Benchmarks

5′-Tailed sequencing primers improve sequencing quality of PCR products

Jonas Binladen, M. Thomas P. Gilbert, Paula F. Campos, and Eske Willerslev

University of Copenhagen, Copenhagen, Denmark

The degraded nature of DNA in many subfossil, archival, and forensic specimens often prevents the PCR amplification of fragments that are >100–200 bp in length (1–4). Unfortunately, conventional dye-labeled Sanger sequencing platforms perform poorly on short DNA fragments such as these when standard sequencing methodologies are applied. Specifically, these methods produce poor quality electropherograms at the 5′ end of the sequences, probably due to the irregular behavior and poor separation of the shortest DNA fragments during subsequent electrophoresis. Two methods have commonly been used to overcome this problem: (*i*) molecular cloning of the PCR products prior to sequencing, enabling sequencing primers located in the flanking vector DNA to increase the length and quality of the sequenced fragment and (*ii*) sequencing both strands in opposite directions, which produces two complementary sequences with good quality dual coverage in the middle and single coverage at the extremes. An alternative approach is the application of new sequencing technologies, such as pyrosequencing, that are tailored for use on small DNA fragments (5,6). However, the use of such methods requires investment in expensive new equipment, which is not always a practical solution for research groups that already have access to conventional gel and capillary sequencing equipment. To address this problem, we introduce a simple technique that in most cases halves the number of sequencing errors in recovered sequences from short PCR products using conventional capillary electrophoresis sequencing equipment through the addition of a nonspecific nucleotide tail to the 5′ end of the sequencing primers.

DNA was extracted from two bone specimens of the extinct Pleistocene woolly rhinoceros (*Coelodonta antiqutatis*) and from three bone specimens of Pleistocene musk oxen (*Ovibos* sp.) (sample details in Supplementary Table S1, available online at www.BioTechniques.com) using conventional silica-based methods (7,8). PCR amplification was performed using conventional primers (see Supplementary Table S2 for primer details) in 25-µL reactions containing 1× PCR High Fidelity PCR buffer, 2.5 mM magnesium sulfate solution (Invitrogen, Carlsbad, CA, USA), 0.4 mM dNTP mix, 1 U Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen), 1 µM each primer,

and 1–5 µL DNA extract. Cycling was performed in an Eppendorf® Mastercycler® gradient thermal cycler (Eppendorf Nordic, Horsholm, Denmark) using denaturation at 94°C for 2 min, followed by 45–50 cycles of 94°C for 30 s, 50°–55°C for 30 s, and 68°–72°C for 30 s, and a final cycle for 7–10 min at 68°–72°C. The PCR products were purified prior to DNA sequencing using the Invisorb® Vacuum Manifold and Invisorb PCR HTS 96 kit (both from Invitek GmbH, Berlin, Germany). To prevent contamination, sample preparation, DNA extractions, and PCR setup were carried out in a dedicated ancient DNA facility physically isolated from other biological laboratories (including post-PCR facilities) with positive air pressure, daily exposure of surfaces to ultraviolet (UV)-irradiation, and where full body suits, face masks, and disposable gloves are worn. Extraction and PCR blank controls were incorporated at ratios of 1:8 and 1:1, respectively. No amplification products were observed in the controls.

Two comparative sets of primers were used for the DNA sequencing. In the first instance, sequencing was
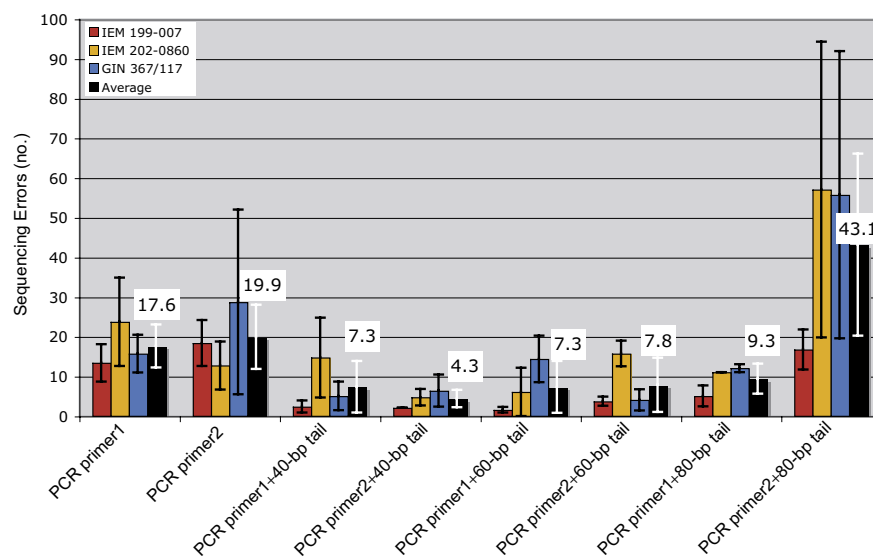
Figure 1. Number of sequencing errors in a 104-bp region of the D-loop of three musk ox (*Ovibos* sp.) samples (IEM 199-007, IEM 202-0860, and GIN 367/117). PCR primers 1 + 2 (original names: L1 and H1) were designed by MacPhee et al. (11). Black columns and numbers are the average number of sequencing errors for each sequencing primer for the three samples. Each sample was successfully sequenced 2–3 times with each primer. Error bars indicate standard deviation. Exact data points, position, and type of sequencing error can be found in the supplementary materials (Supplementary Figures S3–S5). Data for two woolly rhino samples showing the same trend can also be found in the supplementary materials (Supplementary Figures S1, S6, and S7).

