Optimized construction of microsatellite-enriched libraries

Natascha Techen  
*University of Mississippi*

R. S. Arias  
*U.S.D.A. National Peanut Res. Lab.*, renee.arias@ars.usda.gov

Neil C. Glynn  
*USDA-ARS*

Zhiqiang Pan  
*University of Mississippi*

Ikhlas A. Khan  
*University of Mississippi*

*See next page for additional authors*

Follow this and additional works at: [http://digitalcommons.unl.edu/usdaarsfacpub](http://digitalcommons.unl.edu/usdaarsfacpub)

Part of the [Agricultural Science Commons](http://digitalcommons.unl.edu/usdaarsfacpub)

Ttechen, Natascha; Arias, R. S.; Glynn, Neil C.; Pan, Zhiqiang; Khan, Ikhlas A.; and Scheffler, Brian E., "Optimized construction of microsatellite-enriched libraries" (2010). *Publications from USDA-ARS / UNL Faculty*. 897.

[http://digitalcommons.unl.edu/usdaarsfacpub/897](http://digitalcommons.unl.edu/usdaarsfacpub/897)
Authors
Natascha Techen, R. S. Arias, Neil C. Glynn, Zhiqiang Pan, Ikhlaz A. Khan, and Brian E. Scheffler
TECHNICAL ADVANCES

Optimized construction of microsatellite-enriched libraries

NATASCHA TECHEN,* RENÉE S. ARIAS,† NEIL C. GLYNN,‡ ZHIQIANG PAN,§ IKHLAS A. KHAN* and BRIAN E. SCHEFFLER†

*Department of Pharmacognosy, School of Pharmacy, National Center for Natural Products Research and Research Institute of Pharmaceutical Sciences, PO Box 1848, University of Mississippi, MS, USA, †USDA-ARS, Genomics and Bioinformatics Research Unit, 141 Experiment Station Rd, Stoneville, MS, USA, ‡USDA-ARS, 12990 US Highway 441 N, Canal Point, FL, USA, §USDA-ARS-NPURU, PO Box 1848, University of Mississippi, MS, USA

Abstract

The construction of microsatellite-enriched libraries is an indispensable tool to search for molecular markers as complete genome sequences are still not available for the majority of species of interest. Numerous protocols are available in the literature for the construction of these libraries; however, sometimes their low efficiency or lack of optimization in the protocols can restrict their efficacy. We have designed and tested various adapters and ligation methods; we also tested oligo-repeat combinations and hybridization temperatures, and created libraries with this new protocol for four organisms: Ipomoea batatas (L.) Lam, Chionanthus retusus Lindley & Paxton, Rotylenchulus reniformis Linford & Olivera and Puccinia kuehnii W. Krüger. The number of microsatellites detected for these species ranged from 2494 to 3919 per Mb of nonredundant sequence, that was 0.86 and 1.53 microsatellites per contig, with 37–66% of di-nucleotide motifs and 21–49% of tri- to octa-nucleotide repeats combined. A simplified protocol is provided for the successful generation of SSR-enriched libraries.

Keywords: adapter, microsatellites, protocol, simple sequence repeat

Introduction

Simple sequence repeats (SSR) or microsatellites are DNA tandem repeats composed of short repeat motifs (1–8 bp) (Richard et al. 2008), i.e. (AGT)5 consists of a 3-bp repeat motif and a 15-bp repeat length. Microsatellites are distributed throughout the eukaryote genomes (Katti et al. 2001; Richard et al. 2008) and their high variability makes them a powerful tool for genetic studies, such as genome-assisted breeding of animals and plants (Varshney et al. 2005a; Zhang et al. 2006), fingerprinting, authentication of varieties in the industry of natural products (Sharma et al. 2008) and detection of human diseases (Richard et al. 2008). Microsatellites can be found by creating SSR-enriched libraries or by data mining of sequences already available (Katti et al. 2001; Varshney et al. 2005a,b). Although the number of completed genomes is increasing with the availability of new sequencing technologies, only 114 eukaryote genomes have been sequenced until today (http://genomesonline.org), and for the majority of species there are none or few sequences in public databases. Therefore, generating SSR-enriched libraries is still an indispensable tool to find microsatellites as molecular markers for most species. Microsatellites have been known for about 20 years. Several techniques are available to make enriched libraries to find microsatellites; however, these techniques are costly, consist of multiple steps and to be successful the user needs to acquire significant experience (Zane et al. 2002). There is still a constant interest in developing new methods for SSR isolation (Santana et al. 2009) and improve the efficiency of the existing ones (Yue et al. 2009). We do not intend to review all the methods but do mention a few features that helped us develop a more efficient technique, i.e. no dephosphorylation of DNA fragments, T-A
ligation of adapters and emphasis on critical steps to assist the user in obtaining SSR libraries.

