Microsatellites Obtained Using Strand Extension: An Enrichment Protocol

BioTechniques 26:690-697 (April 1999)

ABSTRACT

A new method is described to enrich genomic libraries for clones containing microsatellite repeats. The method involves selection on completed M13 genomic libraries rather than on genomic DNA before library construction. It uses two reactions, in which microsatellite oligonucleotides prime strand extension. The first reaction uses a biotinylated primer allowing vectors with microsatellite-containing inserts to be selected with streptavidin-coated magnetic beads. This reaction may be dependent on the strand displacement activity of the Klenow fragment of DNA Polymerase I. The second strand extension reaction is included to improve the relative transformation efficiency of microsatellite-containing clones. In control experiments starting with 0.7% microsatellite-containing clones, enrichment averaged 99.5%. The method was tested empirically on an estusinus and abalone genomic libraries in which enrichment for (CA)n microsatellites was efficient enough that clones could be sequenced without further screening. This protocol is technically straightforward and permits the isolation of a large number of microsatellite markers in less time than is required to execute traditional protocols involving rounds of filter hybridization.

INTRODUCTION

Microsatellite sequences have been used to study subjects ranging from DNA polymerase function (11) to problems in behavioral and population ecology (5,6). A major technical limitation for application of microsatellites markers to studies of natural populations is that new markers must be developed for each species, or at least each group of closely related species, under study (12). Furthermore, there is a growing appreciation that a large number of markers are required to achieve the resolving power that is needed in many ecological and evolutionary applications (e.g., Reference 4).

The traditional method for isolating microsatellite clones is to create a small-insert, partial genomic library in a plasmid or phage vector and then screen clones by repeated rounds of filter hybridization using an oligonucleotide microsatellite probe (e.g., Reference 14). Enrichment protocols have also been developed to increase the proportion of clones in a given library containing the microsatellite motif of interest. These enrichment protocols typically involve polymerase chain reaction (PCR) amplification of inserts before ligation into a vector (7) and add time and complexity to the cloning procedure (see References 3 and 8 for recent refinements).

This paper describes a new microsatellite enrichment protocol in which the initial library construction is done as in traditional (non-enriching) protocols, and the enrichment involves selection of clones rather than selection of genomic DNA fragments before ligation (Figure 1). The selection process involves two steps in which microsatellite oligonucleotides are sequentially annealed to and extended on clones containing the appropriate microsatellite sequence. After the first strand extension reaction, in which the oligonucleotide used must be biotinylated, clones are selected using streptavidin-coated magnetic beads. Single-stranded (ss) DNA is then eluted from the bound molecules, and the extension reaction is repeated. The second strand extension serves the function of increasing the efficiency with which microsatellite-containing clones are transformed (a principle that has previously been shown to enrich libraries for microsatellites; Reference 10). The procedure takes 1.5 days from the point where it departs from traditional protocols to the point where the enriched library is plated.

MATERIALS AND METHODS

Single-Stranded Library Preparation

Small-insert, partial genomic libraries are constructed in M13 using established methods. In the method I used, for example, genomic DNA (5 µg) is digested with three restriction enzymes (2 U Alu, 13.5 U HaeIII, 10 U Rsal; New England Biolabs, Beverly, MA, USA) and size-fractionated on an agarose gel. DNA is purified from a gel slice corresponding to fragments of 350–550 bp. Ligation is done overnight at 20°C in a volume of 20 µL and contains 40 ng of genomic DNA and 100 ng of double-stranded M13mp18 DNA that has been digested with HincII (New England Biolabs) and treated with calf intestinal alkaline phosphatase (Promega, Madison, WI, USA). The ligation mixture is then ethanol-precipitated in the presence of tRNA carrier (Sigma, Castle Hill, NSW, Australia) and resuspended in
12.5 µL of water. The ligation is titrated with a test transformation and then transformed by electroporation in six aliquots, each using 80 µL of electroporation-competent Epicerian Coli® XL1-Blue Cells (Stratagene, La Jolla, CA, USA). These aliquots are combined, plated on three large LB plates (ca. 10,000 clear plaques/plate) and grown at 37°C for 7 h.

At this point the protocol diverges from traditional methods in which DNA from M13 plaques would be transferred to filter membranes. Instead, M13 phage particles are eluted from the plates by adding 7 mL of LB broth to each plate and incubating at 20°C for 1 h. Single-stranded M13 DNA is then isolated from the broth by polyethylene glycol (PEG) precipitation and organic extraction (13), with an expected yield of several hundred micrograms of DNA.

