Molecular Ecology Notes (2004) **4**, 348–351 doi: 10.1111/j.1471-8286.2004.00693.x

PRIMER NOTE

Characterization of 15 single nucleotide polymorphism markers for chimpanzees (*Pan troglodytes*)

STEVE SMITH,* NICOLA AITKEN,† CARSTEN SCHWARZ‡ and PHILLIP A. MORIN§
Laboratory for Conservation Genetics, Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, 04103 Leipzig, Germany

Abstract

We report the characterization of 15 new single nucleotide polymorphism markers for a threatened species, the chimpanzee (*Pan troglodytes*), developed using a targeted gene approach. These markers are derived from the Y chromosome and autosomal regions of the genome and show frequency differences between chimpanzee subspecies from central and western Africa. These single nucleotide polymorphism markers are the first to be designed for the genotyping of wild chimpanzee populations and will provide a useful addition to the genetic tools employed for the conservation management of this threatened species.

Keywords: chimpanzee, conservation, genetic, *Pan troglodytes*, single nucleotide polymorphism

Received 26 January 2004; revision accepted 7 April 2004

Wild populations of chimpanzees (*Pan troglodytes*) have become increasingly threatened as a result of habitat loss, introduced disease, and hunting. This has resulted in an urgent imperative for research programmes to help conserve the species. Genetic research to date has targeted variation at hypervariable regions of mitochondrial DNA (Morin *et al*. 1994a; Goldberg 1997; Deinard & Kidd 2000), simple sequence repeats in nuclear DNA (Ely *et al*. 1998; Morin *et al*. 1994a,b, 2001; Goossens *et al*. 2000; Vigilant *et al*. 2001; Gusmao *et al*. 2002; Yu *et al*. 2003), and direct sequencing of nuclear loci (de Groot *et al*. 2000; Stone *et al*. 2002; Surridge *et al*. 2002). Each of these approaches, however, suffers from a range of technical and analytical issues that limit their application (see review in Morin *et al*. 2004). Here, we report the isolation and characterization of a set of 15 single nucleotide polymorphism (SNP) markers and their characterization in two subspecies of *P. troglodytes* from central and western Africa.

Correspondence: Steve Smith. *Present address: Molecular Ecology Laboratory, Australian School of Environmental Studies, Griffith University, Kessels Road, Nathan, Qld 4111, Australia. Fax: +61 73875 7459; E-mail: steve.smith@griffith.edu.au
†Present address: Applied Ecology Research Group, University of Canberra, ACT 2601, Australia; ‡Department of Anthropology, McMaster University, Hamilton, Ontario, Canada; §Protected Resources Division, Southwest Fisheries Science Center (SWFSC), National Marine Fisheries Service, 8604 La Jolla Shores Drive, La Jolla, CA 92037, USA.

Potential SNP markers were identified from polymorphism screening conducted in our laboratory on sequences from a set of one central and five western African chimpanzee samples, plus a pool of 19 western and one central African samples (Aitken *et al*. 2004). These loci were selected from previously published primer sets designed from alignments of two species DNA sequences (e.g. human and mouse) to amplify from exons across an intervening intron (CATS loci; Lyons *et al*. 1997). Further SNP assays were developed from previously published polymorphisms (Deinard 1997; Stone *et al*. 2002) in the Y chromosome and autosomal regions of the genome. Assays were designed to target the polymorphic base using a single base extension (SBE) technique via the SNaPshot™ system (Applied Biosystems). The polymerase chain reaction (PCR) primers were designed using the program PRIMER EXPRESS™ (Applied Biosystems) and amplified a target region of 51–138 bp containing the SNP (Table 1).

