Kotze, A., Ehlers, K., Cilliers, D. C., & Grobler, J. P. (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:** 457-562.

Keywords: Acinonyx jubatus/Assignment test/Bayesian/cheetah/Illegal trade/microsatellite/southern Africa

Formerly found in 44 countries in Africa and Asia, cheetahs are currently confined to fragmented populations in 29 African countries, and remnant populations in Iran and Pakistan (Marker 2002). In southern Africa, cheetahs are at present found in Botswana, Namibia, South Africa and Zimbabwe. Trade in cheetah products and live export of cheetah from Namibia and Botswana is stringently controlled (CITES 1992). As a result, conservation authorities are constantly aware of potential illegal trade in cheetah over the Namibian and Botswana borders with South Africa. Where foul-play is involved, identification of source populations of confiscated animals will require implementation of identification techniques based on multilocus genotypes. Manel et al. (2002) demonstrated that genetic methods have high power of resolution to determine the geographic origin of population samples for sufficiently differentiated populations. Forensic science services for domesticated animals are well established in South Africa and have in recent years expanded to include game species, marine fish stock identification and ornamental fish (Grobler et al. 2005). In this paper, we describe the power of resolution of microsatellite markers and assignment tests to determine the geographic origin of cheetah (Acinonyx jubatus) confiscated in South Africa on suspicion of illegal import. Cheetah was formerly thought to be genetically highly monomorphic (presumably following a historic bottleneck), based on allozyme data (O'Brien et al. 1983). Subsequent studies (Menotti-Raymond and O'Brien 1993, 1995) have revealed genetic heterogeneity for microsatellite markers. This has been attributed to accumulated variation since the hypothetical bottleneck, resultant from the high mutation rates of microsatellite markers (Hedrick 1996). The presence of a moderate level of genetic diversity, comparable to other felids for some markers (Menotti-Raymond and O'Brien 1993), suggests that marker-based forensic identification in cheetah is feasible.

Mammalian Biology
Zeitschrift für Säugetierkunde
www.elsevier.de/mambio

SHORT COMMUNICATION

The power of resolution of microsatellite markers and assignment tests to determine the geographic origin of cheetah (*Acinonyx jubatus*) in Southern Africa

Antoinette Kotze[a,b], Karen Ehlers[a], D.C. Cilliers[c], J.P. Grobler[a,*]

[a]*Department of Genetics, University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa*
[b]*National Zoological Gardens of South Africa, P.O. Box 754, Pretoria 0001, South Africa*
[c]*De Wildt Cheetah and Wildlife Trust, P.O. Box 52071, Dorandia 0182, South Africa*

Received 4 September 2007; accepted 23 October 2007

Keywords: Cheetah; Illegal trade; Microsatellite; Assignment test; Bayesian

Formerly found in 44 countries in Africa and Asia, cheetahs are currently confined to fragmented populations in 29 African countries, and remnant populations in Iran and Pakistan (Marker 2002). In southern Africa, cheetahs are at present found in Botswana, Namibia, South Africa and Zimbabwe. Trade in cheetah products and live export of cheetah from Namibia and Botswana is stringently controlled (CITES 1992). As a result, conservation authorities are constantly aware of potential illegal trade in cheetah over the Namibian and Botswana borders with South Africa. Where foul-play is involved, identification of source populations of confiscated animals will require implementation of identification techniques based on multilocus genotypes. Manel et al. (2002) demonstrated that genetic methods have high power of resolution to determine the geographic origin of population samples for sufficiently differentiated populations. Forensic science services for domesticated animals are well established in South Africa and have in recent years expanded to include game species, marine fish stock identification and ornamental fish (Grobler et al. 2005). In this paper, we describe the power of resolution of microsatellite markers and assignment tests to determine the geographic origin of cheetah (*Acinonyx jubatus*) confiscated in South Africa on suspicion of illegal import.

Cheetah was formerly thought to be genetically highly monomorphic (presumably following a historic bottleneck), based on allozyme data (O'Brien et al. 1983). Subsequent studies (Menotti-Raymond and O'Brien 1993, 1995) have revealed genetic heterogeneity for microsatellite markers. This has been attributed to accumulated variation since the hypothetical bottleneck, resultant from the high mutation rates of microsatellite markers (Hedrick 1996). The presence of a moderate level of genetic diversity, comparable to other felids for some markers (Menotti-Raymond and O'Brien 1993), suggests that marker-based forensic identification in cheetah is feasible.

