Conservation Genet Resour (2013) 5:663–666
DOI 10.1007/s12686-013-9877-9

TECHNICAL NOTE

Development of microsatellite markers from 454 transcriptome derived sequences for the scallop *Pecten maximus*

Natalie Hold · Louise Dawnay · Martin I. Taylor

Received: 8 February 2013 / Accepted: 15 February 2013 / Published online: 27 February 2013

Abstract Twelve microsatellite markers were developed for the scallop *Pecten maximus*. The markers were tested in three geographically diverse populations and all markers were polymorphic in all three populations. The mean number of alleles per locus ranged from 2 to 10.67 and the observed and expected heterozygosity ranged from 0.05 to 0.67 and 0.05 to 0.81 respectively. Some loci showed evidence of null alleles and an excess of homozygotes in some populations but 9 loci conformed to Hardy–Weinberg expectations. These new loci can be combined with previously published microsatellites to create a powerful suite of markers for genetic analyses.

Keywords Scallop · *Pecten maximus* · Microsatellites

Pecten maximus supports an important commercial fishery. France and the UK land the largest annual catches with total UK scallop landings of 34,000 tonnes in 2009, equating to a value of £47 million (FAO 2012). Despite its commercial importance, the genetic structure of *P. maximus* stocks remain poorly defined. Watts et al. (2005) developed nine microsatellite markers but to date no studies of population structure have been published using them and our testing found that only four produced reliable results. Consequently, new microsatellite markers were developed using 454 transcriptome sequencing. Since the development of the microsatellites in this study, nine further microsatellite markers have been published from expressed sequence tags (Charrier et al. 2012). The power of many genetic analyses is increased more rapidly by the addition of more loci (and more variable loci) than by increasing the sample size (Felsenstein 2006; Wang and Santure 2009; Landguth et al. 2012). As such, the microsatellite markers from all three studies could be combined to create a very powerful suite of genetic markers.

Microsatellite markers were identified using next generation sequencing of the scallop transcriptome. Briefly, cDNA was isolated from gill, muscle, and mantle tissue from four individual scallops and sequenced on Roche's Titanium Flex 454 sequencer. Sequences were cleaned of vectors and primers using Seqclean (http://compbio.dfci.harvard.edu/tgi/software/) before contig assembly using iAssembler (http://bioinfo.bti.cornell.edu/tool/iAssembler/). Microsatellite containing contigs were then identified using microsatcommander (Faircloth 2008) and primers designed using Primer3 (Rozen and Skaletsky 2000). Following initial testing twelve markers which were polymorphic and amplified reliably were then multiplexed. Primers were labelled with one of four different fluorescent tails (PET®, VIC®, NED®, FAM®). Oligonucleotide tails were added to the 5′ end of the forward primer and fluorescently labelled oligonucleotides complementary to the oligonucleotide tail were then incorporated during PCR with the addition of cycles using an annealing temperature of 50 °C (Schuelke 2000).

The markers were tested on three populations: Galicia in Spain, Norway and the Isle of Man (GAL, NOR, IOM). DNA was extracted using CTAB extraction buffer (Doyle

N. Hold (✉)
School of Ocean Sciences, Bangor University,
Menai Bridge LL59 5AB, UK
e-mail: n.hold@bangor.ac.uk

N. Hold · L. Dawnay
Molecular Ecology and Fisheries Genetics Laboratory,
Environment Centre Wales, Bangor University,
Bangor LL57, UK

M. I. Taylor
School of Biological Sciences, University of East Anglia,
Norwich NR4 7TJ, UK

Table 1 Polymorphic microsatellite markers for *P. maximus* with locus characteristics and multiplex PCR primer concentrations

