# Chapter 14

## Case Study: Using a Nondestructive DNA Extraction Method to Generate mtDNA Sequences from Historical Chimpanzee Specimens*

**Elmira Mohandesan, Stefan Prost, and Michael Hofreiter**

### Abstract

A major challenge for ancient DNA (aDNA) studies using museum specimens is that sampling procedures usually involve at least the partial destruction of each specimen used, such as the removal of skin, pieces of bone, or a tooth. Recently, a nondestructive DNA extraction method was developed for the extraction of amplifiable DNA fragments from museum specimens without appreciable damage to the specimen. Here, we examine the utility of this method by attempting DNA extractions from historic (older than 70 years) chimpanzee specimens. Using this method, we PCR-amplified part of the mitochondrial HVR-I region from 65% (56/86) of the specimens from which we attempted DNA extraction. However, we found a high incidence of multiple sequences in individual samples, suggesting substantial cross-contamination among samples, most likely originating from storage and handling in the museums. Consequently, reproducible sequences could be reconstructed from only 79% (44/56) of the successfully extracted samples, even after multiple extractions and amplifications. This resulted in an overall success rate of just over half (44/86 of samples, or 51% success), from which 39 distinct HVR-I haplotypes were recovered. We found a high incidence of C to T changes, arguing for both low concentrations of and substantial damage to the endogenous DNA. This chapter highlights both the potential and the limitations of nondestructive DNA extraction from museum specimens.

**Key words:** Ancient DNA, Chimpanzees, DNA damage, Genetic diversity, Mitochondrial DNA (mtDNA), Museum collections, Non-destructive DNA extraction, Phylogeography, Population extinction

---

* *Note*: In the case study presented in this chapter, we describe DNA extraction and amplification of mitochondrial DNA from historic chimpanzee samples from museum collections using a method similar to that presented in Chapter 13. We discuss specific challenges associated with nondestructive DNA extraction, including contamination and DNA damage.

Beth Shapiro and Michael Hofreiter (eds.), *Ancient DNA: Methods and Protocols*, Methods in Molecular Biology, vol. 840,
DOI 10.1007/978-1-61779-516-9_14, © Springer Science+Business Media, LLC 2012

## 1. Introduction

Museum specimens represent one of the major sources of ancient DNA. Museum collections are valuable because they often contain rare or extinct species as well as large numbers of conspecific specimens that can be used to reveal the biological history of species and populations. Methods for DNA extraction from bones, teeth, and skin are well established (1, 2). However, for almost all of these, a piece of tooth, bone, or skin has to be removed and dissolved prior to DNA extraction.

To circumvent this limitation, a nondestructive DNA extraction method has been developed, with a reported success rate of 90% for bones up to 164 years old (3). The protocol, described in detail in Chapter 13, involves soaking the sample in GuSCN buffer and subsequently processing the buffer. Because it does not require the removal of a large piece of the specimen, this method prevents significant damage to the specimen, leaving it intact for future analyses. In addition, if necessary, the DNA extraction can be repeated 3–5 times without significant damage occurring to the specimen (3).

Here, we apply this nondestructive DNA extraction method to a large number of museum-preserved chimpanzee specimens. We discuss the success rate of this method, problems that arise during the procedure, and phylogenetic analyses performed subsequent to extraction and sequencing.

Common chimpanzees (*Pan troglodytes*) are traditionally divided into three populations or subspecies based on geographic barriers (mostly rivers): west African *P. t. verus* (4), central African *P.t. troglodytes*, and east African *P. t. schweinfurthii* (5, 6). Additional sampling in northern Cameroon/southern Nigeria has led to the designation of a fourth chimpanzee subspecies, *P. t. vellerosus* (7–11), although the phylogenetic distinctiveness and therefore the validity of this fourth chimpanzee subspecies is still debated (12). A recent analysis of about 300 microsatellites demonstrated convincingly that low levels of gene flow are occurring among the three traditionally accepted chimpanzee subspecies (12). However, due to a lack of captive individuals of *P. t. vellerosus*, the status of this potential subspecies has yet to be ascertained (12). Because chimpanzee populations have declined severely during the last decades (13–15), accessing genetic material from historic chimpanzee specimens should allow a better understanding of the geographical distribution and the population history of chimpanzees.

