

# Technical Challenges in the Microsatellite Genotyping of a Wild Chimpanzee Population Using Feces

LINDA VIGILANT

I will present some preliminary results from a study of habituated communities of chimpanzees (*Pan troglodytes verus*) in the Taï National Park, Ivory Coast. The necessary investment, particularly in terms of time and expense, of generating reliable genotypes from a large number of individuals will be described.

Noninvasive collection of samples from wild mammal populations presents an extraordinary opportunity to conduct genetic analysis without necessarily capturing or even directly observing the subjects of interest. Unfortunately, the hair or feces samples typically collected present formidable challenges in the laboratory. The investment of time and expense required makes careful experimental design a necessity.

Researchers who want to examine the genetics of wild mammal populations have great interest in the use of noninvasively collected sample materials. Questions that can potentially be answered range from population-level inquiries into such topics as population size, sex composition, individual relationship, and dispersal, to comparisons between populations provid-

ing information on geographic differentiation, comparative levels of genetic diversity, and phylogenetic relationships.<sup>1</sup> Effective investigation of these matters relies on an ability to examine genetic variation from the nuclear genome as well as the maternally transmitted mitochondrial DNA (mtDNA). A typical approach to examining the genetic structure of a population involves the assessment of variation at multiple nuclear microsatellite loci. Alleles of these microsatellite loci vary in the length of simple repeat elements, hence allowing for the potential assemblage, when enough loci are examined, of individually distinctive genotypes.<sup>2</sup> I will therefore focus here on microsatellite analyses.

Numerous studies have employed microsatellite analysis of noninvasive samples in the form of feces or shed hair.<sup>3–8</sup> But, as Taberlet et al.<sup>9</sup> recently noted, to date there have been relatively few “comprehensive” studies using nuclear DNA from noninvasive samples. This is surprising in view of the existence of the large body of literature on the topic, and the subjective impression that many field researchers are equipped with collection materials and a plan or at least a hope of arranging for later genetic analysis. The indication is thus that, years after it was heralded as representing a revolution in field studies of primates,<sup>10</sup> noninvasive sampling for genetic analysis either has not yet

reached its potential or has proved disappointing.

Ironically, the simplicity and low costs of the polymerase chain reaction (PCR) that have led to its ubiquitous use in molecular analysis laboratories may have led to an underestimation by laboratory and field researchers alike of the difficulties inherent in the use of noninvasive samples. Such analyses in fact present considerable technical challenges that can only be overcome by investing a great deal of time, expense, and laboratory expertise. Most technical considerations involved in the production of accurate genotypes from noninvasive samples have been thoroughly addressed elsewhere.<sup>9,11</sup> In this contribution, I will present some preliminary results from a study of habituated communities of chimpanzees (*Pan troglodytes verus*) in the Taï National Forest, Ivory Coast. The necessary investment, particularly in terms of time and expense, of generating reliable genotypes from a large number of individuals will be described.

## TAÏ CHIMPANZEES

The chimpanzees of the Taï National Forest have been the subjects of behavioral observation for over 20 years.<sup>12</sup> In the early 1990s, the collection of hair from abandoned night nests was begun for the purpose of microsatellite genotyping. Analysis of the assembled genotypes of some 50 members of a habituated community led to the unexpected inference of a high incidence (7 of 13 cases) of extra-group paternity of infants within the community.<sup>13</sup> The current project on

Linda Vigilant is at the Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, D-Leipzig 04103, Germany (E-mail: vigilant@eva.mpg.de).

Key words: chimpanzees; genotyping; noninvasive sampling; microsatellites; feces; *Pan troglodytes verus*

Evolutionary Anthropology, Suppl 1:162–165 (2002)  
DOI 10.1002/evan.10082  
Published online in Wiley InterScience  
(www.interscience.wiley.com).  
© 2002 Wiley-Liss, Inc.

