Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site

Molecular Ecology Resources
May 2009, Volume 9 Issue 3, Pages 853 - 855
http://dx.doi.org/10.1111/j.1755-0998.2009.02525.x
© 2009 Blackwell Publishing, Inc.

The definitive version is available at www.blackwell-synergy.com

Archimer, archive institutionnelle de l'Ifremer
http://www.ifremer.fr/docelec/

Identification and characterization of 18 novel polymorphic microsatellite makers derived from expressed sequence tags in the Pacific oyster *Crassostrea gigas*

C. Sauvage[1], P. Boudry[1, §], S. Lapègue[1, *]

[1] Ifremer - Laboratoire de Génétique et Pathologie - La Tremblade, France
[§] Present address : Ifremer –UMR M100 Physiologie et Ecophysiologie des Mollusques Marins – Plouzané, France.

*: Corresponding author : S. Lapègue, Tél : +33 (0)5 46 76 26 30 ; Fax : +33 (0)5 46 76 26 11, email address : slapegue@ifremer.fr

Abstract:

We report the development of 18 new polymorphic microsatellite DNA markers derived from Crassostrea gigas expressed sequences tags. Genotyping of 48 wild adult oysters sampled from Marennes-Oléron bay (France) revealed 12 to 48 alleles per locus. Observed and expected heterozygosity levels ranged from 0.64 to 1 and from 0.77 to 0.97, respectively. The development of these new markers creates a useful complementary tool for population genetics studies, parentage analysis and mapping in Pacific oyster, a species of major aquacultural and ecological importance.

Keywords: Crassostrea gigas • EST-SSR markers • null alleles • Pacific oyster

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site

The Pacific cupped oyster is one of the most widely cultivated aquatic species with a world annual production of 4.5 millions metric tons in 2005 (FAO, 2006). It has been introduced into many countries, and may naturalise and become invasive. Most production is based on collection of wild seed and this species has therefore benefited little from genetic improvement by selective breeding. A significant genetic basis has been shown for several traits of aquacultural interest, such as resistance to summer mortality (Degremont et al., 2007), indicating the potential for genetic improvement of cultured stocks. Moreover, the development of genomic resources is increasing (Hedgecock et al., 2005; Jenny et al., 2007; Tanguy et al., 2008). Current linkage maps include around 100 AFLP (Li and Guo, 2004) or 100 microsatellite markers (Hubert and Hedgecock, 2004). Higher density maps, required for QTL or genome scan studies, rely on the development of additional microsatellite (Yu and Li, 2007, Wang et al., 2008) or SNP (Sauvage et al., 2007) markers. However, high frequencies of null alleles (Hedgecock et al., 2004) and segregation distortions (Launey and Hegecock, 2004) have been frequently observed in C. gigas, leading to additional constraints to the development of high density linkage maps.

Here we present the identification and the characterization of a new set of 18 EST-SSR markers, recently reported to present lower frequencies of null alleles than genomic SSRs (Yu and Li, 2008). A total number of 9272 EST contigs assembled from the Public sequences and the Marine Genomics database (http://www.sigenae.org/) were screened to identify every repeated tandem sequence (from penta to dinucleotide) using the Cotton Microsatellite Database Simple Sequence Repeat Server (Clemson University Genomics Institute; http://www.mainlab.clemson.edu/cm/ssr_server/). A total of 22 primers pairs were

designed from the available sequences (see Table 1 for Accession numbers) using the Primer 3 program (Rozen, 2000; http://frodo.wi.mit.edu/).

PCR amplifications were performed as follows: an initial denaturation step of 2 min at 94°c was performed, and then the annealing temperature was decreased, from 65 to 55°C, in increments of 1°C for every subsequent set of 2 cycles, and followed by an extension step of 1 min at 72 °C. When the lowest annealing temperature was reached, it was used for the 25 remaining cycles. The PCR mix was composed of 0.3U *Taq* polymérase (New England Biolabs), 10 mM provided buffer, 1 mM $MgCl_2$, 2 mM dNTP (Eurogentec), 10mM of each primer and 100 ng genomic DNA in a final reaction volume of 10 µL. Capillary electrophoresis was performed on a ABI 3130xl genetic analyser (Applied Biosystems) and alleles were scored using the software Genemapper version 4 (Applied Biosystems).