Six cheetahs held on a farm without the required permits were confiscated by conservation authorities in the North West Province of South Africa. The farm is situated in an area known for a high incidence of illegal trade in cheetah, especially across the border with Botswana, and foul-play was therefore suspected. To assign the six unknown cheetah to their most likely region of origin, the animals were genotyped using microsatellite markers, and then compared to cheetah of known origin. The reference populations consisted of wild cheetah from Botswana, Namibia and the north-western parts of South Africa. Geographical locations of

*Corresponding author.
E-mail address: groblerjp.sci@ufs.ac.za (J.P. Grobler).

1616-5047/$ - see front matter
doi:10.1016/j.mambio.2007.10.011

these sites are indicated in Fig. 1. All wild cheetah sampled were trapped as part of programs which aim to conserve free-roaming cheetah which occur on commercial and communal farmlands, and which relocate animals in cases of cheetah–farmer conflict. The available reference animals were grouped into nominal populations that best represent geographic clusters. Three broad regional populations were identified among the available reference samples: Namibia – cheetah from the Otjiwarongo area in Namibia ($n = 19$); Kalahari – a combination of 17 cheetah in the Jwaneng area in western Botswana, three cheetah from the Sekoma area 50 km further west in Botswana and one individual from the Northern Cape region of South Africa (bordering Botswana); and South Africa – a combination of three individuals from the North-West region of South Africa and eight cheetah from further east within South Africa (Lephalale and Thabazimbe regions in the Limpopo Province). Pooling of animals into nominal population "Kalahari" was considered justified since the animals were collected on both sides of an international border that did not prevent gene flow historically. Care was taken to exclude captive cheetah in South Africa since that group may potentially contain genotypes originating from Namibia. Marker-Kraus (1997) estimated that more than 90% of all cheetahs in captivity worldwide are descendants of Namibian cheetah, with 71% of captive South African cheetah originating from Namibia (Marker 2002).

All 57 animals used during the current study were genotyped at 13 nuclear DNA microsatellite loci originally developed for application in the domestic cat (*Felis catus*), following Menotti-Raymond et al. (1997, 1999). The loci were: FCA08, FCA13, FCA23, FCA43, FCA45, FCA75, FCA77, FCA96, FCA126, FCA191, FCA205, FCA224 and FCA298. DNA was extracted using a QIAGEN© kit and following the standard protocol supplied with the kit. Reaction mixtures and reaction conditions followed Menotti-Raymond et al. (1997, 1999). Amplified fragments were resolved on an ABI377 automated sequencer and analysed using GENESCAN and GENOTYPER software. The microsatellite loci analyzed were highly polymorphic with monomorphic results obtained at only two loci (FCA43 and FCA77). The number of alleles and allele size ranges at each locus are listed in Table 1. All polymorphic loci showed a dinucleotide repeat motif.

Likelihood-based and certain model-based methods of assignment assume lack of linkage – and Hardy –Weinberg (H–W) disequilibrium (Pritchard et al. 2000). Statistical analysis was therefore preceded by testing for linkage and H–W disequilibrium, using POPGENE software (Yeh et al. 1999). Results showed

