males spent more time with females with high reproductive potential. This suggests that relative mating success is higher for males who adopt a territorial strategy as opposed to younger non-territorial males.

In their joint tenure period, a total of 13 offspring were fathered by the two founder bulls (calves born in the period up to 2004). Based on behavioural observations and home range data, management sold the dominant territorial bull to minimise potential risk of inbreeding. Evidence from our study now shows that the behaviourally sub-ordinate bull was in fact the more successful breeding male, fathering 10 of these 13 offspring.

Following the sale of one of the founder bulls, two new bulls were purchased in 2002 from Namibian stock. The source population was chosen in an attempt to minimise the chances of direct relatedness to individuals in the founder population. The two bulls were estimated to be between 6 to 8 years of age upon arrival, and our analysis shows that they were successfully breeding with resident females within six months, well before establishing their own home territories. This finding is very interesting as it is generally assumed that adult male rhinoceros first have to establish territories before breeding successfully (Owen-Smith 1977; White et al. 2007).

The detailed genetic and pedigree results from this study and its continuation will allow management to plan purchases and take-offs to optimize genetic diversity, population health and ultimately promote the conservation of the southern white rhinoceros. This includes selling matched pairs of unrelated males and females to provide founder animals for other managed populations; selling individuals that are not peak breeding animals, such as bulls that do not produce their share of offspring, or females that do not produce calves at regular intervals; and monitoring the diversity of the offspring by generation with a view to minimizing genetic drift and potential problems associated with inbreeding in small isolated populations.

Acknowledgements

We would like to extend a special thanks to the Ongava Anti-Poaching Unit members (Beatus Kangongua, Binius Kamerika, Jesaya Kapona, Henock Shikomba, Habel Hanichab, Jesaya Jafet, Gustav Guchab, Efraim Onamib and Nehemiya Kapona) for sharing their skills, expertise and observations. We would like to express gratitude to the Directors of the Reserve for their support of this project.

We would also like to thank Dr. C. O'Ryan and Dr. N. Coutts from the Department of Molecular and Cell Biology at the University of Cape Town for their contribution. Technical support and assistance was also provided by Dr. J. Rousseau from the Human Genetics Division at the University of Cape Town Medical School. Many thanks also to Mrs. T. Mwaningange and Dr. L. Marker from the Applied Biosystems Genetic Conservation Laboratory at the Cheetah Conservation Fund in Namibia. This manuscript was greatly improved by the comments of two anonymous reviewers.

Ongava Research Centre is funded by charitable donations from the Namibian Wildlife Conservation Trust (UK), West Midlands Safari Park (UK) and the Directors of Ongava Game Reserve. This research was carried out under Namibian Ministry of Environment and Tourism Permit numbers 1011/2006, 1121/2007, 1242/2008, 1371/2009 and 1502/2010.

References

Alderson GW, Gibbs HL, Spencer SG (1999) Parentage and kinship studies of an obligate brood parasitic bird, the Brown-headed Cowbird (*Molothrus ater*) using 5 microsatellite DNA markers. J Hered 90:182-190

Archie EA, Hollister-Smith JA, Poole JH, Lee PC, Moss CJ, Maldonado JE, Fleischer RC, Alberts SC (2007) Behavioural inbreeding avoidance in wild African elephants. Mol Ecol 16:4138-4148

10

Avise JC, Jones AG, Walker D, DeWoody JA (2002) Genetic mating systems and reproductive natural histories of fishes: Lessons for Ecology and Evolution. Ann Rev Genet 36:19-45

15 Banks SC, Hoyle SD, Horsup A, Sunnucks P, Taylor AC (2003) Demographic monitoring of an entire species (the northern hairy-nosed wombat, *Lasiorhinus krefftii*) by genetic analysis of non-invasively collected material. Animal Cons 6:1-10

Berry C, Loutit B (2002) Trees and Shrubs of the Etosha National Park. 3rd Edition. 20 Namibia Scientific Society, Windhoek

Bishop JM, Jarvis JUM, Spinks AC, Benneth NC, O'Ryan C (2004) Molecular insight into patterns of colony composition and paternity in the common mole-rat *Cryptomys hottentotus hottentotus*. Mol Ecol 13:1217-1229

25

Bonin A, Bellemain E, Bronken-Eidesen E, Pompanon F, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261-3273

30 Bothma J du P (2002) Game Ranch Management, 4th Edition, Van Schaik, Pretoria

Brown SM, Houlden BA (1999) Isolation and characterization of microsatellite markers in the black rhinos (*D. bicornis*). Mol Ecol 8:1559-1561