Multiplex	Locus	Primer	Primer sequence	Ta (°C)	Motif	Size range	PCR profile	Fluorescent tail (concentration μM)	Primer concentration (μM)
A	PmNH9	PmNH09F	CTGATTTACCCTCCGCAACG	58	$(CTGT)_4$	199–215	One	VIC (0.2)	0.05
		PmNH09R	TCAGTAATTCTACTCGTGACCC						0.20
	PmNH19	PmNH19F	TCCCTGTTCCTCATTTACTGC	58	$(AATT)_4$	353–369		PET (0.33)	0.08
		PmNH19R	ACCGAATAATACCACTAACCCAG						0.33
	PmNH20	PmNH20F	TGACAGCTGGGTATAATGGTATG	58	$(ATCC)_6$	187–260		FAM (0.2)	0.05
		PmNH20R	AGGTGACCGACAACCTAATC						0.20
B	PmNH11	PmNH11F	GCCATGGTCGGAAATCACC	51	$(ATTT)_4$	304–308	Two	PET (0.2)	0.05
		PmNH11R	CAAACGCGCCAAGTCTACG						0.20
	PmNH23	PmNH23F	AAATGCCGTCAGCTTTCAG	51	$(AT)_8$	253–257		NED (0.2)	0.05
		PmNH23R	ACTGTACAAATCGGCCACG						0.20
C	PmNH59	PmNH59F	CGAAGGTTTGTGCTGTGAATC	58	$(AT)_8$	275–281	One	FAM (0.2)	0.05
		PmNH59R	CCAGCAATGACATCCGATCG						0.20
	PmNH60	PmNH60F	TTGTACAAATGCTGGCGTGG	58	$(AT)_{10}$	186–192		VIC (0.17)	0.04
		PmNH60R	TCTACTCTGGCAGATCATGGG						0.17
	PmNH62	PmNH62F	GGGACCACTGTAAACAATGTG	58	$(AAC)_9$	254–266		PET (0.2)	0.05
		PmNH62R	GCGTGACAGTCGACCATTTC						0.20
	PmNH68	PmNH68F	CCATTGGGTAGGGAGCATCG	58	$(AT)_8$	118–136		PET (0.2)	0.04
		PmNH68R	AAACAAACACTCGCGAACTC						0.17
	PmNH70	PmNH70F	AGTTGTGCTATTGAATGGGAAC	58	$(AAT)_8$	140–146		FAM (0.17)	0.04
		PmNH70R	ATGCACTGCTTGTCCACTTC						0.17
	PmNH73	PmNH73F	CATAGCGATGCAGGACAAGG	58	$(GGT)_8$	213–231		NED (0.2)	0.05
		PmNH73R	ATTCCAATGTCTGCCGTCTG						0.20
	PmNH75	PmNH75F	CGCTGAAAGTGGTTACCGTG	58	$(GT)_8$	116–126		FAM (0.17)	0.04
		PmNH75R	GCACACAGGTCCAAATATAACG						0.17

and Doyle 1987). Each 12 µl PCR contained 1 µl of DNA (~10 ng), 6 µl of Qiagen Type IT multiplex PCR mix, 0.5 µl of BSA (10 mg/ml), 2.5 µl of water and 2 µl of Primer mix. The primer mix contains the forward and reverse primers for all loci in the multiplex plus the appropriate fluorescent tails (PET®, VIC®, NED®, FAM®) (Table 1). There were two different thermocycling profiles for the multiplex reactions. Profile one used a touchdown PCR regime; 95 °C for 5 min, [95 °C 30 s, 64 °C (−1 °C per cycle) 90 s, 72 °C 30 s] × 6 cycles, (95 °C 30 s, 58 °C 90 s, 72 °C 30 s) × 9 cycles, (95 °C 30 s, 50 °C 90 s, 72 °C 30 s) × 15 cycles with a 45 min extension at 72 °C. Profile two; 95 °C 5 min, (95 °C 30 s, 51 °C 90 s, 72 °C 30 s) × 16 cycles, (95 °C 30 s, 50 °C 90 s, 72 °C 30 s) × 19 cycles, final extension of 45 min at 72 °C. PCR products were resolved on an ABI 3130XL sequencer and Genemapper® software (Applied Biosystems) to size alleles.

Data was checked for null alleles, stuttering and large allele dropout using Microchecker (Van Oosterhout et al. 2004). Genepop v4.0 (Rousset 2008) was used to test each pair of loci in each population for evidence for linkage disequilibrium and conformity to Hardy–Weinberg Equilibrium (HWE) using Fisher's exact tests (Markov chain parameters: burn-in = 10,000 followed by 100 batches of 5,000 iterations each). The alpha value was corrected following multiple testing using the False Discovery Rate method (Benjamini and Hochberg 1995). GENALEX 6 (Peakall and Smouse 2006) was used to estimate observed and expected heterozygosity (Ho and He respectively) and the number of alleles.

Marker PmNH9 in GAL had very low amplification success of 43 % and therefore data from this population was not included in the analysis of this marker. No evidence of stuttering or large allele dropout was detected. Marker PmNH20 showed a significant deficit of heterozygotes in all populations. PmNH19 showed a deficit of heterozygotes in IOM and NOR but not GAL, and PmNH68 had a deficit of heterozygotes in IOM only. Microchecker indicated that this could be due to null alleles at these loci/population combinations. All other markers showed no significant deviation from HWE in any population following correction for multiple testing (Table 2). No significant linkage disequilibrium was detected between any pair of loci. The number of alleles ranged from two to 13 and all loci were polymorphic in all populations. The lowest observed heterozygosity was 0.02 at locus PmNH70 in NOR, with the highest observed heterozygosity of 0.77 at locus PmNH59, also in NOR.

