Molecular Ecology Notes (2003) **3**, 324–327 doi: 10.1046/j.1471-8286.2003.00434.x

TECHNICAL NOTE

Fast isolation of microsatellite loci of very diverse repeat motifs by library enrichment in echinoderm species, *Amphipholis squamata* and *Echinocardium cordatum*

A. CHENUIL, M. LE GAC and M. THIERRY
Observatoire Océanologique de Banyuls, BP44, 66 650 Banyuls-sur-mer, France

Abstract

We conceived a microsatellite enrichment protocol in which probes of several repeat motifs are mixed during the whole procedure. The repeats found generally differed from those of the probes and were very diverse, from mononucleotides to pentanucleotides. We tested several modifications with alternative: (i) digestion/ligation buffers; (ii) polymerases; and (iii) purification methods. The simplest methods always worked as well as classical ones or even better, resulting in probably the simplest protocol for isolating microsatellites of diverse motifs from genomes with low microsatellite frequency. The proportion of positive clones, polymorphism levels and cross-specific amplification of microsatellite loci significantly vary between species.

Keywords: *Amphipholis squamata*, *Echinocardium cordatum*, echinodermata, enrichment, microsatellite

Received 10 December 2002; revision received 9 January 2003; accepted 23 February 2003

The species used in this study are the ophiuroid *Amphipholis squamata* (Delle Chiaje, 1828) and the sea urchin *Echinocardium cordatum* (Pennant 1777). Genomic DNA was digested, ligated to adaptors and amplified by polymerase chain reaction (PCR) following the basic steps of Zane *et al*. (2002) with several modifications. For digestion and ligation, four treatments were compared. About 300 ng of DNA was used in treatment T0 and 100 ng in other treatments. For T0, digestion by 1.5 U of *Mse*I was performed for 3 h at 37 °C with 1.5 μg bovine serun albumin (BSA) in 15 μL of buffer NE2 1X (New England Biolabs). After using 7 μL for a control minigel, the remaining digested DNA was purified using a Qiagen minielute column, and the elution product was used in a 15 μL ligation reaction with 20 pmol of the *Mse*I amplified fragment length adaptor (5′-TACTCAGGACTCAT-3′/5′-GACGATGAGTCCTGAG-3′), 0.9 U of T4 ligase (Promega) in T4 ligase buffer 1 × [30 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 10 mM dithiothreitol (DTT), 1 mM ATP], for one night at 8–9 °C. For treatments T1 to T3, digestion and ligation were performed in the same vial (3 h at 37 °C). Combining these reactions is possible because ligation of adaptors to genomic DNA does not restore *Mse*I sites (5′-TTAA-3′), and eliminates the risk of producing chimeric DNA. These reactions were performed in 25 μL using composite buffers, prepared with variable proportions of Y5 Tango buffer [10 ×: 333 mM Tris acetate, 100 mM Mg acetate, 666 mM K acetate, 1 μg/mL BSA (MBI Fermentas)] and T4 ligase buffer. The proportions of Y5 Tango and T4 ligase buffers were, respectively, 4:1 for treatment T1, 2:3 for T2 and 0:1 for T3. We added 0.75 μL of DTT 0.1 M in T1 and 0.5 μL of BSA 10 mg/mL in T3 treatments. Ligated DNA was then diluted 10 times and 5 μL were added into 20 μL PCR mixes containing 20 pmol of each primer *Mse*-N as in Zane *et al*. (2002), 0.5 U of *Taq* polymerase (Promega) and 200 mM of each dNTP. Varying numbers of PCR cycles (15, 20, 25 and 30) were performed on a PTC200 thermal cycler (MJ Research), specifying a speed ramp of 1.4 °C/s and avoiding hot start (2 min 94 °C, 15–30 cycles of 30 s 94 °C, 1 min 53 °C and 1 min 72 °C). Amplicons producing a visible smear on a control minigel (5 μL loaded) and resulting from the smallest number of PCR cycles (15 cycles in this study) were used for subsequent procedures (Zane *et al*. 2002). The simplest treatment, T1, produced high yields of homogeneous smears between 180 and 400 bp such as treatment T0, and appears as the most efficient way of producing amplifiable ligated DNA. Treatment T2 yielded less intense narrower smears, T3 was worse.