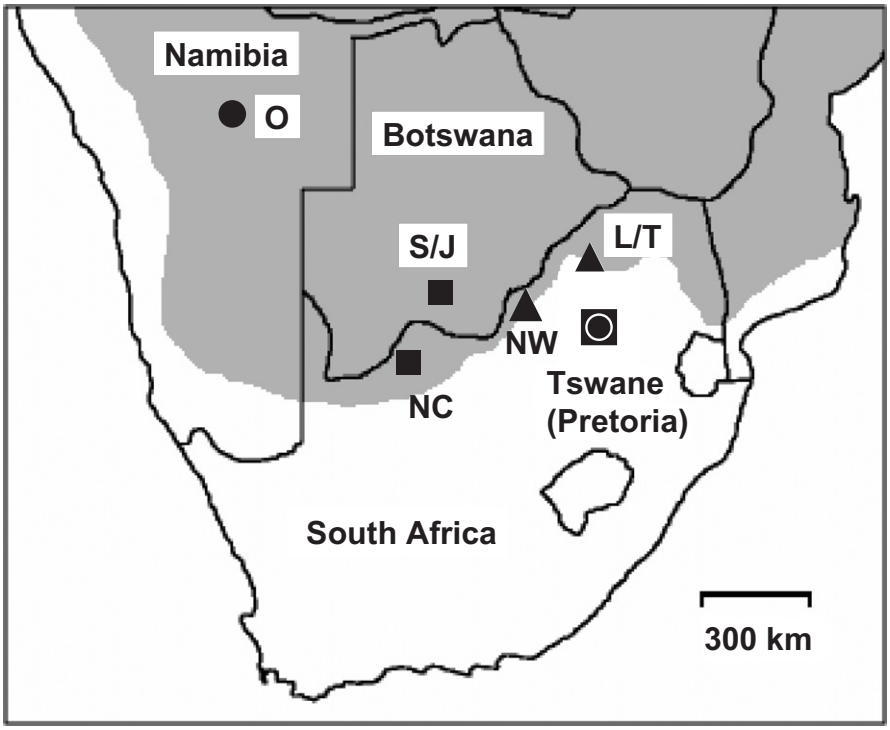

Fig. 1. The distribution of cheetah in southern Africa (shaded area), and localities sampled. Symbols and abbreviations used: (▲) north-western South Africa, (■) Kalahari (Botswana/Northern Cape, South Africa), (●) Namibia; O = Otjiwarongo; J = Jwaneng, S = Sekoma, NC = Northern Cape, NW = North-West, L = Lephalale and T = Thabazimbi.

no significant linkage between alleles at any pair of loci (at the 5% threshold of significance) after Bonferroni correction for multiple pairwise comparisons. The loci FCA75 and FCA126 showed significant ($p = 0.001$) deviations from expected H–W equilibrium within one or more population. The total of 9–11 polymorphic loci finally available for assignment testing is in line with the recommendation by Koskinen (2003) for accurate assignment (10 loci), and in the range of 8–22 used by Manel et al. (2002) for high statistical significance.

It would be ineffectual to attempt assignment of unknown individuals in the absence of measurable geographic genetic differentiation between cheetah from the nominal regional populations. An analysis of molecular variance (AMOVA; using ARLEQUIN software – Excoffier et al. (2005)) indicated that 9.75% of total variation is found among regional populations. F_{ST} values supported the hypothesis of significant differentiation among all populations pairs ($p < 0.001$), with slightly higher F_{ST} values between Namibia and South Africa (0.115) and Namibia and Kalahari (0.108) compared to South Africa and Botswana (0.059). Although these values are somewhat lower than the F_{ST} values of 0.15–0.20 recommended by Manel et al. (2002) and Koskinen (2003) as necessary for high statistical certainty of assignment, the significance of differentiation among populations suggest that accurate assignment may be possible.

Table 1. Number of alleles observed and allele size range at 13 microsatellite loci, in cheetah populations and unknown individuals from the southern African sub-region

Locus code	Number of alleles	Allele size range (bp)
FCA08	10	152–170
FCA13	4	153–159
FCA23	3	134–138
FCA43	1	110
FCA45	5	134–142
FCA75	10	126–146
FCA77	1	122
FCA96	9	197–215
FCA126	6	120–134
FCA191	5	128–142
FCA205	7	101–113
FCA224	8	156–176
FCA298	3	220–228

To track the six unknown individuals to their most likely source populations, the first approach was to assign individuals of unknown origin to the set of predefined populations. An exclusion test (Ranalla and Mountain 1997) and partially Bayesian method Cornuet et al. (1999) was used, as implemented in GENECLASS 2 software (Piry et al. 2004). This method of Cornuet et al. (1999) was used as representative of an approach not based on the assumption that the true population of origin has been sampled. Analysis was based on the nine loci which conformed to expected H–W equilibrium of genotypes in all populations. Results are presented in Table 2. Four individuals were assigned to Namibia with probabilities of 0.753–0.995. Unknown cheetah 04 showed comparable probabilities of resorting to Namibia (0.978), South Africa (0.975) or Kalahari (0.972). The remaining unknown animal nominally resorted to Namibia but with very low probability (0.081).

A second approach was to use a frequency-based method (Paetkau et al. 1995) without simulation, as implemented in GENECLASS (also based on nine loci). Assignment scores are presented in Table 2. Five unknown cheetah were assigned to Namibia (scores = 72.211–99.398%), with the sixth individual (unknown 04) assigned to South Africa (score = 93.424%).