35 Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3:939-943

Chakraborty, R, Jin, L (1993) A unified approach to study hypervariable plymorphisms: Statistical considerations of determining relatedness and population distance. In: DNA 40 Fingerprinting: State of the Science (eds) Pena, SDJ, Chakraborty, R, Epplen, JT, Jeffreys, AJ, Birkhauser Velag, Basel, Switzerland

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nature Rev Genet 10:783-796

45

Comstock KE, Georgiadis N, Pecon-Slattery J, Roca AL, Ostrander EA, O'Brien SJ, Wasser SK (2002) Patterns of molecular genetic variation among African elephant populations. Mol Ecol 11:2489–2498

50 Coutts NJ (2009) Investigating genetic diversity at neutral and adaptive DNA markers in the severely bottlenecked Southern white rhinoceros (Ceratotherium simum simum).

PhD Thesis, University of Cape Town

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

5

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. Heredity 93:504-509

Edmands S (2007) Between a rock and hard place: evaluating the relative risks of
10 inbreeding and outbreeding for conservation and management. Mol Ecol 16:463-475

Emslie R, Brooks M (1999*)* African Rhino Status Survey and Conservation Action Plan. IUCN/SSC African Rhinoceros Specialist Group. IUCN, Gland, Switzerland and Cambridge, UK

15

Florescu A, Davila JA, Scott C, Fernando P, Kellner K, Morales JC, Melnick D, Boag PT, Van Voeverden De Groot P (2003) Polymorphic microsatellites in the white rhinoceros. Mol Ecol Notes 3:344-345

20 Frankham R (2005) Genetics and extinction. Biol Cons 126:131-140

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge

25 Garnier JN, Bruford MW, Goossens B (2001) Mating system and reproductive skew in the black rhinoceros. Mol Ecol 10:2031-2041

Haag T, Santos AS, Sana DA, Morato RG, Cullen Jr L, Crawshaw Jr PG, De Angelo C, Di Bitetti MS, Salzano FM, Eizirik E (2010) The effect of habitat fragmentation on the
30 genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (*Panthera onca*). Mol Ecol 19:4906-4921

Hadfield JD, Richardson DS, Burke T (2006) Towards unbiased parentage assignment:
35 combining genetic, behavioural and spatial data in a Bayesian framework. Mol Ecol 15:3715–3730

Hampton JO, Spencer PBS, Alpers DL, Twigg LE, Woolnough AP, Doust J, Higgs T, Pluske J (2004) Molecular techniques, wildlife management and the importance of
40 genetic population structure and dispersal: a case study with feral pigs. J App Ecol 41:735–743

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

Hedrick PW, Kalinowski ST (2000) Inbreeding depression in conservation biology. Ann Rev Ecol Syst 31:139-162

50 Heller R, Okello JBA, Siegismund H (2010) Can small wildlife conservancies maintain genetically stable populations of large mammals? Evidence for increased genetic drift in

geographically restricted populations of Cape buffalo in East Africa, Mol Ecol 19:1324–1334

Hutchins M, Kreger MD (2006) Rhinoceros behaviour: implications for captive
5 management and conservation. Int Zoo Yearbook 40:150-173

Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations. Mol Ecol 12:2511–2523

10 Jones AG, Small CM, Paczolt KA, Ratterman NL (2010) A practical guide to methods of parentage analysis. Mol Ecol Resources 10:6-30

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment.
15 Mol Ecol 16:1099-1106

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends Ecol Evol 17:230–241

20 Kim M (2009) The utility of DNA microsatellite markers in conservation of a Namibian population of the Black Rhinoceros (*Diceros bicornis*). MSc Thesis, Queen's University, Kingston, Ontario, Canada

Krummenacher TS, Zschokke S (2007) Inbreeding and outbreeding in African
25 rhinoceros species. Pachyderm 42:108-115

Lorenzen ED, Siegismund HR (2004) No suggestion of hybridization between the vulnerable black-faced impala (*Aepyceros melampus petersi*) and the common impala (*A. m. melampus*) in Etosha National Park, Namibia. 2004. Mol Ecol 13:3007–3019
30

Menotti-Raymond M, David VA, Wachter LL, Butler JM, O'Brien SJ (2005). An STR forensic typing system for genetic individualization of domestic cat (*Felis catus*) samples. J Forensic Sci 50:1061-1070

35 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. Cons Genet 9:239-242

40 Nsubuga AM, Robbins MM, Roeder AD, Morin, PA, Boesch C, Vigilant L (2004). Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Mol Ecol 13:2089-2094

Owen-Smith N (1977) On territoriality in ungulates and an evolutionary model. Quart
45 Rev Biol 52:1-38

