CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/136651/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Magliolo, Michelle, Prost, Stefan, Orozco Ter Wengel, Pablo , Kropff, Anna S., Kotze, Antoinette, Groblem, J. Paul and Dalton, Desiree Lee 2021. Unlocking the potential of a validated single nucleotide polymorphism array for genomic monitoring of trade in cheetahs (Acinonyx jubatus). Molecular Biology Reports 48 , pp. 171-181. 10.1007/s11033-020-06030-0

Publishers page: http://dx.doi.org/10.1007/s11033-020-06030-0

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Unlocking the potential of a validated Single Nucleotide Polymorphism array for genomic monitoring of trade in cheetahs (*Acinonyx jubatus*)

Michelle Magliolo[1,2], Stefan Prost[1,4], Pablo Orozco-terWengel[5], Pamela Burger[3], Anna S. Kropff[1], Antoinette Kotze[1,2], J. Paul Grobler[2], Desire Lee Dalton[1,2]

[1]South African National Biodiversity Institute, P.O. Box 754, Pretoria, 0001, South Africa

[2]Department of Genetics, University of the Free State, P.O. Box 339, Bloemfontein, 9300, South Africa

[3]Research Institute of Wildlife Ecology, Department of Interdisciplinary Life Sciences, Vetmeduni Vienna, Vienna, Austria

[4]LOEWE-Centre for Translational Biodiversity Genomics, Senckenberg Nature Research Society, Frankfurt, Germany

[5]School of Biosciences, Cardiff University, Cardiff, United Kingdom

*Corresponding Author, e-mail: d.dalton@sanbi.org.za

Background: Cheetahs (*Acinonyx jubatus*) are listed as vulnerable on the International Union for Conservation of Nature Red List of Threatened Species. Threats include loss of habitat, human-wildlife conflict and illegal wildlife trade. In South Africa, the export of wild cheetah is a restricted activity under the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), however, limited legal trade is permitted of animals born to captive parents. To effectively monitor the legal and illegal trade in South Africa, it was thus essential to develop a validated molecular test.

Methods and Results: Here, we designed a single nucleotide polymorphism (SNP) array for cheetah from Double Digest Restriction Associated DNA sequencing data for individual identification and parentage testing. In order to validate the array, unrelated individuals and 16 family groups consisting of both parents and one to three offspring were genotyped using the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System. In addition, parentage assignments were compared to microsatellite data. Cross-species amplification was tested in various felids and cheetah sub-species in order to determine the utility of the SNP array in other species. We obtained successful genotyping results for 218 SNPs in cheetah (*A. j. jubatus*) with an optimal DNA input concentration ranging from 10 to 30 ng/µl. The combination of SNPs had a higher resolving power for individual identification compared to microsatellites and provided high assignment accuracy in known pedigrees. Cross-species amplification in other felids was determined to be limited. However, the SNP array demonstrated a clear genetic discrimination of two cheetah subspecies tested here.

Conclusions: We conclude that the described SNP array is suitable for accurate parentage assignment and provides an important traceability tool for forensic investigations of cheetah trade.

Keywords: Single Nucleotide Polymorphism; Parentage; Forensics; Cheetah

Introduction

Advances in molecular technologies have resulted in the development of DNA markers that are polymorphic, accurate and robust for parentage analysis [1] and forensic investigations [2]. The most commonly used marker type is microsatellites which have a high mutation rate resulting in high information content. However, these markers have several negative characteristics including size homoplasy and analysis that may be complicated by complex mutation patterns and/or the presence of null alleles, which can introduce ambiguity to data analysis [3-5]. Furthermore, microsatellite analysis is platform dependent, and there are often difficulties with standardization, thus collaboration between laboratories is challenging [3]. Single Nucleotide Polymorphisms (SNPs) are rapidly replacing microsatellites in parentage studies. In contrast to microsatellite analysis, SNPs provide much broader genome coverage [5], they offer higher information recovery from degraded DNA samples making them useful in forensics [4-5] and they can be automatically standardized between laboratories [3]. Due to these qualities we selected a set of SNP markers with the aim of developing a reproducible, low cost, high throughput test that can accurately assign parentage and can be used for individual identification of captive cheetah in order to monitor the legal and illegal trade.

Cheetahs are listed on Appendix I of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), thus, export of animals obtained from the wild for commercial purposes is prohibited. Limited annual exceptions are given to Botswana (n = 5), Namibia (n = 150) and Zimbabwe (n = 50) (https://cites.org/eng/app/appendices.php; accessed 25th of October 2019). Current legislation in South Africa allows for the commercial breeding of predators, including cheetahs, and export permits may be provided for trade if an individual is bred in captivity. The Scientific Authority has been established in South Africa in terms of Section 60 of National Environmental Management: Biodiversity Act (NEMBA). The purpose of the Scientific Authority is to assist with regulating and restricting trade in specimens of Threatened or Protected Species (TOPS) and CITES-listed species. The Scientific Authority has recommended various measures to improve the management of captive-bred cheetahs and to ensure that no wild specimens are traded as “captive-bred”. These measures include: (1) registration of captive cheetah in a studbook, (2) all cheetah in captivity must be identifiable through microchips, photos and DNA fingerprints, (3) all specimens to be exported internationally must first be verified as offspring of captive-bred parents through DNA analyses and (4) any facility exporting internationally must be registered with the Management Authority in compliance with the TOPS and CITES regulations. A management system using forensic principles has been developed in order to monitor the legal trade of cheetah to ensure that only captive bred individuals are exported. This process involves the collection of blood samples from the potential sire, dam and offspring that are placed in a forensic evidence bag in the presence of a conservation official. The entire process is documented with a series of photographs. Additional information provided includes microchip and studbook numbers of each individual and various measurements of the animal. A DNA passport is created following microsatellite profiling using markers derived from domestic cat microsatellites [6-8] once parentage is confirmed. The passport includes photos of the offspring, relevant identification numbers (microchip, studbook etc.) and identifies the dam and sire. All profiles are included in a DNA cheetah trade database that is managed by South