As an alternative to the partially Bayesian method of Cornuet et al. (1999), a fully Bayesian method (Pritchard et al. 2000) was used to identify populations (clusters) and assign individuals probabilistically to each cluster (using STRUCTURE software – Pritchard et al. (2000) and Falush et al. (2003)). This method assumes that the true population of origin has been sampled. A model with assumption of admixture ancestry and correlated allele frequencies was used, and all 11

Table 2. Assignment of unknown cheetah to nominal source populations

	South Africa	Kalahari	Namibia
Unknown 01	0.183 [0.332%]	0.765 [16.592%]	0.995 [83.077%]
Unknown 02	0.300 [2.492%]	0.464 [0.356%]	0.948 [97.153%]
Unknown 03	0.080 [6.146%]	0.225 [2.710%]	0.753 [91.144%]
Unknown 04	0.975 [93.424%]	0.972 [6.523%]	0.978 [0.053%]
Unknown 05	0.044 [23.510%]	0.009 [4.279%]	0.081 [72.211%]
Unknown 06	0.116 [0.340%]	0.324 [0.262%]	0.927 [99.398%]

(i) Results from a partial Bayesian method (Ranalla and Mountain 1997) with probability computation (Cornuet et al. 1999). Values indicate the likelihood that each of the unknown individuals originated from each of three potential source populations. (ii) Scores from a frequency-based method (Paetkau et al. 1995) (in [ ]).

polymorphic loci were included in this analysis. The parameter $\ln \Pr(X|K)$ was calculated for K values (number of populations) of 1–5, with three independent runs for each K, to estimate the true number of populations. All runs consisted of a burn-in period of 100,000 steps, followed by 200,000 iterations. Calculation of the posterior probabilities of K showed the highest probability for a real structure consisting of three cheetah populations ($p = 0.368$).

MCMC simulation was then run for 1,000,000 iterations (with a burn-in period of 100,000 steps). The estimated membership coefficients for all individuals were clustered into a bar plot (Fig. 2). The proportion of membership of each pre-defined population in each of three clusters; and inferred ancestry of pooled and individual unknown cheetah are presented in Table 3. A total of 54.5% of cheetah from north-western South Africa was assigned to cluster 1 and 54.3% of Kalahari cheetahs were assigned to cluster 2, the latter including the single individual originating from the Northern Cape region of South Africa. Cluster 3 contained predominantly cheetah from Namibia, with 69.2% of Namibian cheetah assigned to that cluster. While these values represent most prominent nominal population in each cluster, it should be noted that 45.5% of individuals in nominal population South Africa, 45.7% of Kalahari cheetah and 30.8% of Namibian animals were also assigned to second and third clusters. Among unknown cheetah, three animals have the highest probabilities of belonging to cluster 1 (a cluster

Table 3. (a) Proportion of membership of each pre-defined population in each of three clusters; and (b) inferred ancestry of unknown individuals; from a fully Bayesian clustering approach following Pritchard et al. (2000)

	Inferred clusters		
	Cluster 1	Cluster 2	Cluster 3
(a) Nominal populations			
SA	0.545	0.336	0.119
KAL	0.331	0.543	0.126
N	0.202	0.106	0.692
Unknown	0.469	0.208	0.324
(b) Unknown individuals			
Unknown 01	0.487	0.070	0.443
Unknown 02	0.681	0.083	0.236
Unknown 03	0.242	0.056	0.702
Unknown 04	0.016	0.955	0.029
Unknown 05	0.963	0.013	0.023
Unknown 06	0.423	0.069	0.508

