2014

Multiplexing with Three-Primer PCR for Rapid and Economical Microsatellite Validation

Salla Vartia

Patrick C. Collins

Thomas F. Cross

Richard D. Fitzgerald

David T. Gauthier
Old Dominion University, dgauthie@odu.edu

See next page for additional authors

Authors

Salla Vartia, Patrick C. Collins, Thomas F. Cross, Richard D. Fitzgerald, David T. Gauthier, Philip Mcginnity, Luca Mirimin, and Jens Carlsson

Multiplexing with three-primer PCR for rapid and economical microsatellite validation

SALLA VARTIA[1,2], PATRICK C. COLLINS[2], THOMAS F. CROSS[5], RICHARD D. FITZGERALD[1], DAVID T. GAUTHIER[3], PHILIP MCGINNITY[5], LUCA MIRIMIN[1,4] and JENS CARLSSON[2,5]

[1]*Carna Research Station, Ryan Institute, National University of Ireland, Galway, Carna, Connemara, Co. Galway, Ireland*
[2]*Area 52 Research Group, School of Biology & Environment Science, University College Dublin, Belfield, Dublin, Ireland*
[3]*Department of Biological Sciences, Old Dominion University, Norfolk, VA, USA*
[4]*Molecular Breeding and Biodiversity Lab, Department of Genetics, Stellenbosch University, Stellenbosch, Western Cape, South Africa*
[5]*Beaufort Fish Genetics Programme, School of Biological, Earth and Environmental Sciences/Aquaculture and Fisheries Development Centre, University College Cork, Distillery Fields, North Mall, Cork, Ireland*

The next generation sequencing revolution has enabled rapid discovery of genetic markers, however, development of fully functioning new markers still requires a long and costly process of marker validation. This study reports a rapid and economical approach for the validation and deployment of polymorphic microsatellite markers obtained from a 454 pyrosequencing library of Atlantic cod, *Gadus morhua,* Linnaeus 1758. Primers were designed from raw reads to amplify specific amplicon size ranges, allowing effective PCR multiplexing. Multiplexing was combined with a three-primer PCR approach using four universal tails to label amplicons with separate fluorochromes. A total of 192 primer pairs were tested, resulting in 73 polymorphic markers. Of these, 55 loci were combined in six multiplex panels each containing between six and eleven markers. Variability of the loci was assessed on *G. morhua* from the Celtic Sea (n = 46) and the Scotian Shelf (n = 46), two locations that have shown genetic differentiation in previous studies. Multilocus F_{ST} between the two samples was estimated at 0.067 (P = 0.001). After three loci potentially under selection were excluded, the global F_{ST} was estimated at 0.043 (P = 0.001). Our technique combines three-primer and multiplex PCR techniques, allowing simultaneous screening and validation of relatively large numbers of microsatellite loci.

Salla Vartia, Carna Research Station, Ryan Institute, National University of Ireland, Galway, Ireland. E-mail: salla.vartia@gmail.com

Thirty years after their discovery in the 1980s, microsatellite-based genetic markers are still extensively used in studies of population structure, parentage analysis, genetic mapping, evolutionary processes and forensics (BRUFORD and WAYNE 1993; BROCKMANN et al. 1994; KNAPIK et al. 1998; GOLDSTEIN et al. 1999; PRIMMER et al. 2000). These markers have a wide application due to high allelic diversity and co-dominance of alleles (CHAMBERS and MACAVOY 2000). Many methodologies have been established in order to discover such markers, but it is only following the recent advent of next generation sequencing (NGS) technology that large amounts of markers can be increasingly rapidly and economically developed from non-model organisms. NGS approaches allow the fast discovery of large amounts of microsatellite-containing sequences, however mining such data for suitable DNA fragments and validation of candidate markers are still posing challenges prior to the utilisation of fully operating new markers.

The most common approach to date for de novo microsatellite marker development includes creation of repeat-enriched DNA libraries, fragment replication by cloning, and Sanger sequencing of clones containing potential microsatellites (ZANE et al. 2002). These processes are laborious and time consuming, and typically have low marker yield, with the percentage of positive clones averaging 2–3% (ASHWORTH et al. 2004). The final marker yield is even lower with a large portion of markers discarded during the isolation–characterisation process (SQUIRRELL et al. 2003). Alternatively, microsatellite-containing sequences can be mined from existing molecular data such as genomic DNA or expressed sequence tag (EST) sequences (LI et al. 2004). These approaches are limited by the paucity of data on non-model organisms. EST-linked microsatellites can be relatively easy to identify but have a higher probability of being affected by selective processes, and hence may not be suitable for

population analyses that assume that loci are selectively neutral (Ellis and Burke 2007). Microsatellite markers developed for one species may also be applied to closely related species (Schlötterer et al. 1991). However, this approach is limited by varying levels of successful cross-species amplification between species (Moore et al. 1991). Even when cross-species amplification is successful, levels of variability tend to be lower compared with the species for which the markers were developed (Primmer et al. 1996). Because of these limitations, it may be preferable to develop markers de novo for a species or population of interest to ensure optimal power of newly discovered markers (Carlsson et al. 2013).

Recently, several approaches have been presented for discovery of microsatellites using next generation sequencing (NGS)-generated data (Abdelkrim et al. 2009; Allentoft et al. 2009). Large amounts of sequence data can be generated from either genomic DNA or microsatellite enriched libraries (Guichoux et al. 2011; Malausa et al. 2011) and then mined for microsatellite loci. With this approach, thousands of potential loci can be rapidly identified (Guichoux et al. 2011). Large-scale microsatellite identification has several advantages, including the ability to rigorously screen individual loci for presence of optimal primer-binding sites in flanking regions (Guichoux et al. 2011; Zalapa et al. 2012; Fernandez-Silva et al. 2013). Additionally, deployment criteria (e.g. higher levels of variability, neutrality and low linkage) are study-specific and cannot be assessed until the markers have been validated (Selkoe and Toonen 2006). Validation of a large number of markers enables selection of most suitable loci (Selkoe and Toonen 2006). Even with the NGS approach, validation is labour intensive and a limiting bottleneck in microsatellite marker development (Squirrell et al. 2003; Malausa et al. 2011; Fernandez-Silva et al. 2013).

Typically, methods for validation and genotyping of microsatellites involve capillary gel electrophoresis with fluorescence-based amplicon detection (Edwards et al. 1991; Guichoux et al. 2011). The three-primer PCR method can be used to reduce the expense associated with fluorescently labelled primers (sensu Schuelke 2000; Diniz et al. 2007; Rubin et al. 2009). In three-primer PCR, the primers comprise an unlabelled forward primer with a universal tail attached to its 5′ end, a labelled universal primer matching the tail sequence and an unlabelled reverse primer (Steffens et al. 1993; Oetting et al. 1995; Neilan et al. 1997; Schuelke 2000). The labelled universal primer can be used in combination with any appropriately tailed forward primer, thereby eliminating the need to synthesize a fluorescently labelled forward primer for every unique locus during the validation phase, in which a large proportion of loci may be excluded because of problems with amplification.

