Characterization of eight microsatellite loci for the sea urchin *Meoma ventricosa* (Spatangoida, Brissidae) through Next Generation Sequencing

Quentin Jossart¹, ², Laura B. Geyer³, H.A. Lessios³

¹Laboratoire de Biologie Marine (CP 160/15), Université Libre de Bruxelles (ULB), 50 avenue F. Roosevelt, 1050 Brussels, Belgium

²Biogéosciences, UMR CNRS 6282, Université de Bourgogne, 6 boulevard Gabriel, 21000 Dijon, France

³Smithsonian Tropical Research Institute, PO Box 0843-03092, Balboa, Panama

Abstract

Eight microsatellite loci were characterized for *Meoma ventricosa* (Lamarck, 1816), a burrowing sea urchin that can be afflicted by a bacterial disease causing localized mass mortality. For the analyzed population (29 individuals from St. Croix, US Virgin Islands), we observed 8.125 mean number of alleles, 0.640 mean observed heterozygosity ($H_o$) and 0.747 mean expected heterozygosity ($H_e$). Two loci showed significant deviations from Hardy-Weinberg equilibrium. Overall, the described loci were characterized by a moderately high level of polymorphism suggesting that these markers are useful for a population genetic study in the Caribbean Sea.

Keywords: Microsatellites, Echinoid, Caribbean, Tagged primer method, 454 method

Highlights

- 8 new microsatellite loci were isolated in the sea urchin *Meoma ventricosa*
- The moderately high level of polymorphism of these loci was estimated from a sample of 29 individuals of *M. ventricosa*
- The characterization of these markers permits subsequent population genetic studies
Introduction

*Meoma ventricosa* (Lamarck, 1816) (Spatangoida: Brissidae) is a common echinoid along the Caribbean and neighboring American coasts (Telford 1982; Hendler et al. 1995; De Bruyn et al. 2009). This large burrowing heart urchin (up to 20 cm in length) is found in sandy areas close to sea-grass beds or coral reefs as well as in deeper water. *M. ventricosa* lives in sediments ranging from fine sand to coral fragments (Chesher 1969; Hendler et al. 1995). It spends the daytime burrowed in the sediment and emerges at dusk to forage on the bottom (Kier & Grant 1965; Hammond 1982; Hendler et al. 1995). Chesher (1969) estimated that herds of *M. ventricosa* have a key influence on the disturbance of sand and consequently on the productivity of the areas it inhabits. Like the majority of echinoids, *M. ventricosa* is gonochoric. Fertilization occurs in the water column. The metamorphosis from blastula to first larval stage (echinopluteus) occurs 20 hours later in the pacific subspecies *M. ventricosa grandis* (Mortensen 1921). The larva is planktotrophic, but its pelagic larval duration (PLD) is unknown (Chesher 1969).

Although some molecular markers have been developed for this species (16S rDNA and Cytochrome Oxidase I (COI) mitochondrial genes, 28S rDNA nuclear gene) (Stockley et al. 2005), none has the high mutation rates of microsatellites. Microsatellites could provide information on genetic diversity, information relevant to conservation of this species, which is afflicted by a bacterial disease, sometimes causing localized mass mortality (Nagelkerken et al. 1999; Przeslawski et al. 2008). Microsatellites could also be helpful in evaluating the genetic structure of this species at a fine scale as well as in comparing this genetic structure with that of a parasitic crab of *M. ventricosa*.

Material and methods

A gonad sample from a single individual of *Meoma ventricosa*, collected in San Blas, Panamá and preserved in high salt DMSO buffer (Seutin et al. 1991), was chosen from the research collection of the Smithsonian Tropical Research Institute. We extracted DNA from this sample using the DNeasy Blood & Tissue kit (Qiagen). The DNA was shipped to Genome Québec (Montréal, Canada) for library construction and sequencing using the GS FL Titanium method on a 454 instrument (Roche), at ¼ plate scale. This sequencing run yielded over 179,000 reads. These data were searched for simple sequence repeats (SSRs) with MSATCOMMANDER 1.0.8beta (Faircloth 2008), using a minimum search criterion of 8 di-nucleotide repeats. This search yielded 15,564 potential microsatellite loci. The sequences of these potential loci were
processed with Primer3 (Rozen & Skaletsky 2000) for primer design using a CAG-tag (CAGTCGGGCGTCATC) (Boutin-Ganache et al. 2001), limiting the product size to 100-450 bp. Primer3 identified 913 potential primer pairs. Duplicates were eliminated, as were those primers with potential for hairpin formation, self-annealing and incomplete CAG-tag sequences. The sequences of the remaining 149 loci were examined to eliminate any with sequence motifs likely to create problems during genotyping (sequences with long monomer strings and multiple SSRs). This reduced a potential list to 30 loci for which primers were ordered (Integrated DNA Technology).