Eighteen SSR out of 22 gave PCR products of expected size and showed only two alleles. Descriptive statistics (table 1) on the genotypes from the 48 adult samples were obtained using Genetix 4.1 (Belkhir et al., 1996). The number of alleles per locus ranged from 12 to 48, with mean of 29.9. The expected and observed heterozygosity estimates, ranged from 0.77 to 0.97 and from 0.64 to 1, respectively. These results are similar to those previously published by Huvet et al. (2000). Testing for Hardy-Weinberg Equilibrium (HWE) was carried out using the Markov chain method (100 000 steps; 10 000 dememorisation) using ARLEQUIN Version 3.1 (http://cmpg.unibe.ch/software/arlequin3/). Three out of 18 loci (Cgsili3, 15 and 57) showed a significant deviance from equilibrium. Such deviations are unusually frequent in oysters and were recently observed in a set of EST-SSR developed by Yu and Li (2007) in the Pacific oyster. In most cases, such deviations are due to the high frequency of null alleles (Hedgecock et al., 2004). Moreover, genotypic linkage disequilibrium was detected between three markers (Cgsili3, 6 and 7; $P>0.05$). The null allele frequency (*r)* was computed using the Cervus program (Marshall et al., 1998; http://www.fieldgenetics.com/pages/aboutCervus_Overview.jsp) and results show that this frequency is low ($r>0.5$ for 17% of loci), which is as expected because the frequency of null alleles in EST-derived SSRs is usually lower than in genomic SSRs due to lower mutation at splicing sites (Goldstein and Schlötterer, 1999).

These novel microsatellite loci, developed by an *in silico* screening approach will be useful for further population genetics, parentage analysis and linkage mapping studies in the Pacific oyster.

Acknowledgements

This work was the supported by the EU projects Marine Genomics, Aquafirst and the Bio-Informatics platform SIGENAE. We thank Helen McCombie for editing the English.

Table 1. EST-SSR isolated in *Crassostrea gigas* with: primer sequences, number of individuals genotyped (Ni), observed heterozygosity (Ho), expected heterozygosity (HE), P-Value for Hardy-Weinberg Equilibrium testing (HWE) and calculated frequency of null alleles in 48 individuals.

Locus	Repeat sequence	Primers sequences (5'-3')	Ni	Na (size range, bp)	Ho	He	HWE	Null alleles Frequency (*r*)	Accession no.
Cgsili3	$(TA)_{27}$	F: TGAAATTAACAACCCTTGGT R: TTGAATCTAAAGAATATGGCAA	39	36 (324-403)	0.641	0,9589	0.032*	+0.1995	AM857068
Cgsili4	$(AG)_{26}$	F: GGTGCAGTAGTTGGAAACAT R: TCACATTTAACTAGCGCTCTC	43	45 (267-397)	0.7907	0,9594	0.431	+0.0976	AM854894
Cgsili6	$(GA)_{26}$	F: ATGAACGTCCAAGTTCAGAC R: ACACATTTCCTTATAAAGCC	46	48 (238-410)	0.913	0.97	0.221	+0.0301	AM854296
Cgsili7	$(GA)_{24}$	F: ACAACGCTATCAGAACCATT R: ATCTCCCGGCAAGTATATG	44	32 (209-300)	0.8864	0.9122	0.080	+0.0077	CU682571
Cgsili12	$(CAATG)_{5}$	F: CCATCTTCTGTTTGGTTCAT R: GCACCGGATGGTAGATATT	45	40 (155-242)	1.0000	0.9644	0.057	-0.0183	AM857087
Cgsili15	$(TCCA)_{6}$	F: TCTCTTTCCATCCATCAATC R: GGTGATATTGTAACGGCATT	42	16 (200-230)	0.595	0.763	0.021*	+0.1289	AM856504
Cgsili26	$(GA)_{13}GC(GA)_{7}$	F: CCCACTAAAGCTACCTATTGA R: TGTGGATTCCTTAACTGACC	47	30 (202-278)	1.0000	0.9124	0.689	-0.0522	AM855372
Cgsili29	$(AG)_{20}$	F: TTTATATGCGATTTCTGAAGC R: TTGACTAAGAATTCATCGGG	42	38 (302-400)	0.9524	0.9592	0.744	+0.0030	AM856933
Cgsili37	$(TC)_{15}$	F: TTGCTGGTTGTGATGAATAG R: ATATCTGGCCTAACATGTGC	37	32 (109-190)	1.0000	0.9592	0.451	-0.0210	BQ427164
Cgsili38	$(TG)_{6}C(GT)_{8}$	F: TTTCTGCTCTGTGTGTGTGT R: ACGTGGAGTGACGCTACTAT	34	23 (306-391)	0.9412	0.9325	0.332	-0.0060	AM854072
Cgsili39	$(AG)_{13}$	F: GACCATACAGCTCTGTCCAT R: GCTACTGAATGAGAATGGCT	29	30 (331-445)	1.0000	0.9566	0.094	-0.0237	AM854746
Cgsili43	$(GT)_{10}$	F: AAATGCTGCAGAAATAATCC R: AGATGGCTACAGTGAAATGG	43	15 (340-358)	0.8837	0.8807	0.077	-0.0013	AM854072
Cgsili44	$(AG)_{7}AAA(GA)_{4}$	F: TGGCATTTCATGGTTAATTT R: TGTTGTATGAAATGTCGGAA	39	25 (208-261)	0.9744	0.9395	0.123	-0.0193	AM858556
Cgsili45	$(AG)_{7}AAA(GA)_{4}$	F: GTCGAGAAATAAAGCTGGAA R: CAAGATTCCAAGGAAACAAA	41	20 (184-211)	1.0000	0.898	0.089	-0.0592	AM857706
Cgsili46	$(TG)_{5}A(GT)_{7}$	F: CATGACAATCGAGTCCATAA R: CATGGTGGAGAAAGAGTTGT	39	33 (340-414)	0.9487	0.9481	0.091	-0.0010	AM856490
Cgsili50	$(CA)_{10}$	F: CTATCTGAGCACGCTTCTCT R: TCTCTGTCAGATGATCTCAGG	46	25 (221-250)	0.848	0.913	0.078	+0.0314	AM865904
Cgsili56	$(GTT)_{8}$	F: GTCCCTGTGTTACACTCCTG R: TATCAGTCCAACAAAGGAGG	42	21 (232-289)	1.0000	0.8747	0.239	-0.0720	CU682998
Cgsili57	$(AAG)_{8}$	F: CAGTCCCTCTACGCTACATC R: ACACTACCGCTTTCCTGATA	35	12 (320-374)	1.0000	0.7727	0.046*	-0.1613	AM856592