Owen-Smith N (1988) Megaherbivores: the influence of very large body size on ecology. Cambridge University Press, Cambridge, UK

Piggot MP, Taylor AC (2003) Remote collection of animal DNA and its applications in conservation management and understanding the population of rare and cryptic species. Wildlife Res 30:1-13

5 Rachlow JL, Berkeley EV, Berger J (1998) Correlates of male mating strategies in white rhinos (*Ceratotherium simum*). J Mamm, 79:1317-1324

Rachlow JL, Kie JG, Berger J (1999) Territoriality and spatial patterns of white rhinoceros in Matobo National Park, Zimbabwe. Afr J Ecol 37:295-304

10

Scott CA (2008) Microsatellite variability in four contemporary rhinoceros species: implications for conservation. MSc. Thesis, Queen's University, Kingston, Ontario, Canada

15 Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9:615-629

Seror B, Zorovsky Y, Terkel A, Katcoff, DJ (2002) Genetic kinship and social structure in a herd of square-lipped rhinoceroses (*Ceratotherium simum simum*) at the Zoological
20 Center, Tel Aviv/Ramat-Gan, Israel. Zoo Biol 21:561-571

Skinner JD, Chimimba CT (2005) The mammals of the Southern African Subregion, 3rd Edition. Cambridge University Press, UK

25 Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. PNAS 101:15261–15264

Swart MKJ, Ferguson WH (1997) Conservation Implications of Genetic Differentiation in southern African Populations of Black Rhinoceros (*Diceros bicornis*). Cons Biol
30 11:79-83

Tokarska M, Marshall T, Kowalczyk R, Wojcik JM, Pertoldi C, Kristensen TN (2009) Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison SNP markers in
35 parentage analysis. Heredity 103:326-332

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535-538

40

Walling CA, Pemberton JM, Hadfield JD, Kruuh LB (2010) Comparing parentage inference software: reanalysis of a red deer pedigree. Mol Ecol 10:1914-1928

Westneat DF, Stewart IRK (2003) Extra-pair paternity in birds: causes, correlates and
45 conflict. Ann Rev Ecol Evol Syst 34:365–396

White AM, Swaisgood RR, Czekala N (2007) Differential investment in sons and daughters: do white rhinoceros mothers favor sons? J Mamm 88:632-638

50 Zschokke S, Gautschi B, Baur B (2003) Polymorphic microsatellite loci in the endangered Indian rhinoceros, *Rhinoceros unicornis*. Mol Ecol Notes 3:233-235

5 **Table 1** *Summary Statistics for Each Locus* – k: Number of Alleles, N: Sample count, H_{obs}: Observed heterozygosity, H_{exp}: Expected heterozygosity under Hardy-Weinberg (H-W) equilibrium, PIC: Polymorphic Information Content, F(Null): The F-score for the null hypothesis that the locus is in H-W equilibrium. * Locus Rh8 deviates significantly from H-W

10

Locus	k	N	H_{Obs}	H_{Exp}	PIC	F(Null)
7B (FAM)	3	34	0.53	0.54	0.47	0.01
7C (HEX)	3	29	0.59	0.53	0.43	-0.05
BR6 (HEX)	4	35	0.54	0.55	0.47	0.01
DB1 (FAM)	2	35	0.43	0.34	0.28	-0.12
DB44 (FAM)	2	35	0.40	0.33	0.27	-0.11
DB49 (FAM)	3	34	0.50	0.47	0.41	-0.05
Rh7 (M13)	2	35	0.20	0.23	0.20	0.06
Rh8 (M13) *	3	35	0.17	0.38	0.35	0.37
Rh9 (M13)	3	35	0.57	0.47	0.37	-0.11
WR1 (FAM)	3	35	0.69	0.61	0.53	-0.07
WR2 (HEX)	2	34	0.47	0.50	0.37	0.02

15

Table 2 *Summary of results from the different parentage assignment methods used* - For
20 each of the assignment methods - Manual, Cervus and Colony - the final parent pairs selected are shown. These assignments incorporate all the pedigree constraints: animal gender, tenureship, restricted parent candidates, parameters due to Date-of-Birth (DOB) of calves. Cervus and Colony provide only the most-likely candidate parent pairs while the manual assignments show all possible parent pairs available given the constraints
25 applied. Note: calves #1, #5, #7 and #16 are duplicate samples collected after ear-notching for known-mother calves; the superscript number indicates the corresponding pre-dispersal biopsy dart sample determined after parentage assignment.