Following initial identification and PCR optimisation of successful markers, sets of primers are usually labelled with a fluorescent label either on the forward or the reverse primer (Guichoux et al. 2011). Markers are then amplified in single PCR reactions or combined into a multiplex PCR containing multiple markers (Guichoux et al. 2011). Improvement of the traditional multiplex PCR technique (Missiaggia and Grattapaglia 2006) employed human microsatellite primer sequences as universal tails and combined three universal tails with three dyes in a true multiplex PCR (sensu Guichoux et al. 2011). However, despite the obvious cost benefits of the three primer PCR approach in combination with multiplexing, few studies have employed the method (Langen et al. 2011; Blacket et al. 2012). This may be attributable to poor amplification or poor quality chromatograms, resulting in difficulty in accurate genotyping of individuals (Hagell et al. 2013).

Here we present the development and application of a method for rapid validation and genotyping of novel microsatellites in Atlantic cod, *Gadus morhua,* Linnaeus 1758, using the three primer approach with multiplex PCR. The main aim of this study is the fast and economic development and deployment of microsatellite multiplexes from raw NGS data applicable for studies on a wide range of organisms.

MATERIAL AND METHODS

Sampling

Gadus morhua were obtained by trawling in 2009 and 2011 from the Celtic Sea, south of Ireland (n = 7, n = 46, respectively) and in 1996 from the Scotian Shelf, off Nova Scotia in eastern Canada (n = 46). Previous research has shown that these two populations are genetically differentiated (Hutchinson et al. 2001; O'Leary et al. 2007). Fin clip samples were preserved in 100% ethanol.

DNA extraction

DNA was extracted from fin clips using a Chelex protocol as described in Mirimin et al. (2011). DNA from the Scotian Shelf samples was extracted using a standard phenol–chloroform method (O'Leary et al. 2007). DNA was quantified using a NanoDrop 1000 Spectrophotometer (Thermo Scientific) and normalised to a concentration of 50 ng μl^{-1}.

Microsatellite selection

The unpublished sequence data used here were generated for a previous study (Carlsson et al. 2013), in which microsatellite containing sequences were obtained from five of the 2009 Celtic Sea individuals also used in the

present study, using 454 pyrosequencing of a reduced representation library. Carlsson et al. (2013) identified a total of 11 341 microsatellite containing sequences as suitable for primer design using the Primer3 plug-in (Rozen and Skaletsky 2000) for MISA ver. 1.0 (<http://pgrc.ipk-gatersleben.de/misa>). Of these, 6424 were estimated to be unique. These microsatellite-containing sequences were used in the present study.

To avoid excessive homoplasy (alleles identical in state but not in descent, cf. Estoup et al. 2002) and to ensure ease of genotyping, complex repeat motifs (i.e. compound and imperfect motifs) were excluded. To ensure sufficient space for primer design, reads that had less than 50 bp of sequence before and after the repeat-containing region were removed. In addition, to avoid excessively large allele size ranges, repeat sequences of more than 100 bp, and penta- and hexanucleotide repeats were excluded. A subsample of the remaining microsatellite sequences (n = 1309) were visually inspected for primer design.

Primer design

Primers were designed using Primer3Plus (Rozen and Skaletsky 2000; Untergasser et al. 2007) with optimal primer length as 20bp and optimal T_m at 60°C. Two sets of three size classes were used: the first set of size classes was separated by 30 bp (100–150, 180–250, 280–450 bp), and the second set separated by 50 bp (100–150, 200–250, 300–450 bp). Equal numbers of markers were designed for each size class. Only primer pairs with a T_m difference of less than 1°C were accepted in order to facilitate PCR multiplexing.

Designed primers were cross-referenced with the original sequence data set to identify primers that annealed to multiple regions (not unique) or originated from redundant sequences (different reads of the same sequence). Redundant sequences not detected in the previous steps (due to sequencing error in the primer regions) were identified by performing a de novo assembly with the remaining candidate loci sequences using Geneious ver. 6.1.5 (created by Biomatters; available from <www.geneious.com>), CAP3 plug-in (default settings; i.e. min overlap length = 40 bp, min overlap identity = 90%; Huang and Madan 1999). If two or more reads assembled together they were considered redundant and only one of them was kept for future analysis.

To minimise the risk that primer sequences were derived from contaminants, expressed *G. morhua* gene regions or previously published *G. morhua* microsatellites, microsatellite-containing sequences and primers were subjected to a BLAST search in the GenBank nucleotide database (Altschul et al. 1990). For possible contaminants, we considered a match with ≥ 95% coverage and 100% identity as a threshold for excluding reads. No threshold was employed for matches on *G. morhua* sequences. When such matches were encountered, primers were excluded from further analyses. In addition, validated primer sequences were subjected to BLAST searches against the *G. morhua* genome (Star et al. 2011) in the whole-genome shotgun contigs database in GenBank.

Universal primers

The universal dye-labelled primers used were T3: PET-5′ AATTAACCCTCACTAAAGGG 3′, M13 Reverse: NED-5′ GGATAACAATTTCACACAGG 3′ (Diniz et al. 2007), Hill: 6FAM-5′ TGACCGGCAGCAAAATTG 3′ (Tozaki et al. 2001) and Neomycin rev: VIC-5′ AGGTGAGATGACAGGAGATC 3′. Each forward primer had one of the above universal primer sequences added to its 5′ end. PIG-tails were added to the 5′ end of all the reverse primers. PIG-tailing leads to an addition of a non-templated adenosine nucleotide to the 3′ end on nearly 100% of PCR products which reduces stutter caused by random addition of dATP (Brownstein et al. 1996). The tails were matched with the primers using OligoAnalyzer ver. 3.1 (<www.idtdna.com>) to ensure the least amount of different secondary structures. Equal numbers of primers were paired with each of the four different universal primers.

Microsatellite validation

Primers were combined into twelve multiplex PCR reactions containing 12 markers each (12-plex) and validated using all seven 2009 Celtic Sea individuals. Loci that amplified successfully and showed polymorphism were combined into further multiplexes. The construction of multiplexes was done by means of successive attempts of adding and removing loci from sets of markers that had amplified together in the initial test panels until at least six loci were successfully combined in a panel. When amplified loci were monomorphic, the procedure was repeated on seven Scotian Shelf samples to assess whether they were monomorphic in these individuals.