The PCR was carried out in a 20-μL volume containing 1 × PCR Buffer (10 mM Tris/HCl, 50 mM KCl, pH 8.3), 2 mM $MgCl_2$, 0.8 mg/mL bovine serum albumin, 250 μM each of dA,G,C,T,UTP, 0.025 U/μL *Taq* polymerase, 0.01 U/μL uracil-*N*-glycosylase (Roche), and 200 nM of each primer. Cycling was performed on an MJ Tetrad Thermal Cycler (MJ Research) and conditions were 50 °C for 2 min, 95 °C for 10 min, 35 cycles at 94 °C for 30 s, annealing temperature for 30 s (Table 1) and 72 °C for 30 s and 72 °C for 10 min. The PCR product was purified by incubation at 37 °C for 1 h in the presence of 3 units of Exonuclease1

Table 1 Characterization of 15 single nucleotide polymorphism (SNP) loci in the chimpanzee *Pan troglodytes*

Locus	GenBank Accession no.	SNP identity (amplicon length)	Oligo sequence (5′–3′)	T_a (°C)	Heterozygosity (sample size)†		Rare allele frequency (sample size)†	Source of sequence
					H_E	H_O		
ApoB140	AF245195	A/C (138 bp)	F: CCAAATATTCATTCAGAGATTCCCC R: GGATGTCATGTGTGCATCTAAACA SBE: TCATTCAGAGATTCCCC	57	0.50 (49)	0.20 (49)	0.47 (49)	D
MYH6	AY528405	C/T (80 bp)	F: CCACAAACAAGTGAGAGAGGTAAA R: CTTCATTGCACTTGTATCTGTAAGG SBE: ATAAGCAAATCAGAACATC	57	0.02 (49)	0.02 (49)	0.01 (49)	T
FOS	AY528406	C/T (81 bp)	F: GCCATAGTAAGAATTGGTTCCCC R: AGAGATCCGTCCTGGAGTGC SBE: GTTCATTCTGAGCAACCTCTG	55	0.30 (49)	0.20 (49)	0.18 (49)	T
IGH	AY528407	A/G (79 bp)	F: AGACTTGAGCACAACAAGATAGGA R: CCATTAACAAGCCCCAAATTC SBE: AAGATAGGATCAAACCCA	57	0.50 (49)	0.35 (49)	0.48 (49)	T
ApoB476	AF245196	C/G (100 bp)	F: TGTCATTTACAAGCTCTGCCAG R: GCCCATGACAAACTTTTAGTAACAG SBE: GCTCTGCCAGGGCCAAAT	60	0.44 (49)	0.29 (49)	0.33 (49)	D
CAT	AY528408	C/T (134 bp)	F: CCTCTGTATTTTGTATATGACTTCCCC R: TCCATTCGATCTCACCAAGGT SBE: GACTTCCCCATTATTATAATTG	55	0.19 (49)	0.16 (49)	0.10 (49)	T
TF1*	AY528409	A/G (123 bp)	F: ACCTGCTTCTGTGCCTGTCT R: AGCTGTCCCATCTGTCCACT SBE: TCTCACCGATGTCTCAACCC	55	0.46 (48)	0.19 (48)	0.34 (48)	T
PSUY397	AF244809	C/T (106 bp)	F: AGCTGAGATTGCGCCACT R: GTGCGCATTTATGTAACGACA SBE: CCAGCCTGGGCGACTGAG	55	0.25 (21)	N/A	0.14 (21)	D
sY85	AF440146	C/G (76 bp)	F: GGCATCTGTATTAACAACTCTGGG R: GCTGAACTTTTGGCAATTTGC SBE: GACACAGTAATCTATTTTTC	57	0.44 (16)	N/A	0.31 (16)	S
sY19b	AF440120	C/T (52 bp)	F: CTGCACTTATTTACCTGGTGGCT R: TGTTCCCAAGGTTGCTAGTACAG SBE: TTATTTACCTGGTGGCTCTT	55	0.46 (21)	N/A	0.33 (21)	S
sY19a	AF440119	C/G (70 bp)	F: CTCCTTCCATCCCCAGGG R: AGCAGAGTGGGTTGAAGGTACAA SBE: GGGAGGAGGTTTGACGGG	57	0.25 (21)	N/A	0.14 (21)	S
SMCY	AF440162	C/T (72 bp)	F: GATAGAGTTAGTTGACTGAACCAGAGGC R: AATCTTTCCTAACATGCTTCATGCC SBE: CCAGAGGCAAGGGACTATAA	60	0.46 (21)	N/A	0.33 (21)	D
PSUY658	AF244810	G/T (122 bp)	F: TCGTGATCCGCCCATCTT R: GCTCCTGAGGATGATGCTCAG SBE: ATAGGCCTGAGCCACCGC	57	0.33 (15)	N/A	0.2 (15)	S
sY67	AF440132	C/T (51 bp)	F: AGATGCTAGGGATGAACTCAGAA R: CCGTGCCTGGCTATTATTTC SBE: GCTAGGGATGAACTCAGAAAG	55	0 (21)	N/A	0 (21)	S
sY123	AF440150	G/T (76 bp)	F: GCATAATTTATCTCCTCCTGGG R: GAGGATTGGATTTGTCAGCTTT SBE: CCTGGGAGATGAAGGTTTT	55	0.32 (21)	N/A	0.19 (21)	S