## 2. Materials and Methods

### 2.1. Sample Preparation

We used two rooms during the experiment so that sample preparation could be kept separate from contamination-susceptible steps including buffer preparation and PCR setup. In the second room, we carried out buffer preparation and setup of PCR reagent mix in one fume hood, and DNA extraction and the addition of DNA extract to the PCR in a second hood. In order to prevent modern DNA from potentially contaminating the experiments, we washed all working surfaces with 10–13% sodium hypochlorite solution (bleach) prior to DNA extraction. Both rooms were designated for ancient DNA work, and were spatially separated from all laboratories in which work on modern DNA was performed. The ancient DNA clean rooms were further isolated from any other area by an ante-room, which was used for decontaminating consumables and changing clothes.

We collected teeth from 86 chimpanzee (*Pan troglodytes*) individuals originating from different geographical locations in Africa and that are currently held in different museum collections. The final data set comprised specimens from 35 eastern, 20 central, two western and one western/central (Nigeria-Cameroon) locations.

### 2.2. DNA Extraction and Amplification

Prior to extraction, we prepared TE buffer, extraction solution, binding buffer, washing buffer, and silica suspension as described in Chapter 13. We designed two overlapping primer pairs (A and B; see Table 1) using Primer 3 version 0.4.0 (http://frodo.wi.mit.edu/primer3/). The primers were synthesized in 100 μM stock concentration and stored at −20°C. For use in PCR, we diluted the primers to 10 μM concentration with HPLC-grade water and stored them at −20°C.

**Table 1**
**Primer designed for amplifying the investigated D-loop region of chimpanzee mtDNA**

| | Primer sequence 5–3′ | Product size |
|---|---|---|
| *Primer pair A* | | |
| Outer sense2 (OS2) | 5′-CGC TAT GTA TTT CGT ACA TTA CT-3′ | 210 bp |
| Inner antisense3 (IAS3) | 5′-RTA GGT TTG TTG ATA TYR G-3′ | |
| *Primer pair B* | | |
| Inner sense3 (IS3) | 5′-TCA ACT CTC AAC TRT CRM ACA TA-3′ | 130 bp |
| Outer antisense2 (OAS2) | 5′-GAT TTG ACT GTA ATG TGC TAT G-3′ | |

For extraction, we first cleaned the surface of each specimen using a tissue moistened with HPLC-grade water. Removing dirt from the surface of the samples reduces the amount of substances that might inhibit the DNA extraction and/or the following enzymatic manipulations of DNA extract such as PCR.

We then soaked the samples in 5 mL extraction solution (L6 buffer) and incubated them at room temperature in the dark with constant slow rotation. After 5–7 days, we removed the buffer and rinsed the sample with HPLC-grade water. We dried the samples at room temperature in preparation for return to the museums from which they were obtained.

To continue with the DNA extraction, we transferred the buffer into a new 15-mL centrifuge tube. We added 50–100 μL of silica suspension (after vortexing the silica suspension to be certain that it was adequately mixed) and incubated the mixture for 1–3 h at room temperature with rotation. We then centrifuged the buffer at 1,800 × *g* for 2 min and either discarded the supernatant or stored it at 4°C for later use. Next, we washed the silica pellet with 1 mL L2 buffer by pipetting up and down. We transferred the resuspended mixture to a 2-mL Eppendorf tube. This transfer makes handling more convenient, as 2-mL tubes rather than 15-mL tubes can be used in all the following steps. We pelleted the silica via centrifugation for 5 s at 16,000 × *g*, discarded the supernatant, and carefully removed any remaining liquid using a 200-μL pipette. If the binding solution (L2 buffer) is not completely removed in this step, the salt concentration in the elution buffer will be too high, thus preventing the DNA from being completely released from the silica during elution.