Each of the 30 primer sets was tested for amplification using standard PCR conditions (described below) and an annealing temperature gradient of 48-60 °C. Nine potential loci failed to amplify, and six yielded multiple products. The 15 remaining loci were amplified using a fluorescently labeled CAG-tag and genotyped using an ABI 3130XI Genetic Analyzer. The resulting electropherograms were examined using GeneMapper 4.0 (Applied Biosystems). Seven additional loci were eliminated because they either resulted in uninterpretable multiple peaks or in allele sizes clearly outside the predicted range.

Twenty nine specimens of *Meoma ventricosa* were collected by SCUBA diving (May 2011) at one site in St. Croix, US Virgin Islands (17°39’N, 64°48’W). DNA was extracted from two spines of each individual using the Qiagen DNeasy Blood & Tissue Kit. Eight microsatellite loci (Table 1) were amplified in simplex according to the tagged primer-method (Schuelke 2000). Each reaction (ca. 12 µl) included: 1µl of extracted DNA, 0.030 µM of forward/reverse primers with CAG tail (CAGTCGGGCGTCATC), 0.5 µM of universal dye labelled CAG primer, 0.5 µM 1 of forward/reverse primers without CAG tail, 2.4 µl of Buffer (Promega), 1 µM of dNTP (Promega), 2.5 mM of MgCl₂, 0.8M of Betaine (Affymetrix), 2.9 µl of sterile water and 0.04 U of Taq Polymerase (Promega). For loci CHCM and NLQK, 2.86 µl of water and 0.03 U of Taq were used.

The PCR conditions were: 14 cycles of 20 s at 95 °C (denaturation), 20 s at 50-59 °C (first annealing temperature, (Table 1) and 20 s at 72 °C (elongation); followed by 25 cycles of 20 s at 95 °C, 20 s at 52.9 °C (annealing temperature of the CAG primer) and 20 s at 72 °C. These cycles were preceded by 1 min at 95 °C (first denaturation) and were followed by 5 min at 72 °C (last elongation). Using band intensities on agarose gels as a rough estimate of product concentration, we mixed 0.5 - 2.5 µl of amplified DNA with 0.5 µl of GeneScan 500 LIZ size standard (Applied Biosystems) and 10.5 µl of formamide prior to genotyping with an ABI 3130xL Genetic Analyzer. Genotypes were deduced from electropherograms using the software
Peak Scanner (Applied Biosystems). Allelic binning was done using the Excel macro Autobin (Guichoux et al. 2011) but each genotype was also checked by eye.

We used Micro-Checker (Van Oosterhout et al. 2004) to detect potential genotyping errors. Using GENEPOP 4.2.2 (Rousset 2008), we evaluated the observed ($H_o$) and expected heterozygosities ($H_e$) as well as deviation from Hardy-Weinberg (HW) equilibrium, the presence of linkage disequilibrium and the number of alleles.

**Results and discussion**

In two loci (TCYO, 6SKB), Micro-Checker detected the probable presence of null alleles (Brookfield 1 estimator) with frequencies of 0.15 (TCYO) and 0.22 (6SKB) (Brookfield 1996). Average $H_o$ and $H_e$ over all loci were 0.640 (0.423-0.923) and 0.747 (0.571-0.933), respectively (Table 1). Genotype frequencies in all loci, except for those with inferred null alleles, were not significantly different from HW expectations. For the loci with null alleles, $F_{IS}$ values were equal to 0.338 (TCYO) and 0.503 (6SKB). Two loci (TCYO, CHCM) seem to be at linkage disequilibrium ($p < 0.0018$, Bonferroni-corrected). The number of alleles per locus varied between 3 and 17 with an average of 8.125 (Table 1). All the individuals amplified for most of the eight markers with a percentage of missing genotypes equal to 9% (CHCM: 10%, 9901: 14%, CHTO: 0%, 93XX: 10%, NLQK: 17%, 9U51: 0%, 6SKB: 10%, TCYO: 10%).