*One of the three samples that was genotyped by sequencing and the SNaPshot kit had different genotypes according to the two different methods.

†For all samples combined.

T, this study; D, Deinard (1997); S, Stone *et al.* 2002; T_a, annealing temperature; H_E, expected heterozygosity; H_O, observed heterozygosity; SBE, single base extension.

Table 2 Heterozygosity and rare allele frequencies in each chimpanzee subspecies

	Central African chimpanzees		West African chimpanzees	
Locus	H_E	Rare allele frequency	H_E	Rare allele frequency
ApoB140	0.33	0.20	0.19	0.11
MYH6	0.03	0.02	0.00	0.00
FOS	0.03	0.02	0.51	0.45
IGH	0.38	0.25	0.34	0.21
ApoB476	0.51	0.47	0.19	0.11
CAT	0.00	0.00	0.40	0.26
TF1	0.10	0.05	0.29	0.17
PSUY376	0.16	0.08	0.37	0.22
sY85	0.51	0.40	0.30	0.17
sY19b	0.46	0.33	0.47	0.33
sY19a	0.16	0.08	0.37	0.22
SMCY	0.46	0.33	0.47	0.33
PSUY658	0.55	0.50	0.00	0.00
sY67	0.00	0.00	0.00	0.00
sY123	0.00	0.00	0.52	0.44
Overall	0.24		0.29	

H_E, expected heterozygosity.

(Roche) and 7.5 units of shrimp alkaline phosphatase (Roche) to remove unincorporated primers and dNTPs.

An additional oligo was designed for the SBE reaction to anneal with the 3′ end abutting the SNP site. All primers were synthesized by Eurogentec and are shown in Table 1. The SBE reaction was performed in 10-μL volumes and contained 5 μL SnaPshot™ Ready Reaction Mix (Applied Biosystems), 3 μL PCR product, and 1.2 μM of the SBE primer. Cycling conditions were 25 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 30 s with a final hold at 10 °C. The samples were then purified by the addition of 1 unit of shrimp alkaline phosphatase and incubation for 1 h at 37 °C. Post-extension treatment was required to prevent the interference of unincorporated ddNTPs during capillary electrophoresis. Genotypes were acquired by capillary electrophoresis on an ABI Prism™ 3100 Genetic Analyser, followed by analysis with GENESCAN software (Applied Biosystems).

The assays were tested on 49 individuals: one central and 18 western African chimpanzees; one putative western chimpanzee/bonobo hybrid from the Primate Foundation of Arizona collection; and 29 central African chimpanzee samples from the Centre International de Recherches Medicales Franceville. The results are summarized in Tables 1 and 2. Fifteen SNP markers are reported here. Seven autosomal and seven (of eight) Y chromosome assays produced high-quality, interpretable genotypes with the SNaPshot system and genotypes of three individuals were verified by comparison to previously generated sequences. Several genotypes from two of the assays (sY85 and PSUY658) were not determined due to low signal intensity (see Table 1). One Y chromosome assay (sY67) was monomorphic for the 21 samples in this set but is nonetheless included as it was shown to be a variable site in the source sequence alignment (Stone *et al*. 2002).

