TITLE: Isolation and characterization of microsatellite markers in the portunid crab Liocarcinus depurator using FIASCO and 454 next-generation sequencing

F. PALERO*(1), VH. GARCIA-MERCHAN(2), P. ABELLÓ(3), E. MACPHERSON(4) and M. PASCUAL(2)

(1) Unitat Mixta Genòmica i Salut CSISP-UV, Institut Cavanilles Universitat de Valencia, C/Catedrático Jose Beltran 2, 46980 Paterna, Spain (2) Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, 08028 Barcelona, Spain (3) Institut de Ciències del Mar (CSIC), Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain (4) Centre d'Estudis Avançats de Blanes (CSIC), Carrer d'Accés a la Cala Sant Francesc 14, 17300 Blanes, Spain

KEY WORDS: Portunidae, decapod crustaceans, pyrosequencing, SSRs

Corresponding author: FERRAN PALERO Unitat Mixta Genòmica i Salut CSISP-UV, Institut Cavanilles Universitat de Valencia, C/Catedrático Jose Beltran 2, 46980 Paterna, Spain
e-mail: ferran.palero@uv.es

Running title: Microsatellite variation in *Liocarcinus depurator*

Abstract

The portunid crab *Liocarcinus depurator* is the dominant decapod crustacean present on mud bottoms of the continental shelf of the Mediterranean Sea. In order to better describe its genetic structure, a total of 11 microsatellites were isolated and tested in 28 *L. depurator* individuals from the Mediterranean locality of Tarragona. Initially, microsatellite isolation was carried out using the FIASCO methodology, but only 2 useful primer pairs could be designed. Given the low yield obtained, an alternative method based on 454 next-generation sequencing was assayed, which provided excellent results and from which another 9 primer pairs gave positive genotyping. All loci tested were polymorphic, and 2 to 24 (mean 10.54) alleles per locus were identified. The observed and expected heterozygosities per locus ranged from 0.214 to 0.926 and from 0.232 to 0.962, respectively. One locus (Ldep10) showed the presence of null alleles according to the Brookfield estimator implemented in MICROCHECKER. Nevertheless, none of the markers deviated significantly from Hardy–Weinberg equilibrium or showed evidence of linkage disequilibrium. This set of 11 markers is being used to study the population structure and genetic diversity of *Liocarcinus depurator* populations from both Mediterranean and adjacent Atlantic waters.

The portunid crab *Liocarcinus depurator* is distributed along the coasts of the Eastern Atlantic and around the Mediterranean and is the dominant decapod crustacean on mud bottoms of the continental shelf (Abelló et al., 1988; d'Udekem d'Acoz, 1999). In the present study, isolation of *L. depurator* microsatellite loci was conducted using both the FIASCO protocol and through next-generation sequencing techniques. In the first case, we constructed a partial and enriched genomic library following the FIASCO protocol (Zane et al., 2002). Coxal tissue from a pereiopod of an individual *L. depurator* collected from the Tarragona region (41.075 N − 1.280 E) and preserved in 100% ethanol was used to obtain genomic DNA using the QIAamp DNA extraction kit (Qiagen Inc). DNA was digested with MseI (New England BioLabs Inc.) and ligated to adapters (MseI-A, 5-TACTCAGGACTCAT-3; MseI-B, 5-GACGATGAGGTCCTGAG-3) for 3 hours. Enrichment was performed using streptavidin magnetic beads (Streptavidin Magnesphere paramagnetic Particles, Promega) and four biotinylated probes [$(CA)_{15}$, $(GA)_{15}$, $(CAA)_7$ and $(GATA)_7$]. The recovered DNA was then amplified by PCR using primers for the MseI-N adapter (5-GATGAGTCCTGAGTAAN-3) and subsequently cloned using the pGEM-T Easy Vector System II (Promega) following the supplier's protocol. Positive clones were detected using probes labeled with digoxigenin (http://www.inapg.inra.fr/dsa/microsat/microsat.htm). Approximately 1000 colonies were screened for microsatellites, originating 185 positive clones, 104 of which were sequenced. DNA was extracted with the QIAprep Spin Miniprep Kit (Qiagen Inc.) and two sequencing reactions were performed using the primers T7 (5-TAATACGACTCACTATAGGG-3) and SP6 (5-ATTTAGGTGACACTATAGAA-3) respectively. Sequencing reactions were carried out in a final volume of 10 uL in a GeneAmp PCR System 2700 thermocycler (Applied Biosystems). Sequences were obtained using an automatic sequencer ABI PRISM 3700 from the Technical and Scientific services at the University of Barcelona.