African National Biodiversity Institute (SANBI). In addition to parentage verification, a validated molecular test is required to assist in combating the illegal trade in live cheetah and their body parts. Illegal wildlife trade includes the sale, purchase, or exchange of wildlife species or parts which is prohibited by law [9]. Historically, live cheetahs were removed from the wild and were kept by the aristocracy as exotic pets due to their relatively docile nature or were used in sports for the hunting of antelopes [10]. In addition, cheetahs have been traded internationally for their skins (an estimated 3,000 to 5,000 skins were imported annually in the 1960s) to be used for clothing and ornaments in Europe and North America [11]. Currently, live cheetahs are still illegally taken from the wild to be exploited as pets or as tourist attractions, are illegally collected and introduced into captive breeding facilities [12], killed in illegal trophy hunts, or their body parts sold as garments, ornaments and traditional medicines [13]. The illegal trade, in conjunction with other pressures, is considered a threat to the survival of the remaining populations throughout their range.

In this study, we selected a subset of SNPs from double digest restriction site associated DNA (ddRAD) sequence data. Our aim was to genotype cheetah samples using the developed SNP array on the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System and validate the array for parentage assignment and forensic applications to monitor the illegal and legal trade of cheetah in South Africa. Here, we (1) compare the SNP array to currently available microsatellite markers to assess their performance in resolving parentage, (2) assess the ability of the SNP panel to delineate cheetah subspecies and lastly (3) determine cross-species amplification of the SNP array in other cat species.

Materials and Methods

DNA extraction

The list of samples and species used in this study are provided in Supplementary Table 1. Samples were stored in the SANBI Biobank at -80°C and were thawed at room temperature prior to isolation. Genomic DNA was isolated from blood samples using the Zymo Research Quick-DNA™ Miniprep Plus Kit, following the manufacturer's instructions. The quality (A260/A280) and quantity of DNA samples was determined using the NanoDrop Spectrophotometer ND-1000.

SNP discovery and selection for array design

Animals were selected from a variety of localities and a subset of cheetah included animals of known parentage (parents and offspring) to ensure that generated SNPs were informative for the purposes of parentage assignments and forensic applications. The set of animals chosen for this work was determined on the availability of a wider range of potentially different genetic backgrounds, in light of the constraints of sample availability, and to maximize the chance of finding polymorphic sites with discriminatory resolution between individuals. Double Digest Restriction Associated DNA sequencing (ddRADseq) was conducted by IGATech (Udine, Italy) on 42 samples with minor changes to the protocol outlined in Peterson et al. [14], using the SphI and HindIII enzymes for our samples (inferred using *in silico* digestion to obtain around 45,000 loci). Sequencing was carried out on the HiSeq2500

instrument (Illumina, San Diego, CA) using V4 chemistry and paired-end 125 bp mode. The reads were then mapped against the cheetah reference genome (GCF_001443585.1_aciJub1_genomic.fna) using BWA mem [15] and sorted using SAMtools version 1.9 [16] using default settings and selecting only reads that map to a single place on the reference and show a mapping quality of >40. Coverage and other mapping quality statistics were obtained using qualimap [17]. The covered loci were filtered using *gstacks* and *populations* using Stacks v2.0 software [18]. Loci had to be present in over 75% of all individuals. PLINK [19] was used to remove those markers with a minimum allele frequency of 10%, as above that threshold the alleles had a uniform distribution. Filtering for loci in Hardy-Weinberg Equilibrium was carried out on the 6,404 SNPs that passed the previous filtering. An updated p-value threshold of 0.0535 was applied following Narum [20] which resulted in the removal of 341 loci due to deviation from Hardy-Weinberg Equilibrium. SNPs were sorted based of their expected heterozygosity from higher to lower, and a random set was selected among the 400 most diverse SNPs. Furthermore, an estimation of the linkage disequilibrium (LD) between pairs of SNPs (on the same contig and no more than 1 megabase apart from each other) in order to remove SNPs with high LD and close proximity was carried out. Based on the LD decay curve (Supplementary Figure 1) one SNP from each pair closer than 40Kb from each other and presenting r2 value higher than 7% was removed. These distances and r2 thresholds were chosen as LD stabilizes around 6% from 40Kb onwards. Identity-By-Descent (IBD) analysis was conducted and pi_hat values were examined. However, pi_hat can be upwardly biased due to increased linkage disequilibrium (LD) due to population structure [21] and a reduced genetic variation due to inbreeding [19]. Due to these factors likely playing a role in our dataset, we focused instead in the measurements of sharing of alleles between individuals. Z0 (no alleles shared: IBD=0), Z1 (one allele shared: IBD=1) and Z2 (both alleles shared: IBD=1) values were used to identify the number of shared alleles between pairs of individuals. Related individuals were then removed. We further removed SNPs in repetitive regions of the genome. We also validated the heterozygosity of the inferred SNPs using an independent pipeline, ANGSD [22]. To do so, we generated major and minor allele frequency counts using the genotype likelihood method implemented in ANGSD. We required a heterozygous position to show a p-value lower than 1×10^{-6} and a base quality (phred score) of at least 20.

SNP genotyping

SNP genotyping was performed using the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System (Applied Biosystems, CA, USA) according to the manufacturer's recommended operating conditions. A total of 159 samples were genotyped (Supplementary Table 1) with 240 SNPs (Supplementary Table 2). In short, equal volumes of sample and TaqMan OpenArray MasterMix were added into a 96 well plate and centrifuged at 4,100 revolutions per minute (rpm) and the sample locations were recorded in the OpenArray® Sample Tracker Software. Mixtures were transferred into the 384 well plate and centrifuged at 4,100 rpm. The 384 well plate has a capacity to run 12 samples and 240 SNPs in a single run. The 384 well plate was placed in OpenArray® AccuFill™ System and the QuantStudio™ TaqMan® OpenArray® AccuFill™ Software was set up. Once loaded, SNP arrays were sealed with the OpenArray® Case Lid, using the QuantStudio™ 12K Flex OpenArray® Plate Press 2.0. Immersion fluid was

poured evenly into the array and the SNP array was immediately loaded into the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System and run. The genotype data were analyzed using the TaqMan® Genotyper software (version 1.0.5) (Applied Biosystems, CA, USA). SNP marker genotyping data was exported from TaqMan® Genotyper software as a comma-separated file.