		MANUAL		*CERVUS*		*COLONY*	
Calf ID	DOB	*Mother*	*Father*	*Mother*	*Father*	*Mother*	*Father*
Anne	1998	One Horn	Derek	One Horn	Derek*	One Horn	Derek
Renee	1998	Au Kooi Short Horn Short Horn	Bob Bob Derek	Short Horn	Derek*	Short Horn	Derek
Sharon	2000	One Horn	Bob	One Horn	Bob*	One Horn	Bob
Sonja	2000	Long Horn	Derek	Long Horn	Derek*	Long Horn	Derek
Janine	2000	Au Kooi Short Horn	Bob Bob	Au Kooi	Bob*	Au Kooi	Bob
Lisa	2000	Au Kooi Short Horn Short Horn	Bob Bob Derek	Short Horn	Derek*	Short Horn	Derek
Ivan	2002	Au Kooi Au Kooi	Bob Derek	Au Kooi	Derek+	Au Kooi	Derek
Diane	2003	One Horn	Derek	One Horn	Derek*	One Horn	Derek
#20	2004	Au Kooi	Derek	Au Kooi	Derek*	Au Kooi	Derek
#21	2004	Long Horn	Derek	Long Horn	Derek*	Long Horn	Derek
#22	2004	Anne	Derek	Anne	Derek*	Au Kooi	Derek
#23	2004	Short Horn Short Horn	Bob Derek	Short Horn	Derek+	Short Horn	Derek
#24	2004	Renee Renee	Bob Derek	Renee	Bob*	Renee	Bob
John	2005	One Horn	Derek	One Horn	Derek*	One Horn	Derek
Long Horn calf[1]	2006	Long Horn	Derek	Long Horn	Derek*	Long Horn	Derek
Au Kooi calf[2]	2006	Au Kooi	Derek	Au Kooi	Derek*	Au Kooi	Derek
Short Horn calf[3]	2006	Short Horn	Derek	Short Horn	Derek^	Short Horn	Derek
Renee calf[4]	2006	Renee	Jeff	Renee	Jeff*	Renee	Jeff
Sharon calf / Zac	2006	Sharon	Tony	Sharon	Tony*	Sharon	Tony
#1[4]	2006	Renee	Jeff	Renee	Jeff*	Renee	Jeff
#5[1]	2006	Long Horn	Derek	Long Horn	Derek*	Long Horn	Derek
#6	2006	Janine	Tony	Janine	Tony+	Au Kooi	Tony
#7[3]	2006	Short Horn Short Horn	Derek Jeff	Short Horn	Derek^	Short Horn	Derek
#15	2006	Lisa	Tony	Lisa	Tony+	Au Kooi	Tony
#16[2]	2006	Au Kooi	Derek	Au Kooi	Derek*	Au Kooi	Derek
Max	2006	Sonja	Tony	Sonja	Tony*	Sonja	Tony
Tom	2007	One Horn	Jeff	One Horn	Jeff*	One Horn	Jeff

Cervus assignments: *strict confidence (95%), +relaxed confidence (80%), ^indicates most-likely parent pair assigned at <80% confidence

5 **Table 3** *Heterozygosity for each Locus by Generation* – H_{obs}: Observed, H_{exp}: Expected

	F_0		F_1		F_2	
Locus	H_{Obs}	H_{Exp}	H_{Obs}	H_{Exp}	H_{Obs}	H_{Exp}
7B	0.71	0.56	0.31	0.46	0.71	0.70
7C	0.67	0.59	0.53	0.50	0.60	0.56
BR6	0.63	0.68	0.50	0.45	0.57	0.65
DB1	0.25	0.23	0.50	0.39	0.29	0.26
DB44	0.25	0.23	0.50	0.39	0.29	0.26
DB49	0.43	0.63	0.50	0.41	0.57	0.44
Rh7	0.13	0.33	0.31	0.27	0.00	0.00
Rh8	0.38	0.34	0.19	0.45	0.00	0.26
Rh9	0.75	0.59	0.63	0.48	0.29	0.26
WR1	0.75	0.69	0.75	0.57	0.43	0.62
WR2	0.43	0.54	0.63	0.51	0.14	0.54
Mean	0.49	0.49	0.49	0.44	0.35	0.41
Mean (no Rh7, Rh8)	0.54	0.53	0.54	0.46	0.43	0.48

10

Figure 1 *Assigned Parents by Maternal Generation* – Generation chart showing each matriline by generation (F0, F1 and F2)**.** F0 line represents the four founder cows. For each offspring the assigned father is shown: founder bulls are Derek and Bob and
15 introduced bulls are Tony and Jeff. Offspring are indicated in italics and date of birth and gender are given for each offspring. The single inbred individual (James) is shown in bold, and arose from Derek breeding with Anne, his daughter by One Horn. *Derek (Sire of #7 with Mother Short Horn) was the only parent assigned by Cervus with less than 80% confidence.