We then washed the pellet with 1 mL washing buffer by pipetting up and down. We centrifuged the resuspended mixture for 10 s at 16,000 × *g*. We discarded the supernatant and removed the remaining liquid again carefully with a pipette. We dried the pellets at 56°C for 5 min or approximately 15 min at room temperature with open lids. We then added 100 μL TE (1×) to the pellet, incubated the mixture for 8 min at 65°C, and resuspended the pellet by stirring with the pipette tip and pipetting up and down. Finally, we centrifuged the eluate at 16,000 × *g* for 1 min and transferred the supernatant into a new 2-mL Eppendorf tube, being careful not to leave any trace of silica. For some specimens, second and third extractions starting at the incubation step were subsequently performed (see Subheading 3).

We used the obtained extracts to generate an approximately 225 bp fragment of the HVR-I region of chimpanzee mtDNA by PCR amplifying two overlapping fragments of 210 and 130 bp, respectively, using primer pairs A (OS2/IAS3) and B (IS3/OAS2; see Table 1). PCR was carried out in 20 μL volumes containing 1× PCR buffer (Applied Biosystems), 4 mM $MgCl_2$ (Applied Biosystems), 1 mg/mL BSA (Invitrogen), 0.5 mM mixed dNTPs

(in equal concentrations; Amersham Biosciences), 0.25 μM of each primer (MWG-Biotech AG), 0.5–1 U of Taq Gold DNA polymerase (Applied Biosystems), and 5 μL DNA template (irrespective of DNA concentration). The initial denaturation step (94°C for 4 min) was followed by 60 cycles of denaturation at 93°C for 20 s, binding of primers at 51°C (primer pair A) and 53°C (primer pair B) for 30 s and strand replication at 72°C for 30 s, followed by a final extension at 72°C for 10 min. The PCR products were subjected to electrophoresis in 1.5% agarose, stained with ethidium bromide (50 ng/mL) and visualized over UV light. We included one negative control for every seven PCR reactions. Each fragment was amplified twice for each specimen.

We purified PCR products of the expected length with the QIAquick Gel Extraction Kit (QIAGEN, Germany), and cloned them using the TOPO TA® Cloning Kit (Invitrogen, The Netherlands) according to the manufacturer's instruction. We sequenced the insert sequences for eight clones per sample on an ABI 3700 capillary sequencer after colony PCR and purification on a QIAGEN BioRobot 9600.

### 2.3. Phylogenetic Analysis

We aligned the nucleotide sequences from the HVR-I regions sequenced from 56 chimpanzees in BioEdit version 7.0 (16) using CLUSTAL-W software. We checked the authenticity of obtained DNA sequences using BlastSearch (National Center for Biotechnology Information) (17) and reconstructed the phylogenetic relationship between the recovered sequences as well as extant chimpanzee sequences obtained from GenBank by constructing a serial network (18). The serial network was created using the open-source R script TempNet (available at www.stanford.edu/group/hadlylab/tempnet/). TempNet uses statistical parsimony to illustrate within-species relationships through time.

## 3. Results and Discussion

Using the silica-based nondestructive method, we successfully amplified and sequenced mtDNA sequences from 65% (56 of 86) of the chimpanzee specimens that were stored in different museums. Of these, 53 samples (95%) yielded both PCR products, while the remaining three samples (5%) could only be partially amplified.

All recovered sequences showed between 98 and 100% BLAST similarity to chimpanzee mtDNA sequences archived in GenBank. Analysis of consensus and clone sequences generated from two independent PCRs revealed identical sequences for 29 museum specimens (apart from C to T changes in individual clones, which are almost certainly due to DNA damage; see below) and multiple sequence variants within the remaining 26 (one sample could only

be amplified once and was excluded from further analyses). Thus, just over half of the samples yielded identical sequences across multiple PCRs, although for six of the samples yielding additional sequences, these occurred at such a low frequency that a likely endogenous sequence could be inferred. This overall result most likely indicates that cross-contamination occurred between museums specimens, especially since the sequence variants recovered sometimes belong to different chimpanzee subspecies. To investigate this further, we performed additional nondestructive extractions on 16 of the specimens that had yielded ambiguous sequences. This additional experiment was motivated by the realization that the first extraction may recover not only endogenous DNA but also any potential surface contaminant DNA, including cross-contamination that may have occurred as researchers handled multiple specimens. Additional extractions performed after the first extraction should therefore be less likely to recover surface contaminants.

