Isolation and validation of microsatellite markers from a depleted South African sciaenid species, the dusky kob (Argyrosomus japonicus), by means of the FIASCO/454 approach

L. Mirimin · J. C. Ruiz Guajardo · J. Vervalle ·
Aletta Bester-Van der Merwe · S. Kerwath ·
B. Macey · P. Bloomer · R. Roodt-Wilding

Abstract The dusky kob (Argyrosomus japonicus) is a large, estuarine-dependent sciaenid fish that has been severely depleted in South African waters and that, in recent years, has received considerable attention from the local fish farming industry. Discovery and application of appropriate molecular markers is necessary to improve the understanding of wild population structure, assist the effectiveness of broodstock and breeding programmes, and ensure monitoring of potential interactions between wild and farmed fish. The present study uses a recently tested approach that combines the FIASCO enrichment protocol with 454 GS-FLX Next Generation Sequencing, to identify large numbers of microsatellite-containing sequences at a low cost and high discovery rate from the dusky kob genome. Following the FIASCO enrichment (targeting specifically tetranucleotide repeats), 2,355 potential tetra-nucleotide microsatellites (perfect repeat motifs including eight or more repeat units flanked by regions for primer design) were identified from 1/5th of a single 454 lane. From these sequences, a test panel of 60 potential markers was selected for validation. A total of eight (13 %) markers were successfully amplified from a test sample of wild dusky kob individuals and showed high levels of poly-morphism (observed heterozygosity per locus ranging between 0.375 and 0.905). Cross-species amplification of seven of these markers was also successfully carried out in another closely related and commercially important South African sciaenid species, the silver kob (A. inodorus). The microsatellite markers developed in the present study are readily available tools suitable to address genetic vari- ability of Argyrosomus species of southern Africa.

The dusky kob (Argyrosomus japonicus Temminck and Schlegel 1843) is a marine fish species that occurs in estuarine and coastal waters of southern Africa, between the Cape of Good Hope and Mozambique (Griffiths and Heemstra 1995). Poor management has led to overfishing and severe depletion of South African stocks, with spawner biomass per recruit estimated at only 1.0 4.5 % of pre-exploitation values (Griffiths 1997). Due to wild population declines and to favourable traits such as fast growth and late maturation, this species has been of growing interest to the fish farming industry of South Africa. Hence, the development of a suitable set of molecular tools (such as microsatellite loci) would favour studies of wild population structure, that will result in improved broodstock selection and guarantee the monitoring of potential interactions between wild and farmed fish (Chistiakov et al. 2006). Recent advances in Next Generation Sequencing (NGS) technologies have favoured the rapid and economical

L. Mirimin (&) · J. C. Ruiz Guajardo · J. Vervalle ·
A. Bester Van der Merwe · R. Roodt Wilding
Department of Genetics, Stellenbosch University,
Private Bag X1, Matieland 7602, South Africa
e mail: lmirimin@gmail.com

S. Kerwath · B. Macey
Department of Agriculture, Forestry and Fisheries,
Private Bag X2, Rogge Bay 8012, South Africa

S. Kerwath
Department of Zoology, University of Cape Town,
Private Bag X3, Rondebosch 7700, South Africa

P. Bloomer
Department of Genetics, University of Pretoria,
Private Bag X2, Hatfield 0028, South Africa

discovery of such markers from non-model organisms (Davey et al. 2011; Malausa et al. 2011). The present study reports on the isolation (and partial validation) of a high number of novel microsatellite markers from South African dusky kob, by combining an enrichment protocol (FIASCO) (Zane et al. 2002) with 454 pyrosequencing (Droege and Hill 2008). Reid et al. (2012) recently used this combined approach (hereafter "FIASCO/454") to successfully discover numerous microsatellite loci in sparids.

Simultaneous digestion/ligation of 250 ng of genomic DNA from a dusky kob individual with MseI, Polymerase Chain Reaction (PCR) amplification of MseI-ligated fragments and hybridization to repeat-specific tetranucleotide probes (AAAC, GATA, GCCT, TGTC, GAAA, AGCC) was carried out as detailed in Reid et al. (2012). Success- fully hybridized fragments were recovered by means of streptavidin-coated magnetic beads (Roche) and PCR amplification was carried out as described by Zane et al. (2002) to amplify approximately 2 μg of DNA, which was purified (Roche purification kit) before sequencing was performed by a 454 LifeSciences / Roche Genome Sequencer FLX^{TM} with GS-FLX titanium reagents (approximately 1/5th of a single lane) following manu-facturer specifications. This procedure yielded a total of 28,591 DNA fragments (231 bp average sequence length) suitable for further data mining.