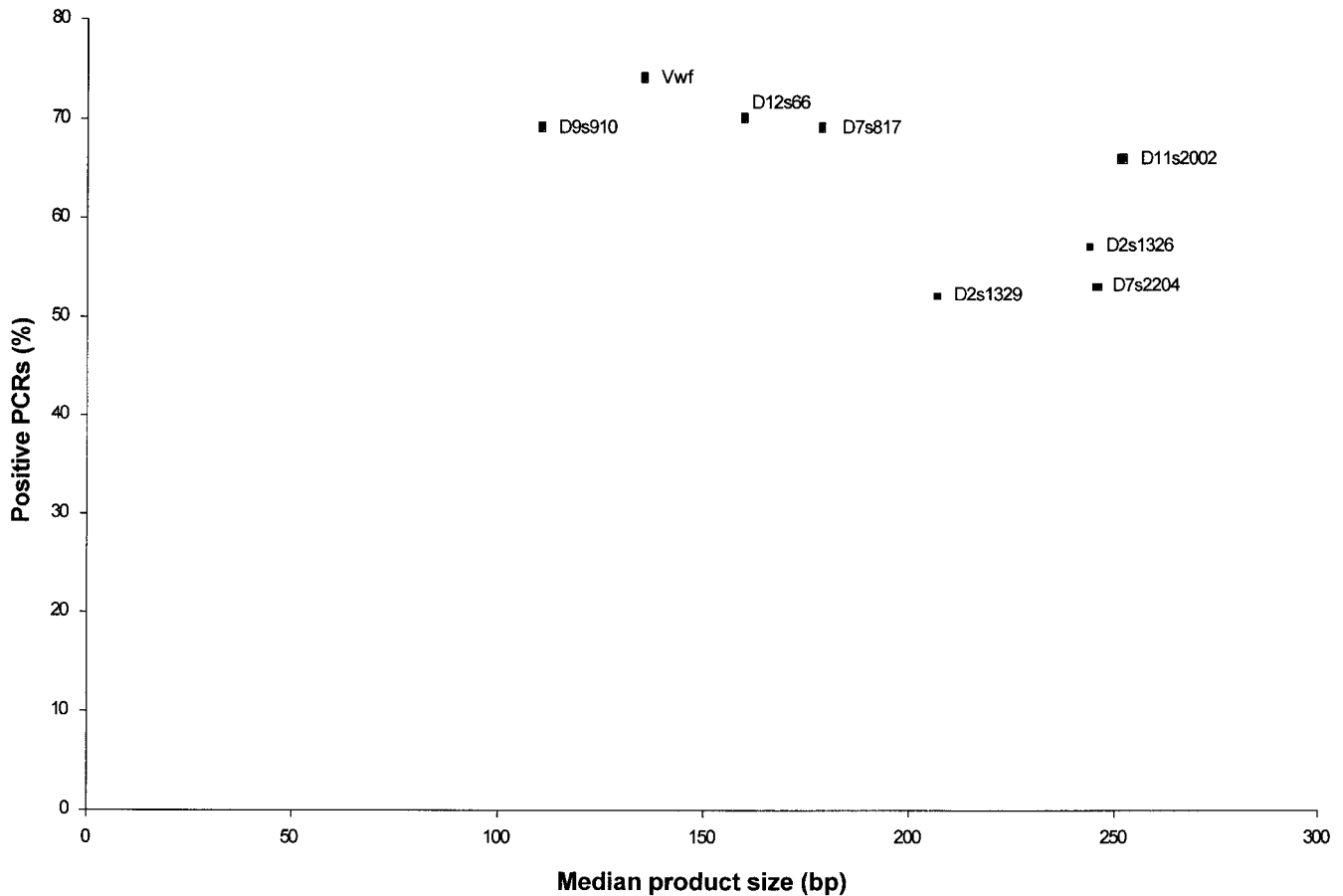


Figure 1. Success of PCR at eight microsatellite loci, using DNA from chimpanzee feces.

the genetics of the Taï chimpanzees involves the reanalysis of all previously examined individuals, plus previously unexamined or new members, as well as members of two more recently habituated communities. The number of individuals to be genotyped thus totals approximately 120. In addition, shed hairs from nests of nonhabituated chimpanzees living at the borders of the territories of habituated groups have been collected continuously for 2 years, and this collection constitutes multiple hairs from some 1,300 nests. Whenever possible, samples from habituated individuals are collected in the form of feces rather than hair, because of the following considerations.

### CHOICE OF SAMPLES

Despite the early work on sex-determination using nuclear DNA from bear feces,<sup>3,14</sup> the use of hair as sample material has predominated over

that of feces. Some of the explanation for this is probably the longer persistence of hair compared to fecal samples under natural conditions. Another factor is that, whereas rapid, simple methods for amplification of DNA from hair exist,<sup>15,16</sup> some methods for the extraction of DNA from feces are much more time-consuming, and success depends heavily on the sample storage method used.<sup>17</sup>

The Taï genotyping project described here used feces with a high success rate. These samples were stored dry on collection using silica gel beads, and DNA was extracted from 100 mg (one-tenth of the total) of the sample material using the QiaAmp DNA Stool Kit (Qiagen). Amplification of a small region of the amelogenin locus for sex determination<sup>18</sup> is successful more than 85% ( $n = 112$ ) of the time. Microsatellite loci amplify in an average of 64% of attempts ( $n = 704$ ), with success for individual loci

ranging from 52–74% (Fig. 1). Loci with a median product size less than 200 bp long generally tend to produce results more readily. Interestingly, samples collected during the same time period from members of the same community exhibited marked variability in success. While some extracts amplified successfully in less than 1 in 5 attempts, the majority of extracts (26 of 31) produced consistently positive results. Multiple samples were collected from each individual to overcome this variation in success rate, which may stem from dietary or other factors such as individual health or physiology.