F0: One Horn

Derek Bob Derek Derek Jeff

F1: *Anne Female 1998* | *Sharon Female 2000* | *Diane Female 2003* | *John Male 2005* | *Tom Male 2007*

Derek Tony

F2: ***James Male 2004*** | *Zac Male 2006*

F0: Long Horn

Derek Derek Derek

F1: *Sonja Female 2000* | *# 21 Female 2004* | *# 5 Male 2006*

Tony

F2: *Max Male 2006*

F0: Au Kooi

Bob Derek Derek Derek

F1: *Janine Female 2000* | *Ivan Male 2002* | *# 20 Male 2004* | *# 16 Female 2006*

Tony

F2: *# 6 Male 2006*

F0: Short Horn

Derek Derek Derek Derek*

F1: *Renee Female 1998* | *Lisa Female 2000* | *# 23 Female 2004* | *# 7 Male 2006*

Jeff Bob Tony

F2: *# 1 Male 2006* | *# 24 Female 2004* | *# 15 Female 2006*

Electronic Supplementary Material

Parentage analysis in a managed free-ranging population of southern white rhinoceros: Genetic diversity, pedigrees and management

Conservation Genetics

Abigail Guerier, Jacqueline M. Bishop, Stuart M. Crawford, Anne Schmidt-Küntzel & Ken J. Stratford

Corresponding Author: Ken Stratford, Ongava Research Centre
Email: kenjstratford@gmail.com

S1 – Extraction Protocol for Trophy Mount Sample

DNA from the trophy mount sample was extracted using a DNeasy Tissue Kit (Qiagen) with a number of modifications to the manufacturer's recommended protocol. Three extractions were performed in parallel to increase chances of success, and one extraction blank was included in order to test for contamination. A first extraction was performed following the standard tissue extraction protocol; the concentration of the resulting DNA aliquot was increased through evaporation. A second extraction was performed with the following modifications: an additional wash step with Buffer AW1 in order to eliminate salts present in preserved skin (according to a recommendation from Qiagen customer service) and the DNA of the three extractions performed in parallel was eluted successively with the same elution buffer aliquot, thereby optimizing the concentration of DNA from the three extractions in a total of 100µl. This step was repeated with a second aliquot of 100µl in order to obtain a maximum yield of DNA, and both elutions were used for PCR.

S2 – Microsatellite Genotyping Protocol

Samples genotyped at the University of Cape Town (UCT) were carried out in 0.2ml, thin-walled PCR tubes and contained the following reagents: 10 to 50 ng genomic DNA, 0.25 units of HotStarTaq™ Flexi DNA polymerase (Promega), 1× reaction buffer (final concentration: 16mM $[NH_4]_2SO_4$; 67mM Tris-HCl pH 8.8; 0.01% Tween-20), 0.3µM fluorescently labelled (FAM™ or HEX™ fluorescent dye) forward and unlabelled reverse primers, 2mM $MgCl_2$ and 0.2mM dNTPs in a final reaction volume of 20µl. The PCR protocol for eight of the loci (See S3) consisted of one denaturing step at 94°C for 2 min, followed by 35 cycles of 94°C for 30 seconds, the specific annealing temperature for 30 seconds (following Coutts, 2009), 72°C for 45 seconds, and followed

by a final extension step of 72°C for 10 minutes. Conditions for the three loci isolated in Indian rhinoceros (Rh7, Rh8, Rh9) contained 0.05µM forward primer with a M13 sequence appended to the 5' end (5'-TGTAAAACGACGGCCAGT-forward primer-3') and 0.2µM reverse primer and M13 primer labelled with either FAM™ or HEX™ fluorescent dye. The PCR protocol consisted of one denaturing step at 94°C for 2 min, followed by 30 cycles of 94°C for 30 seconds, annealing temperature of 55°C or 57°C for 30 seconds, 72°C for 45 seconds, and then eight cycles of 94°C for 30 seconds, 53°C for 30 seconds, 72°C for 45 seconds and then followed by a final extension step of 72°C for 10 minutes. PCR products were visualised by automated fragment size analysis on an ABI Prism™ 373 Genetic Sequencer (Applied Biosystems) by polyacrylamide gel electrophoresis. Fragment size data was collected and analysed using GENESCAN™ 672 Data Collection Software version 1.2.1 (Applied Biosystems). Microsatellite alleles were sized manually relative to an internal size standard (GeneScan™ Rox-350, Applied Biosystems). A positive control of known size was used in each run to ensure consistency of allele fragment size scoring between gels; a negative control was also included on each gel.