undertaken using the regular primers used in the initial PCR amplifications, ranging from 18 to 27 bp in length (Supplementary Table S2). In comparison, we tested modified sequencing primers that were identical at their 3′ ends to the regular primers, but which contained a polynucleotide tail of between 40–80 bp at their 5′ ends (Supplementary Table S2). The tail sequence has previously been described (9,10) and consists of a 40-bp neutral DNA sequence that is not complementary to any published sequence, plus a poly(C) string where extra length was required. Cycle sequencing was performed on the PCR products from a single reaction as recommended by the manufacturer using the BigDye® Terminator kit (Applied Biosystems, Foster City, CA, USA) using each of the sequencing primers relevant to the PCR product. The sequencing products were analyzed by capillary electrophoresis on an ABI Prism® 3130 Genetic Analyzer (Applied Biosystems).

In order to quantify the performance of the different primers, the trace files generated during each sequence run were aligned to the known consensus sequence for the specimens, and the numbers of sequencing errors were quantified (Figure 1 and supplementary materials). The quantification was performed automatically using Sequencher™ 3.1.1 (Gene Codes, Ann Arbor, MI, USA) and counted all sequencing errors (e.g., misreads, missing, and additional nucleotides) to the consensus sequence. No manipulation or selection of sequences was done prior to comparison with the Sequencher program. Each sequencing experiment was repeated between two and four times for each specimen, resulting in a minimum of eight comparisons per primer set.

Our results indicate that the addition of both a 40- and 60-bp tail onto the original sequencing primer enhances the quality of the sequence, 1.2–4.6 times (Figure 1 and supplementary materials) (40 bp, $P \leq 0.01$; 60 bp, $P < 0.01$; paired Student's t-test and Wilcoxon signed rank test, respectively). In most cases (7 out of 8), the products obtained with the tailed primers were found to have less than half the number of sequencing errors

than those obtained by the nontailed primers. In contrast, a comparison of the longest (80-bp tail) primers with the original primers shows no increase in sequencing quality ($P = 0.31/0.84$ paired Student's t-test/Wilcoxon signed rank test). There is no statistical support of a difference in quality between the sequences generated using the 40- and 60-bp tails ($P = 0.85$ paired Student's t-test and Wilcoxon signed rank test). In conclusion, we suggest that the addition of a 40-bp tail is sufficient to greatly increase the sequence quality of short (<150 bp) PCR products and recommend its use in future studies that require the sequencing of short PCR fragments. The method described here is general and should improve the sequencing of the 5′ end of all sizes of PCR products (including those longer than 150 bp).

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

ACKNOWLEDGMENTS

REFERENCES

1. **Hofreiter, M., D. Serre, H.N. Poinar, M. Kuch, and S. Pääbo.** 2001. Ancient DNA. Nat. Rev. Genet. 2:353-359.
2. **Handt, O., M. Richards, M. Trommsdorf, C. Kilger, J. Simanainen, O. Georgiev, K. Bauer, A. Stone, et al.** 1994. Molecular genetic analyses of the Tyrolean Ice Man. Science 264:1775-1778.
3. **Willerslev, E. and A. Cooper.** 2005. Ancient DNA. Proc Biol Sci. 272:3-16.
4. **Pääbo, S., H. Poinar, D. Serre, V. Jaenicke-Despres, J. Hebler, N. Rohland, M. Kuch, J. Krause, et al.** 2004. Genetic analyses from ancient DNA. Annu. Rev. Genet. 38:645-679.
5. **Ronaghi, M., S. Karamohamed, B. Pettersson, M. Uhlen, and P. Nyren.** 1996. Real-time DNA sequencing using detection of pyrophosphate release. Anal. Biochem. 242:84-89.
6. **Ronaghi, M., M. Uhlen, and P. Nyren.** 1998. A sequencing method based on real-time pyrophosphate. Science 281:363-365.
7. **Hoss, M. and S. Pääbo.** 1993. DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Res. 21:3913-3914.
8. **Yang, D.Y., B. Eng, J.S. Waye, J.C. Dudar, and S.R. Saunders.** 1998. Improved DNA extraction from ancient bones using silica-based spin columns. Am. J. Phys. Anthropol. 105:539-543.
9. **Lindblad-Toh, K., E. Winchester, M.J. Daly, D.G. Wang, J.N. Hirschhorn, J.P. Laviolette, K. Ardlie, D.E. Reich, et al.** 2000. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. Nat. Genet. 24:381-386.
10. **Sanchez, J.J., C. Borsting, C. Hallenberg, A. Buchard, A. Hernandez, and N. Morling.** 2003. Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. Forensic Sci. Int. 137:74-84.
11. **MacPhee, R.D.E., A.N. Tikhonov, D. Mol, and A.D. Greenwood.** 2005. Late quaternary range collapse and loss of genetic diversity in muskox (*Ovibos*). BMC Evol. Biol. 5:49.

To purchase reprints of this article, contact: Reprints@BioTechniques.com