

Method

A Modified Enrichment Method to Construct Microsatellite Library from Plateau Pika Genome (*Ochotona curzoniae*)

Jianing Geng^{1#}, Kexin Li^{2,3#}, Yanming Zhang^{2*}, and Songnian Hu^{1*}

¹CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China;

²Northwest Plateau Institute of Biology, Chinese Academy of Sciences, Xining 810001, China;

³Graduate University of Chinese Academy of Sciences, Beijing 100049, China.

Genomics Proteomics Bioinformatics 2010 Mar; 8(1): 72-76. DOI: 10.1016/S1672-0229(10)60007-1

Abstract

A microsatellite-enriched library of plateau pika (*Ochotona curzoniae*) was constructed according to the strong affinity between biotin and streptavidin. Firstly, genomic DNA was fragmented by ultrasonication, which is a major improvement over traditional methods. Linker-ligated DNA fragments were hybridized with biotinylated microsatellite probes, and then were subjected to streptavidin-coated magnetic beads. PCR amplification was performed to obtain double-stranded DNA fragments containing microsatellites. Ligation and transformation were carried out by using the pGEM-T Vector System I and *Escherichia coli* DH10B competent cells. Sequencing results showed that 80.2% of clones contained microsatellite repeat motif. Several modifications make this protocol time-efficient and technically easier than the traditional ones; particularly, composition and relative abundance of microsatellite repeats in plateau pika genome were truly represented through the optimized PCR conditions. This method has also been successfully applied to construct microsatellite-enriched genomic libraries of Chinese hamster (*Cricetulus griseus*) and small abalone [*Haliotis diversicolor* (Reeve)] with high rates of positive clones, demonstrating its feasibility and stability.

Key words: microsatellite, enriched library, magnetic bead, ultrasonication

Introduction

Microsatellites are tandem repeats of 1-6 nucleotides distributed in the nuclear genome of most taxa, also known as simple sequence repeats (SSRs), variable number tandem repeats (VNTRs) and short tandem repeats (STRs). The average length of microsatellite locus typically ranges from 5 to 40 repeats, but longer threads of repeats are also possible. Unlike microsa-

tellite sequences, the flanking regions surrounding the microsatellite locus are usually conserved across individuals of the same species, so microsatellite locus can often be identified by its flanking sequences. Dinucleotide, trinucleotide and tetranucleotide repeats are the most common microsatellites. Due to their high polymorphism, high abundance and co-dominant inheritance, microsatellites are often used as DNA markers for molecular genetic studies (1). To date, microsatellites have been successfully applied in many fields such as population structure analysis, forensic DNA study, linkage map construction, genetic monitoring, strain identification and conserva-

Equal contribution.

* Corresponding authors.

E-mail: zhangym@nwipb.ac.cn; husn@big.ac.cn

© 2010 Beijing Institute of Genomics. All rights reserved.

tion/management of biological resources.

The first protocol for isolation of *de novo* microsatellite markers was described by Rassmann *et al* (2). In this method, SSR-containing clones were picked up through colony hybridization with repeat-containing probes (2). Although it was effective to deal with taxa whose genomes contain high frequency of microsatellites, it turned out to be laborious and costly for species with low microsatellite frequency. A variety of enrichment protocols have been reported in the literature (3-9). These methods use enzyme-digested fragmentation and multi-step PCR amplification to enrich the microsatellite repeats, but they are still very time-consuming due to those additional and repetitive enrichment steps. Here we report a modified method for the construction of plateau pika (*Ochotona curzoniae*) microsatellite jumping clone library by using ultrasonication-induced fragmentation of DNA and single-step PCR amplification. The result shows that this method is very efficient and rapid for the isolation of microsatellite-containing DNA sequences.

