American Journal of Primatology 55:253–259 (2001)

BRIEF REPORT

Characterization and Screening of Microsatellite Loci in a Wild Lemur Population (*Propithecus verreauxi verreauxi*)

RICHARD R. LAWLER[1]*, ALISON F. RICHARD[1], AND MARGARET A. RILEY[2]
[1]*Department of Anthropology, Yale University, New Haven, Connecticut*
[2]*Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut*

Sixteen dinucleotide microsatellite loci were isolated from the genome of *Propithecus verreauxi verreauxi*. All loci were polymorphic when genotyped on a minimum of 16 animals. The number of alleles across these loci ranges from two to 11. Additionally, seven of these loci were genotyped across a minimum of 200 animals in order to estimate heterozygosity and their potential for parentage assignment in this population. Using these seven loci, the mean heterozygosity in this population is 0.705, and the combined probability of these seven loci to exclude a random individual from parentage, when one parent is known, is 0.996. These data suggest that these loci will be useful for estimating a variety of population genetic and genealogical parameters in *P. v. verreauxi* populations. Am. J. Primatol. 55:253–259, 2001.

Key words: microsatellite; *Propithecus*; STR locus; lemur

INTRODUCTION

Microsatellites, or short tandem repeats (STRs), are codominant, highly polymorphic molecular markers that provide information on intra- and interpopulation structure, genetic relatedness, and gene location via linkage maps. They consist of a tandemly repeated motif of one to six nucleotides (e.g., ATT[13] or CA[26], where 26 equals the number of "CA" motifs that are repeated) flanked by a nonrepetitive sequence of nucleotides. Polymerase chain reaction (PCR) primers may be designed in the flanking regions, allowing one to amplify the polymorphic repeat region. To the extent that the flanking sequences and repeat unit are found in neutral portions of the genome and segregate independently, microsatellites will act as single, unlinked, neutral loci that are generally 700 base pairs (bp) or less in size [Scribner & Pearce, 2000]. These properties make microsatellites ideal genetic markers for socioecological studies of wild primate populations that require genetic information on population structure and/or kinship.

Contract grant sponsor: N.S.F.; Contract grant number: DEB-9902146; Contract grant sponsor: N.I.H.; Contract grant number: GM58433.

*Correspondence to: Richard Lawler, Dept. of Anthropology, P.O. Box 208277, New Haven, CT 06520-8277. E-mail: richard.lawler@yale.edu

Received 6 March 2001; revision accepted 20 August 2001

Microsatellites are increasingly characterized and used in population genetic and behavioral studies of numerous nonhuman primates [e.g., Constable et al., 2001; Ellsworth & Hoelzer, 1998; Jekielek & Strobeck, 1999; Von Segesser et al., 1999]. Some of these studies have used cross-specific, or heterologous microsatellite loci (loci characterized in a primate species that is different from the one under study), but it has been demonstrated that some cross-specific microsatellite loci show lowered heterozygosities and may yield inaccurate pedigrees [Smith et al., 2000; Vigilant & Boesch, 2001]. These phenomena are attributed to mispriming and/or template quality, and can lead to potentially serious errors in assigning kin relations and estimating population genetic parameters from cross-specific microsatellite loci [Beaumont & Bruford, 1999; Constable et al., 2001; Gagneux et al., 1997; Pemberton et al., 1995; Smith et al., 2000; Taberlet et al., 1999]. Similarly, cross-specific loci will not always amplify in the species under study; in these cases, it is necessary to isolate species-specific microsatellite loci [e.g., Jekielek & Strobeck, 1999].

In this report we provide information on microsatellite loci isolated from a wild lemur population, the white sifaka (*Propithecus verreauxi verreauxi*) at Beza Mahafaly Special Reserve, southwest Madagascar. We were motivated to isolate these loci after having unsuccessfully tested 20 human and 16 *Eulemur*-specific microsatellite loci on the *P. v. verreauxi* genome. Our goal here is to provide information on these *P. v. verreauxi*-specific microsatellite loci and their potential as estimators of genetic diversity and relationships in the Beza Mahafaly *Propithecus* population.

