Ancient DNA protocols

Tibor Török

Abstract

Ancient DNA work was performed in the specialized ancient DNA (aDNA) facilities of the Department of Genetics, University of Szeged, Hungary with strict clean-room conditions. In order to authenticate the results, we considered the latest recommendations of (Llamas et al. 2017) throughout of the experiments.

DNA extraction:

Note: This protocol is based on (Rohland and Hofreiter 2007), supplemented with partial predigestion (Damgaard et al. 2015) and using a GuHCl Binding Buffer modified after (Gamba et al. 2016).

- 100 mg bone powder from tooth root, petrous bone or other dense bone was predigested in 1 ml 0,5 M EDTA 100 µg/ml Proteinase K for 30 minutes at 48 °C, to increase the proportion of endogenous DNA (Damgaard et al. 2015),
- Then DNA solubilisation was done overnight at 48 °C in 1 ml extraction buffer containing 0.45 M EDTA, 250 µg/ml Proteinase K, 1% Triton X-100, and 50 mM DTT.
- DNA was bound to silica (Rohland and Hofreiter 2007) adding 6 ml Binding Buffer (5,83 M GuHCl, 105 mM NaOAc, 46,8% isopropanol, 0,06% Tween-20) and 150 µl silica suspension to the 1 ml extract, and the pH was adjusted between 4-6 with HCl.
- After 3 hours binding at room temperature silica was pelleted, and washed twice with 80% ethanol, then DNA was eluted in 100 µl TE buffer.

NGS library construction

Note: We used the double stranded library protocol of (Meyer and Kircher 2010) with double indexing (Kircher, Sawyer, and Meyer 2012), except that all purifications were done with MinElute columns . We also applied partial UDG treatment of (Rohland et al. 2015), but decreased the recommended USER and UGI concentrations to half (0.03 U/µL) and at the same time increased the incubation time from 30 to 40 minutes. This modification removed uracils with comparable efficiency to the original method.
DNA free negative control libraries were also made to detect possible contamination during handling or present in materials.

Partial UDG treatment:

reagent	volume (µL) per sample
Buffer Tango (10X)	6
dNTPs (25 mM each)	0,24
ATP (100 mM)	0,6
USER (1 U/µL NEB)	1,8
DNA extract	51,36
Total:	60 µL

The reaction was incubated at 37°C for 40 minutes in PCR machine, with 40 °C lid temerature. Then 1,8 µL UGI (Uracil Glycosylase Inhibitor, 2U/µL NEB) was added to the reaction, which was further incubated at 37°C for 40 minutes.

Blunt-End Repair:

- To each reaction we added 3 µL T4 polynucleotide kinase (10 U/µL) and 1,2 µL T4 DNA polymerase-t (5 U/µL), then incubated in PCR machine at 25°C for 15 minutes, followed by incubation at 12°C for 5 minutes and cooling to 4°C.
- 350 ul MinElute PB buffer (QIAGEN) was added to the reaction, then it was purified on MinElute columns. DNA was eluted in 20 ul EB prewarmed to 55 °C.

Adapter ligation:

reagent	volume (µL) per sample
T4 DNA ligase buffer (10X)	4
PEG-4000 (50%)	4
adapter mix (50 µM each)	0,5
T4 DNA ligase (5 U /µL)	1
H_2O	10,5
DNA	20 ul
Total:	40ul

- We incubated the reaction at 22°C for 30 minutes in PCR machine.
- Then 200 µL PB Buffer was added followed by MinElute purification. DNA was eluted in 20 ul EB.

Adapter Fill-In:

Note: We assembled the reaction below without library DNA, and the elution step from above was centrifuged directly on the fill-in reaction mix.

reagent	volume (µL) per sample
H_2O	14.1
ThermoPol reaction buffer (10X)	4
dNTPs (25 mM each)	0.4
Bst polymerase, large fragment (8 U/µL)	1.5
Library DNA (from adapter ligation)	20
Total:	40ul

- The reaction was incubated at 37°C for 20 minutes, then 200 ul PB was added followed by MinElute purification. Library was eluted in 20 ul EB.

