J Mol Evol (2007) 65:304–307
DOI 10.1007/s00239-007-9012-1

Lessons from a Beetle and an Ant: Coping with Taxon-Dependent Differences in Microsatellite Development Success

Wolfgang Arthofer · Birgit C. Schlick-Steiner · Florian M. Steiner · Dimitrios N. Avtzis · Ross H. Crozier · Christian Stauffer

Received: 15 January 2007 / Accepted: 6 July 2007 / Published online: 29 August 2007
© Springer Science+Business Media, LLC 2007

Abstract Microsatellites are powerful markers often isolated de novo for species yet to be investigated. Enriched genomic libraries are usually used for isolation purposes. We critically evaluate the outcome of an enrichment-based protocol applied to two insect species (the ant *Lasius austriacus* and the beetle *Pityogenes chalcographus*) which yielded contrasting numbers of suitable loci. Our findings of differences in microsatellite isolation are consistent with the available data on differences in genomic characteristics across these taxa. In the beetle repeated isolation of identical motifs, difficulties in primer development, and multibanded products caused loss of most candidate clones. We identified critical steps during marker development.

Keywords Microsatellites · Microsatellite isolation · Enriched genomic library · Molecular marker · Informative locus · Coleoptera · Hymenoptera

Reviewing Editor: Dr. John Oakeshott

W. Arthofer (✉) · B. C. Schlick-Steiner · F. M. Steiner · D. N. Avtzis · C. Stauffer
Institute of Forest Entomology, Forest Pathology and Forest Protection, Department of Forest and Soil Sciences, Boku, University of Natural Resources & Applied Life Sciences, Vienna, Austria
e-mail: wolfgang.arthofer@boku.ac.at

B. C. Schlick-Steiner · F. M. Steiner
Institute of Zoology, Department of Integrative Biology and Biodiversity Research, Boku, University of Natural Resources & Applied Life Sciences, Vienna, Austria

B. C. Schlick-Steiner · F. M. Steiner · R. H. Crozier
School of Marine and Tropical Biology, James Cook University, Townsville, Australia

Introduction

Microsatellites are short DNA stretches in which a one- to six-nucleotide motif is tandemly repeated. Found in all pro- and eukaryotic genomes, microsatellites show high mutation rates and intraspecific length polymorphisms, making them a powerful marker for population genetics. Development of universal, cross-species-amplifying primers for microsatellites is often not possible and the necessity of de novo isolation for taxa yet to be analyzed is a drawback in the applicability of microsatellites (Zane et al. 2002). Most new microsatellites are isolated using enriched genomic libraries. Marker development consists of (i) enrichment of microsatellite-containing fragments of genomic DNA, (ii) construction of a partial genomic library, (iii) screening of the library by colony hybridization, (iv) sequencing of candidate clones, (v) development of flanking primers, and (vi) exclusion of nonamplifying and monomorphic loci. Poor yields of suitable loci may occur due to technical problems in any of the steps but also due to peculiarities of the analyzed organism. Low abundance of microsatellites in the genome (Fagerberg et al. 2001) and occurrence of multicopy microsatellite families with similar flanking regions (Meglécz et al. 2004) have been considered reasons for such difficulties. We suspect, however, that in most cases isolation failure is not reported.

We evaluated the impact of taxon-specific traits on microsatellite development following the FIASCO protocol (Zane et al. 2002). Applying identical laboratory procedures to a beetle, *Pityogenes chalcographus* (Scolytinae), and an ant, *Lasius austriacus* (Formicinae), we address whether (i) numbers of micrsosatellite harboring clones in the enriched libraries, (ii) loss of candidate clones due to subsequent difficulties in amplification, and (iii) rate of polymorphism are comparable between the two species.

We discuss explanations for the contrasting findings and identify critical steps of the isolation process. Finally, we offer advice on some avoidable pitfalls.

Materials and Methods

DNA Preparation and Microsatellite Enrichment

Genomic DNA of 30 individuals was extracted with the GenElute Mammalian DNA kit (Sigma, St. Louis, MO, USA) and quantified photometrically, and 150 ng was used for a one-step digestion-ligation reaction with MseI and AFLP adaptors (Zane et al. 2002). Optimal cycle number for amplification with adaptor primers was determined by gel electrophoresis of 5-µl aliquots of the amplicon after 14, 17, …, 29 cycles. To avoid overamplification leading to high clone redundancy, a cycle number where a fine smear became visible for the first time was selected for the production of >1 µg of amplicon. Products were purified with the QIAquick PCR purification kit, then quantified photometrically, and integrity of the smear was tested on a minigel.