METHODS

Except where noted, we followed the protocol of Hammond et al. [1998]. All tissue samples came from individuals captured and released in the wild [see Richard et al., 1993]. Approximately 0.3 g of pinna (ear) tissue from 19 white sifaka were extracted for DNA following the protocol in Strauss [1998]. On average, each extraction yielded about 70 ng/μl. The DNA samples were pooled and concentrated, digested, and size-selected for a 350–700 bp region.

To screen for microsatellites, a 30 bp oligonucleotide consisting of 15 repeat units of CA was used. Our enrichment process is different from Hammond et al. [1998] [see Fischer & Bachmann, 1998]. Three biotin-tags were attached to the 5′ end and the oligo had a 3′ chain-terminator to prevent concatamers during subsequent PCRs. The addition of a 3′ chain-terminator was a critical component for successful primer design [Koblizkova et al., 1998]. The size-selected DNA was added to 30 μl of 20× SSC and 25 μl of H_2O, and the mixture was denatured (10 min at 95°C) and then placed on ice for 2 min. Then 5 μg of the biotin-labeled CA probes were added, and the mixture was put at 65°C for 20 min. To capture the portions of DNA containing repeat units, we used streptavidin-coated magnetic beads (Promega, Madison, WI). The entire mixture from above was added to the streptavidin beads (suspended in 0.5× SSC) and put at room temperature for 15 min. The beads were separated from the supernatant with a magnetic stand (Promega, Madison, WI) and washed three times with 100 μl of 0.1× SSC, letting them sit 5 min between each wash. The DNA was eluted from the beads by washing two times in 25 μl of ddH_2O. The resulting 50 μl of eluate was concentrated and cleaned using Qiagen (Valencia, CA) purification columns. We performed the enrichment phase two times.

After the enrichment phase, the DNA was ligated and transformed following standard procedures, and 212 colonies were screened for an insert via "colony

PCR." Forty-six positive clones were identified and sequenced using a Perkin Elmer 377 automated sequencer (Applied Biosystems, Foster City, CA). Twenty-one of the 46 clones contained a CA repeat. From this, 16 primer pairs were designed using MacVector software. Primer pairs were then tested on the sifaka template. The forward primer was labeled with a fluorescent dye (6-FAM, HEX, or TET). A PCR reaction with 1× PCR buffer, 1.5 mM of $MgCl_2$, 80 μM of dNTP, 20 pmol of forward primer, 20 pmol of reverse primer, and ~100 ng of DNA template was run and analyzed on a 377 automated sequencer using GeneScan analysis (Applied Biosystems, Foster City, CA) following the prescribed protocols.

RESULTS

We designed 16 primer pairs that yielded easily quantifiable genotypes when visualized on GeneScan software. Initially, we genotyped 16 animals across all loci. Using information from this initial genotyping, loci that revealed high polymorphism were further genotyped on a minimum of 200 animals. For all loci, the primer sequences, number of animals genotyped at the locus, number of alleles, type of repeat unit, and annealing temperature are listed in Table I. The loci range in size from ~150–420 bp, and the number of alleles across all loci range from 2 to 11. Ten of the 16 loci have pure CA repeat units, while the other six have interrupted repeat units. All interrupted repeats contained no more than one or two nucleotides interspersed within the repeat unit. Across all loci, there is a positive association between number of alleles and number of repeats (Spearman's Rho = 0.67, $P = 0.004$). In the text below, we refer to each locus by its number only.

Table II provides population genetic and genealogical data on the seven loci that were genotyped on a minimum of 200 individuals (loci 1, 4, 6, 8, 14, 15, and 16). This information was generated using the program CERVUS 2.0 [Marshall et al., 1998]. All loci except locus 4 conform to Hardy-Weinberg expectations using the chi-square test and the exact test [Haldane, 1954; Raymond & Rousset, 1995]. Additionally, two measures that can be used in parentage analysis are provided in Table II. One exclusionary power, Excl.-A, is the average probability of a locus to exclude a randomly chosen individual from parentage of an offspring given only the offspring's genotype. Excl.-B is the average probability of a locus to exclude a randomly chosen individual from parentage given the offspring's genotype and the genotype of one known parent. Total exclusionary power refers to the combined power of all loci to exclude a random individual from parentage of an offspring [Marshall et al., 1998]. Null allele frequency provides a measure of the *potential* for a locus to possess non-amplifying alleles. A large, positive score relative to other loci indicates an excess of homozygotes, but this score does not necessarily imply that null alleles are present [Summers & Amos, 1997]. The mean expected heterozygosity for the population is 0.747, and the mean observed heterozygosity is 0.705. The total exclusionary power of all seven loci when neither parent is known is 0.957, and it is 0.996 when one parent is known.