Shed hairs are useful for the sampling of nonhabituated groups because of their greater persistence at nest sites. The DNA from single shed hairs does not amplify as readily as DNA from feces, presumably due to factors such as the intrinsically very low DNA content of hair<sup>19</sup> and its pos-

TABLE 1. Microsatellite Loci Used in Genotyping of Tai Chimpanzees

Locus	No. of Individuals	No. of Alleles	Allele Sizes (bp)	Heterozygosity <sup>a</sup>	
				Observed	Expected
D2s1326	19	10	222–263	0.895	0.873
D7s817	19	9	165–202	0.895	0.872
D7s2204	16	8	229–265	0.813	0.835
D9s910	17	7	101–122	0.941	0.861
D2s1329	19	7	171–223	0.947	0.828
D11s2002	17	7	239–262	0.882	0.768
D12s66	22	6	146–178	0.682	0.729
Vwf	20	6	116–148	0.650	0.662

<sup>a</sup> Expected heterozygosity is a measure of the informativeness of a locus and assumes a population in Hardy-Weinberg equilibrium. It is calculated using equation 8.4 in Nei.<sup>23</sup>

sible deterioration under adverse storage conditions (e.g., high humidity). The widely used Chelex method for isolation of DNA from hair is not superior to an equally simple approach with hair digestion and PCR without subsequent purification. In fact, this latter method is likely to produce DNA of greater stability for long-term storage and is therefore preferable.<sup>20</sup>

### PRODUCING ACCURATE GENOTYPES

The ascertainment of genotypes that will allow for the determination of genetic relationships with high confidence requires the typing of individuals at several microsatellite loci that have high heterozygosity, or variability, in the population of interest. The selection of loci is a time-consuming process that requires either the painstaking isolation of candidate loci in the species of interest or the use of loci originally derived in a closely related species. For this project, 67 tri- and tetranucleotide repeat microsatellite loci that were initially characterized in humans and also shown to amplify from macaques (*Macaca mulatta*)<sup>21</sup> or baboons (*Papio hamadryas anubis*, *P. h. cynocephalus*)<sup>22</sup> were screened using pooled high-quality genomic DNA derived from the blood of 10 unrelated West African chimpanzees (*P. t. verus*). Only eight of these loci failed to amplify from chimpanzee DNA. Loci that amplified only weakly, produced products in excess of 250 bp, or seemed invariable were eliminated from further consideration. The set of 10 chimpanzees was then individually genotyped at 29 loci.

Eight of these loci were subsequently used in the genotyping of the Tai chimpanzees (Table 1). The high heterozygosity of these loci in this population allows for the determination of paternity exclusions at the 99% significance level.

A significant challenge in the use of DNA from feces is the low concentration of DNA in the final extract. The recommendation for such samples is to employ a “multiple tubes” procedure, in which each result is repeated up to a total of seven times to achieve a result reliable by statistical considerations.<sup>11,24</sup> This necessity, coupled with the fact that not all amplifications produce a product, means that a large number of PCRs must be generated in order to produce a sufficient number for further characterization.

In practical terms, this means that a large amount of repetition is necessary to generate accurate genotypes. This is possible, however, only with a considerable amount of time and expense. A look at the effort involved in the production of genotypes using fecal samples from a typical sampling of 9 chimpanzees clearly illustrates this. All samples were extracted in duplicate, and PCRs were set up in triplicate from each extract. This means that for the genotyping of 9 samples at eight loci, a total of almost 500 PCR reactions (including controls) was produced. This resulted in assignment with high confidence of genotypes at seven or more loci for 6 of the 9 individuals. Tentative results were obtained for the other loci, but required confirmation by additional PCRs. The 3 remaining samples produced DNA of poor quality, and these individuals

had to be completed using DNA extracted from new samples. Using the system established, it is conservatively estimated that a month is needed for one person to produce complete genotypes of five individuals. This deliberate pace is coupled with a nontrivial cost per individual. The cost of genotyping a single sample at eight microsatellite loci, in terms of reagents alone, totals about \$100. This is without counting either the costs of the necessary equipment or the disposable items used.

Fortunately, steady technical improvements decrease the difficulty and increase the reliability of genotyping using noninvasive samples. For example, the use of tetranucleotide repeats and heat-activated polymerases (e.g., AmpliTaq Gold™, PE Biosystems) reduces the occurrence of spurious products that can cause genotyping errors. Automated PCR product analysis reduces both experimental variability and dependence on human interpretation of results. Success rates are improved by continuing small adjustments in storage and DNA-extraction methods. The sensitivity of methods has also improved, such that less of each sample needs to be used for each PCR, and the danger of exhausting the sample is thus reduced.