For the construction of SSR-enriched libraries, an ingenious adapter had been reported that allowed the simultaneous cut of adapter duplexes and blunt-end ligation of the adapter to dephosphorylated DNA fragments to avoid concatamerization (Hamilton et al. 1999); however, there is always risk of concatamerization of DNA fragments despite the presence of restriction enzyme during ligation. Another interesting improvement to the construction of these libraries was done by using DNA polymerase to extend the hybridized DNA (SSR-oligo to genomic DNA) these libraries was done by using DNA polymerase to extend the hybridized DNA (SSR-oligo to genomic DNA fragment) (Hayden et al. 2002). This step increases the length of the hybrid molecule and therefore augments its melting temperature, which allows for a more stringent removal of nonspecific hybrids (Hayden et al. 2002).

In one recent study Yue et al. (2009) reported that 36% of the clones after enrichment had SSRs. On a different approach to the detection of microsatellites, using second-generation deep sequencer Roche 454 on SSR-enriched libraries the percentage of sequences containing satellites ranged from 25% to 97% (Santana et al. 2009). Comparing the efficiency of different SSR enrichment protocols is generally difficult and should be done with scepticism because of differences in study organisms used, criteria used to identify SSRs as well as other variance among laboratories and researchers.

To increase the quality, efficiency and success rate of making SSR-enriched libraries, we changed from blunt-end ligation of dephosphorylated DNA fragments to sticky-end ligation without dephosphorylation. Blunt-end ligation is known to be between one and two orders of magnitude less efficient than cohesive-end ligation (Ausubel et al. 1987). We designed and tested new adapters to maximize amplification of the desired fragment size range and minimize background amplification. We incorporated the extension after hybridization proposed by Hayden et al. (2002), and optimized the time, buffers and conditions for each step, which resulted in increased recovery of SSRs. We tested the method on DNA from various organisms including a fungus, a crop plant, an animal (nematode) and a tree. Finally, as we noticed that certain steps in the process, sometimes not detailed in the literature, were crucial to avoid PCR artefacts and to reduce the number of clones that do not harbour repeats, we indicated those steps and elaborated an easy-to-follow protocol for the successful generation of SSR-enriched libraries.

Materials and methods

Adapter design

Two types of random adapters 20–23 nucleotides in length were designed with the program Primo Random (http://www.changbioscience.com/primo/primor.html) with the settings: TM 60–64 °C, TM formula AT2 CG4, 60% CG and any 3’-end. Individual sequences were further analysed for potential dimerization using the program Clone Manager Suite (Sci-Ed Central Software, Cary, NC, USA). Adapters 1 and 2 were designed with a 3’-T overhang to be ligated to genomic DNA fragments having an A overhang at the 3’-end. Adapters 4 and 5 were designed for blunt-end ligation using dephosphorylated blunt-end DNA fragments, and a half Smal restriction site (CCC) was added to perform ligation in the presence of Smal restriction enzyme, similar to the method reported by Hamilton et al. (1999). Adapter 3 is the same as adapter 4 after adding a 3’-T overhang. Reverse oligos of the adapters were phosphorylated at the 5’-end and were added an AAA overhang at the 3’-end to prevent concatamers. The designed adapters and their corresponding oligonucleotides are shown in Table 1.

