Developing Methods for Access to High Quality Genome Sequences from Wild Ape Populations

Dylan Koundakjian, William J. Nielsen, & Dr. Greg Wray
Duke University Biology Department

Abstract

Modern evolutionary study of wild ape populations requires the collection of genomic DNA from individuals living in their natural habitat. Our aim is to explore routes of noninvasive DNA collection for use in these studies. A method for isolating genomic DNA from chimpanzee (*Pan troglodytes*) hair follicle cells is developed and tested for limit of detection. Validation of the method is then established through the determination of the frequency of polymorphisms due to genomic amplification error by comparing sequences obtained from three identically handled samples. From this, a basis is formed for the use of this method for collecting genomic DNA from great apes and other mammals.

Background

- Investigators in ecology and evolutionary biology have known the importance of studying wild ape populations in their natural habitats since the seminal work of Jane Goodall in 1960.
- The collection of DNA samples from wild ape populations was essential to Beatrice Hahn and colleagues' discovery of the origin of HIV and AIDS.
- Modern studies have made major strides in elucidating the evolutionary history of chimpanzees and other apes using genomic information from wild subjects.
- Conservation efforts require that DNA be collected in a noninvasive manner due to the critically endangered status of great apes.

*Figure 1: A common chimpanzee (*Pan troglodytes*) and a chimpanzee's nest*

- The traditional method of noninvasive DNA collection is extraction from feces. This method is compared to collection from other sources.

Table 1: Comparison of DNA Extraction Sources

	Advantages	Disadvantages
Feces	Established method, extraction kits commercially available	Fragmented DNA and contamination from sequences of gut microbiota
Urine	Collection method already developed, little microbial contamination	No method currently developed for DNA extraction
Hair	Smaller collection volume, longer DNA reads, easy handling	Methods still in development, transport permits unavailable

Methods

Limit of Detection

- DNA was extracted from hair follicles of chimpanzees "Lance" and "Ebi" using a custom protocol for the Dneasy® Blood and Tissue Kit (QIAGEN).
  - 6 extractions were made using a decreasing number of hairs: "Lance" (30, 13, 7 hairs), "Ebi" (13, 8, 5 hairs).
- Extracted DNA was amplified using a REPLI-g Whole Genome Amplification (WGA) Kit (QIAGEN).
- Genomic DNA was run through PCR amplification using two primers, D4 and D18, developed in-house.
- Samples were visualized on a 1% agarose gel to determine quality of DNA amplification for each extraction.

Method Validation

- Three WGA's were created simultaneously from a single DNA extraction from "Lance."
- Each WGA was amplified by PCR using both primers D4 and D18.
- PCR fragments were purified using a QIAquick® PCR purification kit (QIAGEN) and cloned into competent *E. coli* cells using the PCR®4 Blunt TOPO® vector (Invitrogen).
- Plasmid DNA was isolated from broth cultures of transformed *E. coli* using a Wizard® Plus SV Miniprep kit (Promega).
- Purified plasmid DNA was run through a Big Dye v3.1 sequencing reaction and submitted to Eton Bioscience Inc. for sequence visualization.
- Sequences were analyzed for error rates to determine consistency of amplification for each extraction.

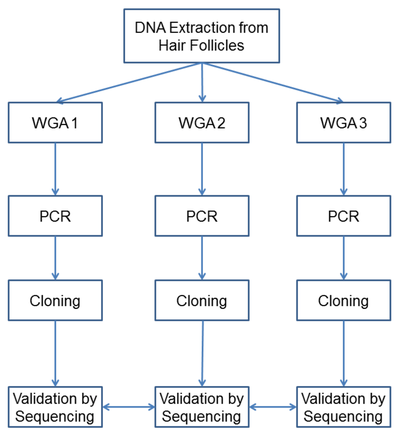

Figure 2: Validation Protocol

Validation of this method is established using the frequency of polymorphisms due to amplification error. Frequencies are calculated by comparison between replicates from the same WGA and between replicates from different WGAs.

Results

Limit of Detection

- DNA fragments were observed at expected band sizes: 270 bp for D4 primer, 290 bp for D18 primer.