Correspondence: Anne Chenuil. E-mail: chenuil@obs-banyuls.fr

Our enrichment procedure follows Perrin & Roy (2000), except that (i) we used 10 probes corresponding to di-, tri-, and tetranucleotidic repeats, (ii) only one round was performed, (iii) one purification step was removed and (iv) the amount of HCl necessary to neutralize the NaOH solution used to elute DNA was previously determined by pH paper and pH of neutralized DNA was checked by putting 1 μL on pH paper. Probes biotinylated in 5′ were separated in two groups according to their Tm, the group containing $(GA)_{12}$ $(CA)_{12}$ $(AAT)_9$ $(GTG)_7$ $(GAG)_7$ and $(AAAT)_8$ was hybridized at 60 °C and the group containing $(GATA)_8$ $(CATA)_8$ $(GACA)_8$ and $(GAAA)_8$ at 70 °C; all probes of one group were mixed together. Hybridization was performed by a one cycle PCR reaction (2 min 94 °C, 1 min 60 °C or 70 °C, and 1 min 72 °C) in 20 μL containing 5 pmol of each probe, about 5 μg of amplified ligated DNA previously purified on a minielute column (Qiagen), 0.2 mM of each dNTP, 0.5 U of *Taq* polymerase (Promega). After enrichment washes, 40 cycles of PCR, either with *Taq* (Promega) or with Qbiotaq (Qbiogen) polymerase were performed, either directly from 1 μL of this 110 μL solution of enriched ssDNA, or from 1 μL of isopropanol precipitated DNA (80 μL of enriched ssDNA were precipitated in cold isopropanol, washed in cold 70% ethanol, dried and suspended in 10 μL), or from 1 μL of purified DNA (20 μL of enriched ssDNA was column purified and eluted in 10 μL). Other PCR conditions were as for the first PCR but a final extension step of 7 min was added. Amplicon amount was higher with Qbiotaq than with *Taq* polymerase. Nonpurified DNA amplified as well as isopropanol precipitated and much better than column purified DNA. In most cases, a smear was visible between 100 bp and 400 bp, and more intense between 150 or 200 bp and 300 bp. The microsatellite isolation procedure was then completed only for DNA enriched with 70 °C probes (all tetranucleotidic), amplified without purification using Qbiotaq, and cloned using the TOPO-TA cloning kit (Invitrogen).

White colonies were screened by two PCR per colony, containing all repeat probes plus one of the vector primers (PucF: 5′-CGCCAGGGTTTTCCCAGTCACGAC-3′ and PucR: 5′-TCACACAGGAAACAGCTATGAC-3′). Conditions were those of the first PCR except that 40 cycles were performed and annealing temperature was 55 °C. Using Qbiotaq instead of *Taq* polymerase produced smears, except when the annealing temperature was increased to 65 °C (PucF primer only). About 52% of clones were positive (32 and 31 positive out of 60 colonies screened for each species). PCR with the two vector primers and under the same conditions as screening, but with Qbiotaq, yielded high amounts of DNA. One sequencing reaction generally gave the whole insert sequence.

For *Amphipholis* and *Echinocardium*, respectively, 23 and 28 positive clones were sequenced (GenBank Accession nos: AY143609–AY143659), eight and 20 clones contained true microsatellites (Table 1), all the other clones displayed traces of sequence simplicity (e.g. sequences of highly biased nucleotidic composition, or degenerate microsatellites). These proportions significantly differ among species

Table 1 Positive clones from a PCR screening of *Amphipholis squamata* and *Echinocardium cordatum* enriched libraries. Sequences containing more than eight perfect tandem repeats for mononucleotides and more than five for other motifs were considered true microsatellites. Total repeat number per motif is indicated; dashes separate motifs of a same clone, commas separate distinct clones

	True microsatellites (size of repeat unit, in bp)				
Positive clones (without twins)	1	2	3	4	≥ 5
AG8, AG10, AG11, AG12, AH2 = AH10, AH5, AH6, AH7, AH12, AI2, AJ2, AJ4, AJ5, AJ9	—	—	—	—	
AG7, AI3, AI9, AJ1	—	> 45, 34, 14, 17	—	—	—
AI8	—	8	—	—	—
AI1	—	—	7–6–28	—	6
AH1, AJ8	—	—		> 10, > 58	—
Total *A. squamata*: 22	0	5	1	2	1
EL4, EL7, EL8 = EN7, EM6, EN10, EO4	—	—	—	—	—
EM3, EM12	13, 9	—	—	—	—
EL3, EL5, EL10, EL11, EM1, EM4, EN2, EN5, EN9, EO3, EO9	—	30, 7, 14, 28, 33–8, 15, 13, 6, 12, 14, 12	—	—	—
EM2, EO12	—	—	12– > 50, 56	—	—
EM8, EM9, EN6	—	—	—	> 20, 35, > 80	—
EN8 = EL12, EO8	—	—	—	—	12, > 15
Total *E. cordatum*: 26	2	11	2	3	2

Table 2 PCR conditions, motif, size ranges and numbers of alleles of microsatellite loci