For empirical testing of the method, such ssDNA libraries were prepared for the agile antechinus (Antechinus agilis), a recently described (2) "marsupial mouse", and for the tropical abalone (Haliotis asinina), a commercially important shellfish. For protocol development and optimization, ssDNA was isolated from a single clone containing an uninterrupted (CA)20 sequence and from unmodified M13mp18. These DNAs were mixed at a ratio of up to 1:149, allowing enrichment to be monitored by counting the proportion of white plaques when the final products were plated in the presence of isopropyl-β-D-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). The following methods are the result of that optimization process.

### Enrichment

Table 1 shows temperature profiles, reaction volumes and dNTP and primer concentrations for all reactions. In the first reaction, microsatellite-containing clones are labeled with biotin by annealing and extending a 5′-biotinylated oligonucleotide primer [1 used (AC)18, (ATGG)4 and AG(ATA)4]. The reaction buffer is 1× E. coli DNA Polymerase I/Klenow Buffer (New England Biolabs) [10 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 7.5 mM dithiothreitol (DTT)] supplemented with 50 mM NaCl. The reaction contains 300 ng of single-stranded library DNA. It is placed in a heat block prewarmed to 90°C and carried to a 37°C room where condensation is spun down in a microcentrifuge. One unit of Klenow Polymerase (Promega), diluted in reaction buffer to a volume of 1 µL, is then added. After 3 h, 1 µL of 500 mM EDTA is added before removing the reaction from the 37°C room.

While the extension reaction is in progress, 5 µL of Dynabeads® M-280 Streptavidin (Dynal, Oslo, Norway) are prepared using a magnetic particle concentrator. This involves washing once with 50 µL Wash Buffer [2 M NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA] and once with 25 µL kilobase-BINDER™ Binding Buffer (Dynal) before resuspending in 20 µL binding buffer. An equal volume of the extension reaction is then added to the prepared beads, and the mixture is placed on a rotating wheel for 3 h at 20°C. Following this incubation, the beads are washed three times in 40 µL Wash Buffer and once in 20 µL of Elution Buffer [10 mM Tris-HCl (pH 9.0), 100 mM EDTA] before ssDNA is eluted at 85°C for 2 min in a 20 µL volume and then immediately separated from the beads using the magnetic particle concentrator.

The microsatellite-containing ssDNA eluted from the Dynabeads is made double-stranded (ds) by a second round of primer extension. The reaction buffer (Expand™ 1) and polymerase mixture (Taq/Pwo) are from the Expand Long Template PCR Kit (Boehringer Mannheim GmbH, Mannheim, Germany). The reaction mixture contains 0.5 µL (1.75 U) of polymerase mixture. Before the completed reaction is allowed to cool below 68°C, 1 µL of 500 mM EDTA is added.

### Testing and Sequencing

After the second extension reaction, an aliquot (0.5–1 µL) is transformed and grown overnight. Ethanol precipitation with tRNA carrier can be used to concentrate the sample and improve transformation efficiency, but this is not normally required. Well-separated plaques are placed in 1 mL of LB and incubated at 20°C for 1 h. Fifty microliters of this phage suspension and 50 µL of an overnight bacterial culture are then added to 800 µL LB in a 1.5-mL microcentrifuge tube and incubated with shaking for 6 h at 37°C. Cells are
pelleted by centrifugation, and the supernatant is transferred to a fresh tube. One microliter of this phage stock is used as template in a PCR to test for the presence of long, uninterrupted microsatellite sequences. In the screens for (CA)$_k$ microsatellites, amplification is with a (GT)$_k$ oligonucleotide and an M13 forward primer (5′-GGCCAGGGTTTCCAGTCACCATGAC-3′). When screening for tetrmeric microsatellites, the amplification primers are the one used for the primer extension reactions and a reverse sequencing primer (5′-CGCCAGGGTTTCCAGTCACCATGAC-3′). When screening for tetrameric microsatellites, amplification primers and Biotin-21-dUTP (CLONTECH Laboratories, Palo Alto, CA, USA) were similarly unsuccessful. A model that explains the above results is that the biotin label is only bound efficiently when it is physically exposed by being located on the displaced 5′ end of a ssDNA molecule (Figure 1). The Klenow fragment of DNA Polymerase I differs from the other enzymes that were tried in that it can show strand displacement activity (9). Under the proposed model, the polymerase would extend the annealed primer along the circular M13 template until it encountered the 5′ (biotinylated) end of the primer again. At this point, the enzyme would displace the primer and continue the process of synthesis and strand displacement, placing the biotin label on the end of an increasingly long length of ssDNA.

Salt concentrations were also optimized for the Klenow primer extension reaction. Varying the concentration of MgCl$_2$ from 2–5 mM had little impact, but at 10 mM specificity was considerably reduced. This is important because the Klenow buffers provided by some manufacturers contain 10 mM MgCl$_2$. NaCl concentrations of 0, 25 and 50 mM caused a stepwise improvement in specificity. Addition of KCl at 50 mM had a negative effect on efficiency.