Library preamplification:

Libraries were preamplified in 2 x 50 µl reactions containing 800 nM each of IS7 and IS8 primers, 200 µM dNTP mix, 2 mM $MgCl_2$, 0,02 U/µl GoTaq G2 Hot Start Polymerase (Promega) and 1X GoTaq buffer,

followed by MinElute purification. PCR conditions were 96 °C 6 min, 11 cycles of 94 °C 30 sec, 58 °C 30 sec, 72 °C 30 sec, followed by a final extension of 64 °C 10 min. Libraries were eluted from the column in 50 µl 55 °C EB buffer (Qiagen), and concentration was measured with Qubit (Termo Fisher Scientific). Libraries below 5 ng/µl concentration were reamplified in the same reaction for additional 5-12 cycles, depending on concentration, in order to obtain 50 µl preamplified library with a concentration between 10-50 ng/µl.

Double indexing:

50 ng preamplified libraries were double indexed according to (Kircher, Sawyer, and Meyer 2012) in a 50 µl PCR reaction containing 1 x KAPA HiFi HotStart ReadyMix (Kapa Biosystems) and 1000 nM each of P5 and P7 indexing primers. PCR conditions were 98 °C 3 min, 6 cycles of 98 °C 20 sec, 66 °C 10 sec, 72 °C 15 sec followed by a final extension of 72 °C 30sec. Indexed libraries were MinElute purified and their concentration was measured with Qubit, and size distribution was checked on Agilent 2200 TapeStation Genomic DNA ScreenTape.

Note: Lately we use 30 ng preamplified libraries and just 5 PCR cycles to avoid overamplification.

Mitochondrial DNA capture and sequencing:

Biotinilated mtDNA baits were prepared from three overlapping long-range PCR products as described in (Maricic, Whitten, and Pääbo 2010), but using the following primer pairs, L14759-H06378, L10870-H14799, L06363-H10888, described in (Haak et al. 2010).

Capture was done according to (Maricic, Whitten, and Pääbo 2010) with the following modifications: Just four blocking oligos, given below were used in 3 µM (each) final concentration:
BO1.P5.part1F: AATGATACGGCGACCACCGAGATCTACAC-Phosphate,
BO2.P5.part2F ACACTCTTTCCCTACACGACGCTCTTCCGATCT-Phosphate,
BO4.P7.part1 R GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-Phosphate,
BO6.P7.part2 R CAAGCAGAAGACGGCATACGAGAT-Phosphate.

- For one capture 300 ng biotinilated bait was used with 30 µl Dynabeads MyOne Streptavidin C1 magnetic beads (Thermo Fisher Scientific).
- Double indexed libraries of 20 samples (300 ng each) were mixed and concentrated on MinElute columns, then captured together in a 64 µl hybridization reaction. When fewer samples were enriched, we used proportionally smaller amounts of baits.
- After washing, bead-bound enriched libraries were resuspended in 20 µl water and released from the beads in a 60 µl PCR reaction containing 1 X KAPA HiFi HotStart ReadyMix and 2000 nM each of IS5- IS6 library primers. PCR conditions were: 98 °C 1 min, 10 cycles of 98 °C 20 sec, 60 °C 30 sec, 72 °C 30 sec, followed by a final extension of 72 °C 30 sec. The captured and amplified library mix was purified on MinElute column and eluted in 15 µl EB.
- Before sequencing, libraries were quantified with Qubit, and quality checked and Agilent 2200 TapeStation Genomic DNA ScreenTape. Sequencing was done at the SeqOmics Biotechnology Ltd., using MiSeq sequencer with MiSeq Reagent Kit v3 (Illumina, MS-102-3003) generating 2x150bp paired-end sequences.

Note: Lately we perform low coverage shotgun sequencing prior to enrichment to estimate endogenous DNA content. Then 5-7 libraries with similar endogenous content are enriched together.

Data analysis

The adapters of paired-end reads were trimmed with the cutadapt software (Martin 2011) in paired end mode. Read quality was assessed with FastQC (S. Andrews 2016). Sequences shorter than 25 nucleotide were removed from this dataset. The resulting analysis-ready reads were mapped to the GRCh37.75 human genome reference sequence using the Burrows Wheeler Aligner (BWA) v0.7.9 software (Li and Durbin 2009) with the BWA mem algorithm in paired mode and default parameters. Aligning to the GRCh37.75 human reference genome that also contains the mtDNA revised Cambridge Reference Sequence (rCRS, NC_012920.1) (R. M. Andrews et al. 1999) helped to avoid the forced false alignment of homologous nuclear mitochondrial sequences (NumtS) to rCRS, though the proportion of NumtS, derived from low copy nuclear genome, is expexted to be orders of magnitudes lower than mtDNA in aDNA libraries. Samtools v1.1 (Li et al. 2009) was used for sorting and indexing BAM files. PCR duplicates were removed with Picard Tools v 1.113 (Broad Institute 2016). Ancient DNA damage patterns were assessed using MapDamage 2.0 (Jónsson et al. 2013), and read quality scores were modified with the rescale option to account for post-mortem damage. Freebayes v1.02 (Garrison and Marth 2012) was used to identify variants and generate variant call format (VCF) files with the parameters -q 10 (exclude nucleotids with <10 phred quality) and -P 0.5 (exclude very low probability variants). Each variant call was also inspected manually. From VCF files FASTA format was generated with the Genom Analysis Tool Kit (GATK v3.5) FastaAlternateReferenceMaker walker (McKenna et al. 2010).