We performed second and in some cases third DNA extractions from 16 of the samples with variant sequences. Each extraction yielded less amplifiable DNA than the previous extraction, as judged by the number of failing PCRs and the strength of the product when amplifications were successful. However, the amount of DNA contamination was also reduced to some extent, and a likely endogenous DNA sequence could eventually be deduced for 9 of these 16 samples, while the remaining seven samples could not be resolved. Thus, in total we were able to recover reproducible sequences from 44 samples, resulting in a total of 39 distinct haplotypes.

This result is in stark contrast to previous experience with this protocol when no evidence for contamination was observed (3, 19, 20). However, while it should be noted that two of these previous studies were performed on small mammal specimens, where both storage conditions and, due to the fragile nature of the specimens, extraction kinetics might be different, the initial study introducing this method used both chimpanzee and hyena teeth. It is not clear why the results of this study differ so much from those of previous studies. One potential cause may lie in differences in museum storage and handling conditions that might have facilitated cross-contamination among the samples used in this study, but it is impossible to ascertain this possibility. Another fact worth mentioning is a high incidence of C to T changes, indicative of DNA damage (21) in our results. Thus, of the 29 samples that yielded unambiguous sequences, 26 showed C to T changes in individual clones. This observation suggests not only high DNA damage but also low DNA concentrations in these samples, making them more susceptible to contamination. Independent of the eventual cause for the high contamination rate on the samples used, our results show that studies on museum specimens face similar problems as those using fossil DNA, at least when using this extraction method. Therefore, similar precautions such as multiple

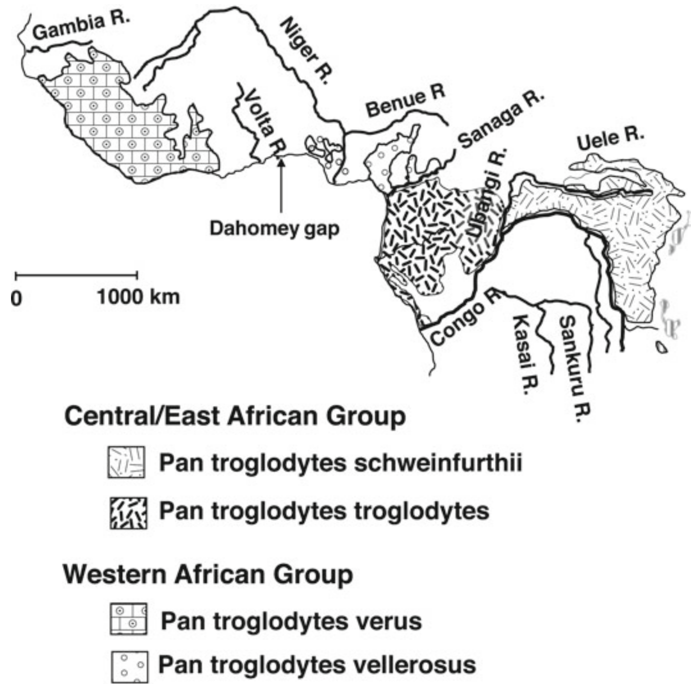


Fig. 1. Geographical distribution of chimpanzee subspecies.

extractions and amplifications as well as obtaining multiple clonal sequences are an absolute requirement in such studies.

Chimpanzee subspecies are divided into two geographically and genetically defined groups: a central/eastern African group (*P. t. schweinfurthii* and *P. t. troglodytes*) and a western African group (*P. t. verus* and *P. t. vellerosus*) with a significant phylogeographic break at the Sanaga River in central Cameroon (Fig. 1). A temporal network (18) reconstructed from our historical sequences and modern chimpanzee sequences obtained from GenBank shows that all historical haplotypes are closely related to modern ones (Fig. 2), although some of them have not (yet) been found in the extant gene pool.