Samples genotyped at the Cheetah Conservation Fund genetics laboratory contained 10 to 50 ng genomic DNA, 5µl of AmpliTaq Gold ® MasterMix, 0.4µM fluorescently labelled (FAM™ or HEX™ fluorescent dye) forward and unlabelled reverse primers and 0.2µg/µl of BSA in final reaction volume of 10µl. The PCR protocol for all loci except WR2, consisted of one initial step at 95 °C for 10 minutes, followed by 10 touchdown cycles, for which annealing temperature is reduced by 1 °C at each cycle: 95 °C for 15 seconds, 60 °C - 51 °C annealing temperature for 30 seconds, 72 °C for 45 seconds, followed by 30 cycles consisting of one denaturing step at 95 °C for 15 seconds, 50 °C for 30 seconds, 72 °C for 45 seconds, and followed by a final extension step of 72 °C for 30 minutes. For microsatellite locus WR2, 12 touchdown cycles were performed with annealing temperature varying from 60 °C to 49°C and an annealing temperature of 48 °C for the 30 remaining cycles. For the trophy sample, a modified touchdown PCR program was used: the protocol consisted of one initial step at 95 °C for 10 minutes, followed by 10 touchdown cycles, for which annealing temperature is reduced by 1 °C at each cycle: 95 °C for 15 seconds, 60 °C - 51 °C annealing temperature for 30 seconds, 72 °C for 45 seconds, an additional 10 cycles consisting of 95 °C for 15 seconds, 50 °C for 30 seconds, 72 °C for 45 seconds, followed by 30 cycles consisting of 95 °C for 15 seconds, 48 °C for 30 seconds, 72 °C for 45 seconds and followed by a final extension step of 72 °C for 30 minutes. PCR products were visualised on an ABI Prism™ 310 Genetic Analyser (Applied Biosystems) by capillary electrophoresis. The fragment size data was collected using Data Collection software version 3.1 (Applied Biosystems) and analysed using

Genemapper software version 4.0 (Applied Biosystems). Microsatellite allele sizes were sized against internal size standard (GeneScan™ Liz-500, Applied Biosystems). For each microsatellite locus an individual with known genotype was run to calibrate the results to the initial genotype data obtained at UCT and allele sizes adjusted accordingly.

S3 – Primer details - Specification for the 11 microsatellite loci used in this study.

Locus	Repeat motif/microsatellite motif	Primer pair sequence (5'-3')	Reference
7B	$(TG)_{16}$	F: CCTCTGTGATTAAGCAAGGC R: ATGAACAGGAAGGAAGACGC	Florescu *et al* (2003)
7C	$(CT)_{14}(AT)_{11}$	F: TGAACTCTGATGGAAATGAG R: AACAGGTCTTGATTAGTGC	Florescu *et al* (2003) adapted by Coutts (2009)
DB1	$(CA)_{14}$	F: AGATAATAATAGGACCCTGCTCCC R: GAGGGTTTATTGTGAATGAGGC	Brown & Houlden (1999)
DB44	$(CA)_4G(CA)_{16}$	F: GGTGGAATGTCAAGTAGCGG R: CTTGTTGCCCCATCCCTG	Brown & Houlden (1999)
DB49	$(CA)_{14}$	F: GTCAGGCATTGGCAGGAAG R: CAGGGTAAGTGGGGGTGC	Brown & Houlden (1999)
BR6	$(CA)_{15}$	F: TCATTTCTTTGTTCCCCATAGCAC R: AGCAATATCCACGATATGTGAAGG	Cunningham *et al* (1999)
Rh7	$(TG)_{17}$	F: CCGTCACATATGACAGTGTGC R: GGGCAGCTTATGCTCAAGTC	Zschokke *et al (*2003)
Rh8	$(TG)_{22}(AG)_2ANACA(GA)_{28}$ $CA(GA)_3CA(GA)_9CG(TA)_5$	F: ACACACCTTTATAAACAATATGGTCAC R: AGTCTCTAGTCAAAAGGGATTGC	Zschokke *et al (*2003)
Rh9	$(TG)_4TT(TG)_{17}TA\ (TG)_5$	F: TCTGGTACCACCAAATGTAGC R: ACGATTACGTCTTTCAGTTGC	Zschokke *et al (*2003)
WR1 / AY138542	$(AC)_{18}$	F: GGCAAAACTAAGAGAACTTG R: GATACCAAACTGGAAATGG	Nielsen *et al* (2008) redesigned Florescu *et al* (2003)
WR2 / AY138545	$(TA)_8(CA)_4$	F: ACAGCTAGAATCACCAAAAC R: TCCTGCTGCATAAATCTC	Nielsen *et al* (2008) redesigned from Florescu *et al* (2003)