Microsatellite analysis

Individual profiles were obtained for 24 microsatellites, previously confirmed as polymorphic in Southern African cheetahs [7-8, 23], for performance comparisons between conventional microsatellites and the newly developed SNP markers (Supplementary Table 3). Polymerase Chain Reaction (PCR) amplification was conducted in a 12.5 µl reaction volume consisting of Ampliqon Taq DNA Polymerase Master Mix RED, forward and reverse primers (0.5 µM each), and 50 ng genomic DNA template. The conditions for PCR amplification were as follows: 5 min at 95°C denaturation, 30 cycles for 30 sec at 95°C, 30 sec at 55-60°C (depending on the marker amplified), and 30 sec at 72°C, followed by extension at 72°C for 40 min in a T100 Thermal Cycler (Bio-Rad Laboratories, Inc.). Based on size range and fluorescent dye of the markers, PCR products were pooled and analyzed on an ABI 3130 Genetic Analyzer (Applied Biosystem Inc.). At each locus, allele sizes were estimated by comparison with 500-LIZ® (Applied Biosystem Inc.). One standard individual was run in each genotyping series, in parallel with negative controls. Data were collected and analyzed using GeneScan® 1.2.2-1 and Genotyper® 3.1 (Applied Biosystem Inc.).

Array performance

Array performance was tested using 76 *A. j. jubatus* samples from South Africa (Supplementary Table 1) and a number of approaches were used to validate the set of SNPs. First, the repeatability of the marker set was assessed by analysis of ten samples in duplicate with reactions carried out by two different technicians on different days. Secondly, the optimal concentration range was determined by comparing the results obtained for five samples (in duplicate) at six dilutions (40 ng/µl, 30 ng/µl, 25 ng/µl, 20 ng/µl, 10 ng/µl and 5 ng/µl). The minor allele frequency (MAF) for each SNP was calculated and deviations from Hardy-Weinberg equilibrium (HWE) were tested using Genepop version 4.2.2 [24] on the 76 *A. j. jubatus* samples with analysis calculated per population (for populations that included more than eight individuals). Lastly, in order to evaluate the information content of each marker, the Polymorphic Information Content (PIC) for each locus was calculated using CERVUS [25]. In addition, to determine the discriminatory power of each marker, the Probability of Identity (PI), Probability of Identity for Sibs (PISib) and Probability of Exclusion (Pe1, one parent known and Pe2, both parents known) was calculated in GenALEx [26]. Mean number of alleles per locus (N_a), number of effective alleles (N_e), average observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased heterozygosity (H_z = expected heterozygosity adjusted for unequal sample sizes; [27]) and inbreeding coefficient (F_{IS}) were used to estimate the level of genetic diversity and heterozygote deficit or excess (GenAlEx 6.5; [26]) in cheetah using microsatellite (number of samples = 78) and SNP data (number of samples = 76). The PI of the overall marker sets was assessed in GenAlEx 6.5 to calculate the minimum number of microsatellite or SNP markers necessary to achieve a reliable individual genotyping for unrelated (PI) and related samples (PISib).

Parentage assignment

To validate the SNP array for parentage assignment and to determine the general resolving power of SNPs and microsatellites, we used the genotype data of 16 captive-bred cheetah families (n = 53), which included both parents and one to three offspring, as well as 20 unrelated samples. The logarithm of the odds (LOD) score of the marker set was evaluated using the cheetah samples of known parentage to assess the ability of the marker set to correctly determine the true sire/dam and to estimate the impact of allowing a range of mismatches. Parentage of known pedigrees was confirmed using the likelihood-based approach in CERVUS version 3.0.7 [25]. Allele frequency data were used to run the simulations and to calculate the mean proportion of loci typed and the mean PIC. The mean proportion of loci typed was used as an input value for the simulation of parentage analysis. To simulate parentage analysis, we needed to estimate which proportion of the candidate parents were sampled, as well as the proportion of loci that were mistyped. Parentage was simulated assuming 1% of loci were mistyped due to genotyping errors and that only 50% of the candidate parents had been sampled. For the simulation and subsequent parentage analysis, it was assumed that the gender of the parents was known. This was conducted for both the microsatellite and SNP data. Further, we conducted Discriminant Analysis of PCs (DAPC) [28] using the adegenet package [29] in R to assign individuals to family groups. To reduce over- or under-discrimination, we selected the number of retained PCs by predicting the maximum α-score with the optim.a.score function (20 replicate α-scores were calculated), following Jombart et al [28]. Lastly, the number of SNPs required to determine parentage accurately using actual data was assessed. In order to do this randomly, a custom python code was written that uses the built-in randint function. The accuracy of parentage assignments was determined using the following percentages of missing data: 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and 100%.

Cross-species and subspecies amplification

Cross-(sub)species amplification of the developed SNP array was assessed by genotyping two subspecies of cheetah: 76 samples of *A. j. jubatus* and 21 samples of *A. j. soemmeringii*, as well as nine lion (*Panthera leo*), eight leopards (*P. pardus*), seven domestic cats (*Felis catus*) and seven African wildcats (*F. silvestris lybica*). We further investigated the marker set's ability to accurately assign samples to their respective species/sub-species applying Principal Component Analysis (PCA). To do so, we first carried out an *in silico* PCA by extracting the 240 SNPs from the data set of Prost et al. [30] and running ANGSD and PCangsd [31]. This data set includes individuals of all five classically described subspecies (*A. j. jubatus*, *A. j. raineyi*, *A. j. soemmeringii, A. j. hecki* and *A. j. venaticus*). Next, we validated the discrimination *in vitro* using the 76 samples of *A. j. jubatus* and 21 samples of *A. j. soemmeringii* and subsequent PCA analysis using the R stats package (R Core Team, 2013).