DISCUSSION

We feel the primary factors contributing to our success in microsatellite characterization were the initial isolation of high molecular-weight DNA for subsequent digestion, the use of 3′ chain-terminators on the CA biotin-labeled probes during the enrichment phase, and repeating the enrichment phase two times. Six of the seven loci conformed to Hardy-Weinberg expectations when screened

TABLE I. Primer Name, Primer Sequences, Number of Animals Genotyped (n), Size in Base Pairs, Number of Alleles, Repeat Type, and Annealing Temperature for the 16 Microsatellite Loci Used in This Study.

Primer	Sequence (5′—> 3′)	(n)	Size (bp)	No. of Alleles	No. of repeats (Type)	Temp.
P.V. 1 F	GTTTCTTTTTCTTGCAGC	228	156-174	10	CA-17 (pure)	54°C
P.V. 1 R	CTTCTCTGGCTTCACATC					
P.V. 2 F	GAAGGTAAGTTTTCTGGCAG	16	273-289	4	CA-15 (interrupted)	58°C
P.V. 2 R	AGTGTTTTATCGTATGGATGC					
P.V. 3 F	GAAAGAAATGCTAGACCTAGAACGC	16	409-423	9	CA-12 (pure)	54°C
P.V. 3 R	GGGATCAGGACTTCAACATACTGC					
P.V. 4 F	TCATTAGTGCCACGCAGTATGG	208	296-339	7	CA-15 (interrupted)	57°C
P.V. 4 R	TGGAAGAACACGCTGACGACAG					
P.V. 5 F	CCCTTCTTCTCTCTGTGAGTGG	16	266-274	5	CA-16 (pure)	55°C
P.V. 5 R	TTGGGTTTGCTGCTGTCCTG					
P.V. 6 F	CAAGTGCTAGTCTAAACCTGGGTG	258	260-278	10	CA-21 (pure)	55°C
P.V. 6 R	CACAGAAGCCTGATGTAACAACAG					
P.V. 7 F	TTCTCCCACTACTGAGCGAG	16	253-261	4	CA-13 (pure)	55°C
P.V. 7 R	TCTGGAGGGCTGGAACAAAG					
P.V. 8 F	CTCAAAGACATTTTCCTTCAGCC	241	211-227	6	CA-16 (interrupted)	53°C
P.V. 8 R	TTTCTACTCACCCCACAGTCATTAG					
P.V. 9 F	TTTCCTCCTCCAGGGAGTCCAAAC	16	222-226	2	CA-12 (pure)	58°C
P.V. 9 R	GGACATCTGCACCATTGACCTAAC					
P.V. 10 F	ACGACCAACCCTATCTCTTAAAC	16	237-241	3	CA-11 (interrupted)	50°C
P.V. 10 R	TGTCTTAGGATTGCGTGGG					
P.V. 11 F	GGAAGGGATTTGGGTACACAGAGAG	16	334-338	3	CA-9 (interrupted)	58°C
P.V. 11 R	CATTCGTGGAGGTCAGTTCCATC					
P.V. 12 F	GCCCCTAATAATTTGAGCCAC	16	334-353	6	CA-8 (interrupted)	53°C
P.V. 12 R	ATCAAGCTGCTGTCCAACAAGCCC					
P.V. 13 F	CCTGTGTATGAATCGCAAAGGCAAG	16	229-235	4	CA-15 (pure)	57°C
P.V. 13 R	GCAGAGAAGAGTAGGTGAAAGGAAG					
P.V. 14 F	GGCTCAAGACTGATGCTTCAGGTC	241	301-325	11	CA-20 (pure)	60°C
P.V. 14 R	GTTTCCAATAGGACAATCACTGGC					
P.V. 15 F	CCTTCATTCCTTTTCATTTCTTGG	227	247-267	11	CA-16 (pure)	50°C
P.V. 15 R	TTTTGTATTAGACTAAGCTGCC					
P.V. 16 F	TGAGGGTGGTGAGCTTTAGC	243	270-293	10	CA-15 (pure)	55°C
P.V. 16 R	GGGCTGGGGAAAAAATATAAC					

TABLE II. Population Genetic and Genealogical Information for the 7 Loci Genotyped on a Minimum of 200 Animals.