Using this FIASCO method, a total of 136 microsatellite loci were isolated (83.1% perfect, 9.5% imperfect and 7.4% compound) with an average repeat length of 15.08 ± 10.11. Among the clones, dinucleotides were much more abundant (79%) than trinucleotides (15.4%) or tetranucleotides (5.6%). AG dinucleotides were more frequent (56.5%) than those with AC repeats (22.4%). A total of 17.2% of the clones showed too short flanking regions to design primers. Finally, a total of 34 microsatellite loci were evaluated and their amplification conditions subjected to various combinations of temperature. However, due to low quality and resolution of bands amplified and observed on agarose gels, only 2 markers (Ldep06 and Ldep10) were kept. Given the low yield obtained with the FIASCO methodology, an alternative method based on pyrosequencing or 454 next-generation sequencing was assayed (Margulies et al., 2005). In this case, total genomic DNA was extracted from a pleopod of *L. depurator* using the QIAamp extraction kit (Qiagen), adding 4 ul (100 mg / ml) RNAse and modifying step 8 of the supplier's protocol (the resuspension of the DNA pellet). In this step, three resuspensions were performed, each in 20 ul of distilled water, maintained at room temperature for 15, 10 and 10 minutes respectively, then centrifuged at 8000 rpm and finally mixed to obtain a volume of 60 ul at a concentration of approximately 21 ng/ul. DNA quality was checked by agarose gel electrophoresis. After 454 pyrosequencing at the scientific and technical services of the University of Barcelona, a total of 94 196 DNA sequences (of which 90 584 were unique) were obtained. The N50 was 325 bp (N75 = 229 bp, N25 = 388 bp) allowing the capture of microsatellites with enough flanking region to design specific primers. The microsatellite detection, sequence size selection and primer design, was carried out using a custom PERL pipeline along with the QDD software (Meglecz et al., 2010). Given the large number of sequences obtained, we only kept those reads longer than 400 bp and for which it was feasible to design primers amplifying fragments larger than 150 bp. With these criteria, most microsatellites recovered were trinucleotides (70%), followed by dinucleotides (9%), tetranucleotides (13%) and some penta- and hexanucleotides (8%). Primers were designed for 171 sequences, including mostly microsatellites with tri- or tetranucleotide repeats. Of these, 16 microsatellites were finally evaluated and nine of those amplified correctly (Ldep01-Ldep05, Ldep07-Ldep09 and Ldep11).

Finally, the set of 11 microsatellite markers were amplified in 28 individuals from the Tarragona

population using two multiplex reactions (Table 1). Each PCR reaction consisted of an initial temperature of 95 °C for 10 minutes, followed by 27 cycles of denaturation at 94 °C for 40 seconds, primer binding at 54 °C for 30 seconds, extension at 70 °C for 30 seconds, and a final extension at 60 °C for 20 minutes. Each reaction was carried out in a final volume of 11 ul, containing master mix multiplex PCR kit (Qiagen) at 0.45x concentration, Primer Mix (2 mM of each primer) at 0.16x concentration and ~30-60 ng DNA. The forward primers were labeled with the fluorochromes NED, HEX or 6FAM (Table 1). The PCR products were analyzed on an automatic sequencer ABI PRISM 3700 of the scientific and technical services at the University of Barcelona. The sizes of the alleles were assigned with an internal size marker CST ROX 70-500 (BioVentures, Inc.) and using Microsatelight (Palero et al., 2011).