* Significance level of 0.05

References

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. and Bonhomme, F.: GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France), 1996.

Degremont, L., Ernande, B., Bedier, E. and Boudry, P., 2007. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). I. Estimation of genetic parameters for survival and growth. Aquaculture. 262, 41-53.

FAO (2006) Yearbook of Fisheries statistics: Summary Tables. World aquaculture production of fish, crustaceans, molluscs, etc by principal species in 2005. http://www.fao.org/fi/statist/summtab/default.asp.

Goldstein, D. and Schlotterer, C.: Microsatellites: Evolution and Applications. Oxford University Press, 1999.

Hedgecock, D., Gaffney, P.M., Goulletquer, P., Guo, X., Reece, K. and Warr Gregory, W., 2005. The case for Sequencing the Pacific Oyster Genome. Journal of Shellfish Research. 24, 429-441.

Hedgecock, D., Li, G., Hubert S., Bucklin K., Ribes V., 2004. Widespread null alleles and poor cross-species amplification of microsatellite DNA cloned from the Pacific oyster Crassostrea gigas. Journal of Shellfish Research 23: 379-385.

Hubert, S. and Hedgecock, D., 2004. Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. Genetics. 168, 351-62.

Jenny, M.J., Chapman, R.W., Mancia, A., Chen, Y.A., McKillen, D.J., Trent, H., Lang, P., Escoubas, J.M., Bachere, E., Boulo, V., Liu, Z.J., Gross, P.S., Cunningham, C., Cupit, P.M., Tanguy, A., Guo, X., Moraga, D., Boutet, I., Huvet, A., De Guise, S., Almeida, J.S. and Warr, G.W., 2007. A cDNA microarray for *Crassostrea virginica and C. gigas*. Marine Biotechnology (NY). 9, 577-91.

Launey, S. and Hedgecock, D., 2001. High genetic load in the Pacific oyster *Crassostrea gigas*. Genetics. 159, 255-65.

Li, L. and Guo, X., 2004. AFLP-based genetic linkage maps of the pacific oyster *Crassostrea gigas* (Thunberg, 1793). Marine Biotechnology. 6, 26-36.

Marshall, T. C., J. Slate, L.E.B. Kruuk and J. M. Pemberton, 1998. "Statistical confidence for likelihood-based paternity inference in natural populations." Molecular Ecology 7(5): 639-655.

Rozen, S. and H. Skaletsky, 2000. "Primer3 on the WWW for general users and for biologist programmers." Methods Molecular Biology 132: 365-86.

Sauvage, C., Bierne, N., Lapegue, S. and Boudry, P., 2007. Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. Gene. 406, 13-22.

Tanguy, A., Bierne, N., Saavedra, C., Pina, B., Bachere, E., Kube, M., Bazin, E., Bonhomme, F., Boudry, P., Boulo, V., Boutet, I., Cancela, L., Dossat, C., Favrel, P., Huvet, A., Jarque, S., Jollivet, D., Klages, S., Lapegue, S., Leite, R., Moal, J., Moraga, D., Reinhardt, R., Samain, J.-F., Zouros, E. and Canario, A., 2008. Increasing genomic information in bivalves through new EST collections in four species: Development of new genetic markers for environmental studies and genome evolution. Gene. 408, 27-36.

Yu, H. and Li, Q.I., 2007. EST-SSR markers from the Pacific oyster, *Crassostrea gigas*. Molecular Ecology Notes. 7, 860-862.

Yu, H. and Li, QI., 2008. Exploiting EST Databases for the Development and Characterization of EST-SSRs in the Pacific Oyster (*Crassostrea gigas*). Journal of Heredity. 99: 208-214.

Wang, Y., Ren, R. and Yu Z., 2008. Bioinformatic mining of EST-SSR loci in the Pacific oyster, *Crassostrea gigas.* Animal Genetics (in press).