Multiplex PCRs were performed in 5 µl reactions with 50 ng template DNA, 1 × Multiplex PCR Master Mix (Qiagen), 0.2 µM of each reverse primer, 0.05 µM of each unlabelled forward primer (modified with the appropriate universal tail) and 0.2 µM of labelled universal primer for each forward primer labelled with matching universal tail. Further adjustments made to optimize concentrations of primers in the PCR reactions are given in Table 1. PCR thermal cycling conditions were as follows: 1 × 95°C (15 min); 30 × 94°C (30 s), 60°C (90 s), 72°C (60 s); 8 × 94°C (30 s), 53°C (90 s), 72°C (60 s); 1 × 60°C (30 min). No-template controls were included to monitor for potential contamination.

A total of 1 µl of the multiplex PCR product was added to 9 µl of Super-DI Formamide (MCLAB) with 0.01 µl of Orange DNA Size Standard (MCLAB) and run on an ABI 3130xl Genetic Analyzer according to manufacturer's recommendations. GeneMarker ver. 1.97 (<www.soft genetics.com>) was used for fragment length analysis.

Microsatellite genotyping

All 46 Celtic Sea and 46 Scotian Shelf samples were genotyped with multiplex panels. Genotype data were inspected with Micro-Checker ver. 2.2.3 for genotyping errors and presence of null-alleles (van Oosterhout et al. 2004) using default settings. The 99% confidence interval was used when checking for null alleles to avoid false positives resulting from multiple tests. MSAnalyser ver. 4.05 (Dieringer and Schlötterer 2003), using default settings, was used to assess the number of alleles, allelic richness, allele size ranges, F_{ST} estimates and expected and observed heterozygosity. Data were analysed for possible departure from Hardy–Weinberg equilibrium, linkage disequilibrium, and excess and deficit of heterozygotes using Genepop ver. 4.2 with default settings (Raymond and Rousset 1995; Rousset 2008). False discovery rate (FDR) was used to correct for multiple comparisons (Benjamini and Yekutieli 2001) with initial $\alpha = 0.05$. Lositan (Antao et al. 2008) was used to detect loci that could be under positive or balancing selection (settings "Neutral' mean F_{ST}' and 'Force mean F_{ST}' with 10 000 simulations were used under both the infinite allele model and stepwise mutation model).

F_{ST} replicate sampling

The current study purposefully aimed to validate more markers than required for accurate evaluation of population differentiation (i.e. multilocus F_{ST}). To estimate the number of microsatellite loci future studies on cod population structure may require we investigated how many markers were needed to accurately estimate multilocus F_{ST}. Data sets were generated by randomly drawing 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 loci from the 55-locus dataset with each condition (number of loci) replicated ten times. Average F_{ST} and 95% confidence interval of the ten replicates were calculated and plotted to visualise the variability of average F_{ST} estimates as a function of numbers of markers (Fig. 1).

RESULTS

Primer design

Of 1309 candidate microsatellite loci, 559 were determined to be suitable for primer design upon visual inspection. A total of 349 primer pairs were rejected based on

T_m difference, sequence redundancy or secondary structure with the 5′ tail. Another 18 were excluded due to a BLAST match (one match to a *Gadus morhua* microsatellite, nine to a *G. morhua* gene and eight to a possible contaminant). The remaining 192 primer pairs were chosen for validation. Of those, 51 failed to amplify and 45 were excluded due to low scorability. Unambiguous amplification of PCR products in the expected size range was successful in 96 of the 192 markers tested (50%), of which 73 showed polymorphism (38%). The 73 polymorphic loci were used to build multiplex panels. Of these 13 were not included in the final multiplexes because of incompatible size, associated fluorochrome or failure to amplify with the other markers in a panel. As a result 60 markers were combined into six multiplex panels ranging between eight and twelve loci. Five markers were not used in the final analysis due to ambiguous genotyping leading to high chance of scoring errors, resulting in a final panel of 55 polymorphic markers combined into six multiplexes (Table 1). The results of the BLAST search on the validated primers against the *G. morhua* genome are presented in the Supplementary material Appendix 1 Table A1.

Application of markers to test populations

The mean allelic richness (R_S) was 7.1 (SD = 4.11) in the Celtic Sea sample and 7.2 (SD = 4.24) in the Scotian Shelf sample. The minimum number of alleles was two for both the Celtic Sea and Scotian Shelf samples; the maximum number of alleles was 21 and 25, respectively (Supplementary material Appendix 1 Table A2). Micro-Checker analyses indicated no genotyping errors. However, ten loci had a different repeat pattern than the motif originally identified from the raw sequence (Supplementary material Appendix 1 Table A3). Null alleles were observed in 13 loci in the Celtic Sea sample and 11 in the Scotian Shelf sample (Supplementary material Appendix 1 Table A4). Twelve and eleven loci deviated significantly from Hardy–Weinberg equilibrium (after FDR correction) in Celtic Sea and Scotian Shelf samples, respectively. Linkage disequilibrium was observed (after FDR correction) in locus pair A43_T3 x C01_M13 in the Celtic Sea sample; and locus pairs A43_T3 x B19_T3 and C15_Hill x C17_M13 in the Scotian Shelf sample. Lositan identified loci A11_Hill, C40_M13, C42_M13 and D14_Hill as being potentially affected by positive selection, (both under IA and SMM). Only C40_M13 (global $F_{ST} = 0.581$), C42_M13 (global $F_{ST} = 0.301$) and D14_Hill (global $F_{ST} = 0.246$) remained significant after correction for multiple comparisons.