Assembly of adapters and test for nonspecific amplification

To assemble each adapter, equimolar quantities of each complementary oligonucleotide were mixed in the presence of buffer 4 from New England Biolabs (NEB, Ipswich, MA, USA) (Fig. 1I), heated for 30 s at 90 °C, brought to 65 °C and slowly cooled down over 2–3 h to 30 °C (Fig. 1I), aliquots were kept at −20 °C until use. One of our objectives was to make SSR-enriched libraries for plants used as dietary supplements; thus, we chose DNA of Achillea filipendulina Lam. (fernleaf yarrow) as an example. Two sets of gradient PCR reactions (annealing 55–72 °C) were set in the presence or absence of 100 ng of blunt-end DNA fragments of A. filipendulina. A mix of forward and reverse oligos of each adapter, 12 picomols each, were used in 20 μL volume PCR reactions with Taq DNA polymerase (Promega, Madison, WI, USA), 96 °C for 3 min, 50 °C for 30 s, 72 °C for 90 s, 40 cycles, and final extension 72 °C for 7 min to test for nonspecific/background amplification.

Ligation of adapters and amplification of various DNA sizes using a DNA ladder

Ligation efficiency of the designed adapters to fragmented DNA was tested by ligating the adapters to the six equimolar blunt-end DNA fragments, 0.1, 0.2, 0.4, 0.8, 1.2 and 2 kb of the low-mass DNA ladder (LMDL) (Invitrogen, Carlsbad, CA, USA) using high-concentration T4-DNA ligase (NEB). For adapters 1–3 LMDL was first treated with Taq DNA polymerase in the presence of deoxyadeno sine to add A overhangs (A-tailing) at the 3’-end (Fig. IIII) and then T–A ligated to the adapters (for details, see Fig. IIII, IV). For adapter 5, the ligation was done using...
dephosphorylated LMDL DNA fragments in the presence of the restriction enzyme SmaI, similar to the method reported by Hamilton et al. (1999). Gradient PCR was used to determine the efficiency of the ligations and performance of the adapters amplifying the LMDL fragments.

**Genomic DNA preparation, A-tailing and ligation to adapter 2**

Simple sequence repeat-enriched libraries were prepared using genomic DNA of four species, *Rotylenchulus reniformis* Linford & Olivera, *Puccinia kuehnii* W. Krüger, *Ipomoea batatas* (L.) Lam. and *Chionanthus retusus* Lindley & Paxton (Table 2). To obtain DNA fragments in the range of 200–1000 bp, 2 µg of DNA were digested in 40 µL reactions using the following combinations of pairs of restriction enzymes that generate blunt ends: *Alu*I, *Hae*III, *Alu*I+*Rsa*I, *Dra*I+*Xmn*I, *Hpy*CH4V+*Xmn*I (NEB) (Fig. 1II) for 1–2 h. The digested DNA was combined after enzyme inactivation. Without cleaning these reactions, genomic DNA fragments were A-tailed as indicated in Fig. 1III. After A-tailing the reaction was incubated at 72 °C for 2 h, clean with MinElute, elute twice with 25 µL 0.5×EB.

**Table 1** Adapters designed and their corresponding oligonucleotides

| Adapter 1 | SSRLIBF1 | 5’-CATCCTGGGCTTGCTTCGAGT-3’ |
| Adapter 2 | SSRLIBR1 | 5’-/Phos/CTGACGAAGCAAGCCACGTCAGTAAAA-3’ |
| Adapter 3 | SSRLIBF3 | 5’-CGGAGAGCGAAGAGAGACT-3’ |
| Adapter 3 | SSRLIBF3 | 5’-/Phos/CTGACGAAGCAAGCCACGTCAGTAAAA-3’ |
| Adapter 4 | SSRLIBF4 | 5’-TGATTGGCCGCTGCTGACCCCT-3’ |
| Adapter 5 | SSRLIBR4 | 5’-/Phos/GGCGTCAGAAGCGCGCAATCAAAA-3’ |
| Adapter 6 | SSRLIBR5 | 5’-/Phos/GGCGTCAGAAGCGCGCAATCAAAA-3’ |
| Adapter 7 | SSRLIBF5 | 5’-CGGAGATGTCAGCAGCTGCTGACCCCT-3’ |
| Adapter 8 | SSRLIBR5 | 5’-/Phos/GGCGTCAGAAGCGCGCAATCAAAA-3’ |