### CONCLUSIONS

Noninvasive samples are often the only type of genetic material available to address interesting questions from wild animal populations. The technical difficulties and limitations are considerable, but can be overcome by a

rigorous technical approach to ensure both optimal success rates and reliable results. This, coupled with the significant expenditure of both time and money, means that genetic analyses of wild populations must represent a carefully planned separate project and not an ancillary attempt to “do some genetics” as an addition to a field study. As the number of studies published increases, there is a clear danger that the data from these will not be directly comparable, because of the choice of different loci by different workers. This can, for instance, result in the inability to make comparisons of genetic diversity between populations or subspecies. Because comprehensive projects on populations are slow to reach completion, it is unfortunate if workers do not choose to make methodological details available prior to publication. Details of the methods used and loci tested for the project on chimpanzees (*Pan troglodytes verus*), as well as information on ongoing collaborative projects on gorillas (*Gorilla gorilla gorilla*, *G. g. beringei*) and gibbons (*Hyllobates lar*), are available from the author on request.

#### ACKNOWLEDGMENTS

The project described here would not have been possible without the commitment and efforts of Christophe Boesch and numerous field workers. I also thank Heike Siedel and Anette Abraham for excellent technical assistance, and Brenda Bradley,

Phillip Morin, and Svante Pääbo for constructive comments on the manuscript.

#### REFERENCES

1. Luikart G, England PR. 1999. Statistical analysis of microsatellite DNA data. *Trends Ecol Evol* 14:253–256.
2. Queller DC, Strassmann JE, Hughes CR. 1993. Microsatellites and kinship. *Trends Ecol Evol* 8:285–288.
3. Taberlet P, Mattock H, Dubois-Paganon C, Bouvet J. 1993. Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. *Mol Ecol* 2:399–403.
4. Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS. 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* 265:1193–1201.
5. Constable JJ, Packer C, Collins DA, Pusey AE. 1995. Nuclear DNA from primate dung. *Nature* 373:393.
6. Gagneux P. 1998. Population genetics of West African chimpanzees. Ph.D. thesis, University of Basel, Basel, Switzerland.
7. Gerloff U, Hartung B, Fruth B, Hohmann G, Tautz D. 1999. Intracommunity relationships, dispersal pattern and paternity success in a wild living community of bonobos (*Pan paniscus*) determined from DNA analysis of faecal samples. *Proc R Soc Lond [Biol]* 266:1189–1195.
8. Kohn MH, York EC, Kamradt DA, Haught G, Sauvajot RM, Wayne RK. 1999. Estimating population size by genotyping faeces. *Proc R Soc Lond [Biol]* 266:657–663.
9. Taberlet P, Waits LP, Luikart G. 1999. Noninvasive genetic sampling: look before you leap. *Trends Ecol Evol* 14:323–327.
10. Woodruff DS. 1993. Non-invasive genotyping of primates. *Primates* 34:333–346.
11. Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res* 24:3189–3194.
12. Boesch C, Boesch-Achermann H. 2000. The chimpanzees of the Tai forest. Oxford: Oxford University Press.
13. Gagneux P, Woodruff DS, Boesch C. 1997. Furtive mating in female chimpanzees. *Nature* 387:358–359.
14. Kohn M, Knauer F, Stoffella A, Schröder W, Pääbo S. 1995. Conservation genetics of the European brown bear—a study using excremental PCR of nuclear and mitochondrial sequences. *Mol Ecol* 4:95–103.
15. Walsh PS, Metzger DA, Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10:506–513.
16. Allen M, Engström AS, Meyers S, Handt O, Saldeen T, von Haeseler A, Pääbo S, Gyllenstein U. 1998. Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. *J Forensic Sci* 43:453–464.
17. Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR. 1997. Techniques for application of faecal DNA methods to field studies of ursids. *Mol Ecol* 6:1091–1097.
18. Sullivan KM, Mannucci A, Kimpton CP, Gill P. 1993. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *Biotechniques* 15:636–641.
19. Gagneux P, Boesch C, Woodruff DS. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Mol Ecol* 6:861–868.
20. Vigilant L. 1999. An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. *Biol Chem* 380:1329–1331.
21. Nürnberg P, Saueremann U, Kayser M, Lanfer C, Manz E, Widdig A, Berard J, Bercovitch FB, Kessler M, Schmidtke J, Krawczak M. 1998. Paternity assessment in rhesus macaques (*Macaca mulatta*): multilocus DNA fingerprinting and PCR marker typing. *Am J Primatol* 11:1–18.
22. Morin PA, Mahboubi P, Wedel S, Rogers J. 1998. Rapid screening and comparison of human microsatellite markers in baboons: allele size is conserved, but allele number is not. *Genomics* 53:12–20.
23. Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.
24. Navidi W, Arnheim N, Waterman MS. 1992. A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *Am J Hum Genet* 50:347–359.