Materials and Methods

DNA extraction and oligo synthesis

Plateau pika samples were captured in Dawu town, Guoluo autonomous prefecture, Qinghai Province in June 2006 (E100.3°, N34.5°, altitude 3,800 m). Genomic DNA was extracted from muscle tissues of four individuals with standard protocols (10). The SNX forward oligonucleotide (5'-CTAAGGCCTTGCTAG CAGAAGC-3') and the SNX reverse oligonucleotide (5'-pGCTTCTGCTAGCAAGGCCTTAGAAAA-3'), devised by Smithsonian Institute's Molecular Genetics Laboratory (7), and fourteen types of biotin-probes, which covered all dinucleotide and trinucleotide repeat motifs (Table 1), were synthesized by commercial company (Sangon Co., Ltd.).

Genomic DNA fragmentation and size selection

Genomic DNA was sheared with ultrasonicator into fragments ranging in size from 500 to 1,000 bp. The fragments were end-blunted by T4 DNA polymerase

(Promega) at 37°C for 1 h and precipitated with 1/10 volume of 3 M sodium acetate and 2 volumes of cold absolute ethanol. The re-suspended fragments were loaded in 1% agarose gel and run for overnight at 4°C. The 500 to 1,000 bp fragments were recovered from agarose gel and purified with a QIAquick gel column (Qiagen). Finally, the DNA was eluted in 50 µL of ddH₂O.

Table 1 Fourteen types of dinucleotide/trinucleotide biotinylated probes for constructing enrichment library

No.	Probe sequence	Temperature (°C)
1	5'-biotin-ATAGAATAT(AT) ₁₂ -3'	48.39
2	5'-biotin-ATAGAATAT(AAT) ₈ -3'	48.39
3	5'-biotin-ATAGAATAT(AAG) ₈ -3'	58.33
4	5'-biotin-ATAGAATAT(AAC) ₈ -3'	58.33
5	5'-biotin-ATAGAATAT(ATC) ₈ -3'	58.33
6	5'-biotin-ATAGAATAT(ACT) ₈ -3'	58.33
7	5'-biotin-ATAGAATAT(AC) ₁₂ -3'	63.30
8	5'-biotin-ATAGAATAT(AG) ₁₂ -3'	63.30
9	5'-biotin-ATAGAATAT(AGC) ₈ -3'	68.27
10	5'-biotin-ATAGAATAT(AGG) ₈ -3'	68.27
11	5'-biotin-ATAGAATAT(ACC) ₈ -3'	68.27
12	5'-biotin-ATAGAATAT(ACG) ₈ -3'	68.27
13	5'-biotin-ATAGAATAT(GCC) ₈ -3'	78.21
14	5'-biotin-ATAGAATAT(GC) ₁₂ -3'	78.21

Ligation of genomic DNA to SNX linker

The fragmented genomic DNA was ligated to SNX linker with conditions as follows: 240 ng of genomic DNA, 1.5 µL linker (50 µM/µL) and 3 units of T4 DNA ligase (Promega) in 10 µL volume at 14°C for 16 h. Five reactions were carried out separately and then the resulting ligation products were mixed together in order to get enough SNX-ligated DNA complex. The success of linker ligation reaction was confirmed by PCR using SNX forward oligonucleotide as primer.

Biotinylated probe hybridization and two-round capture of microsatellite-containing DNA fragments

First-round capture

Hybridizations were carried out at different tempera-

tures, which were about 5°C below temperature of probes (Table 2). Totally, five reactions were required for all dinucleotide and trinucleotide microsatellite motifs. For each reaction, 10 µg of SNX linker-ligated DNA was denatured by heating at 98°C for 10 min, and then hybridized with 2 pmol of each biotin-labeled probe in 6× SSC (20× SSC: 3 M NaCl, 0.3 M sodium citrate) with overnight incubation at specific temperature (Table 2) of total volume of 100 µL. An amount of 20 µL (4 mg/mL) of magnetic beads (Dynabeads M-280, Dynal Biotech, Norway) were re-suspended and washed four times with 150 µL of BW buffer (10 mM Tris, pH=7.5, 1.0 mM EDTA, 1.0 M NaCl). Moreover, 150 µL of hybridization buffer (6× SSC, 0.1% SDS) and 100 µL of probe-DNA mixture were gently agitated at 43°C for 3 h in gentle shaking bed. After removing supernatant from tubes placed in magnetic field, the bead-probe-DNA complex was washed twice with 200 µL of 2× SSC, 0.1% SDS at room temperature, followed by twice wash with 200 µL of 1× SSC, 0.1% SDS, one at 45°C and another at 60°C. The captured DNA was eluted with 40 µL of pure water by denaturing the beads at 90°C for 10 min, then placed in magnet quickly and the supernatant was pipetted to a new tube. At this stage, the products released from the probe-bead complex were single-stranded microsatellite-containing fragments.