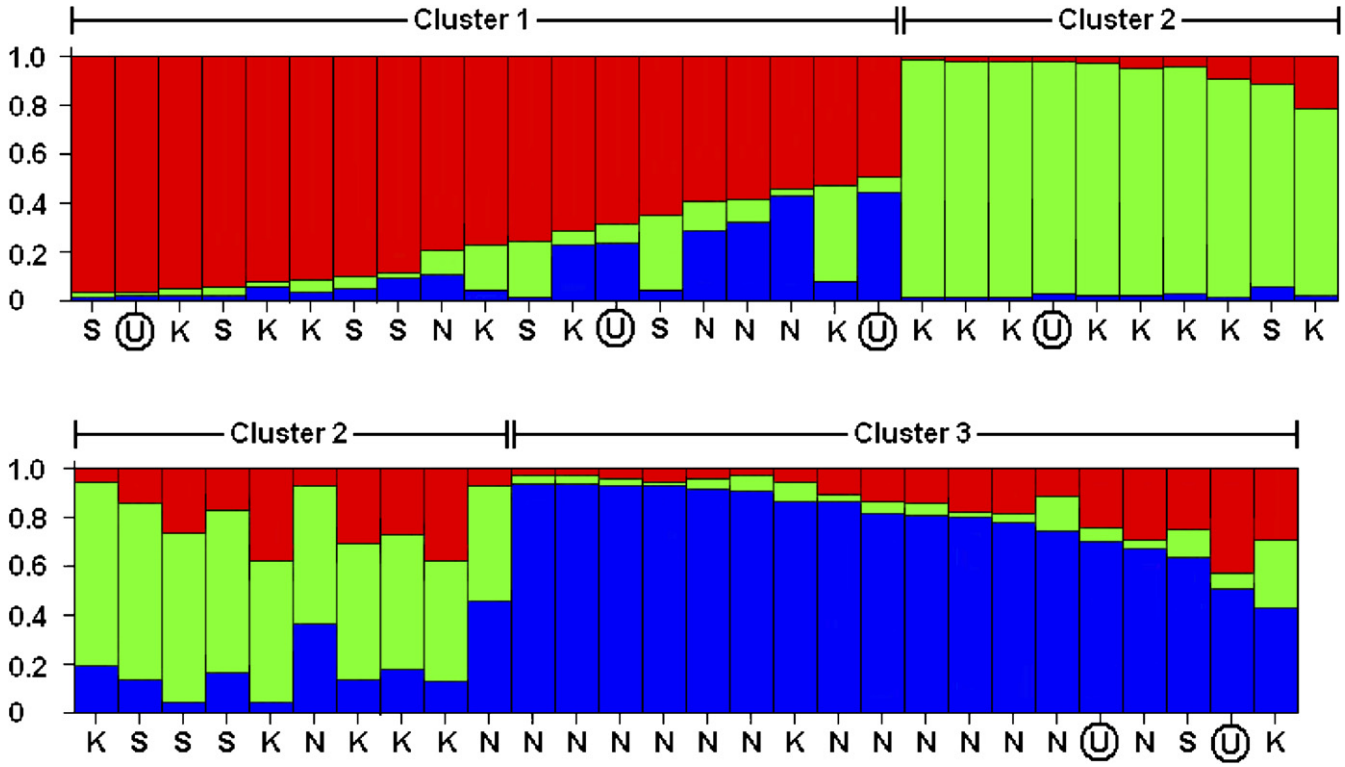

Fig. 2. Bar plot (from STRUCTURE) with three clusters of cheetah. Each individual is represented by a single vertical line broken into 1–3 coloured segments, with lengths proportional to each of the three inferred clusters. Abbreviations used: K = Kalahari, S = South Africa, N = Namibia and U = unknown.

with a high proportion of South African cheetah), although one animal is assigned to South Africa and Namibia with almost similar probabilities (0.487 and 0.443). Two of the remaining unknown cheetahs were assigned to cluster 3, dominated by cheetah from Namibia, and one animal to cluster 2 (with 54.3% of Kalahari animals).

The overall result of this attempt to assign cheetah to a specific region of origin is inconclusive. Nevertheless, the results provide significant support for a Namibian or South African origin for five or all of the unknown animals, whereas only one individual may possibly originate from the Kalahari/Botswana region. A simple frequency-based test assigned five out of six unknown individuals to Namibia and the sixth to South Africa. Cornuet et al. (1999) showed that frequency-based methods are less accurate compared to Bayesian methods but nevertheless results from a partially Bayesian exclusion test (Cornuet et al., 1999) supported the frequency-based results, with all cheetah showing the highest probability of being of Namibian origin. Furthermore, the fully Bayesian method used assigned two unknown cheetah to the Namibian region. The latter method also assigned three cheetahs to South African as the most likely source population. Since fully Bayesian methods have high power of resolution, as demonstrated by Manel et al. (2002) and Koskinen (2003), the assignment of five unknown cheetah to clusters with a high proportion of true Namibian and South African cheetah constitute strong support for a hypothesis that illegal translocation from Botswana was not involved in the current investigation.