## RESULTS AND DISCUSSION

### Optimization with Known DNA Mixtures

The crux of this protocol is the binding of large (>7 kb) circular DNA molecules to magnetic beads. The only enzyme that was successfully used in the first extension reaction was the Klenow fragment of DNA Polymerase I. When this enzyme was replaced with T4 (Promega), Taq (PE Biosystems), rTth (PE Biosystems) or a combination of Taq and Pwo (Boehringer Mannheim) polymerases, the efficiency (number of plaques recovered) and specificity (proportion of white, microsatellite-containing plaques) of the procedure were no greater than in controls without enzyme. Interestingly, the Taq/Pwo combination was actually preferable to Klenow in the second strand extension reaction because (i) the higher annealing temperature improved specificity and (ii) the reaction time was shorter. Another important variable was the binding buffer. When the washing buffer was placed in place of the kilobaseBINDER buffer, which is standard practice for binding smaller molecules, efficiency and specificity were little better than in controls with unlabeled primers. Finally, attempts to label by incorporation using unlabeled microsatellite primers and Biotin-21-dUTP (CLONTECH Laboratories, Palo Alto, CA, USA) were similarly unsuccessful.

### Table 2. Results from Empirical Tests of the Method

<table>
<thead>
<tr>
<th>Species</th>
<th>Motif</th>
<th>No. of Blue$^a$</th>
<th>No. of PCR$^b$</th>
<th>No. of Seq$^c$</th>
<th>No. of Uninterrupted Repeats</th>
<th>No. of Clones$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antechinus</td>
<td>AC</td>
<td>3/587</td>
<td>24</td>
<td>24</td>
<td>0 0 10 14</td>
<td>24</td>
</tr>
<tr>
<td>Antechinus</td>
<td>ATGG</td>
<td>11/57</td>
<td>12</td>
<td>2</td>
<td>0 2 0 0</td>
<td>1</td>
</tr>
<tr>
<td>Antechinus</td>
<td>ATAG</td>
<td>4/38</td>
<td>24</td>
<td>6</td>
<td>1 1 4 0</td>
<td>5</td>
</tr>
<tr>
<td>Abalone</td>
<td>AC</td>
<td>0/34</td>
<td>14</td>
<td>14</td>
<td>3 2 4 5</td>
<td>13</td>
</tr>
<tr>
<td>Abalone</td>
<td>ATGG</td>
<td>95/473</td>
<td>12</td>
<td>1</td>
<td>0 1 0 0</td>
<td>1</td>
</tr>
<tr>
<td>Abalone</td>
<td>ATAG</td>
<td>27/99</td>
<td>0</td>
<td>-</td>
<td>- - - -</td>
<td>-</td>
</tr>
</tbody>
</table>

$^a$The proportion of blue plaques after transforming aliquots of the enriched library. The initial ss libraries produced approximately 30% blue plaques before enrichment. With the combination of restriction enzymes used it would be possible to eliminate most of these blue plaques by digesting ligation products with Hincll, but in this case the proportion of blue plaques was useful as a preliminary indication of enrichment success. The abalone ATAG screen was abandoned because this proportion was too high.

$^b$The number of plaques screened for microsatellite motifs by PCR.

$^c$The number of clones sequenced (equal to the number of positives in the PCR test except for the abalone AC screen were five negatives were also sequenced).

$^d$The number of unique clones identified by sequencing.
After conditions were optimized, the protocol was repeated five times on a 1:149 mixture of microsatellite (white plaques) to non-microsatellite (blue plaques) clones. The final proportion of white plaques in test transformations from these experiments ranged from 75/76–89/89, with a mean frequency across the five experiments of 99.5%. In one trial, the enrichment was checked before and after the second extension. In this case, approximately 1/4 of the eluate from the Dynabeads was ethanol-precipitated, resuspended in water and transformed, yielding 79 white plaques out of 145 (54.5%). After the second strand extension, an aliquot corresponding to 1/25th of the starting mixture was transformed without ethanol-precipitation, yielding 203 white plaques out of 204 (99.5%). These results show that both phases of strand extension contribute to the enrichment process.

**PCR Screen for Positives**

The specificity of the PCR-based test for microsatellite-containing clones was assessed by experiments with known positives and negatives and by sequencing both positives (nine) and negatives (five) in a screen for \((AC)_n\) repeats in abalone (Table 2). In the latter experiment, the largest number of tandem AC repeats in the five negative clones was five, whereas the smallest number of tandem repeats among the positive clones was 7.5, and those 7.5 repeats were embedded in a larger interrupted repeat sequence. Clearly this screening procedure provides a rapid way to eliminate clones that do not contain a reasonably large tract of microsatellite sequence.