References:

Andrews, R M et al. 1999. “Reanalysis and Revision of the Cambridge Reference Sequence for Human Mitochondrial DNA.” *Nature genetics* 23(2): 147. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&>

Andrews, S. 2016. “FastQC: A Quality Control Tool for High Throughput Sequence Data.” *babraham bioinformatics*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Broad Institute. 2016. “Picard Tools.” *https://broadinstitute.github.io/picard/*. http://broadinstitute.github.io/picard/.

Damgaard, Peter B et al. 2015. “Improving Access to Endogenous DNA in Ancient Bones and Teeth.” *Scientific Reports* 5: 11184. http://dx.doi.org/10.1038/srep11184%5Cn10.1038/srep11184%5Cnhttp://www.nature.com/articles/srep11184#supplementary-information.

Gamba, Cristina et al. 2016. “Comparing the Performance of Three Ancient DNA Extraction Methods for High-Throughput Sequencing.” *Molecular Ecology Resources* 16(2): 459–69.

Garrison, Erik, and Gabor Marth. 2012. “Haplotype-Based Variant Detection from Short-Read Sequencing.” *arXiv preprint arXiv:1207.3907*: 9. http://arxiv.org/abs/1207.3907.

Haak, Wolfgang et al. 2010. “Ancient DNA from European Early Neolithic Farmers Reveals Their near Eastern Affinities.” *PLoS Biology* 8(11).

Jónsson, Hákon et al. 2013. “MapDamage2.0: Fast Approximate Bayesian Estimates of Ancient DNA

Damage Parameters." In *Bioinformatics*, , 1682–84.
Kircher, Martin, Susanna Sawyer, and Matthias Meyer. 2012. "Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform." *Nucleic Acids Research* 40(1).
Li, Heng et al. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics (Oxford, England)* 25(16): 2078–79. http://bioinformatics.oxfordjournals.org/cgi/content/long/25/16/2078 (August 20, 2016).
Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics (Oxford, England)* 25(14): 1754–60. http://bioinformatics.oxfordjournals.org/cgi/content/long/25/14/1754 (August 22, 2016).
Llamas, Bastien et al. 2017. "From the Field to the Laboratory: Controlling DNA Contamination in Human Ancient DNA Research in the High-Throughput Sequencing Era." *STAR: Science & Technology of Archaeological Research* 3(1): 1–14. http://dx.doi.org/10.1080/20548923.2016.1258824.
Maricic, T, M Whitten, and S Pääbo. 2010. "Multiplexed DNA Sequence Capture of Mitochondrial Genomes Using PCR Products." *PLoS ONE* 5: e14004. http://dx.doi.org/10.1371/journal.pone.0014004.
Martin, Marcel. 2011. "Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads." *EMBnet.journal* 17(1): 10–12.
McKenna, Aaron et al. 2010. "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data." *Genome Research* 20(9): 1297–1303.
Meyer, Matthias, and Martin Kircher. 2010. "Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing." *Cold Spring Harbor Protocols* 5(6).
Rohland, Nadin et al. 2015. "Partial Uracil-DNA-Glycosylase Treatment for Screening of Ancient DNA." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 370(1660): 20130624. http://www.ncbi.nlm.nih.gov/pubmed/25487342.
Rohland, Nadin, and Michael Hofreiter. 2007. "Ancient DNA Extraction from Bones and Teeth." *Nature Protocols* 2(7): 1756–62. http://www.nature.com/nprot/journal/v2/n7/full/nprot.2007.247.html%5Cnhttp://www.nature.com/nprot/journal/v2/n7/pdf/nprot.2007.247.pdf.

Citation: Tibor Török Ancient DNA protocols. **protocols.io** dx.doi.org/10.17504/protocols.io.rmvd466
Published: 12 Jul 2018

Protocol