The consistent assignment a portion of unknown cheetah to Namibia could suggest illegal importation from that country, but could equally indicate that foul-play did not occur. Since 71% of legally kept captive South African cheetahs originate from Namibia (Marker 2002), the presence of nominally Namibian genotypes in a captive South African population could be typical. The unknown group could thus represent a characteristic profile of South African captive cheetah, with animals of South African origin and animals descendent from legally imported Namibian animals.

The efficacy of assignment testing in cheetah will most likely be improved through more comprehensive sampling. Sample sizes attainable during the current study (maximum 20 animals per population) were slightly below levels of 30–50 individuals recommended by Manel et al. (2002) and Koskinen (2003). Furthermore, a sampling strategy that includes more subpopulations of cheetah in Botswana, Namibia and South Africa will be beneficial. We therefore recommend that current efforts to sample cheetah from across southern Africa continue, to develop a database that reveal fine-grained structure of cheetah populations in the sub-region and that can be used for future forensic application. The practical difficulties of sampling cheetah, and overall low numbers, may however, dictate that such a database will not be completed in the short term.

References

CITES, 1992. Quotas for trade in specimens of cheetah. Eighth meeting of the convention of International Trade in Endangered Species of Wild Fauna and Flora. Document 8.22 (rev.), 1–5.

Cornuet, J.-M., Piry, S., Luikart, G., Estoup, A., Solignac, M., 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153, 1989–2000.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin, Version 3.0: An Integrated Software Package for Population Genetics Data Analysis. Computational and Molecular Population Genetics Laboratory (CMPG), Institute of Zoology, University of Berne, Switzerland.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics 164, 1567–1587.

Grobler, J.P., Kotze, A., Swart, H., Hallerman, E.M., 2005. The application of microsatellite DNA markers for forensic analysis of koi carp (*Cyprinus carpio*). S. A. J. Sci. 101, 19–21.

Hedrick, P.W., 1996. Bottleneck(s) or metapopulation in cheetahs. Conserv. Biol. 10 (3), 897–899.

Koskinen, M.T., 2003. Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. Anim. Genet. 34 (4), 297–301.

Manel, S., Berthier, P., Luikart, G., 2002. Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conserv. Biol. 16 (3), 650–659.

Marker, L.L., 2002. Aspects of cheetah (*Acinonyx jubatus*) biology, ecology and conservation strategies on Namibian farmlands. Ph.D. Thesis, University of Oxford, 459pp.

Marker-Kraus, L., 1997. History of the cheetah (*Acinonyx jubatus*) in zoos 1829–1994. Int. Zoo Yearbook 35, 27–43.

Menotti-Raymond, M., O'Brien, S.J., 1993. Dating the genetic bottleneck of the African cheetah. Proc. Natl. Acad. Sci. USA 90, 3172–3176.

Menotti-Raymond, M., O'Brien, S.J., 1995. Evolutionary conservation of ten microsatellite loci in four species of Felidae. J. Hered. 86, 319–322.

Menotti-Raymond, M., David, V.A., Stephens, V.C., O'Brien, S.J., 1997. Genetic individualization of domestic cats using feline STR loci for forensic analysis. J. Foren. Sci. 42, 1039–1051.

Menotti-Raymond, M., David, V.A., Lyons, L.A., Schaffer, A.A., Tomlin, J.F., Hutton, M.K., O'Brien, S.J., 1999. A

genetic linkage map of microsatellites in the domestic cat (*Felis catus*). Genomics 57, 9–23.

O'Brien, S.J., Wildt, D.E., Goldman, D., Merril, C.R., Bush, M., 1983. The cheetah is depauperate in genetic variation. Science 221, 459–462.

Paetkau, D., Calvert, W., Stirling, I., Strobeck, C., 1995. Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. 4, 347–354.

Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, L., Estoup, A., 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. J. Hered. 95, 536–539.

Pritchard, J.K., Stephens, M., Rosenberg, N.A., Donnelly, P., 2000. Association mapping in structured populations. Am. J. Hum. Genet. 67, 170–181.

Ranalla, B., Mountain, J.L., 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA 94, 9197–9201.

Yeh, F.C., Yang, R., Boyle, T., 1999. POPGENE ver. 3.31. Microsoft Window-based freeware for population genetic analysis.