One microgram of purified PCR product was diluted with sterile water to 250 µl, denatured for 5 min in a boiling water bath, and cooled on ice. Twenty picomoles of each $(AC)_8$ and $(GA)_8$ biotinylated oligoprobe and 13 µl of 20× SSC were added and the final volume was set to 500 µl. The mixture was hybridized for 1 h at 50°C under constant agitation. Three hundred microliters of Streptavidine MagneSphere Paramagnetic Particles (PMPs; Promega, Madison, WI, USA) were washed three times with 0.5× SSC, resuspended in 100 µl of 0.5× SSC, and, after addition of the DNA-probe mixture, incubated for 20 min at room temperature under constant agitation. Afterward, PMPs were washed four times with 300 µl 0.1× SSC for high stringency or two times with 0.5× SSC and two times with 0.2× SSC for low stringency. DNA was eluted with 100 µl pre-warmed sterile water at 50°C. Recovery PCR was performed using adaptor primers.

Library Construction and Screening

DNA from recovery PCR was ligated into the pGEM vector (Promega) and used for transformation of competent JM109 cells. White colonies were transferred to master plates and, after 16 h, probed with sterile nylon membranes (Roche, USA). Membranes were hybridized overnight at 60°C in Church buffer containing 0.01% denatured salmon sperm DNA (Roche) and 1 µg of each digoxigenin labeled $(AC)_x$ and $(GA)_x$ oligoprobes. After washing, the presence of probe-target hybrids was detected using the DIG luminescent detection kit (Roche) following the manufacturer's instructions. Colonies showing the darkest signals on film were transferred to liquid culture for plasmid extraction.

PCR Pretest for Microsatellite Inserts

Two-tenths microliters plasmid DNA was used as template in a PCR containing 1.5 mM magnesium chloride, 50 µM dNTPs (Fermentas, Lithuania), 0.2 µM of each $(AC)_8$ and $(GA)_8$ oligonucleotide, 0.2 µM SP6 vector primer, 0.4 U Biotherm *Taq* DNA polymerase (Genecraft, Germany), and the reaction buffer provided by Genecraft. Amplification was performed with an initial denaturation at 94°C (3 min) followed by 32 cycles at 94°C (30 s), 59°C (45 s), and 72°C (45 s). Plasmids showing no bands were subjected to PCR with SP6 replaced by T7 vector primer. Plasmids with positive reaction in any of the two PCRs were classified as potential microsatellite isolates and sequenced.

Population Screening

Flanking primers of microsatellite carrying inserts were designed and tested on plasmid and genomic DNA. DNA of single individuals was extracted with the Sigma GenElute Mammalian DNA kit and used for screening. For loci yielding strong bands of expected size, forward primers were labeled with either 5'-FAM, 5'-HEX, or 5'-TET fluorophores. Retrieved microsatellite amplicons were supplemented with GeneScan 500-TAMRA size standard (Applied Biosystems, USA) and separated on an ABI 310 analyzer (Applied Biosystems). Allele sizes were determined using Genotyper software (Applied Biosystems).

Results

We constructed three libraries for *P. chalcographus* (P1, high stringency; P2 and P3, low stringency) and one for *L. austriacus* (L1, high stringency). After first adaptor-primed PCR, smears were comparable for all libraries, without bands indicative of accumulation of preferentially amplified loci. Three hundred two to 410 white colonies per library were picked for screening (Table 1). Sixty-two and one-half to 92.7% of the plasmids positive in both hybridization and PCR pretest harbored microsatellite inserts. No significant differences in the number of microsatellite-containing clones between the libraries were found. Sequencing revealed extreme redundancy in library P1, where all 15 plasmids carried in total five different alleles of a locus with the motif $(TC)_{16}$-TTCT-$(TC)_4$. In

Table 1 Microsatellite isolation by screening of genomic libraries enriched for $(AC)_x$ and $(GA)_x$ motifs of *Lasius austriacus* L1 and *Pityogenes chalcographus* P1–P3

Library	L1	P1	P2	P3
Stringency	High	High	Low	Low
Colonies screened	410	302	410	410
Positive pre-PCR	28	24	27	25
Inserts with microsatellite motif	26	15	24	20
Unique motifs	25***	1**	11	7
Defined loci	20**	1*	9	7
Informative loci	13***	1	2	1