Markers PmNH20, PmNH19 and PmNH68, whilst highly polymorphic, show evidence of null alleles and should therefore be used with caution. Marker PmNH9 did not amplify reliably in the Galicia samples but was reliable in the other populations. The remaining eight markers conformed to HWE in all three populations and amplified reliably. When combined with previously developed markers (Watts et al. 2005; Charrier et al. 2012) these loci will constitute a large and powerful suite of markers for *P. maximus*.

Table 2 Number of alleles (N_a), observed (H_o) and expected heterozygosity (H_e)

Locus	Measure	Populations			
		CHK N = 48	NOR N = 48	GAL N = 29	Mean
PmNH9	H_e	0.31	0.25	N/A	0.28
	H_o	0.15	0.18	N/A	0.16
	Na	6.00	4.00	N/A	5.00
PmNH11	H_e	0.29	0.06	0.25	0.20
	H_o	0.27	0.07	0.21	0.18
	Na	2.00	2.00	2.00	2.00
PmNH19	H_e	0.28	0.53	0.12	0.31
	H_o	**0.10**	**0.30**	0.13	0.18
	Na	5.00	7.00	2.00	4.67
PmNH20	H_e	0.81	0.80	0.80	0.81
	H_o	**0.31**	**0.31**	**0.40**	0.34
	Na	13.00	8.00	11.00	10.67
PmNH23	H_e	0.24	0.43	0.18	0.28
	H_o	0.19	0.51	0.20	0.30
	Na	3.00	3.00	2.00	2.67
PmNH59	H_e	0.71	0.65	0.71	0.69
	H_o	0.65	0.77	0.59	0.67
	Na	5.00	5.00	8.00	6.00
PmNH60	H_e	0.29	0.50	0.32	0.37
	H_o	0.33	0.47	0.32	0.37
	Na	5.00	3.00	2.00	3.33
PmNH62	H_e	0.69	0.68	0.65	0.67
	H_o	0.71	0.59	0.67	0.65
	Na	6.00	5.00	3.00	4.67
PmNH68	H_e	0.78	0.76	0.83	0.79
	H_o	**0.38**	0.70	0.63	0.57
	Na	9.00	9.00	8.00	8.67
PmNH70	H_e	0.04	0.02	0.08	0.05
	H_o	0.04	0.02	0.08	0.05
	Na	3.00	2.00	3.00	2.67
PmNH73	H_e	0.23	0.18	0.28	0.23
	H_o	0.21	0.17	0.24	0.21
	Na	6.00	4.00	4.00	4.67
PmNH75	H_e	0.57	0.54	0.53	0.55
	Ho	**0.42**	0.47	0.50	0.46
	Na	7.00	4.00	4.00	5.00

The results in bold show significant deviation from HWE using Fishers' exact test in Genepop (dememorization 10,000 iterations; 100 batches of 5,000 iterations) following correction for multiple testing using the False Discovery Rate

References

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300

Charrier G, Morvezen R, Calves I, Laroche J (2012) Development of new microsatellite markers derived from expressed sequence tags for the great scallop (*Pecten maximus*). Conserv Genet Resour 4:931–934

Doyle JJ, Doyle JL (1987) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Faircloth BC (2008) msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8:92–94

FAO (2012) Global capture production. Available at: http://www.fao.org/fishery/statistics/global-capture-production/en. Accessed 15 May 2012

Felsenstein J (2006) Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? Mol Biol Evol 23:691–700

Landguth EL, Fedy BC, Oyler-Mccance SJ et al (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Mol Ecol Resour 12:276–284

Peakall R, Smouse P (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Res 8:103–106

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments: a poor man' s approach to genotyping for research and high-throughput diagnostics. PRism 18:1–2

Van Oosterhout C, Hutchinson W, Wills DP, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Wang J, Santure AW (2009) Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181:1579–1594

Watts PC, Mallanaphy PJ, McCarthy C, Beukers-Stewart BD, Mosley MWJ, Brand AR, Saccheri IJ (2005) Polymorphic microsatellite loci isolated from the great scallop, *Pecten maximus* (Bivalvia: Pectinidae). Mol Ecol Notes 5:902–904