Global multilocus F_{ST} was estimated at 0.067 (P = 0.001). After the three outlier loci identified by Lositan were excluded, the global F_{ST} was estimated at
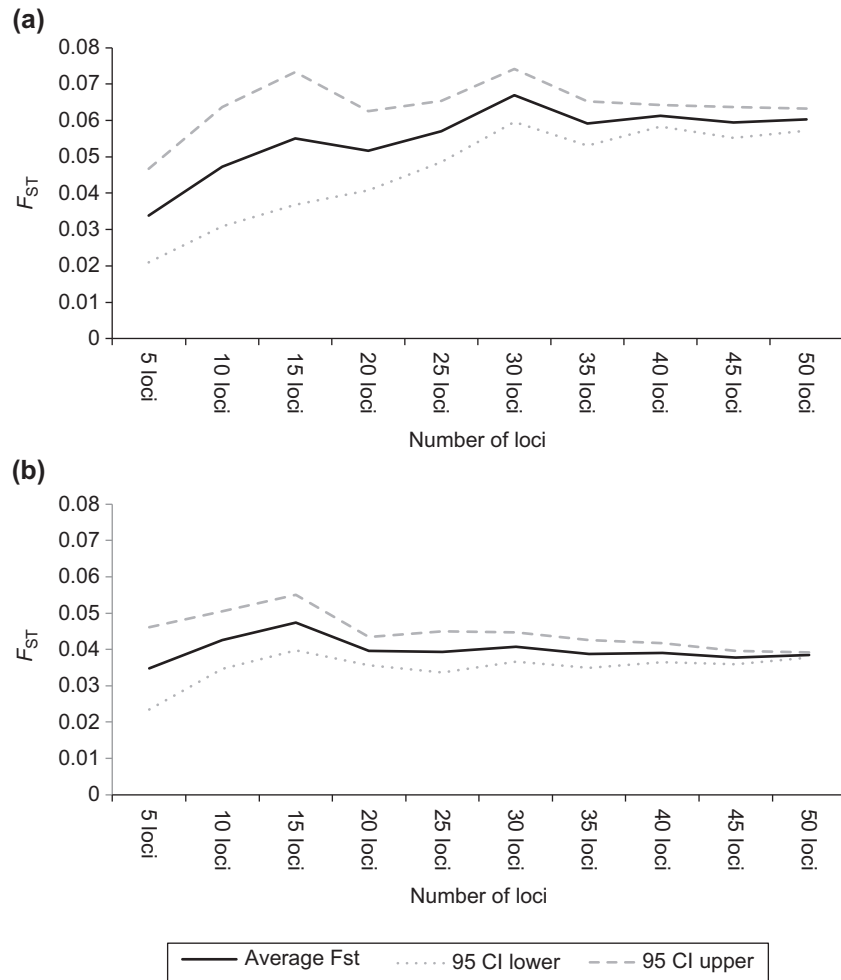
Fig. 1a–b. F_{ST} replicate sampling with all 55 loci (**a**) and with outliers excluded (**b**). 95% confidence interval is displayed.

0.043 (P = 0.001). Global F_{ST} values of individual loci are presented in Supplementary material Appendix 1 Fig. A1.

Replicate sampling of loci to visualise the effect of increasing numbers of loci on F_{ST} estimates and their variances is presented in Fig. 1. This was done in order to see how many loci were needed to reach F_{ST} point estimates with low variances to accurately describe the level of genetic variability between the Celtic Sea and Scotian Shelf samples. The procedure was performed both with and without loci under potential selection. In both cases, increased number of markers reduced the variation in multilocus F_{ST} estimates.

DISCUSSION

Since the initial reports of three-primer PCR (Steffens et al. 1993; Oetting et al. 1995; Neilan et al. 1997; Schuelke 2000) the approach has gained wide acceptance, particularly for initial validation while using conventional

two-primer PCR for genotyping (Guichoux et al. 2011; Hunter and Hart 2013; Olafsdottir et al. 2013; Skirnisdottir et al. 2013). Similarly, multiplex amplification of microsatellites is now commonly employed. In a few instances, these two techniques have been combined for microsatellite deployment (Missiaggia and Grattapaglia 2006; Langen et al. 2011; Blacket et al. 2012). However, the combined three primer/multiplex PCR approach, as used here for both microsatellite development and deployment, has not to our knowledge been previously reported. The lack of such studies may reflect conservative views on multiplexing and/or the limited availability of suitable universal primers (Guichoux et al. 2011; Blacket et al. 2012). The three-primer/multiplex PCR approach for validation and genotyping has several characteristics that facilitate cost savings (consumables and labour) relative to other approaches. Fluorescently labelled primers are typically an order of magnitude more expensive than unlabelled primers. Therefore, direct modification of locus specific

Table 1. Six multiplex panels for amplifying 55 loci. Underlined sequence in the forward primer signifies the universal tail sequence, whereas underlined sequence in the reverse primer signifies the PIG-tail sequence. Adjustment to primer concentration applies for all three primers used to amplify a locus (forward, reverse and universal primer). SRA accession number for the sequences is SRP041380.

Multiplex	Name	Motif	Size range	Adjustment to primer concentration	Forward primer	Reverse primer
1	A08_T3	tetra	163–187		AATTAACCCTCACTAAAGGGATCTCGAGTGGCGCAGTAG	GTTTCTTGCACGCTGACAAGTGAAGAG
	A16_M13	tetra	156–194		GGATAACAATTTCACACAGGCGTTCCCTCAGCTTGTCTCT	GTTTCTTCGTCCAACAGATCTATGCAATC
	A43_T3	tetra	309–369		AATTAACCCTCACTAAAGGGCCACTTTAACCTGCGGTTTC	GTTTCTTGCCTGCTTGTATACGCTGA
	B19_T3	tri	242–320		AATTAACCCTCACTAAAGGGAAAAGAATCTGCTTGCCTCA	GTTTCTTCGTGAGCTCAGTTTTGGCTA
	B38_Neo	tri	388–415		AGGTGAGATGACAGGAGATCGAATTGAGGAGGCATGGGTA	GTTTCTTGGTTAATTCCAGCCGTAGAGG
	C15_Hill	tetra	169–205		TGACCGGCAGCAAAATTGCCTTTCGTTCTCTCCGTCAG	GTTTCTTGAGGATTGGTGGGATGAT
	C28_Neo	tri	255–282		AGGTGAGATGACAGGAGATCCAGCACAAGTGGTAGGGTCA	GTTTCTTGCGATCAGAAGTTGTGCTT
	C36_Hill	tri	319–352		TGACCGGCAGCAAAATTGGTTGGCTCACACAATCATCG	GTTTCTTATCCTTCAAACAGCCCTCAA
	C40_M13	tri	276–277		GGATAACAATTTCACACAGGGGTCTTTGGGAGGTCTTCCT	GTTTCTTCATCCTGCTTGCGGACTTAT
	D14_Hill	tetra	228–276		TGACCGGCAGCAAAATTGGGCGATAATCTGCCATTTTG	GTTTCTTCGTGACAAGTGTCGATTTGC
	D30_M13	tetra	317–333		GGATAACAATTTCACACAGGTTCACAAACGGGAACTACGA	GTTTCTTGACAAGTCCAAGATGTGTCA TCA
2	A18_M13	tetra	203–243		GGATAACAATTTCACACAGGGACTGTCCGTTGAG GGTGTT	GTTTCTTGGGTCGAACTGGTCTGGTTA
	A19_T3	tetra	242–274		AATTAACCCTCACTAAAGGGTCCTGGTTCCAACACATGAC	GTTTCTTCTAGCCAATGGTGCAAGT
	A34_M13	tetra	297–321	2x	GGATAACAATTTCACACAGGTCCTTAACGACAGGCACCTT	GTTTCTTCCGACTTGTGTCGTTCCAG
	A37_Neo	tetra	265–489		AGGTGAGATGACAGGAGATCCCGGCAGTACAGC TAATGAA	GTTTCTTAAATGCTCAACCCATTGGAC
	B03_T3	di	147–161		AATTAACCCTCACTAAAGGGGCGATAATAGCGTTCCCATCC	GTTTCTTGGGGTACCTTGTGACCTGT
	B12_T3	di	346–398		AATTAACCCTCACTAAAGGGGCTTTGGCAACACTGTTTGA	GTTTCTTGTCGAGCAGACCAGAAGACC
	B30_Neo	tri	151–169		AGGTGAGATGACAGGAGATCTTGTACGGACAGG AAGTCCA	GTTTCTTGCGAACAGTGTGTAAATTGAA
3	C01_M13	tetra	115–195		GGATAACAATTTCACACAGGACCAGGAGGTTGGATCAGTG	GTTTCTTCCATTATTCATTCGTCATCCA
	C14_Hill	tetra	146–166		TGACCGGCAGCAAAATTGAGGTTCAGCCAGAAGCTGAT	GTTTCTTCCATTGGTTGTCGGTGATTA
	C20_Neo	tri	227–239		AGGTGAGATGACAGGAGATCCCGCCTATCACCCTA AATCTG	GTTTCTTCGTCTACATGTCGTGGTAGGG
	D12_Hill	tetra	251–291		TGACCGGCAGCAAAATTGAACGGCTCCTCAAGACAAAC	GTTTCTTAGGCATCTGCGTCCATACTC
	A33_M13	tetra	233–297		GGATAACAATTTCACACAGGAGACACTGAGCTC GACAGCA	GTTTCTTATCAATGATCCCAGGCAAAC
	A39_Hill	tetra	356–416		TGACCGGCAGCAAAATTGCCTGTCCAAATGCACACAAG	GTTTCTTAGTGCTTGGATGGTGTGATG
	B01_Hill	di	129–213		TGACCGGCAGCAAAATTGTAGACTCTGGGGCTGGGTAA	GTTTCTTGCATCCGAGACTCTGTTCC
	B15_T3	di	170–178		AATTAACCCTCACTAAAGGGCGATGCGATTCTT GGTAAATG	GTTTCTTGGTGTCTCATCCCTCCTTCA