Results

SNP discovery

Samples submitted for ddRADseq had an average A260/A280 of 1.7 and were 50 ng/µl. Sequencing on the HiSeq2500 instrument yielded a total of 110,142,763 reads across the 42 samples. A total of 13,053 SNPs were

identified in 1,196 contigs. Removal of markers with a minimum allele frequency of 10% resulted in 6,404 SNPs and a further 341 loci were removed due to deviations from Hardy-Weinberg Equilibrium. A total of 400 diverse SNPs were selected based on expected heterozygosity. Lastly, further markers were removed due to high LD or those that occurred in repetitive regions in the genome. Based on these parameters, we selected 240 SNP for the SNP array development.

Array performance

To obtain a robust SNP marker set for reliable identification and parentage testing in cheetahs, we excluded a total of 22 SNPs that were monomorphic (Supplementary Figure 2a), produced diffuse clusters that could not be accurately genotyped (Supplementary Figure 2b) or were heterozygous for all samples in cheetah (Supplementary Figure 2c). Thus, the final dataset consisted of 219 polymorphic loci that could successfully be amplified (Supplementary Figure 2d). We generated SNP data for 83 cheetah samples of which seven samples failed to amplify and were thus excluded from further analysis. In the remaining 76 samples, 92% of the samples had a genotyping success rate between 91-100% while the remaining 8% had a genotyping success rate between 77-90% (Supplementary Figure 3). The analysis of the 10 samples in duplicate resulted in identical genotypes for each individual. In addition, the five samples that were analyzed at various concentrations displayed identical genotypes. Sample concentrations as low as 5 ng/µl and as high as 40 ng/µl still produced usable results. Low concentration samples (5 ng/µl) resulted in an average loss of 4.5 SNPs whereas at higher concentrations (40 ng/µl) an average loss of 6.5 SNPs was observed. Thus, the optimal DNA concentration for successful amplification was determined to be 10-30 ng/µl (Supplementary Figure 4).The PI of the SNPs varied from 0.375 to 0.634 (average = 0.383), PISib values varied from 0.594 to 0.799 (average = 0.603), PIC values varied from 0.195 to 0.375 (average = 0.368), and Pe1 and Pe2 varied from 0.024 to 0.125 (average = 0.119) and 0.167 to 0.281 (average = 0.277) respectively. The MAF for SNPs varied from 0.125 to 0.5 (mean = 0.44, Supplementary Table 4). Only one SNP (2648128-2648228) had a MAF below 0.25 and was excluded from further analysis. Among the 218 selected SNPs, 42 SNPs had H_o values below 0.1 or above 0.9 which may results in an excess of heterozygotes or homozygotes. A total of 37 of the 42 markers deviated from HWE ($p<0.001$) when significance levels were adjusted using the False Discovery Rate (FDR). In addition, five of the 42 markers deviated from HWE with lower significance values ($p<0.01$, $p<0.05$). Further downstream analysis was conducted where the above SNPs were excluded (n = 176 SNPs) and included (n = 218 SNPs) and results obtained were similar.

To examine the effectiveness of the SNP array, samples analyzed using SNPs were re-genotyped using microsatellites. Population summary statistics using a subset of individuals obtained from SNPs and microsatellites were compared (Table 1). The mean H_o and H_e per locus for SNPs was 0.57 and 0.49 respectively whereas the mean H_o and H_e per locus for microsatellites was 0.60 and 0.74 respectively. SNPs indicated slightly too moderately outbred populations (-0.166 to -0.001) while microsatellites indicated that populations were slightly inbred (0.20). The combination of 218 SNPs (PI = 1.9×10^{-91} and PISib = 1.2×10^{-48} for full siblings and non-relatives respectively) and the combination of 176 SNPs (PI = 5.6×10^{-74} and PISib = 2.2×10^{-39}) had a higher resolving

power for individual identification in comparison to a panel of 24 microsatellites (PI = 4.0 x 10^{-27} and PISib = 2.5 x 10^{-10} for full siblings and non-relatives respectively). A PI cutoff of 0.0001 was selected as it is considered sufficiently low for most application in natural populations [32]. Based on these criteria, our results indicate that, in cheetah, approximately 10 SNPs (PI = 6.5×10^{-5}) provide a resolution in identifying individuals comparable to that provided by approximately four microsatellites (PI = 1.9×10^{-4}). The more conservative estimate PISib indicate that in cheetah, approximately 18 SNPs (PISib = 1.1×10^{-4}) provide a resolution in identifying individuals comparable to that provided by approximately 10 microsatellites (PISib = 6.9×10^{-5}) (Figure 1).

Parentage assignments

As an initial application of the generated data, we examined the utility of the developed SNP array in comparison to microsatellites for parentage analyses. When both the dam and sire SNP genotypes were included in the analysis, the average LOD score was 46.95. In this case, all 22 offspring (100%) were assigned to the correct parent pair. When paternity was assigned to the progeny using only the sire information the average LOD score was reduced to an average of 22.61. Here, one offspring failed to be assigned to its known sire (LOD score = -5.65). In addition, when all known parents were removed from the analysis, no parent pairs were assigned to any offspring at 95% confidence (Figure 2). Further, DAPC analysis identified 16 groups corresponding to the 16 families analyzed here with 21 out of 22 offspring being correctly assigned (Figure 3). Contrastingly, using the microsatellites marker test, when both known parents were included, the average LOD score was 27.68 and 18 offspring (81%) were assigned to the correct parent pair. A total of four offspring were not assigned to their known parents with high confidence due to mismatches at four to six loci (average LOD = -7.78). When paternity was assigned to the progeny using only the sire information the average LOD score was reduced to an average of 11.34. Here, five offspring failed to be assigned to its known sire (average LOD = -5.22) (Figure 2). The number of SNPs required to determine parentage accurately using actual data was assessed (Supplementary Figure 5). Parentage was assigned correctly for all 22 offspring when 10 to 50% of the SNPs were missing. A total of 19 to 20 offspring were correctly assigned when 60 to 80% of the data was missing and at 90% missing data, the number of offspring correctly assigned was further reduced to only 10 offspring correctly assigned.