The program MSATCOMMANDER 0.8.2 (Faircloth 2008) was used to identify sequences that contained repeat motifs. A total of 13,054 sequences (45.7 % of total sequences obtained) included microsatellite-like regions, including 1,006 tri- and 6,222 tetranucleotides. Redundant fragments were ruled out by sequence alignment using CLUSTAL X 1.64b (Thompson et al. 1997). A search for sequences containing perfect repeat motifs with eight or more repeat units and flanking regions for the design of primers led to the identification of 146 and 2,355 sequences containing tri- and tetranucleotide repeats, respectively. Primer pairs were successfully designed with PRIMER3 1.1.1 (Rozen and Skaletsky 2000) for a test panel of 60 markers, including 12 and 48 sequences containing tri- and tetranucleotide repeat motifs, respectively. Following PCR optimization and preliminary screening of amplified products, eight out of 60 (13 %) primer pairs successfully amplified microsatellite markers. Presence of core repeat regions was confirmed in all eight newly discovered markers by Sanger sequencing, using the BigDye$^{®}$ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Sequencing products were purified with SigmaSpin^{TM} post-reaction Clean-Up Columns (Sigma-Aldrich) and sequence compo-sition was resolved on an ABI 3730xl DNA Analyzer (performed by Central Analytical Facility DNA sequencing unit, www.sun.ac.za/CAF). Individual sequence data for

each newly discovered marker were deposited in GenBank (Accession Numbers JY503142 JY503149).

Genomic DNA was isolated from 24 dusky kob and 24 silver kob tissue samples (randomly selected along the South African coast) following a proteinase K digestion and chloroform extraction (modified from Saghai-Maroof et al. 1984). The quality and quantity of extracted DNA were evaluated with a NanoDrop^{TM} ND 1,000 spectro-photometer (Thermo Fisher Scientific). PCR amplification for each of the eight newly discovered markers was carried out in a total volume of 10 μl, including 1X GoTaq flexi Buffer (Promega), 1 mM MgCl$_2$, 0.25 mM dNTPs, 0.2 units of GoTaq polymerase (Promega), 0.5 μM of each primer (fluorescently labelled as detailed in Table 1) and 60 ng of genomic DNA. Thermal cycling conditions included an initial denaturation step (5 min at 95 °C), followed by 35 3-step cycles (30 s at 95 °C, 45 s at 56 °C, 45 s at 72 °C) and a final extension step (10 min at 72 °C). Genotyping was conducted on an ABI 3730xl DNA Ana-lyzer, where the size of amplified fragments was resolved by comparison to GeneScan^{TM} 600 LIZ$^{®}$ (Applied Bio-systems) with the aid of GeneMapper$^{®}$ 4 software (Applied Biosystems). Overall, no evidence of null alleles, large allele drop out or genotyping errors due to stuttering (based on 95 % confidence intervals of 1,000 Monte Carlo simu- lations) were detected by approaches as implemented in MICROCHECKER 2.2.3 (van Oosterhout et al. 2004). Levels of polymorphism were high in both species with average observed heterozygosities of 0.66 and 0.70 for dusky and silver kob, respectively (Table 1). Locus Ajap05 failed to amplify in silver kob, probably due to mutations in the primer region, whereas some other loci (in particular Ajap02) were amplified only in a portion of genotyped individuals (see Table 1), indicating that further PCR optimization may be required for these markers. No evi-dence of linkage disequilibrium between pairs of loci or deviation from expected Hardy Weinberg proportions were detected in the two species using ARLEQUIN 3.5.1.2 (Excoffier et al. 2005). Overall, non-significantly positive F_{IS} estimates were detected using FSTAT 2.9.3.2 (Goudet 1995) (Table 1). These results indicate that the newly discovered markers are suitable to address broader popu-lation genetic questions of Argyrosomus species in South Africa, and provide a significant addition to previously developed markers from Australian dusky kob (Archangi et al. 2009). Furthermore, the large amount of repeat-containing sequence data hereby discovered (i.e. 146 tri- and 2,355 tetranucleotides) have the potential for the development of many additional markers (approximately 298, assuming a 13 % success rate), which will be funda-mental tools to study wild populations, assist breeding programmes and monitor potential interactions between wild and farmed fish.