Table 1. *Continued.*

Multiplex	Name	Motif	Size range	Adjustment to primer concentration	Forward primer	Reverse primer
	B29_Neo	tri	262–281		AGGTGAGAGATGACAGGAGGAGATCGGGAAAGAGAGCCGGAAAAGTA	GTTTCTTGCTAATGTTGGCAGAAACCA
	C13_Neo	tri	158–176		AGGTGAGAGATGACAGGAGGAGATCGGGTGATTGAGGTTGCGATA	GTTTCTTGCTCACACATCCTACGAGCA
	C17_M13	tetra	154–190		GGATAACAATTTCACACAGGCTTCTCGATGGCATGTTTCC	GTTTCTTCCTGCACAATGATCTGCAT
	C22_Hill	tetra	225–273		TGACCGGCAGCAAAATTGGGCTTGTCGTTGGTTCCTT	GTTTCTTGTTGAATGCAACCCTCAGT
	C30_Neo	tri	337–343		AGGTGAGAGATGACAGGAGATCAGGTGGTCGCAGTGAAGAAG	GTTTCTTGGGTGAATGCCTCTTAAITCG
	C42_M13	tri	345–358		GGATAACAATTTCACACAGGGCTGAGGGGATGCGATAATA	GTTTCTTAGCCAAGGGTGAAGTGTTGT
	D37_T3	tetra	313–333		AATTAACCCTCACTAAAGGGATGTGACACCGAATCACAGC	GTTTCTTACCCGTCCTGTACGTGAACT
4	A11_Hill	tetra	163–175		TGACCGGCAGCAAAATTGCGACAGGGAGGCATAAAGAC	GTTTCTTGTTCACCTCCTGGCTCTT
	A22_Neo	tetra	214–243		AGGTGAGATGACAGGAGGAGATCGGTGAGGTTCTTGAGGGTCA	GTTTCTTGATTATTTCCCCCTGCTG
	A31_T3	tetra	314–386		AATTAACCCTCACTAAAGGGGGGATATGTGGGGATGAGCAC	GTTTCTTATGGGTCCTTCTTCCTTTGGT
	B33_M13	tri	120–135	0.5x	GGATAACAATTTCACACAGGCTACAGCAGGGGTTCCTCAG	GTTTCTTGTTTGTTGCTCCGATGGACT
	C08_T3	tetra	145–185		AATTAACCCTCACTAAAGGGGCTCGGACCCAGAGATCAAAA	GTTTCTTGCAGCATCTGAACTGAAACG
	D15_Hill	tetra	244–256		TGACCGGCAGCAAAATTGTGACTCAACGGAGGTACGTG	GTTTCTTCCATCAGGATCAGGACCACT
	D21_Neo	tetra	337–389		AGGTGAGATGACAGGAGATCAACACGCTTGCTGGGACTAC	GTTTCTTCACTGGAGTGTACGGTCTCTGA
	D46_Neo	tetra	139–155		AGGTGAGATGACAGGAGATCCTCCTAATACCATGTCACCA	GTTTCTTCGTCTGTTCACGGATGCAC
5	A04_Neo	tetra	143–211		AGGTGAGAGATGACAGGAGGAGATCACAATCAACCCTCCAACTCG	GTTTCTTCAGGTCCGAATATCAAGG
	B07_M13	di	163–187		GGATAACAATTTCACACAGGTGACAATTACATTGAAAATCACAG	GTTTCTTCCTGAACTGCCTGTCAATG
	B28_Neo	tri	226–238		AGGTGAGATGACAGGAGGAGATCCCCACCCCTTAAIGTTTCAA	GTTTCTTGGCGTCATTCTCTTTGATG
	D10_Neo	tetra	308–312	2x	AGGTGAGAGATGACAGGAGGAGATCCGCCAATGCAAAITCTCTTTT	GTTTCTTATCTGAGTGCGTCGAGTGC
	D35_M13	tetra	316–360		GGATAACAATTTCACACAGGTCCACACTTGGTCGATGAAA	GTTTCTTGACGAGTGTCAGCAGGTGTG
	D43_T3	tetra	260–292	2x	AATTAACCCTCACTAAAGGGGTGCCGCTCACGCTACTAAT	GTTTCTTCGTGATCGCTCTCGATTC
6	A03_Hill	tetra	158–170		TGACCGGCAGCAAAATTGGAGCGTGTTGAAACGACTTGA	GTTTCTTCCTGAGCAGTGGAGTGACAA
	A20_Neo	tetra	164–182		AGGTGAGAGATGACAGGAGATCCGAGGCTACACAGCCTGTAA	GTTTCTTACTGTGGGCATGTAAACAGCA
	A30_T3	tetra	245–269		AATTAACCCTCACTAAAGGGAGTTGAACTGCGGGTTCTGT	GTTTCTTGCACGATGTCACAGCTGATT
	B36_Hill	tri	384–396		TGACCGGCAGCAAAATTGCCGCCCAGACATAATAAGA	GTTTCTTCAGTCTCAGCCTCCACATCA
	C31_Neo	tri	332–341		AGGTGAGAGATGACAGGAGATCGCCAAGACAAGCAITTCCAT	GTTTCTTCGAGCCAGCGTTTACTTCTC
	C35_M13	tri	233–348		GGATAACAATTTCACACAGGGCAATGTCGTACACCTCAA	GTTTCTTCTGAACGGCAACACTTCGTA
	D05_Hill	tetra	230–270		TGACCGGCAGCAAAATTGACTGCCCCTGAIAACAAIGC	GTTTCTTAGGCATCGACCATTTGTAGC
	D16_T3	tetra	424–436		AATTAACCCTCACTAAAGGGCCAGCAGCTTTCTGGGTAGT	GTTTCTTCGAAGCGTTACTGCAGACAG