Note. Absolute numbers of clones are shown; positive pre-PCR summarizes all clones positive in colony hybridization that also gave a signal in PCR with microsatellite and vector primers; inserts with microsatellite motif give the number of positive clones verified by sequencing, with multiple isolations of the same motif still included; defined loci are all loci where primers were developed; informative loci showed allele polymorphism after preliminary population screening. Deviations from expected frequencies were computed line by line using the chi-square test: $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

L1, only 1 of 26 motifs was isolated twice; P2 and P3 showed intermediate levels of redundant clones. Again, the above motif appeared in 40.8% and 44.0% of the plasmids, respectively. In total this locus was isolated 37 times.

After development of 20 primer pairs for *L. austriacus* 65.0% of the loci genotyped were polymorphic. Of 17 primer pairs developed for *P. chalcographus*, 23.5% gave polymorphic amplicons. The overall number of informative loci per screened colony was 0.032 for *L. austriacus* and 0.004 for *P. chalcographus*, respectively.

The informative loci of both species were further compared for quality traits (Table 2): PCR failure over all samples, repeat count *rc* of the cloned allele, and allelic richness *A* over all samples as estimated by using a sample coverage method (Huang and Weir 2001). The beetle microsatellites differ significantly from their ant counterparts concerning *rc* and *A*. PCR failure rate was comparable for both species.

Discussion

We modified the FIASCO protocol by adapting enrichment stringency and included a PCR pretest before plasmid sequencing, which excluded almost all false-positive clones. The frequency of microsatellite-containing clones across all libraries indicates comparable enrichment efficiency with both species.

Repeated isolation of identical motifs is a common feature of enriched libraries (Meglécz et al. 2004), with low microsatellite abundance in the genome (Fagerberg et al. 2001) hypothesized as the cause. Combined with high enrichment efficiency, it may dramatically increase redundancy. While there were almost no redundant sequences in the ant library, repeated isolation of identical motifs frequently occurred with the beetle. Finding different allelic forms of redundant loci in the libraries indicates that redundancy is due to enrichment bias rather than multiple ligation of identical PCR products.

Requirements necessary to turn unique microsatellite sequences into functional markers include suitability of the flanking regions for primer development, availability of sufficient single-specimen PCR product, and interpretable, polymorphic patterns in fragment analysis. The ant met these demands. Thirteen informative loci resulted from one enrichment procedure after excluding five loci because of short, unpromising motifs. As redundancy was negligible (1 of 26 motifs), we expect a scaled-up library screening to result in a multiple number of informative loci. In contrast, in the beetle the 19 unique loci were reduced by flanking regions unsuitable for primer development (2 loci), insufficient yield of PCR product (8 loci), multibanded products (3 loci), and monomorphism (2 loci).

The two species differ not only in the availability of polymorphic microsatellites, but also in their repeat count and allelic richness. While the latter is influenced by population structure, only in the ant was it possible to isolate long, uninterrupted motifs. Hymenopterans have a microsatellite-rich genome (Thoren et al. 1995). This contrasts with the situation in noncarabid coleopterans,

Table 2 Quality traits of informative loci isolated from the ant *L. austriacus* and the beetle *P. chalcographus*

Species	Loci	*n*	PCR failure				*rc*				*A*			
			Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
L. austriacus	13	30	0.27	0.14	0.07	0.50	28.0***	10.6	15	46	9.3*	4.1	2.0	16.6
P. chalcographus	4	288	0.25	0.12	0.14	0.42	11.5***	1.7	10	14	4.7*	1.5	2.6	6.2

Note. Thirteen and four informative loci for the ant and the beetle, respectively, were used in a preliminary screening of *n* individuals; PCR failure shows the portion of individual samples (1=100%) not yielding sufficient product for fragment analysis; *rc* is the repeat count of the cloned alleles; *A* refers to the allelic richness over all samples as estimated using the sample coverage method. Two-sided Student's *t*-test was used for taxon comparison: $^{***}p < 0.001$; $^{*}p < 0.05$.

where short motifs containing <10 repeat units as well as low allele numbers are common (e.g., Sallé et al. 2003) and loci variability often is insufficient for population genetics (e.g., Sallé et al. 2007). Similar results are reported for lepidopterans (Meglécz et al. 2004; Zhang 2004).