With 51%, the DNA extraction success rate in this study is lower than in previous studies reporting the method (3, 19, 20), but still sufficiently high to obtain DNA from about half of the investigated specimens. Similarly, the length of the obtained PCR products is large enough to obtain, by using several overlapping fragments, DNA sequences sufficiently long for phylogeographic and phylogenetic analyses. However, the high incidence of contaminating sequences found also indicates that a substantial failure rate has to be taken into account when planning a study, although there seem to be large differences among collections and species, probably depending on storage and handling.

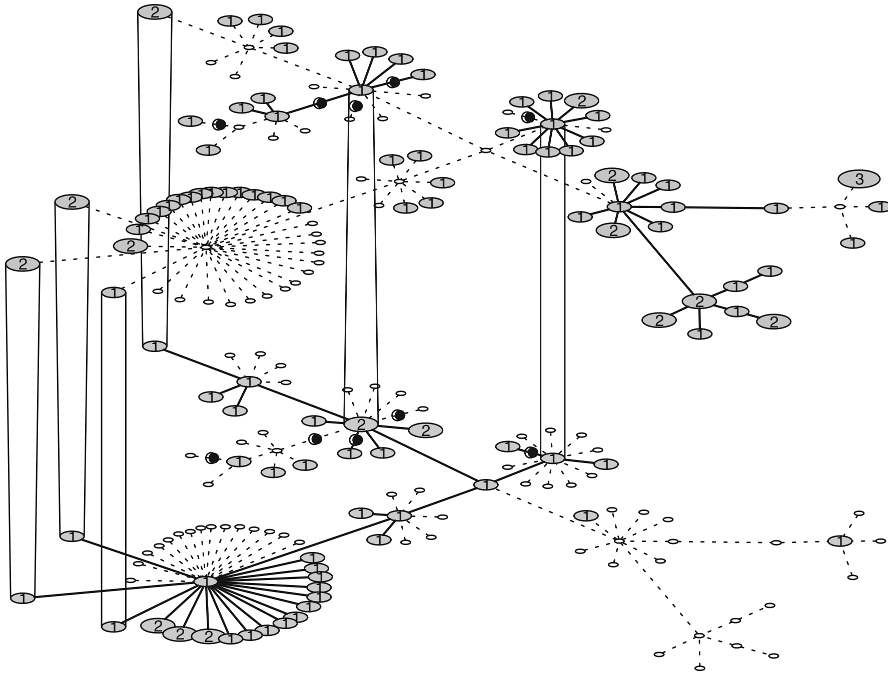

Fig. 2. Temporal statistical parsimony network of modern and ancient chimp sequences. The upper layer comprises modern-day sequences obtained from GenBank, whereas the lower layer consists of ancient DNA samples generated in this study. Haplotypes sampled in a given time layer are represented as *gray ellipses*. Those present in the overall network, but not in the individual time layer are shown as *small white ellipses*. Haplotypes shared between the two layers are connected by *vertical lines*. Haplotypes present in a time-horizon are connected by *solid lines*, whereas lines connecting at least one unsampled haplotype for this time-horizon are *dotted*. Those separated by more than one mutation are indicated by one *small black circle* for each additional mutation. Please note that for graphical reasons, not all modern sequences available were used in the network. Therefore, a larger proportion of museum sequences than shown in this figure are actually still present in the modern population.

Both success rate and total length of the DNA sequences that can be obtained should increase considerably when using DNA hybridization capture methods (22–25) rather than PCR for targeting specific DNA regions. These methods have recently been used successfully for targeting both mitochondrial (up to complete mtDNA genomes (25, 26)) and nuclear DNA (27). Due their ability to target very short DNA fragments, they are ideally suited for the analysis of fragmented DNA such as that recovered from museum specimens. It needs to be noted, though, that measures used to distinguish endogenous ancient DNA obtained from Pleistocene specimens from contaminating modern DNA such as fragmentation or nucleotide substitution patterns may not be applicable to museum specimen DNA for several reasons. First, due to their younger age, museum specimens may not have accumulated DNA damage to the extent that fossil DNA dating to the Pleistocene has. Second, and perhaps even more importantly, the sequences contaminating museum specimens probably originate

quite frequently from cross-contamination with DNA from other museum specimens, which is likely to display highly similar damage patterns. However, as our results show, this problem can be addressed at least partially by performing two consecutive extractions and by preferential use of the second extract.