primers substantially increases project costs, especially when markers must be excluded due to low scorability and/or bias (Selkoe and Toonen 2006). The use of fluorescently labelled universal primers avoids these potential complications and further decreases project costs, as a limited number of these primers can be purchased at large synthesis scales. Further multiplexing six to eleven amplicons per ABI capillary lane reduces PCR and genotyping costs as well as labour effort.

Primers were designed to amplify loci in three non-overlapping allele size ranges per dye as reported by Neff et al. (2000), however, larger size separations between ranges were employed. Overlapping size ranges have a disadvantage in that only one marker can be used per dye (Miller et al. 2013a, 2013b). Because actual allele sizes were not known in advance of capillary separation, two gap sizes (30 bp and 50 bp) between marker class size ranges were used to minimize overlap chances within a dye set. In practice, only a single overlap between markers was observed in the combined set of gap sizes, while the remaining markers were separated by at least 8 bp. However, microsatellites generated using the 50 bp gap size were more easily combined in multiplex PCR. We therefore recommend that marker size classes are set apart by at least 50 bp during primer design (e.g. 100–150, 200–250, 300–450 bp) to facilitate combining loci in multiplexes.

Both raw reads and contigs have been used with similar success for microsatellite discovery (Table 2). Contigs can yield more robust primers because increased sequencing depth can be used to detect sequencing errors or genetic variation in the primer binding region (Fernandez-Silva et al. 2013; Zalapa et al. 2012). However, repeat-containing reads may fail to assemble during contig construction, preventing discovery of some valid microsatellite loci (sensu Cavagnaro et al. 2010). Also, if the assembler is not able to distinguish the repeat and uses it as the basis for alignment, the unique flanking regions can easily be erroneously collapsed (Treangen and Salzberg 2012). We used raw reads to maximise microsatellite yield in this study and were able to achieve a 50% amplification success rate for trialled primers.

The design of a multiplex panel usually starts with evaluation of loci in single locus PCR reactions (Neff et al. 2000; Guichoux et al. 2011). For this study, evaluating 192 loci in single PCR reactions would have required 192 additional PCRs on the validation panel of seven individuals and a negative control, and the analyses of the resulting 1536 amplicons via capillary electrophoresis. The elimination of this step reduced primer validation

Table 2. *Comparison of previous studies using either contigs or raw reads in microsatellite discovery via 454 pyrosequencing. The studies employed 454 GS-FLX Titanium chemistry, apart from the publications marked with *which used the 454 GS-FLX chemistry.*

Data	Taxon name	Primers screened	Polymorphic	Primer-to-polymorphic marker proportion	Reference		
Contig	*Neophoca cinerea*	28	12	0.43	Ahonen et al. 2013	Average	0.45
	Cyanoramphus malherbi	35	18	0.51	Andrews et al. 2013*	Median	0.43
	Catha edulis	63	27	0.43	Curto et al. 2013*		
	Stylissa carteri	96	12	0.13	Giles et al. 2013		
	Python molurus bivittatus	26	18	0.69	Hunter and Hart 2013		
	Popenaias popeii	28	20	0.71	Inoue et al. 2013		
	Isoodon obesulus	46	9	0.20	Li et al. 2013		
	Antilocapra americana sonoriensis	100	14	0.14	Munguia-Vega et al. 2013		
	Scomber scombrus	80	30	0.38	Olafsdottir et al. 2013		
	multiple species	16–81	8–25	0.15–0.88	Schoebel et al. 2013		
	Unio crassus	77	11	0.14	Sell et al. 2013*		
	Cyclopterus lumpus	48	22	0.46	Skirnisdottir et al. 2013		
	Kunzea pulchella	27	10	0.37	Tapper et al. 2013		
Raw	*Gadus morhua*	15	6	0.40	Carlsson et al. 2013	Average	0.44
	Mulloidichthys flavolineatus	24	23	0.96	Fernandez-Silva et al. 2013	Median	0.38
	Pleuromamma xiphias	15	8	0.53	Fernandez-Silva et al. 2013		
	Brachyptera braueri	30	5	0.17	Geismar and Nowak 2013		
	Euastacus bispinosus	40	15	0.38	Miller et al. 2013a*		
	Neophema chrysogaster	55	14	0.25	Miller et al. 2013b*		
	Prionace glauca	100	12	0.12	Taguchi et al. 2013		
	Silurus asotus	70	47	0.67	Xu et al. 2013		

time and lowered consumable and labour costs. The present approach will yield markers for use in multiplex panels. However we recognise that some markers that would amplify in single locus PCR may fail in multiplex, therefore potentially lowering the conversion proportion from tested loci to polymorphic loci. Nevertheless, we contend that the increase in speed outweighs the possible loss of potential markers.

The conversion proportion from tested loci to polymorphic loci in this study was 38% (73/192). This value was similar to the conversion proportion of 40% (6/15) observed in the initial small scale validation by Carlsson et al. (2013) using the same data, and is consistent with recent studies using raw reads from 454 GS-FLX sequencing of genomic DNA (Table 2). Considerable variability of primer-to-polymorphic marker proportion has been observed among studies (Table 2). A portion of this variation can be attributed to differences in the genome composition of the study organisms (Schoebel et al. 2013), for example, PCR amplification success is lower in organisms with comparatively large genomes (Garner 2002; Schoebel et al. 2013). This can be due to larger genomes typically harbouring more repetitive elements (Hancock 2002). Microsatellite discovery strategies, such as differences in search parameters and algorithms, or using contigs versus raw sequence reads, can possibly affect the conversion from tested loci to polymorphic loci as well. Variation is also likely caused by different strategies in selection of loci to be validated (Fernandez-Silva et al. 2013).