Locus 7C was not completed for all individuals as the published sequence was incorrect; hence we were unable to obtain genotypes for this locus in the second phase of our analysis

S4 - Microsatellite Genotyping and Error Rates

Initially a total of 321 genotypes were obtained for 32 samples (28 individuals plus 4 duplicates) across the 11 loci. In order to assess the validity of these results a sub-set of 102 genotypes was independently scored. This process resulted in 11 new genotypes (9 'blanks' filled in and 2 errors in calls corrected – 0.6%), while independent scoring supported the remaining results. These results were incorporated into the genotype table. At CCF three new samples were added to the study set (see above Supplementary Material S2 for details) providing an additional 26 genotypes (3 individuals across 10 loci – 7C excluded). A preliminary review of the genotype table in light of known pedigree information allowed us to identify possible errors occurring in the data table. These 'alerts' included occurrences such as known calf-mother mismatches and instances where calves could not be assigned to any of the candidate fathers in the population and an absence of match between duplicate samples. In order to investigate these anomalies further and evaluate error rates in the data set repeat PCR's were performed for a sub-set of 128 samples across targeted and random loci and individuals. The results of these PCR's gave rise to 29 new or altered genotypes including 17 results for previous 'blanks' and 12 corrections (3.75%) to previous incorrect results. Results that were altered were repeated again to confirm the new results. A total of 154 genotypes were obtained from PCR's performed at CCF. Corrected genotypes were incorporated into the final data set.

S5 – Manual Parentage Assignment

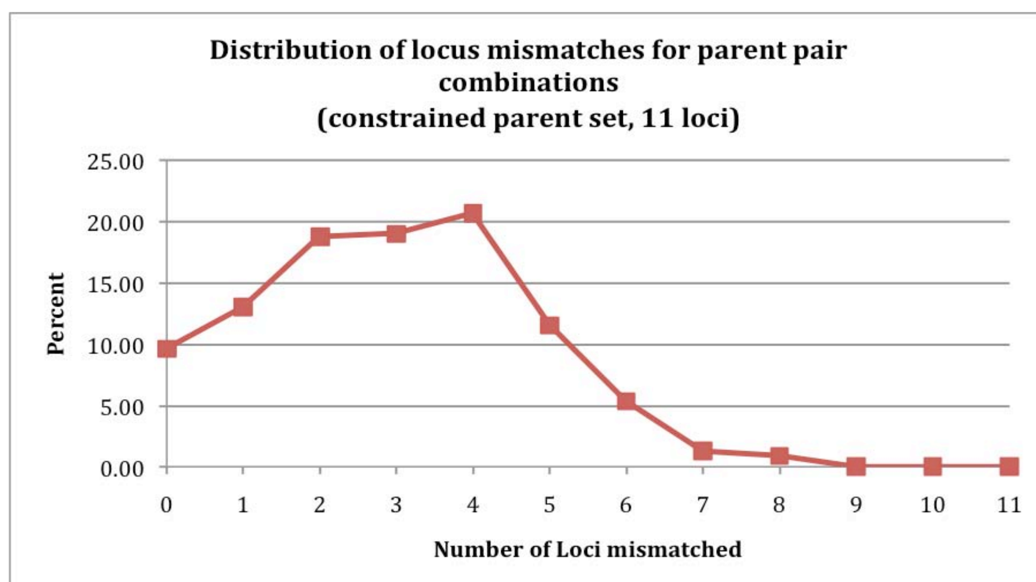
The first step in the manual assignment of parentage was to calculate the number of possible parent pairs in the candidate data set based on all males and females in the total population; this was done using the in-house software (available on request K. Stratford). The next step was to narrow the possible parents selection down to only those male and female candidates which would have been of breeding age and present on Ongava Game Reserve at the time of conception for all the offspring. For this step each offspring's genotype was compared to the afore-mentioned parent candidates' genotypes and individuals were discarded based on mismatches at one (or more) locus. The influence that the presence of a possible null allele at Rh8 was also investigated by repeating step 2 and removing Rh8 from all the genotypes. The fourth and final step in refining the parentage assignments was to narrow the candidate parent options down according to the calf's date of birth (i.e. only mothers which gave birth in that particular year available as candidates) and also to exclude mothers which had been previously unequivocally assigned in step 2 and 3 (e.g. if three calves are born in a particular year, and for one of these calves a single mother has been assigned in step 2 and 3, then that mother is excluded as a candidate for the remaining two calves). A similar process was used for the computational assignment of parentage; the first step included all males

and females in the total population as candidate father and mothers respectively to allow for unrestricted assignments based solely on genotypes. The second step provided only candidate parents to the program based on their breeding age and residency on the Reserve relative to the times of conception. Step 3 excluded Rh8 as per the manual assignments. Step 4 further refined the candidates from step 2 by narrowing that candidate selection down to only provide those parents available based on the calf's date of birth. The results from this analysis were then recorded as the set of possible parent pairs for each offspring.