**Fig. 1** Microsatellite-enriched library – short protocol.

Published 2009. This article is a US Government work and is in public domain in the USA.
cleaned with a MinElute Kit (Qiagen, Valencia, CA, USA) and ligated to the assembled Adapter 2 (SSRLIB3) (Fig. 1I) in 80-µL reactions in the presence of high-concentration DNA Ligase (NEB) Fig. 1IV. The ligation reaction was cleaned with the MinElute Kit and eluted twice with 25 µL of 0.5XEB buffer (diluted from MinElute Kit). PCR amplification of the ligated DNA was carried out in a final volume of 200 µL for 20 cycles with High-Fidelity DNA Polymerase (Invitrogen), the conditions were 95 °C for 2 min 30 s, 94 °C for 45 s, 60 °C for 30 s, 68 °C for 40 s, and a final extension of 5 min at 68 °C (Fig. 1V). Then 195 µL of the 200-µL reaction were removed from the reaction tube and the concentration of PCR product was determined using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The remaining 5 µL of PCR reaction was run for an additional 20 cycles, using the same conditions indicated in the previous paragraph, and run on a 0.7% agarose gel to determine if effective amplification occurred within the desirable size range (200–1000 bp). Amplified DNA fragments were purified using MinElute and eluted twice with 25 µL of 0.5XEB buffer.

Hybridization to biotinylated oligo repeats

The amplified products were hybridized to four groups of 5'-biotinylated oligo repeats similar to the ones described by Glenn & Schable (2005): group 1 [(AC)13, (AACC)5, (AACG)5, (AAAG)5, (ATCC)5], group 2 [(AG)12, (AAC)5, (AAG)5, (ACT)12, (ATC)5], group 3 [(AAAC)5, (AAGAC)5, (AATT)5, (ACAG)5, (ACCT)5, (ACTC)5, (ACTG)5] and group 4 [(AAAT)5, (AAGT)5, (AGAC)5, (AGAT)5], primers were purchased from MWG-Biotech (Huntsville, CA, USA). Then the plasmids were sequenced using primers M13Forward (GTAAAACGACGGCCAGT) and primer R271 (GCAGGTTTAAACGAATTCGC). We designed primer R271 to replace M13Reverse for sequencing.

Results

Design and testing of adapters

To determine the appropriate adapters for library construction, we designed five adapters for the experiment to eliminate possible production of artefacts during ligation and PCR amplification. The results of a gradient PCR showed that adapter 1 and 4 generated undesirable background amplifications especially in the presence of genomic DNA (Fig. 2a). We presume that the smear...
amplification in adapters 1 and 2 could have been primer dimer caused by the presence of complementary oligos; however, for adapter 4 it seemed that the primers specifically bound to regions of the genomic DNA and amplified separate bands, even at 72 °C annealing temperature, posing a more serious problem than primer dimer. Therefore, after the first testing, adapter 4 was no longer used.

Performance test of the adapters after ligation

Efficient sequencing by current technologies requires the use of fragments of 500 bp (+ or −300 bp) in length. To test amplification of fragments in the desired range, adapters 1, 2, 3 and 5 were ligated to LMDL. The ligated DNA was used in a gradient PCR to amplify the range of DNA fragment sizes from 0.1 to 2.0 kb. The smallest DNA fragment amplified (0.4 kb), was only observed using adapters 1 and 2 (Fig. 2b), with adapter 2 showing more efficient amplification, especially at annealing temperatures higher than 60 °C. There was a very weak amplification of the small fragment when using adapters 3 and 5 (Fig. 2b). Based on these results we chose adapter 2 for the construction of the SSR libraries in this work.