Cross-species and subspecies amplification

The SNP array was applied to closely related cheetah sub-species. In *A. j. jubatus*, 100% of the SNPs amplified of which 91% were polymorphic and in *A. j. Soemmeringii*, 91% of the SNPs amplified and 82% were polymorphic (Supplementary Figure 6). In lion, 90% of the SNPs amplified of which 35% were polymorphic and in domestic cat, 88% of the SNPs amplified of which 30% were polymorphic. Lastly, in African wildcat, 92% of SNPs amplified with 34% being polymorphic and in leopard, 81% of the SNPs amplified with 30% of the genotyped SNPs being identified as polymorphic (Supplementary Figure 6). Genetic differentiation among the five different species was further resolved via PCA (Figure 4a). The cheetah (*A. jubatus*) samples were distinct and clustered separately whereas African wildcat and domestic cat grouped together and lion and leopard were in a single cluster. For the subspecies distinction, the *in silico* analysis using the full 240 SNPs extracted from the data of Prost et al. [30]

showed a clear distinction between the five classically recognized subspecies (*A. j. jubatus*, *A. j. raineyi*, *A. j. soemmeringii, A. j. hecki* and *A. j. venaticus;* Supplementary Figure 7). Overall, the clustering was highly similar to that obtained with 3,743 SNPs in Prost et al. 2020. The *in vivo* test applying the array using available samples of *A. j. jubatus* and *A. j. soemmeringii* also showed a strong discrimination between the two cheetah subspecies (Figure 4b).

Discussion

The aim of this study was to develop an accurate, reproducible and cost effective genetic test that can be implemented for parentage assignments in cheetah in order to monitor the South African legal trade. Due to the advantages SNPs have over other molecular markers [4], they are becoming the marker of choice for parentage. Here, the novel SNP array for cheetah (*A. jubatus*) was developed using a subset of 218 SNP loci identified from ddRADseq data. This method has been successfully used previously to identify a panel of SNPs in order to characterize African forest elephant [33]. In order to validate the assay, cheetah DNA was analysed using Taqman® genotyping on the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System. This system was selected as it has high throughput, reduced hands-on time and is cost effective ($23 per sample in order to genotype 240 SNPs). Taqman® assays have been successfully used to determine parentage in steelhead (*Oncorhynchus mykiss*) [34] and Pacific lamprey (*Entosphenus tridentatus*) [35].

In this study, 218 (91%) out of 240 SNPs genotyped were identified as being polymorphic in *A. jubatus*. Similarly, Roques et al [36] reported homozygosity in 10% of SNPs genotyped by MassArray technology following the selection of polymorphic SNPs from ddRADseq data. The authors attributed this discrepancy to low ddRADseq coverage. Increasing the coverage threshold (5x to 30x) is reported to decrease genotyping errors and genomic breadth of coverage [37]. A total of 20 SNPs deviated from Mendelian expectations which may be attributed to low number of samples, genotyping errors, duplicated genomic regions, deleterious allelic variants or unusual meiotic segregation distortions [38]. The percentage of SNPs (10%) that deviated from Mendelian patterns is considered low in comparison to 22% reported in turbot [39] or 15% identified in Atlantic cod [38]. Here, insufficient samples may be responsible for the observed deviations. All analysis was conducted using both the full set of markers (218) and the subset of markers (176) that did not deviate from HWE and similar results were obtained. The array was determined to be both reproducible and reliable with duplicate samples providing identical results. Further, the SNP array demonstrated the ability to amplify samples with low DNA concentrations (sample concentrations as low as 5 ng/µl), a requirement for forensic investigations.

Here, we compared microsatellite and SNP data and identified that genetic diversity was lower in samples analyzed via the SNP array (H_o = 0.57, H_e = 0.49) in comparison to microsatellite markers (H_o = 0.60, H_e = 0.74). This finding can be attributed to the biallelic nature of SNPs resulting in lower information content in comparison to microsatellites which have multiple alleles [40]. However, the average level of SNP heterozygosity observed here is considered satisfactory for future genetic monitoring of cheetah populations as they are similar to those previously

reported. Other studies based on microsatellite analysis reported heterozygosity values from 0.46 [6] to 0.52 [41]. Subsequent studies where more polymorphic markers with more alleles per marker were selected reported higher diversity values [41, 8, 7]. In this study, SNPs did not detect inbreeding that was apparent with microsatellites (Table 1). However, this observation may be due to sample selection of unrelated individuals to validate the SNP data which would result in the inbreeding estimates being more consistent with outbreeding compared to the microsatellite dataset. Currently, cheetah parentage testing relies on the use of microsatellite markers. However, it has been reported that microsatellite markers have a lower ability to assign parentage with high confidence in comparison to SNPs, when a sufficient number of SNPs are used [42]. Due to lower PIC values reported for SNPs, theoretically more SNPs are required than microsatellites to obtain the same power of exclusion (10 microsatellites versus 16 SNPs). In this study, the overall combination of 218 SNPs had a higher resolving power for individual identification and the SNP array had a greater ability to correctly assign known parent-offspring trios. The observed difference in the power of the markers for parentage reconstruction may be attributed to genotyping errors. Microsatellites have been reported to have a higher genotyping error (1 to 5%) and a higher frequency of null alleles in comparisons to SNPs [43-44] resulting in errors in parentage assignments. Lastly, we empirically demonstrate that accurate parentage (>95%) can be achieved with as few as 109 SNPs (Supplementary Figure 5). The number of SNPs required for accurate parentage assignment depends on the MAF of the markers. Anderson & Garza [45] determined that fewer loci with higher MAFs have a similar power to more loci with lower MAFs. Here, our panel of 218 SNPs had relatively high MAFs ranging from 0.125 to 0.5 (Supplementary Table 4).