Overall, we selected 8 loci from 15,564 potential microsatellite loci. This reflect the classical high post-sequencing selection associated with Next Generation Sequencing (Fernandez-Silva et al. 2013). The eight loci showed moderately high polymorphism, which suggests that the microsatellites can be useful for a broader population genetics study. The two loci associated to the possible presence of null alleles should not be discarded for further analyses but some precautions should be taken to account for the null allele frequency (eg. adjusted $F_{ST}$ values calculated with the software FreeNA, Chapuis & Estoup 2007).

The characterization of microsatellites of *Meoma ventricosa* permits the comparison of genetic structure of this species with that of its parasite, the crab *Dissodactylus primitivus* that may have different capacity of dispersal than its host (Yednock & Neigel 2011; Jossart et al. 2014). Microsatellite markers have already been characterized and validated for this crab (Anderson et al. 2010; Jossart et al. 2013). A comparative study could reveal the factors that influence dispersal of each species and provide information about their potential for local adaptation (Greischar & Koskella 2007).
Acknowledgements

This work was supported by a PhD grant (FRF-FNRS, Fonds de la Recherche Scientifique 1.E047.11F), two travel grants (Fédération Wallonie-Bruxelles and FNRS) to Quentin Jossart as well as a grant from the Conseil régional de Bourgogne (PARI 2012-0780) for 454-sequencing. We thank Génome Québec for their help in the first steps of development; Ligia Calderón for her inestimable help in this characterization; Giomar Borrero for helpful advice; Chantal De Ridder, Bruno David, Sébastien Motreuil and Jean-Philippe Maréchal for assistance in sampling; C. De Ridder for comments on the manuscript. We also thank Thierry Rigaud and Rémi Wattier for interesting discussions on molecular aspects of the work.

References


Kier PM, Grant RE (1965) Echinoid distribution and habits, Key Largo coral reef preserve, Florida. Smithsonian Miscellaneous Collections 149, 1-68.
Mortensen T (1921) Studies of the development and larval forms of echinoderms. G.E.C. GAD.
<table>
<thead>
<tr>
<th>Locus</th>
<th>Motif</th>
<th>Primer sequence</th>
<th>GenBank accession no.</th>
<th>Ta</th>
<th>Size range (bp)</th>
<th>Na</th>
<th>Ho</th>
<th>He</th>
</tr>
</thead>
</table>
| CHCM   | (CA)$_{13}$ | F: CCTGACAAGTTGACCACACG  
R: CAGTCGGGCGTCATCAGGGGAACGAGCATAGAACCAGG | KJ875804 | 50 °C | 202-234 | 17 | 0.923 | 0.933 |
| 9901   | (AC)$_{10}$ | F: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC | KJ875805 | 51 °C | 235-243 | 5 | 0.720 | 0.657 |
| CHTO   | (GT)$_{8}$  | F: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC  
R: AACGGAGCTAAGCCCTTCTG | KJ875806 | 59 °C | 314-326 | 5 | 0.517 | 0.612 |
| 93XX   | (AT)$_{15}$ | F: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC  
R: AGCTCAGCGTGCTGAATGTGAAG | KJ875807 | 57 °C | 156-174 | 7 | 0.885 | 0.784 |
| NLQK   | (CT)$_{11}$ | F: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC  
R: AGCTCAGCGTGCTGAATGTGAAG | KJ875808 | 50 °C | 138-146 | 5 | 0.625 | 0.661 |
| 9U51   | (AC)$_{9}$  | F: GGCATTCGAGTTCTGACAGC  
R: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC | KJ875809 | 56 °C | 169-173 | 3 | 0.448 | 0.571 |
| 6SKB   | (AC)$_{10}$ | F: GTCCATGGTTCGAGTTCTGACAGC  
R: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC | KJ875810 | 59 °C | 321-341 | 10 | 0.423 | 0.851* |
| TCYO   | (CA)$_{17}$,TA(CA)$_{13}$,CT(CA)$_{4}$  
F: CAGTCGGGCGTCATCAGGATGCTACACGCTAAGCCGCTC | KJ875811 | 59 °C | 322-348 | 13 | 0.615 | 0.930* |

Table 1. Characterization of eight microsatellite loci for *Meoma ventricosa*. Ta = annealing temperature, Na = Number of alleles, Ho = Observed heterozygosity, He = Expected heterozygosity, * indicates significant deviation from Hardy-Weinberg equilibrium. CAGTCGGGCGTCATC of primer sequence corresponds to the CAG-tail.