Table 2 Hybridization temperatures for dinucleotide/trinucleotide repeat motifs

Probe	T _m (°C)	First Hyb T _m (°C)	Second Hyb T _m (°C)
(AG) ₁₂	63.30	60	62
(AC) ₁₂	63.30	60	62
(AAC) ₈	58.33	53	55
(AAG) ₈	58.33	53	55
(ATC) ₈	58.33	53	55
(ACT) ₈	58.33	53	55
(ACC) ₈	68.27	65	67
(AGG) ₈	68.27	65	67
(AGC) ₈	68.27	65	67
(ACG) ₈	68.27	65	67
(GC) ₁₂	78.21	70	72
(GCC) ₈	78.21	70	72
(AT) ₁₂	48.39	40	43
(AAT) ₈	48.39	40	43

Second-round capture

The first-round affinity-captured DNA was hybridized with biotinylated repeat oligonucleotide and captured by magnetic beads as the first round, but hybridization temperature was increased by 2°C than first-round capture.

PCR amplification and repeat-enriched DNA purification

After two rounds of hybridization, the captured fragments were amplified by PCR in a Perkin-Elmer thermal cycler 2720. The reactions were carried out in a volume of 50 µL mixture with 35 µL of single-stranded captured DNA, 20 pm of SNX forward primer, 5 units of AmpliTaq DNA polymerase (Promega), 4 µL of 2.5 mM dNTPs, 5 µL of 10× PCR buffer with the program of 94°C for 5 min, 25 cycles of 94°C for 30 s, 62°C for 30 s, 72°C for 45 s, and 72°C extension for 10 min. Then the PCR products were separated by 1% agarose gel (LE, Analytical Grade, Promega) and further purified with QIAquick gel extraction kit (Qiagen).

Ligation, *E. coli* competent cell transformation and positive clone selection

The ligation reaction was performed as follows: 1 µL of 10× T4 DNA ligase buffer (Promega), 0.5 µL of pGEM-T vector (20 ng/µL) (Promega), about 50 ng of PCR products and 3 units of T4 DNA ligase (Promega) were mixed and incubated overnight at 4°C water bath. The ligation products were electro-transformed into *Escherichia coli* DH10B competent cells (Invitrogen) and were spreaded on Luria-Bertani (LB) agar plates including ampicillin, IPTG and X-gal (10). White colonies were picked out and cultivated on LB liquid culture medium. The extracted plasmids were sequenced on 3730 DNA Analyzer (Applied Biosystems, Inc.).

Results

The majority of DNA fragments generated by ultrasonication were concentrated in the size ranging from 500 to 1,000 bp, which is a desired length distribution

for constructing the microsatellite-enriched library. After two rounds of hybridization and affinity capture, single-stranded DNA containing the selected microsatellite DNA was used as a template for PCR amplification to obtain double-stranded targeted fragments. A successful PCR reaction is presented in **Figure 1**, showing that the size of major band ranges from 500 to 1,000 bp.