The SNP assay in this study was validated for parentage determination in South African cheetah (*Acinonyx jubatus*). However, cross-species amplification observed in *A. j. Soemmeringii* provides support that this panel may be successful internationally on all sub-species of cheetah for parentage verification. In addition, the potential of the SNP assay to assign cheetah sub-species is supported by the *in silico* analysis which showed a clear distinction between the five classically recognized subspecies. However, this would have to be further investigated using a larger dataset which includes cheetah sub-species throughout their distibution range. Potential further applications for this assay include identification of the source of cheetah, population genetic analysis and individual identification.

In conclusion, we identified and developed a panel of SNP markers that can be used for accurate individual identification and parentage verification for monitoring of legal and illegal trade in cheetahs. The developed tool will greatly aid efforts to create passports for all captive cheetahs that are traded and thus aid efforts to fight illegal trafficking of the species. In addition, the array can be used in the future to address several management, conservation and research questions in cheetahs. This method could be used in routine monitoring to assess genetic diversity and differentiation in cheetah population. A larger dataset would be needed in order to determine the ability of the SNP marker to differentiate clearly between wild and captive South African populations. Here, we demonstrate the potential use of the SNP array to distinguish between cheetah sub-species. However, additional sampling would be required in order to validate this method in sub-species identification.

Tables

Table 1: Characterization of microsatellites and single nucleotide polymorphisms (SNPs) in cheetah. N = number of samples; Nm = number of markers; N_a = number of alleles; N_e = number of effective alleles; H_o = observed heterozygosity; H_e = expected heterozygosity; H_z = unbiased expected heterozygosity and Fis = inbreeding coefficient.

	N	**Nm**	**N_a**	**N_e**	**H_o**	**H_e**	**H_z**	**Fis**
Microsatellites	78	24	11.08	4.72	0.60	0.74	0.75	0.20
SNPs	76	218	2.00	1.95	0.57	0.49	0.49	-0.166
SNPs	76	176	2.00	1.94	0.48	0.49	0.49	-0.001

Figures

Figure 1: Relationship between probability of identity, probability of identity between siblings (PIDsib) and the number of STR and SNPs genotyped.

Figure 2: Trio Logarithm of the odds (LOD) scores per cheetah offspring obtained for microsatellite and single nucleotide polymorphisms (SNPs). Blue bars indicate LOD scores obtained for SNPs and orange bars indicated LOD scores obtained for microsatellites.

Figure 3: Discriminant analysis of principal components (DAPC) clusters for (A) 218 SNPs and (B) 176 SNPs with individual assignment probabilities for each cheetah family displayed in a bar graph.

Figure 4: Genetic relationships among (A) felid species and (B) cheetah sub-species, based on the first two principal components (PC1, PC2) derived from a multivariate analysis (PCA).

Supplementary material

Supplementary Table 1: List containing all samples used for this study, what they were used for and where samples were from. 1 = samples with more than 50 missing genotypes. 2 = Samples that did not amplify. * indicates samples repeated in duplicate. # indicates samples repeated at various concentrations.

Supplementary Table 2: Primer and probe sequences for SNPs used in cheetah.

Supplementary Table 3: Primer sequences of cross species microsatellites used in cheetah.

Supplementary Table 4: Details of single nucleotide polymorphisms used in the study. The SNP identified to have a MAF value below 0.25 is indicated in blue. The 42 SNPs with Ho values below 0.1 or above 0.9 are indicated in red.

Supplementary Figure 1: Linkage disequilibrium (LD) decay curve.

Supplementary Figure 2: Examples of genotype plots. The fluorescence of the two alleles are plotted along the x- and y- axis. Blue and red dots represent homozygous genotypes whereas green dots indicate heterozygous genotypes. Yellow dots indicate no amplification and black dots indicate genotypes that could not be assigned. (a) Genotype plot with monomorphic SNPs, (b) genotype plot with diffuse clusters, (c) genotype plot with one genotype (heterozygous) for all samples and (d) genotype plot of a polymorphic marker.

Supplementary Figure 3: Assessment of genotyping call success rate of the SNP array.

Supplementary Figure 4: Average number of missing SNPs as DNA concentration (ng/ul) is increased.

Supplementary 5: Relationship between the number of SNPs genotyped and the number of offspring that could be successfully assigned to the correct parents.

Supplementary Figure 6: Bar chart indicating genotyping success across species using the full 240 set of SNPS. Blue bars indicate percentage of polymorphic loci, maroon bars indicate percentage of monomorphic loci and green bars indicate percentage of SNP markers with no amplification.

Supplementary Figure 7: *In silico* PCA analysis using the full 240 SNPs extracted from the data of Prost et al. (2020). Purple: *A. j. jubatus*, orange: *A. j. raineyi*, gray: *A. j. soemmeringii*, green: *A. j. hecki* and blue: *A. j. venaticus*.

References

[1] Flanagan SP, Jones AG (2019) The future of parentage analysis: From microsatellites to SNPs and beyond. Molecular Ecology 28:544-567. DOI: 10.1111/mec.14988.

[2] Alacs EA, Georges A, FitzSimmons NN, Robertson J (2010) DNA detective: a review of molecular approaches to wildlife forensics. Forensic Science, Medicine, Pathology 6:180-194. DOI: 10.1007/s12024-009-9131-7.

[3] Glover KA, Hansen MM, Lien S, Als TD, Hoyheim B, Skaala O (2010) A comparison of SNP and STR loci for delineating population structure and performing individual genetic assignment. BMC Genetics 11(2). https://doi.org/10.1186/1471-2156-11-2.

[4] Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. Trends in Ecology and Evolution 19(4):208-216. DOI: 10.1016/j.tree.2004.01.009.

[5] Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: present problems, paradoxes and future challenges. Biodiversity Conservation 16:4147-4163. DOI: 10.1007/s10531-007-9212-4.

[6] Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: A test in wild cats. Genome Research 12(3):414-423. DOI: 10.1101/gr.185702.