Optimized construction of SSR-enriched libraries

We modified the technique of Hamilton et al. (1999) to simplify the experimental procedure as well as improve the efficiency for the generation of SSR-enriched libraries. The enzyme combinations suggested here, resulted mostly in DNA fragments between 300 and 800 bp suitable for cloning and sequencing, not requiring a gel purification step. Adding the extension step proposed by Hayden et al. (2002) significantly improved the recovery of SSR-containing clones. As the melting temperature of the hybrid molecules is significantly increased with the extension after hybridization, it allowed increasing the stringency of the washes without concern of losing the DNA fragments. In our preliminary experiments, before using this step, a much lower percentage of clones contained repeats (12–50%). We have successfully used this technique to isolate SSRs from numerous species. Here, we present examples of microsatellites obtained from four of them: Ipomoea batatas, Chionanthus retusus, Rotylenchulus reniformis and Puccinia kuehni (Table 2). The complete protocol has been summarized in Fig. 11–VIII to facilitate its use in the laboratory. In addition, an overview of the procedure is shown in Fig. 3.

Calculating efficiency of the method

A variable number of clones for each organism were chosen for sequencing depending on the project. The number of contigs assembled was about 50–70% of the number of sequences obtained (Table 2), singletons
were not included in the analysis to avoid designing primers on unreliable sequences. To compare the efficiency of the method to other techniques from the literature, we calculated the number of repeats per Mb of contig sequence, where 1 Mb corresponds to 2500 non-overlapping clones with an average insert size of 400 bp of sequence (Zane et al. 2002). According to this calculation, our method detected between 2500 and 3900 repeats per Mb (Table 2), which corresponded to an efficiency of 0.86–1.53 repeats per 400 bp of contig sequence across the four species tested. Repeat motif lengths are shown in Table 2, with di-nucleotides ranging from 37% to 66%, and tri- plus tetra-nucleotides ranging from 17% to 43% depending on the species. The total number of clones that were sequenced for each species, number of assembled contigs and number of repeats detected by SSRfinder and Sputnik combined for each species are summarized in Table 2. Although the frequency and type of repeat motifs isolated depend on the species, in general, we observed that most tetranucleotide motifs were isolated with group 3 of oligo repeats, less frequently with group 4 and few with group 1. Sequences of microsatellites were submitted to GenBank with accession numbers: C. retusus GQ117288–GQ118148, R. reniformis FJ906198–FJ906620, I. batatas GU171483–GU172144 and P. kuehnii GU171394–GU171482.

Efficiency was calculated as $C = \frac{B}{A}$.

**Table 2** Summary of species used to test the protocol, and percentage of simple sequence repeats detected

<table>
<thead>
<tr>
<th>Species</th>
<th>Clones sequenced</th>
<th>Contigs assembled (nonsingletons) (A)</th>
<th>Repeats detected 1–8 bp motif and &gt;12 bp length (B)</th>
<th>Mb of sequence in contigs (C)</th>
<th>Repeats per Mb of contig sequence (efficiency) (D)</th>
<th>Repeats per contig (efficiency) (E)</th>
<th>Percentage (%) of repeats detected for each motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chionanthus retusus</td>
<td>2208</td>
<td>1079</td>
<td>1647</td>
<td>0.420290</td>
<td>3919</td>
<td>1.53</td>
<td>13.4 66.0 13.4 3.8 3.4</td>
</tr>
<tr>
<td>Ipomoea batatas</td>
<td>4608</td>
<td>2593</td>
<td>2638</td>
<td>1.056420</td>
<td>2497</td>
<td>1.02</td>
<td>3.3 62.2 21.6 8.6 4.3</td>
</tr>
<tr>
<td>Puccinia kuehnii</td>
<td>768</td>
<td>513</td>
<td>440</td>
<td>0.138460</td>
<td>3178</td>
<td>0.86</td>
<td>13.9 36.8 32.4 10.9 5.9</td>
</tr>
<tr>
<td>Rotylenchulus reniformis</td>
<td>1248</td>
<td>694</td>
<td>941</td>
<td>0.259930</td>
<td>3620</td>
<td>1.36</td>
<td>18.8 48.1 18.2 12.2 2.7</td>
</tr>
</tbody>
</table>

Efficiency was calculated as $C = \frac{B}{A}$.