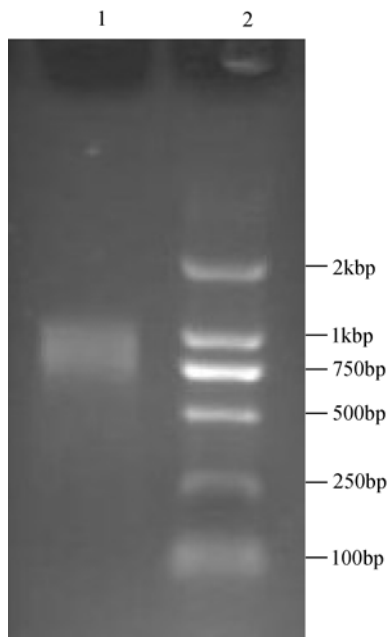


Figure 1 Hybridization-enriched DNA band. The microsatellite-containing DNA fragments were captured by magnetic beads, and single-stranded DNA was amplified in a single-primer PCR. PCR products (Lane 1) were checked up on 1% agarose gel with a DL2000 (Lane 2) marker as standard.

The sequencing results showed that 80.2% (231/288 clones) of the high-quality reads (Phred20 quality; >99% accuracy) contained microsatellite repeat motifs, representing various microsatellite loci on plateau pika genome. Characterization of these loci revealed that perfect repeat sequences constitute the major (67.5% of total) microsatellite class, followed by imperfect (28.5%) and compound repeat sequences (4%). Of the 231 microsatellite repeat clones, AC/GT, AG/CT, A/T, AAC/GTT, ACC/GGT, AGG/CCT and AAG/CCTT repeats represented 64.6%, 22.5%, 3.3%, 2.1%, 1.25%, 0.83% and 0.83% of the total microsatellite repeats, respectively. The percentage of the rest microsatellite repeat motifs (C/G, CG/GC, AAG/CTT, ACG/CTG, ACT/ATG, AGC/CGT, CCG/

CGG, ACTC/AGTG, AGAT/ATCT, AGGG/CCCT and ACACCC/GGGTGT) is 0.42% in total. Additionally, 120 pairs of primers were designed with primer 3.0 software and 13 polymorphic microsatellite markers (GenBank accession No. EU518184-EU518196) suitable for further population structure analysis were characterized (11).

Discussion

By conventional methods, the fragmentation of genomic DNA was often managed by restriction enzyme digestion. When DNA is fragmented with restriction enzymes, the average length of fragments depends largely on genomic G+C composition and endonuclease recognition sites, causing genomic region sampling bias. Hamilton *et al* attempted to use multiple restriction enzymes to overcome this limitation, but resulted in a smaller average size of fragments (7). Ultrasonication, as a method of mechanical breakage of genomic DNA, is an effective way for shearing DNA. The length of DNA fragments produced by ultrasonication is less dependent on the genomic DNA composition, therefore it ensures the random distribution of DNA fragments (5, 8). Besides, ultrasonication also shows several advantages over enzyme digestion. Firstly, the length of DNA fragments can be managed by changing sonication time and power. Secondly, the requirements for DNA quality, reaction temperature and ion concentration are not stringent. Thirdly, the pre-experimental works for ultrasonication parameter adjustment is easier than enzyme selection and enzyme reaction optimization. With this ultrasonication sheared protocol (**Figure 2**), candidate microsatellite locus identification could be accomplished as soon as one week.

Multi-step PCR amplifications are performed for DNA fragment recovery and enrichment in traditional methods, resulting in changes of the distribution of DNA fragments across the genome. Alternatively, we used two rounds of capture to get enough DNA fragments for the double-stranded DNA recovery (Figure 2). It could completely avoid over-amplification and keep the average size of enriched fragments in good consistence with unprocessed genomic fragments.