**Empirical Testing**

Five screens of antechinus and abalone libraries (Table 2) were pursued through the stages of PCR testing and sequencing. The most successful screen was the one for antechinus \((AC)_n\) microsatellites, where 24 randomly chosen clones all contained long uninterrupted microsatellite sequences. The worst outcome was the screen for abalone \((ATGG)_n\) repeats, where only one of 12 clones produced a (very weak) product in the PCR test. Sequencing showed this clone to contain only three perfect repeats.

Although the method showed promise for enriching libraries with clones containing tetranucleotide repeats, isolating many of these sequences would still require further screening. One option would be to plate a large number of clones that have been passed through the enrichment protocol and then use filter hybridization to identify microsatellite-containing clones. Another alternative would be to plate the clones as above, but then isolate ssDNA and perform another one or two rounds of strand extension to further enrich the library. A third possibility would be to grow the initial ssDNA library in a bacterial strain \((dut and ung)\) that allowed dUTP to be incorporated so that single-stranded transformants would be destroyed after the final transformation (10). This would improve the efficiency of the second enrichment phase, but depends on the availability of an F\(^{+}\) bacterial strain (assuming M13 is used) that can be made to transform with the efficiency required for the initial library plating.

A general problem with enrichment protocols is that they require amplification of the DNA molecules being screened. This causes many copies of each clone to be present, often in unequal proportions, and becomes a problem when the same clones are isolated and sequenced multiple times. In the protocol I describe here, the amplification happens in vivo during the preparation of the single-stranded genomic library. This did not cause a problem in the screen for antechinus \((AC)_n\) microsatellites, but duplicate clones were sequenced in three of the other screens (Table 2). This suggests that isolating a large number of clones containing rare motifs could involve considerable...
duplication of effort. A combination of using larger libraries, improving methods to encourage even representation of individual clones and sequencing a small number of clones from each library will minimize this duplication.

CONCLUSIONS

Overall, results from tests with known and unknown mixtures of DNAs demonstrate that the technique is efficient and reproducible. From the point where it departs from traditional methods, the protocol requires several hours for the isolation of ssDNA and then takes a full day (ca. 8 h) to perform the strand-extension reactions and transform the enriched library. This is less time than is normally required for a single round of filter hybridization. Furthermore, the method requires little specialized equipment or technical expertise. These qualities make this protocol an attractive alternative to existing methods for isolating microsatellite markers.

REFERENCES

Selection of an Anti-CD20, Single-Chain Antibody by Phage ELISA on Fixed Cells

BioTechniques 26:697-702 (April 1999)

ABSTRACT

Cloning the correct genes that code for antibody-variable domains from hybridomas is often complicated by the presence of several immunoglobulin transcripts, some of them arising from a myeloma cell line. For the rapid functional evaluation of recombinant antibody fragments against cell-surface antigens, we established an efficient expression and screening system using phagemid antibodies and fixed cells. \( V_L \) and \( V_H \)-polymerase chain reaction (PCR) products, amplified from hybridoma cDNA, were cloned into the phagemid vector pSEX81. After transduction into E. coli and phage rescue, clones were tested for antigen binding using a phage-enzyme-linked immunosorbent assay (ELISA) procedure with whole cells fixed to ELISA wells. This procedure facilitated the successful cloning of a functional anti-CD20, single-chain antibody from hybridoma cDNA. The CD20 B-lymphocyte surface antigen expressed by B-cell lymphomas is an attractive target for cancer treatment using immunoconjugates or bi-specific antibodies.

INTRODUCTION

CD20 is a B-cell integral membrane protein capable of initiating growth-modulating signals in human B lymphocytes upon its engagement with monoclonal anti-CD20 antibodies (5). CD20 is mainly expressed during pre-B-cell development (19). It is present on both resting and activated B-cells, but is lost before terminal B-cell differentiation into plasma cells. CD20 is expressed on about 90% of B-cell lymphomas in an approximately 200-fold higher amount than on normal B cells (10). It is therefore an attractive target for the therapy of B-cell lymphomas with immunoconjugates or bi-specific antibodies. The B cells can be replaced after treatment by their antigen-negative precursors. Monoclonal anti-CD20 antibodies have been used either alone or labeled with \( ^{131}I \) for the treatment of non-Hodgkin’s lymphoma patients (4,18). One problem using conjugates with the relatively large monoclonal antibody is suboptimal tumor penetration. Typically, less than 0.01% of the injected dose of the administered monoclonal antibody (MAb) localizes to tumors in patients (16). The use of single-chain Fv fragments (scFvs) as