In addition to higher-level systematic differences, also closely related species may differ in microsatellite content (Ross et al. 2003). Microsatellites probably originate from substitutions (Zhu et al. 2000) and evolve by replication slippage (Rose and Falush 1998). Microsatellite content may mirror an equilibrium between slippage events increasing allele size and point mutations breaking up repeat sequences and ultimately leading to loci losses (Ross et al. 2003). While the role of genome size—microsatellites are frequent in noncoding regions and large genomes offer plenty of space to harbor them (Dieringer and Schlötterer 2003)—is not consistent for ants and beetles which show comparable C-values, base composition might be important. Equal abundance of the four bases should produce few microsatellites (Dieringer and Schlötterer 2003). Base composition can differ even between congeners (Lockhart et al. 1994), so that differences between species may reflect changes in mutational processes over a short evolutionary time.

Multicopy microsatellite families with similar flanking regions are characteristic of taxa refractory for microsatellite development (Meglécz et al. 2004). Zhang (2004) suggested a genomic dispersal of new microsatellites via transposition of mobile elements followed by accumulation of point mutations in the flanking regions leading toward single-copy loci. Low isolation efficiency and multibanded PCR amplicons support the hypothesis of a recent propagation of early-stage microsatellites in the apparently microsatellite-poor genome of *P. chalcographus*.

We demonstrate that poor results in microsatellite isolation are not necessarily a methodology problem. While a protocol might perform well for one group, switching to a new organism may cause problems. Based on the literature and our own experiences, the following critical steps in isolation are of particular importance. (a) Using a standard species with a high microsatellite content may help calibrating isolation efficiency. (b) Scaling-up will only compensate for low yield in the case of high rates of negative clones with some unique positives but not when low yield is caused by clone loss due to redundancy. Controlling redundancy might include stepwise adaptations of stringency as well as careful choice of capture probes for enrichment. (c) Attention often focuses on the outcome of plasmid sequencing, but awareness that as many as half of the isolated loci might never become functional markers is also important. Causes for difficulties include unsuitable primer binding sites, library contamination, suboptimal PCR amplification, and lack of polymorphism.

Acknowledgments We thank A. Stradner, H. Konrad, and K. Moder for valuable assistance, Editor-in-Chief, M. Kreitmann, Associate Editor, J. G. Oakeshott, and two anonymous referees for constructive criticism and the FWF for funding this project.

References

Dieringer D, Schlötterer C (2003) Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. Genome Res 13:2242–2251

Fagerberg AJ, Fulton RE, Black WC (2001) Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. Insect Mol Biol 10:225–236

Huang S, Weir BS (2001) Estimating the total number of alleles using a sample coverage method. Genetics 159:1365–1373

Lockhart PJ, Steel MA, Hendy MD, Penny D (1994) Recovering evolutionary trees under a more realistic model of sequence evolution. Mol Biol Evol 11:605–612

Meglécz E, Petenian F, Danchin E, Coeur d'Acier A, Rasplus JY, Faure E (2004) High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. Mol Ecol 13:1693–1700

Rose O, Falush D (1998) A threshold size for microsatellite expansion. Mol Biol Evol 15:613–615

Ross CL, Dyer KA, Erez T, Miller SJ, Jaenike J, Markow TA (2003) Rapid divergence of microsatellite abundance among species of *Drosophila*. Mol Biol Evol 20:1143–1157

Sallé A, Kerdelhué C, Breton M, Lieutier F (2003) Characterization of microsatellite loci in the spruce bark beetle *Ips typograhus* (Coleoptera: Scolytinae). Mol Ecol Notes 3:336–337

Sallé A, Arthofer W, Lieutier F, Stauffer C, Kerdelhué C (2007) Phylogeography of the spruce bark beetle *Ips typographus* in Europe, as revealed by microsatellites and mitochondrial DNA sequences. Biol J Linn Soc 90:239–246

Thoren PA, Paxton RJ, Estoup A (1995) Unusually high frequency of $(CT)_n$ and $(GT)_n$ microsatellite loci in a yellowjacket wasp, *Vespula rufa* (L.) (Hymenoptera: Vespidae). Insect Mol Biol 4:141–148

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

Zhang DX (2004) Lepidopteran microsatellite DNA: redundant but promising. Trends Ecol Evol 19:507–509

Zhu Y, Strassmann JE, Queller DC (2000) Insertions, substitutions, and the origin of microsatellites. Genet Res 76:227–236