All loci were polymorphic, with the number of alleles ranging from 2 to 24 and the observed heterozygosity from 0.214 to 0.926 (Table 1). The allelic richness and the observed heterozygosity were similar to those found in other decapod crustaceans such as *Eriocheir sinensis* (Sui et al., 2009) or *Palinurus elephas* (Palero and Pascual, 2008). Null alleles were detected for the Ldep10 locus when using the Brookfield estimator as implemented in MICROCHECKER (Van Oosterhout et al., 2004). The Hardy–Weinberg equilibrium (HWE) and the genotypic linkage disequilibrium between pairs of loci were tested using GENEPOP v.4.1 (Rousset, 2008). Our results showed that most loci were in HWE, with loci Ldep02 and Ldep09 deviating marginally and not-significantly after Bonferroni correction. The exact test for genotypic linkage disequilibrium was not significant among any pair of loci and thus can be considered independent. The microsatellite markers reported here provide robust, high resolution tools for efficient genetic studies, e.g. to contrast current levels of genetic variation and population differentiation among *Liocarcinus* from Atlantic and Mediterranean localities.

Acknowledgements

Research was funded by projects BIOCON08-187 from Fundación BBVA and CTM2010-22218 from the Ministerio de Educación y Ciencia. The authors are part of the research group 2009SGR-636, 2009SGR-655 and 2009SGR-1364 of the Generalitat de Catalunya.

References

Abelló P, Valladares FJ, Castellón A (1988) Analysis of the structure of decapod crustacean assemblages off the Catalan coast (North-West Mediterranean). *Marine Biology,* **98**, 39–49.

d'Udekem d'Acoz C (1999) Inventaire et Distribution des Crustacés Décapodes de l'Atlantique Nord-Oriental, de la Méditerranée et des Eaux Continentales Adjacentes au Nord de 25° N. *Patrimoines naturels (M.N.H.N./S.P.N.)*, **40**, 1–383.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* **437**, 376–380.

Meglecz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF (2010) QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics,* **26**, 403–404.

Palero F, Pascual M (2008) Isolation and characterization of microsatellite loci in *Palinurus elephas*. *Molecular Ecology Resources,* **8**, 1477–1479.

Palero F, González-Candelas F, Pascual M. 2011. MICROSATELIGHT – pipeline to expedite microsatellite analysis. *Journal of Heredity* 102: 247–249.

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources,* **8**, 103–106.

Sui LY, Zhang FM, Wang XM, Bossier P, Sorgeloos P, Hanfling B (2009) Genetic diversity and population structure of the Chinese mitten crab *Eriocheir sinensis* in its native range. Marine Biology, **156**, 1573–1583.

Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

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

Table 1. Characteristics of 11 polymorphic microsatellite loci obtained from *Liocarcinus depurator* and used for genotyping 28 individuals: locus name, primer sequences, repeat motif, multiplex group (both at 54ºC), size range of alleles (in base pair), allele richness (AR), observed (Ho) and expected (He) heterozygosities and GeneBank accession numbers.

Locus	Primer sequence (5'-3')	Motif	Group	Fluorochrome	n	Range	AR	Ho	He	Accession
Ldep01	F: CATTCTCTGTGCTGGTGTGAAT R: AAAAGAGCGTGGACCAATTTTA	$(ACAG)_{11}$	A	NED	28	113 - 223	9.7	0.893	0.887	JX104535
Ldep02	F: TGGTAGTGGCTATGGTAACGAG R: CGGTCAGGAATTAAAGGATTTG	$(ACC)_5$	A	NED	28	269 - 314	3.9	0.286	0.350	JX104536
Ldep03	F: TACACGCTCCCTACCTTCGTAT R: GCGGATCATGTGGTGATTATTA	$(AGT)_{12}$	A	6FAM	28	243 - 294	9.1	0.821	0.875	JX104537
Ldep04	F: ATGTGGTACTCGATGTCTTCCC R: AACGCTCACTCTTCCTTGTCTC	$(ACC)_8$	A	6FAM	28	110 - 131	6.3	0.750	0.811	JX104538
Ldep05	F: TGTCTTGAGGCACGATAAACAC R: CATCCTTTCTTGCTTTTGCTCT	$(ACG)_8$	A	HEX	28	106 - 139	6.4	0.536	0.618	JX104539
Ldep06	F: ACCCTCCCTCTATTCACATCGTC R: AGTCACACCTGCCCGTTGAAGG	$(CA)_{10}$	B	NED	28	148 - 185	8.6	0.786	0.858	JX104540
Ldep07	F: AGGAAGTTGCAGACGAGTAAGC R: GAACACAAGAGAAAATCCTCGC	$(ACC)_7$	B	NED	28	258 - 268	7.5	0.857	0.872	JX104541
Ldep08	F: CGTTGAAAGTCACTATGGGTCA R: TGTTATGTAATGCTGAGGTGGC	$(ACT)_8$	B	6FAM	28	145 - 154	2.8	0.214	0.232	JX104542
Ldep09	F: ATGAAGCATGGCTGGTTAATTT R: TTCACAGTTACAATATTCAAAGCAA	$(AGT)_5$	B	6FAM	28	274 – 277	2.0	0.250	0.456	JX104543
Ldep10	F: TCACTGCAGAACAGGACGAGC R: TGAGAGTGTAACCAACGCGGTC	$(CA)_{24}$	B	HEX	27	270 - 341	14.4	0.926	0.962	JX104544
Ldep11	F: ATATTTTCCTTGCCTCCCTTTC R: GTATTGCACCTCAGGACACGTA	$(AAT)_9$	B	HEX	28	144 - 156	4.4	0.464	0.529	JX104545