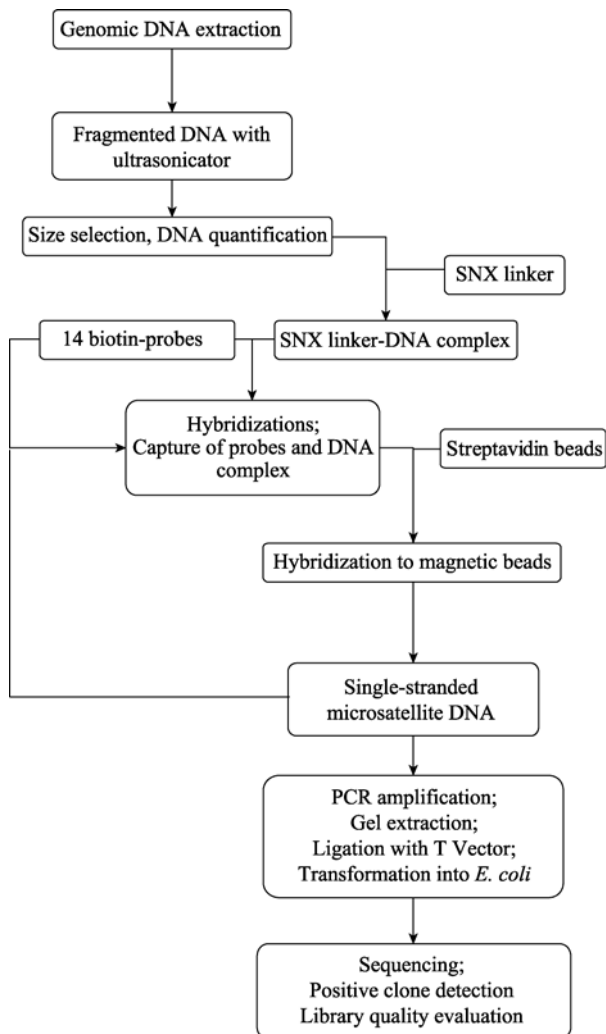


Figure 2 Experimental flow for microsatellite-enrichment library construction.

We also adopted pGEM-T Vector System (Promega) to clone the PCR amplicons, which compensated for the inefficiency of blunt-end ligation and tendency of cohesive end ligation. The positive clone rate was about 80%, which is similar to the traditional enriched methods (3). Subsequently, two microsatellite libraries for Chinese hamster (*Cricetulus griseus*) and small abalone [*Haliotis diversicolor* (Reeve)] (12) were successfully constructed with this method and the positive clone rate was up to 75% and 84%, respectively.

In summary, the modified microsatellite-enrichment method reported here is more time-efficient and technically easier than other methods. The independence of species and the original microsatellite frequency in the genome makes this method suitable for microsatellite library construction in a large variety of taxa.

Authors' contributions

SH and YZ conceived and designed the experiments. JG and KL carried out the experiments and drafted the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors have declared that no competing interests exist.

References

- Selkoe, K.A. and Toonen, R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* 9: 615-629.
- Rassmann, K., et al. 1991. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis* 12: 113-118.
- Zane, L., et al. 2002. Strategies for microsatellite isolation: a review. *Mol. Ecol.* 11: 1-16.
- Fischer, D. and Bachmann, K. 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *Biotechniques* 24: 796-800, 802.
- Karagyozov, L., et al. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acids Res.* 21: 3911-3912.
- Ostrander, E.A., et al. 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proc. Natl. Acad. Sci. USA* 89: 3419-3423.
- Hamilton, M.B., et al. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques* 27: 500-502, 504-507.
- Kandpal, R.P., et al. 1994. Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proc. Natl. Acad. Sci. USA* 91: 88-92.
- Brown, S.M., et al. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.* 93: 190-198.
- Sambrook, J. and Russell, D.W. 2001. *Molecular Cloning: A Laboratory Manual* (third edition). Cold Spring Harbor Laboratory Press, New York, USA.
- Li, K.X., et al. 2009. Isolation and characterization of 13 microsatellite loci in the plateau pika (*Ochotona curzoniae*). *Conserv. Genet.* 10: 785-787.
- Zhan, X., et al. 2009. Isolation and characterization of eleven microsatellite loci in small abalone, *Haliotis diversicolor*. *Conserv. Genet.* 10: 1185-1187.