**Fig. 3** Overview of genomic DNA SSR-enriched library protocol.

were not included in the analysis to avoid designing primers on unreliable sequences. To compare the efficiency of the method to other techniques from the literature, we calculated the number of repeats per Mb of contig sequence, where 1 Mb corresponds to 2500 non-overlapping clones with an average insert size of 400 bp of sequence (Zane et al. 2002). According to this calculation, our method detected between 2500 and 3900 repeats per Mb (Table 2), which corresponded to an efficiency of 0.86–1.53 repeats per 400 bp of contig sequence across the four species tested. Repeat motif lengths are shown in Table 2, with di-nucleotides ranging from 37% to 66%, and tri- plus tetra-nucleotides ranging from 17% to 43% depending on the species. The total number of clones that were sequenced for each species, number of assembled contigs and number of repeats detected by SSRfinder and Sputnik combined for each species are summarized in Table 2. Although the frequency and type of repeat motifs isolated depend on the species, in general, we observed that most tetranucleotide motifs were isolated with group 3 of oligo repeats, less frequently with group 4 and few with group 1. Sequences of microsatellites were submitted to GenBank with accession numbers: C. retusus GQ117288–GQ118148, R. reniformis FJ906198–FJ906620, I. batatas GU171483–GU172144 and P. kuehnii GU171394–GU171482.

**PCR amplification using the designed markers**

From a total of 384 primer sets designed and tested on 12 Chionanthus-related taxa (Arias et al. 2009), only 11 (3%) did not result in amplification. In another example, of 192 primer sets designed and tested on six populations of R. reniformis, only 14 (7%) showed no amplification. Puccinia and Ipomoea are still being tested, but so far 72 (75%) of 96 primer sets amplified two isolates of Puccinia, and 450 (62%) of 725 primer sets amplified two Ipomoea cultivars, although the DNA used for making the libraries has not been tested yet.
Discussion

After applying various protocols available in the literature to construct SSR-enriched libraries, we explored the possibility of improving their efficiency and providing an easy-to-follow procedure for the successful generation of these libraries. Advantages of our protocol compared with other published methods are: first we designed adapters that favour the amplification of fragments within the desired size range. To maximize annealing of the complimentary oligos of the adapters we used a slow cooling over 2–3 h, as it has been shown for 19–37mer oligos that 10–40% of complementary strands of nucleic acids remained unhybridized even after 1 h at 40°C (Schwille et al. 1996). Second, we do not need to dephosphorylate the DNA fragments before ligating the adapters. In Hamilton et al. (1999) the DNA fragments are dephosphorylated to prevent self-ligation; however, ligation of adapters to dephosphorylated DNA is less efficient than to phosphorylated DNA (Sambrook & Russell 2001). Third, we use cohesive-end ligation (T–A ligation) of adapters to the DNA fragments. Most of the previously published methods use blunt-end ligation of adapters (Hamilton et al. 1999; Glenn & Schable 2005); normally, blunt-end ligation has between one and two orders of magnitude lower efficiency than cohesive-end ligation (Ausubel et al. 1987). Although cohesive-end ligation of adapters is used in FIASCO method (Zane et al. 2002), this is in combination with a single restriction enzyme, which results in biased fragments. The group of Hamilton et al. (1999) suggested the use of several restriction enzymes, whereas Glenn & Schable (2005) suggested the use of only one enzyme (BstUI or Rsal) with the use of additional enzymes being optional. We strongly recommend the use of multiple restriction enzymes for two reasons. First, a single restriction enzyme applied to diverse genomes usually does not result in the desirable size range of DNA fragments (200–800 bp). Second, single enzymes result in biased cuts; thus, if a large number of fragments are sequenced, the use of more enzymes in separate reactions and then combining the fragments allows assembling longer contigs and designing a larger number of markers. We digest the DNA for 1–2 h depending on the DNA quality, as the method of extraction can influence the activity of the restriction enzymes (Do & Adams 1991).