We used the program GDA (Lewis & Zaykin 2001) to calculate allele frequencies and expected and observed heterozygosities and to test for linkage disequilibrium (LD) within each population. The LD among markers was observed between ApoB140 and ApoB476 in both subspecies. In the western chimpanzees there was also significant LD between TF1 and both ApoB140 and ApoB476. For the central African chimpanzees significant LD was noted between ApoB140 and IGH and between MYH6 and FOS. Four loci (MYH6, CAT, PSUY658 and sY123) appear to show fixed differences with the loci being variable in one population but fixed in the other population in each case. Further, we observed a high overall F_{ST} value for these SNP markers (0.33) between the subspecies. The markers reported here will prove useful in monitoring genetic variation of natural and actively managed populations of this rare species and provide an alternative and complementary set of genetic markers for population genetics studies of chimpanzees.

Acknowledgements

This work was funded by the Max Planck Institute for Evolutionary Anthropology. We thank Jo Fritz from the Primate Foundation of Arizona and Jean Wickings of the Centre International de Recherches Medicales Franceville for chimpanzee samples.

References

Aitken N, Smith S, Schwarz C, Morin PA (2004) Single nucleotide polymorphism (SNP) discovery and genotyping in mammals: a targeted-gene approach. *Molecular Ecology*, **13**, 1423–1431.

Deinard A (1997) *The evolutionary genetics of the chimpanzees*. PhD Thesis, Yale University.

Deinard AS, Kidd K (2000) Identifying conservation units within captive chimpanzee populations. *American Journal of Physical Anthropology*, **111**, 25–44.

Ely JJ, Gonzalez DL, Reeves-Daniel A, Stone WH (1998) Individual identification and paternity determination in chimpanzees (Pan troglodytes) using human short tandem repeat (STR) markers. *International Journal of Primatology*, **19**, 255–271.

Goldberg TL (1997) Inferring the geographic origins of 'refugee' chimpanzees in Uganda from mitochondrial DNA sequences. *Conservation Biology*, **11**, 1441–1446.

Goossens B, Latour S, Vidal C *et al*. (2000) Twenty new microsatellite loci for use with hair and faecal samples in the chimpanzee (Pan troglodytes troglodytes). *Folia Primatologica*, **71**, 177–180.

de Groot NG, Otting N, Arguello R *et al*. (2000) Major histocompatibility complex class I diversity in a West African chimpanzee

population: implications for HIV research. *Immunogenetics*, **51**, 398–409.

Gusmao L, Gonzalez-Neira A, Alves C *et al.* (2002) Genetic diversity of Y-specific STRs in chimpanzees (Pan troglodytes). *American Journal of Primatology*, **57**, 21–29.

Lewis PO, Zaykin D (2001) GENETIC DATA ANALYSIS: Computer program for the analysis of allelic data. Free program available from http://lewis.eeb.uconn.edu/lewishome/software.html.

Lyons LA, Laughlin TF, Copeland NG *et al.* (1997) Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nature Genetics*, **15**, 47–56.

Morin PA, Moore JJ, Chakraborty R *et al.* (1994a) Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science*, **265**, 1193–1201.

Morin PA, Moore JJ, Wallis J, Woodruff DS (1994b) Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular Ecology*, **3**, 469–478.

Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (Pan troglodytes verus). *Molecular Ecology*, **10**, 1835–1844.

Morin PA, Luikart G, Wayne RK, SNP Workshop Group (2004) Applications of single nucleotide polymorphisms (SNPs) in ecology, evolution, and conservation. *Trends in Ecology and Evolution*, **19 (4)**, 208–216.

Stone AC, Griffiths RC, Zegura SL, Hammer MF (2002) High levels of Y-chromosome nucleotide diversity in the genus Pan. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 43–48.

Surridge AK, Smith AC, Buchanan-Smith HM, Mundy NI (2002) Single-copy nuclear DNA sequences obtained from noninvasively collected primate feces. *American Journal of Primatology*, **56**, 185–190.

Vigilant L, Hofreiter M, Siedel H, Boesch C (2001) Paternity and relatedness in wild chimpanzee communities. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 12 890–12 895.

Yu N, Jensen-Seaman MI, Chemnick L *et al.* (2003) Low nucleotide diversity in chimpanzees and bonobos. *Genetics*, **164**, 1511–1518.