The multiplex groups A and B are defined so that the fragments labeled with the same fluorochrome are non-overlapping. Allelic richness values are based on a minimum sample size of 10 diploid individuals.

Sheet1

Locus	Primer sequence (5'-3')	Motif	Group
Ldep01	F: CATTCTCTGTGCTGGTGTGAAT	$(ACAG)_{11}$	A
	R: AAAAGAGCGTGGACCAATTTTA		
Ldep02	F: TGGTAGTGGCTATGGTAACGAG	$(ACC)_5$	A
	R: CGGTCAGGAATTAAAGGATTTG		
Ldep03	F: TACACGCTCCCTACCTTCGTAT	$(AGT)_{12}$	A
	R: GCGGATCATGTGGTGATTATTA		
Ldep04	F: ATGTGGTACTCGATGTCTTCCC	$(ACC)_8$	A
	R: AACGCTCACTCTTCCTTGTCTC		
Ldep05	F: TGTCTTGAGGCACGATAAACAC	$(ACG)_8$	A
	R: CATCCTTTCTTGCTTTTGCTCT		
Ldep06	F: ACCCTCCCTCTATTCACATCGTC	$(CA)_{10}$	B
	R: AGTCACACCTGCCCGTTGAAGG		
Ldep07	F: AGGAAGTTGCAGACGAGTAAGC	$(ACC)_7$	B
	R: GAACACAAGAGAAAATCCTCGC		
Ldep08	F: CGTTGAAAGTCACTATGGGTCA	$(ACT)_8$	B
	R: TGTTATGTAATGCTGAGGTGGC		
Ldep09	F: ATGAAGCATGGCTGGTTAATTT	$(AGT)_5$	B
	R: TTCACAGTTACAATATTCAAAGCAA		
Ldep10	F: TCACTGCAGAACAGGACGAGC	$(CA)_{24}$	B
	R: TGAGAGTGTAACCAACGCGGTC		
Ldep11	F: ATATTTTCCTTGCCTCCCTTTC	$(AAT)_9$	B
	R: GTATTGCACCTCAGGACACGTA		

Sheet1

Fluorochrome	n	Range	AR	Ho	He	Accession
NED	28	113 - 223	9.7	0.893	0.887	JX104535
NED	28	269 - 314	3.9	0.286	0.350	JX104536
6FAM	28	243 - 294	9.1	0.821	0.875	JX104537
6FAM	28	110 - 131	6.3	0.750	0.811	JX104538
HEX	28	106 - 139	6.4	0.536	0.618	JX104539
NED	28	148 - 185	8.6	0.786	0.858	JX104540
NED	28	258 - 268	7.5	0.857	0.872	JX104541
6FAM	28	145 - 154	2.8	0.214	0.232	JX104542
6FAM	28	274 – 277	2.0	0.250	0.456	JX104543
HEX	27	270 - 341	14.4	0.926	0.962	JX104544
HEX	28	144 - 156	4.4	0.464	0.529	JX104545