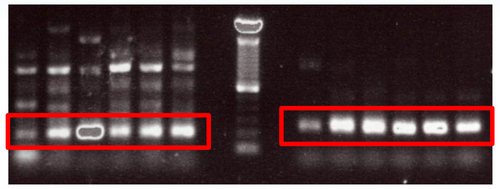

Figure 3: Gel image showing amplification of targets at expected band sizes. Sample names indicate the source individual and number of hairs used in the extraction

- Shows high-quality amplification of the same gene product across multiple extractions, using a range of 30 hairs down to 5 hairs.

Method Validation

- Sequences obtained were not high enough quality to align replicates, both within and between WGAs. Without proper alignment, sequences could not be compared for polymorphism frequency.

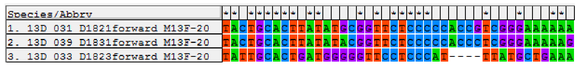

Figure 4: Highest quality alignment of three experimental sequences is only 21/42 bases (50%)

- A hypothetical sequence set based on experimental data can be used to illustrate the different types of polymorphisms expected to be found during validation.

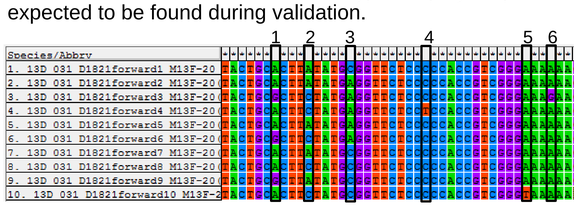

Figure 5: Ten hypothetical replicates from the same WGA showing different categories of polymorphism. Sequence 1 is taken directly from experimental data. Validation requires comparison between replicate sets from all three WGAs.

- *Polymorphism 1 occurs at 30%. This would be classified as due to error early in amplification only if it is not present in the other WGAs.*
- *Polymorphisms 2 and 3 occur at 50%. These would be classified as normal allelic variation if this percentage is also present in the other WGAs.*
- *Polymorphisms 4, 5, and 6 occur at 10%. These would be classified as isolated errors from late in amplification.*
- *The relative frequencies of polymorphisms type 1 and type 4-6 would be used to validate the consistency of this method.*

Conclusions

- DNA extraction from hair follicles is a robust process capable of returning high quality amplification products.
- This high quality amplification can be accomplished using a very small sample volume, down to 5 hairs.
- DNA extracted using this method can return replicable sequence data.
- Sequencing must be re-performed to obtain data of higher quality than that found in this study before the method can be validated.

Implications

- A method of DNA extraction from hair follicles can be further developed for field collection of genomic information from wild apes and other mammals.
- Further development and use of this method would confer several advantages over traditional noninvasive DNA collection methods, including smaller collection volumes and lower contamination levels.
- The next step in investigating access to high quality genome sequences from wild populations will be to develop a similar method for DNA extraction from urine samples.
- Both of these approaches should be incorporated into the suite of available DNA collection methods to allow maximum adaptability to sample availability in the environment.

References

1. Bjork, A. *et al.* (2011). "Evolutionary History of Chimpanzees Inferred from Complete Mitochondrial Genomes". *Mol. Biol. Evol.* 28(1): 615–623.
2. Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior.* Cambridge, Mass.: Belknap Press of Harvard University Press.
3. Hahn, B. and Sharp, P. (2011). "Origins of HIV and the AIDS Pandemic". *Cold Spring Harb Perspect Med*. 1(1): a006841.
4. Knott, C.D. (1997) "Field Collection and Preservation of Urine in Orangutans and Chimpanzees". *Tropical Biodiversity* 4(1): 95-102.
5. Morin, P.A. and D.S. Woodruff. (1992) "Paternity exclusion using multiple hypervariable microsatellite loci amplified from nuclear DNA of hair cells". In: *Paternity in Primates: Genetic Tests and Theories*. Martin, R.D., A.F. Dixon and E.J. Wickings, eds. Basel, Karger, pp. 63-81.
6. http://upload.wikimedia.org/wikipedia/commons/6/6d/Pan_troglodytes_(male).jpg
7. http://2.bp.blogspot.com/-OjvshAC_YmQ/TtLYbz4UddI/AAAAAAAABQk/b4MAeeVb8xo/s1600/11-5-2011+Kibale+National+Park%252C+Uganda+%252815%2529.JPG

Acknowledgements

We'd like to thank Dr. Ian Gilby and the North Carolina Zoological Park for the donation of hair samples. We'd also like to thank Eton Biosciences Inc. for their help in sequencing our validation samples.