Locus	k	Het(obs)	Het(exp)	Excl.-A	Excl.-B	Null alleles
P.V. 1	10	0.706	0.753	0.367	0.550	0.0317
P.V. 4*	7	0.659	0.768	0.367	0.546	0.0715
P.V. 6	10	0.729	0.781	0.401	0.581	0.0322
P.V. 8	6	0.618	0.628	0.219	0.382	0.0095
P.V. 14	11	0.722	0.734	0.348	0.532	0.0072
P.V. 15	11	0.749	0.750	0.365	0.546	–0.0005
P.V. 16	10	0.749	0.814	0.453	0.629	0.0416

Mean number of alleles per locus:	9.29
Mean expected heterozygosity:	0.747
Mean observed heterozygosity	0.705
Total exclusionary power (no parent known):	0.957
Total exclusionary power (one parent known):	0.996

* deviates from Hardy-Weinberg expectations.
k, number of alleles; Het(obs), observed heterozygosity; Het(exp), expected heterozygosity; Excl.-A, the exclusion probability of the locus when no parents are known; Excl.-B, the exclusion probability of the locus when one parent is known.

on a minimum of 200 individuals (Table II). This suggests that these six loci will be useful for estimating a variety of population-genetic parameters (e.g., population substructure) that require neutral markers. Locus 4 deviates from Hardy-Weinberg expectations. Several factors may account for this: linkage, selection, pooling samples across families or age-cohorts, and/or the presence of null-alleles (the null allele frequency was highest in locus 4, Table II).

In addition to providing estimates on genetic variability, microsatellite loci can provide a powerful means to assess parentage and kin relations among individuals [Luikart & England, 1999]. The method of assigning individuals to parents with exclusion equations is a function of the number and frequency of alleles at a locus (assuming all candidate parents are sampled). While the average exclusion probabilities with or without the genotype of a known parent are calculated differently [Chakravarti & Li, 1983; Marshall et al., 1998], all exclusion probabilities are generally maximized when there are numerous alleles at relatively equal frequencies at the locus [Evett & Weir, 1999]. When several unlinked loci are used, the total exclusion probability is the complement of the product of the single-locus inclusion probabilities; that is, the total exclusion probability is 1 minus the combined probability of the set of loci to include a random individual [Evett & Weir, 1999; Marshall et al., 1998]. Exclusion equations rely on allele frequencies, not genotype frequencies, to assign parents to offspring [Evett & Weir, 1999]; for this reason, locus 4 can still contribute some information to parent–offspring relationships. As Table II shows, the exclusion probabilities for locus 4 fall within the range of the other loci, although samples typed at this locus should be checked for the potential for null alleles. The seven loci we have genotyped above can reliably exclude a random individual from parentage with a probability of 95%, when the other parent is unknown. If there is a known parent that is genotyped, the probability of reliably excluding a random individual from parentage is 99%. Overall, the above data suggest that a number of population genetic (e.g., substructure and effective population size) and genealogical (e.g., reproductive success and kinship) parameters may be estimated from the sifaka population at Beza Mahafaly using these loci.

ACKNOWLEDGMENTS

We thank David Watts, Michael Jensen-Seaman, Kristin Saltonstall, Jeffrey Huckaby, and members of the Riley/Dorit labs (in the Department of Ecology and Evolutionary Biology, Yale University) for comments and assistance. We appreciate the detailed and constructive comments of three anonymous reviewers; their input improved the final version of this manuscript. We are grateful to the government of Madagascar for permission to undertake this research, and we thank the Director and faculty of the School of Agronomy for their institutional support and collaboration. We particularly appreciate the contributions of the Beza Mahafaly Monitoring Team, led by Dr. Joelisoa Ratsirarson, who collected much of the data from the wild on which this study is based. This study was supported by a N.S.F. doctoral dissertation improvement grant (DEB-9902146) to R.R.L., and a N.I.H. grant (GM58433) to M.A.R.