Although some methods may use double enrichment of microsatellites to generate libraries (Glenn & Schable 2005), in our method a single enrichment step is sufficient to obtain between 2500 and 3900 microsatellites per Mb of contig sequence, or 0.86–1.53 microsatellites per 400-bp contig. A reduction in the number of PCR cycles has been suggested for the preparation of microsatellite-enriched libraries to avoid biased amplification (Zane et al. 2002). For our method, we recommend that if a large number of clones will be sequenced to isolate microsatellites, only a 15-cycle PCR be used to amplify the ligation and a 10-cycle PCR be used for the final amplification of single-strand DNA to avoid obtaining redundant sequences. Using these numbers of cycles, if we start with 10 µg of DNA for cut and ligation, and work with biotinylated oligo repeat groups 2–4, we obtain between 1 and 2 µg of microsatellite-containing DNA for sequencing.

We optimized the conditions required for each step and tested oligo-repeat mixtures and annealing temperatures that resulted in larger numbers of clones harbouring a variety of repeat motifs. From all the species on which we tested this method so far, we have isolated a much higher proportion of di-nucleotides compared with tri- or tetra-nucleotides. Eukaryote genome-wide screenings have shown that di-nucleotide repeats are more abundant than tri- or tetra-nucleotide motifs (Katti et al. 2001; Anwar & Khan 2005). Although less abundant, tri- and tetra-nucleotide motif repeats, however, are often more desirable for SSR analysis than di-nucleotides because of their ease to score (Thiéry & Mugniéry 2000; Kumar et al. 2002). If the users prefer to isolate tri- and tetra-nucleotides they can use groups 3 and 4 for maximum efficiency. The method presented here allowed recovery of 21–49% of tri- to octa-nucleotide repeat motifs.

Our method rendered between 57% and 75% of the clones (not including singletons) harbouring microsatellites without performing a preliminary PCR screening or colony hybridization. Examples of efficiencies obtained by other methods are: i.e. libraries made using FIASCO (Zane et al. 2002) with 25–67% of clones containing repeats (Zhang et al. 2008; Yang et al. 2009 respectively). In addition, libraries made using the method of Glenn & Schable (2005) had efficiencies of 8–66% (Markwith & Scanlon 2006; Byrne et al. 2009 respectively), whereas for other methods it was as low as 16% (Cuc et al. 2008). However, we would like to point out that for the studies cited, the values obtained by those methods required preliminary PCR screening or colony hybridizations before sequencing.

We show results of microsatellite-enriched libraries for four organisms, a nematode, a fungus, a crop plant and a tree, to demonstrate that our protocol worked effectively across phyla. Some of the species we tested are polyploids (Ipomoea batatas is a hexaploid) which result in large numbers of slightly different sequences assembled in single contigs, probably resulting in an apparent overall reduced efficiency. The protocol presented here for the construction of microsatellite-enriched libraries is an effective tool for generating molecular markers from a variety of species. By following...
the steps described in Fig. 1, the user will be able to successfully construct SSR-enriched libraries. Comparing the results of the method presented here with those of Santana et al. (2009) using FIASCO, we believe that our method can be a better alternative to make microsatellite-enriched libraries for pyrosequencing. Additionally, the methods presented here will avoid problems associated with concatamers created by the use of blunt-end ligation approaches in the current Roche (454) library construction protocols.

Acknowledgements

This work was funded in part by the Food and Drug Administration, grant number FD-U-002071-01 and by USDA Agricultural research Service Specific Cooperative Agreement No. 58-6408-2-0009. The authors thank Ms Marilyn Ruscoe and Ms Xiaofen (Fanny) Liu for technical support and to Dr O. Perera for suggestions.

References


Published 2009. This article is a US Government work and is in public domain in the USA.