[7] Kotze A, Ehlers K, Cilliers DC, Grobler JP (2008) The power of resolution of microsatellite markers and assignment tests to determine the geographic origin of cheetah (*Acinonyx jubatus*) in Southern Africa. Mammalian biology 73:457-462. https://doi.org/10.1016/j.mambio.2007.10.011.

[8] Marker LL, Wilkerson AJ, Sarno RJ, Martenson J, Breitenmoser-Wursten C, O'Brien SJ, Johnson WE (2008) Molecular genetic insights in cheetah (*Acinonyx jubatus*) ecology and conservation in Namibia. Heredity 99(1):2-13. https://doi.org/10.1093/jhered/esm081.

[9] Nellemann C, Henriksen R, Raxter P, Ash N, Mrema E (2014) The environmental crime crisis: threats to sustainable development from illegal exploitation and trade in wildlife and forest resources, Nairobi: United Nations Environment Programme (UNEP).

[10] Tricorache P, Nowell K, Wirth G, Mitchell N, Boast LK, Marker L (2018) Pets and pelts: Understanding and combating poaching and trafficking in cheetahs. In: Marker L, Boast LK, Schmidt-Kuentzel A, Nyhus PJ (ed) Cheetahs: Biology and conservation: Biodiversity of the world: Conservation from genes to landscapes, 1st edn.

Academic Press, pp. 191-205.

[11] Nowell K, Jackson P (1996) Wild cats : status survey and conservation action plan, Cambridge: IUCN, Gland, Switzerland.

[12] Buk KG, van der Merwe VC, Marnewick K, Funston PJ (2018) Conservation of severely fragmented populations: lessons from the transformation of uncoordinated reintroductions of cheetah (*Acinonyx jubatus*) into a managed metapopulation with self-sustained growth. Biodiversity and Conservation 27(13):3393-3423. https://doi.org/10.1007/s10531-018-1606-y.

[13] Nowell K (2014) An assessment of conservation impacts of legal and illegal trade in cheetah *Acinonyx jubatus*, Geneva,Switzerland: Report to the 65th meeting of the CITES Standing Committee.

[14] Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7(5):37135. https://doi.org/10.1371/journal.pone.0037135.

[15] Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows_Wheeler transform. Bioinformatics 25(14):1754-1760. https://doi.org/10.1093/bioinformatics/btp324.

[16] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078-2079. DOI: 10.1093/bioinformatics/btp352.

[17] Garcia-Alcalde F, Okonechnikov K, Carbonell J, Cruz L, Gotz S, Tarazona S, Dopazo J, Meyer T, Conesa A (2012) Qualimap: evaluating next-generation sequencing alignment data. Bioinformatic*s* 28(20):2678-2679. DOI: 10.1093/bioinformatics/bts503.

[18] Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Molecular Ecology 22(11). https://doi.org/10.1111/mec.12354.

[19] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, Maller J, Sklar P, de Bakker P, Daly M, Sham P (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics 81:559-575. DOI: 10.1086/519795.

[20] Narum S (2006) Beyond Bonferroni: Less conservative analyses for conservation genetics. Conserv Genet

7:783-787. https://doi.org/10.1007/s10592-005-9056-y.

[21] Huang Q, Shete S, Amos CI (2004) Ignoring linkage disequilibrium among tightly linked markers induces false-positive evidence of linkage for affected sib pair analysis. American Journal of Human Genetics 75(6):1106-1112. DOI: 10.1086/426000.

[22] Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of next generation sequencing data. BMC Bioinformatics 15:356. https://doi.org/10.1186/s12859-014-0356-4.

[23] Charruau P, Fernandes C, Orozco-terWengel P, Peters J, Hunter L, Ziaie H, Jourabchian A, Jowkar H, Schaller G, Ostrowski S, Vercammen P, Grange T, Schlotterer C, Kotze A, Geigl E-M, Walzer C, Burger PA (2011) Phylogeography, genetic structure and population divergence time of cheetahs in Africa and Asia: evidence for long-term geographic isolates. Molecular Ecology 20(4):706-724. DOI: 10.1111/j.1365-294X.2010.04986.x.

[24] Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106. DOI: 10.1111/j.1471-8286.2007.01931.x.

[25] Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16(5):1099-1106. DOI: 10.1111/j.1365-294X.2007.03089.x.

[26] Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28(19):2537-2539. DOI: 10.1093/bioinformatics/bts460.

[27] Nei M (1987) In: Molecular evolutionary genetics, Chichester, West Sussex, Columbia University Press, New York, pp. 177-179.

[28] Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genomics 11:94. https://doi.org/10.1186/1471-2156-11-94.

[29] Jombart T, Ahmed I (2011) Adegenet 1.3.1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071. https://doi.org/10.1093/bioinformatics/btr521.

[30] Prost S, Machado A, Zumbroich J, Preier L, Mahtani-Williams S, Meissner R, Guschanski K, Brealey J, Fernandes C, Vercammen P, Hunter L, Abramov A, Godsall-Bottriell L, Botriell P, Dalton D, Kotze A, Burger PA (2020) Conservation genomic analyses of African and Asiatic cheetahs (*Acinonyx jubatus*) across their

current and historic species range. BioRxiv. https://doi.org/10.1101/2020.02.14.949081

[31] Meisner J, Albrechtsen A (2018) "Inferring population structure and admixture proportions in low-depth NGS data Genetics 210(2):719-731. https://doi.org/10.1534/genetics.118.301336.

[32] Waits LP, Luikart G, Taberlet P (2001), "Estimating the probability of identity among genotypes in natural populations: cautions and guidelines," *Molecular Ecology,* vol. 10(1). DOI: 10.1046/j.1365-294X.2001.01185.x.

[33] Bourgeois S, Senn H, Kaden J, Taggart J, Ogden R, Jeffrey K, Bunnefeld N, Abernethy K, McEwing R (2018) Single-nucleotide polymorphism discovery and panel characterization in the African forest elephant. Ecol Evol 8:2207-2217. DOI: 10.1002/ece3.3854.