Locus	Primer sequences (5′–3′) (F and R primer)	°C	Motif	Size range	Allele no./ individual no.	GenBank Accession no
AG7	ACGCTTTTCTTGAATAACGTCAT ACAATGTAGCTCAAACAGACTTAG	64	CA	110–200	≥ 2/100	AY143629
AI8-a	TTAATGCAGTACCACATGAGTC GTAGGAGAGACCTTATGGCTGC	64	CA	300	4/300	AY143621
AI8-b*	TTAATGCAGTACCACATGAGTC GTAGGAGAGACCTTATGGCTGC	64	CA	190	2/300	AY143621
AJ1	TAACAATCCTAGAACCTAACACATA CATTTGTGCATGTCAGAGTGGTTTG	64	CT	110	6/250	AY143623
AJ5	ATCATCAAACATTGACATTCACTC AGCACTGCTCCCAGGAAATGA	64	Rich in T & C	250	1†/31	AY143626
AI9	GCTTGATAGTCGTTTGCTCATC GTGGGACCCCTGGAAATTGGTCA	64	CA, CT	170	1†/50	AY143622
EL3	ATGATGCTTTGAATACCAGTCC CATTGATAGGAACTACATATACC	50–54	CA	150–220	3/45	AY143633
EL10	TGTATAGACAACATTTGTGGTCA GTGAACTGTGATCATTATCATC	55	CA	140	≥ 3‡/80	AY143638
EM9	AATACAGATGTCAATCATTTTGG AATCAACGAATCATGTTCAGTTC	55	GATA	130–340	8‡/100	AY143647
EN5	GAGTGTCTGTATATAATCGTATGT GCATCTCAAGTGATCTCATATAG	50	CT	220	≥ 6‡/30	AY143650
EN8	CCAAGTCTAACTCTTGTTTGTAG AAGCAAGTAGCCAACAGAGACA	58	GTTTT	260–300	≥ 3/70	AY143639

*Segregation of alleles from loci AI8a and AI8b appear independent.
†Allele size assessed by 2% agarose gels only.
‡Null alleles are present.

[exact test P-value < 0.008, GENEPOP (Raymond & Rousset 1995)]. The efficiency of our enrichment procedure is therefore 19% for *Amphipholis* and 37% for *Echinocardium*. Surprisingly, only one of the 26 clones containing microsatellites corresponded to a repeat motif used for enrichment. Most were dinucleotides (16 clones, mostly CA) but 11 repeats of larger unit size were also found [trinucleotides (CAT, CTT), tetranucleotides (very long CAGT, TACT, medium GATT, GTTT, and GATA) and pentanucleotides (GTTTT, CATAT, TAATC)]. For *Amphipholis* and *Echinocardium*, respectively, only four and five clones contained at least two tandem repeats of a searched motif. Several dispersed repeats of at least one searched motif were identified in most sequences and seem slightly in excess relative to expectations based on binomial probability and equi-frequency of bases. The single negative clone sequenced (GenBank Accession no.: AY 143632) also displayed simplicity with two tandem plus two isolated repeats of (TACTA). These results suggest that our procedure efficiently enriched for a wide variety of repeats.

All primer pairs tested in *Amphipholis (*clones AJ1, AI3, AI9, AI8, AJ5 and AG7) and *Echinocardium* (clones EM9, EL10, EN8, EN5 and EL3) except one gave successful amplification in 20-μL reactions containing 0.5 U Qbiotaq polymerase, supplied buffer (1 ×: Tris-HCl 50 mM pH 9.1, NH_4SO_4 16 mM, $MgCl_2$ 3.5 mM, BSA 0.15 mg/mL), 5–100 ng of chelex DNA extracts (Chenuil & Féral 2003), 20 pmol of each primer, 0.33 mM of each dNTP, using cycling conditions: 3 min 95 °C, 35–40 cycles of 40 s 94 °C, 1 min at annealing temperature, 1.5–2 min 72 °C according to locus on a PTC200 thermal cycler (Table 2). Within *A. squamata*, although hundreds of individuals belonging to several cryptic but highly diverging species were screened (Sponer & Roy 2002), few alleles were detected (one to six) and null amplifications were not observed (Le Gac 2002), whereas in *E. cordatum*, all five loci are polymorphic, null alleles occur at three loci and their geographical distribution supports the existence of partially sympatric cryptic taxa (Chenuil & Féral 2003). These contrasted properties between species are probably related to life history traits and suggest that our protocol did not yield a biased population of loci concerning polymorphism and cross-amplification levels.

Because many potential motifs of more than two or three bp (which offer the advantage of being more easily characterized) exist, and most of them are probably rare, our protocol appears very useful to isolate them, particularly from genomes poor in microsatellites.

Acknowledgements

Sigurd Von Boletzky kindly corrected the english. This work was supported by a grant from the IFB (Institut Français de la Biodiversité).

References

Chenuil A, Féral J-P (2003) Sequences of mitochondrial DNA suggest that *Echinocardium cordatum* is a complex of several sympatric or hybridizing species. A pilot study. In: *Echinoderm Research 2001, Proceedings of the 6th Eur Conference Echinoderm, Banyuls-Sur-Mer, France* (eds Féral J-P, David B), pp. 15–21. Swets & Zeitlinger, Lisse, NL.

Le Gac M (2002) *Détermination de l'échelle de différenciation spatiale et génétique de l'ophiure hermaphrodite incubante Amphipholis squamata (Delle Chiaje, 1828)*. DEA. University Pierre et Marie Curie, France.

Perrin C, Roy MS (2000) Rapid and efficient identification of microsatellite loci from the sea urchin, *Evechinus chloroticus*. *Molecular Ecology*, **12**, 2221–2223.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Sponer R, Roy M (2002) Phylogeographic analysis of the brooding brittle star *Amphipholis squamata* (Echinodermata) along the coast of New Zealand reveals high cryptic genetic variation and cryptic dispersal potential. *Evolution*, **56**, 1954–1967.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.