REFERENCES

Beaumont MA, Bruford MW. 1999. Microsatellites in conservation genetics. In: Goldstein DB, Schlotterer C, editors. Microsatellites: evolution and applications. Oxford: Oxford University Press. p 165–182.

Chakravarti A, Li CC. 1983. The effect of linkage on parentage calculations. In: Walker RH, editor. Inclusion probabilities in parentage testing. Arlington, VA: American Association of Blood Banks. p 411–422.

Constable JL, Ashley MV, Goodall J, Pusey AE. 2001. Noninvasive paternity assignment in Gombe chimpanzees. Mol Ecol 10:1279–1300.

Ellsworth JA, Hoelzer GA. 1998. Characterization of microsatellite loci in a New World primate, the mantled howler monkey (*Alouatta palliata*). Mol Ecol 7:657–658.

Evett IW, Weir BS. 1998. Interpreting DNA evidence: statistical genetics for forensic scientists. Sunderland: Sinauer Press. 278 p.

Fischer D, Bachmann K. 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa L.).* Biotechniques 25:796–802.

Gagneux P, Boesch C, Woodruff DS. 1997. Microsatellite scoring errors associated with non-invasive genotyping based on nuclear DNA from shed hair. Mol Ecol 6:861–868.

Haldane JBS. 1954. An exact test for randomness of mating. J Genet 52:631–635.

Hammond RL, Saccheri IJ, Ciofi C, Coote T, Funk SM, McMillan WO, Bayes MK, Taylor E, Bruford MW. 1998. Isolation of microsatellite markers in animals. In: Karp A, Isaac PG, Ingram DS, editors. Molecular tools for screening biodiversity. London: Chapman and Hall. p 279–285.

Jekielek J, Strobeck C. 1999. Characterization of polymorphic brown lemur (*Eulemur fulvus*) microsatellite loci and their amplification in the family Lemuridae. Mol Ecol 8:901–903.

Koblizkova A, Dolezel J, Macas J. 1998. Subtraction of 3′ modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. Biotechniques 25:32–38.

Luikart G, England PR. 1999. Statistical analysis of microsatellite DNA data. Trends Ecol Evol 14:253–256.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655.

Pemberton JM, Slate J, Bancroft DR, Barrett JA. 1995. Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol Ecol 4:249–252.

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249. http://wbiomed.curtin.edu.au/genepop/

Richard AF, Rakotomanga P, Schwartz M. 1993. Dispersal by *Propithecus verreauxi* at Beza Mahafaly, Madagascar. Am J Primatol 30:1–20.

Scribner KT, Pearce JM. 2000. Microsatellites: evolutionary and methodological background and empirical applications at individual, population, and phylogenetic levels. In: Baker AJ, editor. Molecular methods in ecology. Oxford: Blackwell. p 235–273.

Smith KL, Alberts SC, Bayes MK, Bruford MW, Altmann J, Ober C. 2000. Cross-species amplification, non-invasive genotyping, and non-Mendelian inheritance of human STRPs in savannah baboons. Am J Primatol 51:219–227.

Strauss WM. 1999. Preparation of genomic DNA from mammalian tissue. In: Ausebel FM, Brent R, Kingston RE, Moore DD,

Seidman JG, Smith JA, Struhl K, editors. Current protocols in molecular biology. New York: John Wiley and Sons. Unit 2.2.

Summers K, Amos W. 1997. Behavioral, ecological and molecular genetic analyses of reproductive strategies in the Amazonian dart-poison frog, *Dendrobates vertrimaculatus*. Behav Ecol 8:260–267.

Taberlet P, Waits LP, Luikart G. 1999. Noninvasive genetic sampling: look before you leap. Trends Ecol Evol 14:323–327.

Vigilant L, Boesch C. 2001. Reproductive strategies of West African chimpanzees (*P. t. verus*) of the Tai National Park: a reduced role for extra-group paternity. Am J Phys Anthropol 32:156.

Von Segesser F, Menard N, Gaci B, Martin RD. 1999. Genetic differentiation within and between isolated Algerian subpopulations of Barbary macaques (*Macaca sylvanus*): evidence from microsatellites. Mol Ecol 8:433–442.