[34] Steele C, Anderson E, Ackerman M, Hess M, Campbell N, Narum S, Campbell M (2013) A validation of parentage-based tagging using hatchery steelhead in the Snake River basin. Canadian Journal of Fisheries and Aquatic Sciences 70:1046-1054. https://doi.org/10.1139/cjfas-2012-0451.

[35] Hess J, Campbell N, Docker M, Baker C, Jackson A, Lampman R, McIlraith B, Moser M, Statler D, Young W, Wildbill A (2015) Use of genotyping by sequencing data to develop a high-throughput and multifunctional SNP panel for conservation applications in Pacific lamphrey. Mol Ecol Res 15:187-202. https://doi.org/10.1111/1755-0998.12283.

[36] Roques S, Berrebi P, Rochard E, Accolas M (2018) Genetic monitoring for the successful restocking of the species with low diversity: The case of the critically endangered European sturgeon, Acipenser sturio. Biol Conserv 221:91-102. https://doi.org/10.1002/ece3.5268.

[37] Fountain ED, Pauli JN, Reid BN, Palsbøll PJ, Peery MZ (2016) Finding the right coverage: the impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. Mol Ecol Res, 16: 966-978. https://doi.org/10.1111/1755-0998.12519.

[38] Hubert S, Tarrant Bussey J, Higgins B, Curtis BA, Bowman S (2009) Development of single nucleotide polymorphism markers for Atlantic cod (*Gadus morhua*) using expressed sequences. Aquaculture 296: 7-14. DOI: 10.1016/j.aquaculture.2009.07.027.

[39] Vera M, Álvarez-Dios JA, Millán A, Pardo BG, Bouza C, Hermida M, Fernández C, de la Herrán R, Molina-

Luzón MJ, Martínez P (2011) Validation of single nucleotide polymorphism (SNP) markers from an immune Expressed Sequence Tag (EST) turbot, *Scophthalmus maximus,* database. Aquaculture 313: 31-41. DOI : 10.1016/j.aquaculture.2011.01.038.

[40] Aitken N, Smith S, Schwarz C, Morin PA (2004) Single nucleotide polymorphism (SNP) discovery in mammals: a targeted‐gene approach. Mol Ecol 13: 1423-1431. DOI: 10.1111/j.1365-294X.2004.02159.x.

[41] Luo S-J, Kim J-H, Johnson WE, van der Walt J, Martenson J, Yukhi N, Miquelle DG, Uphyrkina O, Goodrich JM, Quigley HB, Tilson R, Brady G, Martelli P, Subramaniam V, McDougal C, Hean S, Huang S-Q, Pan W, Karanth UK, Sunquist M, Smith JLD, O'Brien SJ (2004) Phylogeography and genetic ancestry of tigers (*Panthera tigris*). PLOS Biol 2: 2275-2293. https://doi.org/10.1371/journal.pbio.0020442.

[42] Heaton MP, Leymaster KA, Kalbfleisch TS, Kijas JW, Clarke SM, Mcewan JC, Maddox JF, Basnayake V, Petrik DT, Smith T, Chitko-Mckown CG (2014) SNPs for parentage testing and traceability in globally diverse breeds of sheep. PloS one 9(4). DOI: 10.1371/journal.pone.0094851.

[43] McGoldrick DJ, Hedgecock D, English LJ, Baoprasertkul P, Ward RD (2000) The transmission of microsatellite alleles in Australian and North American stocks of the Pacific oyster (*Crassostrea gigas*): selection and null alleles. J Shellfish Res 19:779–788. http://hdl.handle.net/102.100.100/208972?index=1.

[44] Lorenz H, Melissa B, Ray H, Lisa W, Jamese S (2011) An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (Oncorhynchus nerka) population. Mol Ecol Res 11:150–161. DOI: 10.1111/j.1755-0998.2010.02961.x.

[45] Anderson EC, Garza JC (2006) The power of single nucleotide polymorphisms for large-scale parentage inference. Genetics 172(4):2567-2582. https://doi.org/10.1534/genetics.105.048074.

[46] CITES (2017) Appendices. [Online] Available at: https://cites.org/eng/app/appendices.php [Accessed 21 October 2020].

[47] Scientific Authority of South Africa (2019) Non-detriment for Acinonyx jubatus (cheetah). [Online] [Accessed 8 April 2019].

[48] Menotti-Raymond M, David VA, Lyons LA, Schaffer AA, Tomlin JF, Hutton MK, O'Brien SJ (1999) A genetic linkage map of microsatellites in the domestic cat (*Felis catus*). Genomics 57:9-23. DOI:

10.1006/geno.1999.5743.

[49] Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ (2003) Molecular genetics and evolution of melanism in the cat family. Current Biology 13:448-453. DOI: 10.1016/s0960-9822(03)00128-3.

Ethics declarations

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

Ethical approval was obtained from the Animal Research Ethics Committee, University of the Free State, South Africa (Student Project Number: UFS-AED2018/0040) and the South African National Biodiversity Institute's Research Ethics and Scientific Committee (Reference: SANBI/RES/P18/20). The ethics committees use the following guiding documents to approve projects: SANS10386:2008 for the care and use of animals for scientific purposes, SANS 10379:2005 Zoo and aquarium practice, WAZA Code of Ethics and Animal Welfare, 2003, Veterinary and Para-Veterinary Professions Amendment Act, No. 19 of 1982, Professional Code of Ethics of the African Association of Zoos and Aquaria 2007 and the National Research Foundation Act 1998, Additional permitting for research use of the samples was approved under a Section 20 permit from the Department of Agriculture, Forestry and Fisheries, South Africa (Reference: 12/11/1/1/18) that permits research on animals under Section 20 of the Animal Diseases Act, 1984 (Act No 35 of 1984):.

Informed consent

All authors consent to publication.

Funding

D.L.D. and P.B. acknowledge the Scientific & Technological Cooperation Austria, from the Austrian Agency for International Cooperation in Education and Research and the National Research Foundation (NRF) of South Africa. This work is based on the research supported in part by the National Research Foundation of South Africa (Grant Number 116055).