Table 1 Descriptive statistics for eight microsatellite markers for dusky and silver kob

| Locus name | Primer sequence (5' 3') | Repeat motif | Dusky kob | | | | | | Silver kob | | | | | |
			Allele size range (bp)	n	NA	Ho	Hs	F_{IS}	Allele size range (bp)	n	NA	Ho	Hs	F_{IS}
Ajap02	F:TCTGAACACACCACAATGG	CTTT	244-256	7	3	0.571	0.560	−0.021	244-284	14	7	0.429	0.537	0.208
	R: FAM-CATGGGAGAGGTATTTTCC													
Ajap05	F: PET-TATGGGGTGCCGTTTGTAA	AGAT	131–179	21	12	0.905	0.893	−0.013						
	R: AGGGCGAATGAGGAACAG													
Ajap06	**F:** *FAM-GGTTCTGTCCTACTTTGGC*	GOAT	177–197	24	6	0.750	0.777	0.035	161–189	24	*5*	0.667	0.645	−0.034
	R: AAATCCTCAACCACTTCAG													
Ajap12	F: PET-TCTCTCTCACACACTTACCACTGA	ATCT	113-185	22	11	0.818	0.812	−	117–225	24	15	0.792	0.911	0.134
	R:TGATACTTTACATCCTCCCATTTTT													
Ajap14	**F:** FAM-TGAGTGGCTTCATCAGAAATG	ATCT	120-208	23	13	0.826	0.911	0.095	112-196	24	19	0.917	0.925	0.009
	R: GAGACGGAGGAGAGGGTGA													
Ajap24	F:TTCAGGGGCGTAGTTTAGTTT	AGAT	258-286	11	7	0.727	0.866	0.167	246-278	3	4	1.000	0.867	−
	R: PET-ATTTATCTTGTGTATCTTGTGTCCAT													
Ajap35	*F:NED-GTCACATCCACGCATACTG*	ACT	137–185	23	6	0.565	0.619	0.089	146-158	20	3	0.050	0.099	0.500
	R: ATGGAGTGAGGAACGAGTG													
Ajap31	*F:NED-ATCATCCTCCGAACCTGGA*	AGC	!58-!82	24	6	0.375	0.456	0.!80	167–215	24	10	0.875	0.787	−
	R: GCTTCTCTCTGTCACGGTCTG													
Mean				19	8	0.692	0.737	0.062		19	9	0.675	0.682	0.003

Fluorescently labelled primers are denoted by *FAM-, NED-, VIC-* or *PET-*

n number of individuals, *NA* number of alleles, H_O observed heterozygosity, *HE* expected heterozygosity, F_{IS} inbreeding coefficient

References

Archangi B, Chand V, Mather PB (2009) Isolation and characteriza tion of 15 polymorphic microsatellite DNA loci from Argyros omus japonicus (mulloway), a new aquaculture species in Australia. Molecular Ecology Resources 9:412 414

Chistiakov DA, Hellemans B, Volckaert FAM (2006) Microsatellites and their genomic distribution, evolution, function and applica tions: a review with special reference to fish genetics. Aquacul ture 255:1 29

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome wide genetic marker discovery and geno typing using next generation sequencing. Nat Rev Genet 12:499 510

Droege M, Hill B (2008) The Genome Sequencer FLX System longer reads, more applications, straight forward bioinformatics and more complete data sets. J Biotechnol 136:3 10

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data anal ysis. Evolutionary Bioinformatics Online 1:47 50

Faircloth BC (2008) MSATCOMMANDER: detection of microsat ellite repeat arrays and automated, locus specific primer design. Molecular Ecology Resources 8:92 94

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F statistics. J Hered 86:485 486

Griffiths MH (1997) Management of South African dusky kob Argyrosomus japonicus (Sciaenidae) based on per recruit mod els. S Afr J Mar Sci 18:213 228

Griffiths MH, Heemstra PC (1995) A contribution to the taxonomy of the marine fish genus Argyrosomus (Perciformes: Sciaenidae), with descriptions of two new species from southern Africa. Ichthyol. Bull. J. L. B. Smith Inst. Ichthyol 65:40

Malausa T, Gilles A, Meglécz E, Blanquart H, Duthoy S, Costedoat C, Dubut V, Pech N, Castagnone Sereno P, Délye C, Feau N, Frey P, Gauthier P, Guillemaud T, Hazard L, Le Corre V, Lung Escarmant B, Malé P JG, Ferreira S, Martin J F (2011) High throughput microsatellite isolation through 454 GS FLX Tita nium pyrosequencing of enriched DNA libraries. Molecular Ecology Resources 11:638 644

Reid K, Hoareau TB, Bloomer P (2012) High throughput microsat ellite marker development in two sparid species and verification of their transferability in the family Sparidae. Molecular Ecology Resources 12:740 752

Rozen S, Skaletsky H (2000) PRIMER 3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365 386

Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphisms in barley: men delian inheritance, chromosomal location, and population dynamics. Proceedings of the National Academy of Science USA 81:8014 8018

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acid Res 25:4876 4882

van Oosterhout C, Hutchinson WF, Willis DPM, Shipe P (2004) MICRO CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:536 538

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