S6 – Analysis of Genotype Mismatches

We explored the potential problem of 'false' parent pair exclusions due to genotype mismatches. For both the unconstrained case (all possible parent pair matches) and the constrained case (candidate parents restricted according to prior pedigree information) we computed the frequency distribution of locus mismatches (see Figure S6-1 below shows this distribution for the constrained case). In the constrained case, possible matches were obtained in 9.6% of pairings. Relatively few parent-pair rejections occurred as the result of single locus mismatches (13%) while approximately three-quarters (77.4%) of candidate parent pairs were excluded based on genotype mismatches at more than one locus.

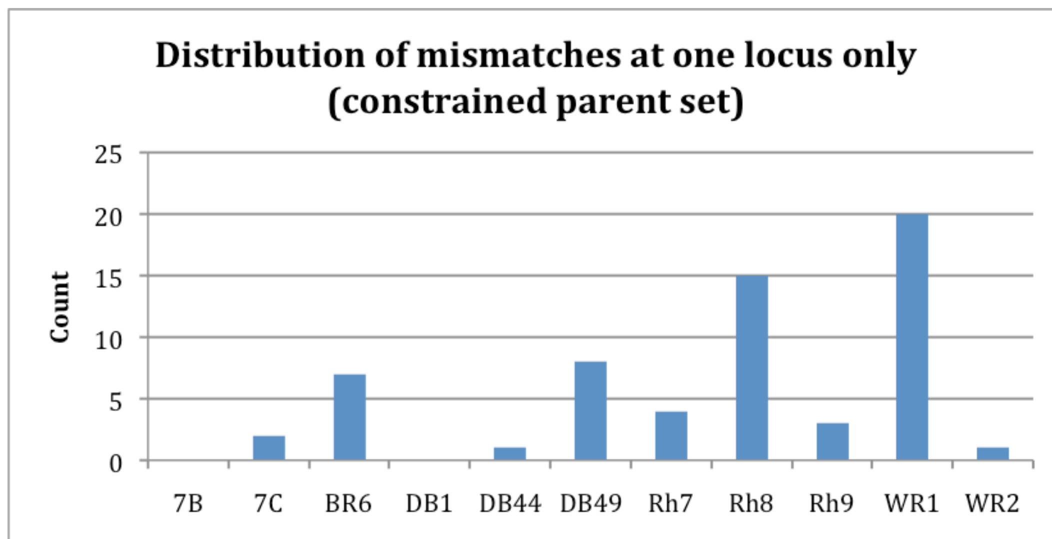
Figure S6-1 – Distribution of locus mismatches for the constrained case of candidate parent pairs



We then analysed further the distribution of loci for those parent pairs excluded based on single locus mismatch. Figure S6-2 shows the frequency distribution of these single locus exclusions by locus. Locus WR1 (the most variable locus, PIC=0.53) contributed most instances of these exclusions (20 of 61), with locus Rh8 (a locus with a null-allele) contributing 15 of 61 mismatches. Exclusions based on single locus mismatches for these loci combined led to about 7% of the total

number of exclusions for the constrained parent set. Given the low genotype error rate, in the context of the full study we do not think that any erroneous exclusions at these loci will have significantly biased the assignments.

Figure S6-2 – Distribution by locus of single locus mismatches for the constrained case of candidate parent pairs



References

Brown SM, Houlden BA (1999) Isolation and characterization of microsatellite markers in the black rhinos (*D. bicornis*). Mol Ecol 8:1559-1561

Coutts NJ (2009) Investigating genetic diversity at neutral and adaptive DNA markers in the severely bottlenecked Southern white rhinoceros (*Ceratotherium simum simum*). PhD Thesis, University of Cape Town

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

Florescu A, Davila JA, Scott C, Fernando P, Kellner K, Morales JC, Melnick D, Boag PT, Van Voeverden De Groot P (2003) Polymorphic microsatellites in the white rhinoceros. Mol Ecol Notes 3:344-345

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. Cons Genet 9:239-242

Zschokke S, Gautschi B, Baur B (2003) Polymorphic microsatellite loci in the endangered Indian rhinoceros, *Rhinoceros unicornis*. Mol Ecol Notes 3:233-235