Implementation in G. morhua

As a proof of concept we applied the described approach to samples of *G. morhua* from the Celtic Sea and the Scotian Shelf. *G. morhua* from the Celtic Sea were used for initial microsatellite development (Carlsson et al. 2013) and the Scotian Shelf *G. morhua* form a genetically distinct population from eastern Atlantic *G. morhua* (Hutchinson et al. 2001; O'Leary et al. 2007). The present study estimated F_{ST} between Celtic Sea and Scotian Shelf *G. morhua* at 0.067 when 55 loci were employed. After exclusion of three loci that were potentially under positive selection, F_{ST} was estimated at 0.043. The reduction in F_{ST} is consistent with previous studies that have demonstrated that inclusion of outlier loci that are potentially under selection can markedly affect F_{ST} estimates (Nielsen et al. 2006; Allendorf et al. 2010). The presence and scale of population structure between Celtic Sea and Scotian Shelf *G. morhua* in the present study concurs with previous studies that examined these populations (Hutchinson et al. 2001; O'Leary et al. 2006, 2007; Pampoulie et al. 2008), and with additional studies that demonstrated population differentiation

between the eastern and western Atlantic *G. morhua* (i.e. allozymes, Mork et al. 1985; minisatellite, Galvin et al. 1995; nuclear RFLPs, Pogson et al. 1995, 2001; microsatellites, Bentzen et al. 1996; Hutchinson et al. 2001; O'Leary et al. 2007; Pampoulie et al. 2008; single nucleotide polymorphisms, O'Leary et al. 2006; Nielsen et al. 2009; Bradbury et al. 2010).

The combined microsatellite validation and genotyping approach presented here was designed to be a fast and cost-effective means for developing and deploying large numbers of microsatellite markers. Using larger numbers of genetic markers confers considerable advantages of increased precision and statistical power when assessing intra- and inter-population genetic parameters such as population structure and gene flow, as well as when inferring demographic parameters, such as effective population size, population expansions and bottlenecks (Nei and Tajima 1981; Ryman et al. 2006). This improved precision allows for more robust and trustworthy management advice based on genetic data. In the present case, the rate of reduction of multilocus F_{ST} variability decreased after 20–30 loci suggesting that this is the point where using more loci only slightly improves the precision of the multilocus F_{ST} estimate. The point of diminishing returns may not be the same for other populations, other geographic scales or other organisms. The advantage of the method presented here is that more loci can be effectively genotyped, ensuring that the point of diminishing returns has in fact been reached and the most precise estimate of population genetic parameter acquired.

Conclusions

The current study combines three-primer PCR with multiplexing to allow for more economical, rapid development and deployment of microsatellite markers discovered from high throughput sequencing data. Fifty-five polymorphic *G. morhua* microsatellites were combined into six PCR multiplexes, which allowed for determination of F_{ST} between two populations with high precision. This approach is transferable to any species, including those for which extensive sequence resources are not available, and will allow for large and robust population genetic studies while minimising expensive and labour intensive capillary sequencing runs.

REFERENCES

Abdelkrim, J., Robertson, B., Stanton, J. A. et al. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. – BioTechniques 46: 185–192.

Ahonen, H., Lowther, A. D., Goldsworthy, S. D. et al. 2013. Characterization of 12 novel microsatellite loci and cross-amplification of four loci in the endangered Australian sea lion (*Neophoca cinerea*). – Conserv. Genet. Resour. 5: 283–285.

Allendorf, F. W., Hohenlohe, P. A. and Luikart, G. 2010. Genomics and the future of conservation genetics. – Nat. Rev. Genet.11: 697–709.

Allentoft, M., Schuster, S. C., Holdaway, R. et al. 2009. Identification of microsatellites from an extinct moa species using high-throughput (454) sequence data. – Biotechniques 46: 195–200.

Altschul, S. F., Gish, W., Miller, W. et al. 1990. Basic local alignment search tool. – J. Mol. Biol. 215: 403–410.

Andrews, B. J., Hale, M. L. and Steeves, T. E. 2013. Characterisation of microsatellite loci in the critically endangered orange-fronted kākāriki (*Cyanoramphus malherbi*) isolated using genomic next generation sequencing. – Conserv. Genet. Resour. 5: 235–237.

Antao, T., Lopes, A., Lopes, R. J. et al. 2008. LOSITAN: a workbench to detect molecular adaptation based on a F_{st}-outlier method. – BMC Bioinf. 9: 323.

Ashworth, V. E. T. M., Kobayashi, M. C., De La Cruz, M. et al. 2004. Microsatellite markers in avocado (*Persea americana* Mill.): development of dinucleotide and trinucleotide markers. – Sci. Hort. 101: 255–267.

Benjamini, Y. and Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. – Ann. Stat. 29: 1165–1188.

Bentzen, P., Taggart, C. T., Ruzzante, D. E. et al. 1996. Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. – Can. J. Fish. Aquat. Sci. 53: 2706–2721.

Blacket, M. J., Robin, C., Good, R. T. et al. 2012. Universal primers for fluorescent labelling of PCR fragments – an efficient and cost-effective approach to genotyping by fluorescence. – Mol. Ecol. Resour. 12: 456–463.

Bradbury, I. R., Hubert, S., Higgins, B. et al. 2010. Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. – Proc. Biol. Sci. 277: 3725–3734.

Brockmann, H. J., Colson, T. and Potts, W. 1994. Sperm competition in horseshoe crabs (*Limulus polyphemus*). – Behav. Ecol. Sociobiol. 35: 153–160.

Brownstein, M. J., Carpten, J. D. and Smith, J. R. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. – BioTechniques 20: 1004–1010.

Bruford, M. W. and Wayne, R. K. 1993. Microsatellites and their application to population genetic studies. – Curr. Opin. Genet. Dev. 3: 939–943.

Carlsson, J., Gauthier, D. T., Carlsson, J. E. et al. 2013. Rapid, economical single-nucleotide polymorphism and microsatellite discovery based on de novo assembly of a reduced representation genome in a non-model organism: a case study of Atlantic cod *Gadus morhua*. – J. Fish Biol. 82: 944–958.

Cavagnaro, P. F., Senalik, D. A., Yang, L. et al. 2010. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). – BMC Genomics 11: 569.

Chambers, G. K. and MacAvoy, E. S. 2000. Microsatellites: consensus and controversy. – Comp. Biochem. Physiol. B Biochem. Mol. Biol. 126: 455–476.

Curto, M. A., Tembrock, L. R., Puppo, P. et al. 2013. Evaluation of microsatellites of *Catha edulis* (qat; Celastraceae) identified using pyrosequencing. – Biochem. Syst. Ecol. 49: 1–9.

Dieringer, D. and Schlötterer, C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. – Mol. Ecol. Notes 3: 167–169.

Diniz, F. M., Iyengar, A., Sarmanho da Costa Lima, P. et al. 2007. Application of a double-enrichment procedure for microsatellite isolation and the use of tailed primers for high throughput genotyping. – Genet. Mol. Biol. 30: 380–384.