## Acknowledgments

We thank the lab members of the Research Group Molecular Ecology at the Max Planck Institute for Evolutionary Anthropology, and especially Tim Heupink, for their assistance in laboratory work, museum curators for providing us with chimpanzee samples, and the Max Planck Society for financial support.

### References

1. Krings M, Stone A, Schmitz RW et al (1997) Neandertal DNA sequences and the origin of modern humans. Cell 90:19–30
2. Hadly EA, Kohn MH, Leonard JA et al (1998) A genetic record of population isolation in pocket gophers during Holocene climatic change. Proc Natl Acad Sci USA 95:6893–6896
3. Rohland N, Siedel H, Hofreiter M (2004) Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. Biotechniques 36(814–816):818–821
4. Schwarz E (1934) On the local races of chimpanzees. Ann Mag Nat Hist Lond 13: 576–583
5. Hill WCO (1969) In: Bourne GH (ed) The chimpanzee; a series of volumes on the chimpanzee, vol. 1. S. Karger AG, Basel, NY, pp. 22–49
6. Groves CP (2001) Primate taxonomy. Smithsonian Institution Press, Washington, DC, 350p
7. Gonder MK, Oates JF, Disotell TR et al (1997) A new West African chimpanzee subspecies? Nature 388:337
8. Gonder MK, Disotell TR, Oates JF (2006) New genetic evidence on the evolution of chimpanzee populations and implications for taxonomy. Int J Primatol 27:1103–1127
9. Gonder MK (2000) Evolutionary genetics of chimpanzees (*Pan troglodytes*) in Nigeria and Cameroon. Ph.D. Dissertation, City University of New York, New York, 338pp
10. Gonder MK, Disotell TR (2006) Contrasting phylogeographic histories of chimpanzees in Nigeria and Cameroon: a multilocus analysis. In: Lehman S, Fleagle J (eds) Primate biogeography. Springer, New York, pp 129–161
11. Gonder MK, Disotell T, Oates JF (2006) New genetic evidence on the evolution of chimpanzee populations and implications for taxonomy. Int J Primatol 27:1103–1127
12. Becquet C, Patterson N, Stone AC et al (2007) Genetic structure of chimpanzee populations. PLoS Genet 3:e66
13. Campbell G, Kuehl H, N'Goran KP et al (2008) Alarming decline of West African chimpanzees in Côte d'Ivoire. Curr Biol 18:903–904
14. Walsh PD, Abernethy KA, Bermejo M et al (2003) Catastrophic ape decline in western equatorial Africa. Nature 422:611–614
15. Greengrass E (2009) Chimpanzees are close to extinction in Southwest Nigeria. Prim Cons 24:77–83
16. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98
17. Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs (Review). Nucleic Acids Res 25: 3389–3402
18. Prost S, Anderson CNK (2011) TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. Methods Ecol Evol 2:663–667. doi:10.1111/j.2041-210X.2011.00129.x

19. Asher RJ, Hofreiter M (2006) Tenrec phylogeny and the noninvasive extraction of nuclear DNA. Syst Biol 55(2):181–194
20. Asher RJ, Maree S, Bronner G, Bennett NC, Bloomer P, Czechowski P, Meyer M, Hofreiter M (2010) A phylogenetic estimate for golden moles (Mammalia, Afrotheria, Chrysochloridae). BMC Evol Biol 10:69
21. Hofreiter M, Jaenicke V, Serre D, Haeseler Av A, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res 29(23):4793–4799
22. Hodges E, Xuan Z, Balija V et al (2007) Genome-wide in situ exon capture for selective resequencing. Nat Genet 39:1522–1527
23. Hodges E, Rooks M, Xuan Z et al (2009) Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nat Protoc 4:960–974
24. Gnirke A, Melnikov A, Maguire J et al (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol 27:182–189
25. Briggs AW, Good JM, Green RE et al (2009) Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325: 318–321
26. Krause J, Briggs AW, Kircher M et al (2010) A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol 20:231–236
27. Burbano HA, Hodges E, Green RE et al (2010) Targeted investigation of the Neandertal genome by array-based sequence capture. Science 328:723–725