Edwards, A., Civitello, A., Hammond, H. A. et al. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. – Am. J. Hum. Genet. 49: 746–756.

Ellis, J. R. and Burke, J. M. 2007. EST-SSRs as a resource for population genetic analyses. – Heredity 99: 125–132.

Estoup, A., Jarne, P. and Cornuet, J. M. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. – Mol. Ecol. 11: 1591–1604.

Fernandez-Silva, I., Whitney, J., Wainwright, B. et al. 2013. Microsatellites for next-generation ecologists: a post-sequencing bioinformatics pipeline. – PLoS ONE 8: e55990.

Galvin, P., McGregor, D. and Cross, T. 1995. A single locus PCR amplified minisatellite region as a hypervariable genetic marker in gadoid species. – Aquaculture 137: 31–40.

Garner, T. W. 2002. Genome size and microsatellites: the effect of nuclear size on amplification potential. – Genome 45: 212–215.

Geismar, J. and Nowak, C. 2013. Isolation and characterisation of new microsatellite markers for the stonefly *Brachyptera braueri* comparing a traditional approach with high throughput 454 sequencing. – Conserv. Genet. Resour. 5: 413–416.

Giles, E. C., Saenz-Agudelo, P., Berumen, M. L. et al. 2013. Novel polymorphic microsatellite markers developed for a common reef sponge, *Stylissa carteri*. – Mar. Biodivers. 43: 237–241.

Goldstein, D. B., Roemer, G. W., Smith, D. A. et al. 1999. The use of microsatellite variation to infer population structure and demographic history in a natural model system. – Genetics 151: 797–801.

Guichoux, E., Lagache, L., Wagner, S. et al. 2011. Current trends in microsatellite genotyping. – Mol. Ecol. Resour. 11: 591–611.

Hagell, S., Whipple, A. V. and Chambers, C. L. 2013. Population genetic patterns among social groups of the endangered Central American spider monkey (*Ateles geoffroyi*) in a human-dominated landscape. – Ecol. Evol. 3: 1388–1399.

Hancock, J. M. 2002. Genome size and the accumulation of simple sequence repeats: implications of new data from genome sequencing projects. – Genetica 115: 93–103.

Huang, X. and Madan, A. 1999. CAP3: a DNA sequence assembly program. – Genome Res. 9: 868–877.

Hunter, M. E. and Hart, K. M. 2013. Rapid microsatellite marker development using next generation pyrosequencing to inform invasive Burmese python – *Python molurus bivittatus* –management. – Int. J. Mol. Sci. 14: 4793–4804.

Hutchinson, W. F., Carvalho G. R. and Rogers, S. I. 2001. Marked genetic structuring in localised spawning populations of cod *Gadus morhua* in the North Sea and adjoining waters, as revealed by microsatellites. – Mar. Ecol. Prog. Ser. 223: 251–260.

Inoue, K., Lang, B. K. and Berg, D. J. 2013. Development and characterization of 20 polymorphic microsatellite markers for the Texas hornshell, *Popenaias popeii* (Bivalvia: Unionidae), through next-generation sequencing. – Conserv. Genet. Resour. 5: 195–198.

Knapik, E. W., Goodman, A., Ekker, M. et al. 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). – Nat. Genet. 18: 338–343.

Langen, K., Schwarzer, J., Kullmann, H. et al. 2011. Microsatellite support for active inbreeding in a cichlid fish. – PLoS ONE 6: e24689.

Li, Y. C., Korol, A. B., Fahima, T. et al. 2004. Microsatellites within genes: structure, function, and evolution. – Mol. Biol. Evol. 21: 991–1007.

Li, Y., Lancaster, M. L., Cooper, S. J. B. et al. 2013. Characterization of nine microsatellite loci from the endangered southern brown bandicoot (*Isoodon obesulus*) using 454 pyrosequencing. – Conserv. Genet. Resour. 5: 105–107.

Malausa, T., Gilles, A., Meglécz, E. et al. 2011. High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. – Mol. Ecol. Resour. 11: 638–644.

Miller, A. D., Van Rooyen, A., Sweeney, O. F. et al. 2013a. The development of 10 novel polymorphic microsatellite markers through next generation sequencing and a preliminary population genetic analysis for the endangered Glenelg spiny crayfish, *Euastacus bispinosus*. – Mol. Biol. Rep. 40: 4415–4419.

Miller, A. D., Good. R. T., Coleman. R. A. et al. 2013b. Microsatellite loci and the complete mitochondrial DNA sequence characterized through next generation sequencing and de novo genome assembly for the critically endangered orange-bellied parrot, *Neophema chrysogaster*. – Mol. Biol. Rep. 40: 35–42.

Mirimin, L., O'Keeffe, D., Ruggiero, A. et al. 2011. A quick, least-invasive, inexpensive and reliable method for sampling *Gadus morhua* postlarvae for genetic analysis. – J. Fish Biol. 79: 801–805.

Missiaggia, A. and Grattapaglia, D. 2006. Plant microsatellite genotyping with 4-color fluorescent detection using multiple-tailed primers. – Genet. Mol. Res. 5: 72–78.

Moore, S. S., Sargeant, L. L., King, T. J. et al. 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. – Genomics 10: 654–660.

Mork, J., Ryman, N., Stahl, G. et al. 1985. Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. – Can. J. Fish. Aquat. Sci. 42: 1580–1587.

Munguia-Vega, A., Klimova, A., Culver, M. 2013. New microsatellite loci isolated via next-generation sequencing for two endangered pronghorn from the Sonoran Desert. – Conserv. Genet. Resour. 5: 125–127.

Neff, B. D., Fu, P. and Gross, M. R. 2000. Microsatellite multiplexing in fish. – Trans. Am. Fish. Soc. 129: 584–593.

Nei, M. and Tajima, F. 1981. Genetic drift and estimation of effective population size. – Genetics 98: 625–640.

Neilan, B. A., Wilton, A. N. and Jacobs, D. 1997. A universal procedure for primer labelling of amplicons. – Nucleic Acids Res. 25: 2938–2939.

Nielsen, E. E., Hansen, M. M. and Meldrup, D. 2006. Evidence of microsatellite hitch-hiking selection in Atlantic cod (Gadus morhua L.): implications for inferring population structure in nonmodel organisms. – Mol. Ecol. 15: 3219–3229.

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A. et al. 2009. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). – BMC Evol. Biol. 9: 276.

Oetting, W. S., Lee, H. K., Flanders, D. J. et al. 1995. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. – Genomics 30: 450–458.

Olafsdottir, G., Olafsson, K., Skirnisdottir, S. et al. 2013. Isolation and characterization of thirty microsatellite loci for Atlantic mackerel (*Scomber scombrus* L.). – Conserv. Genet. Resour. 5: 491–494.

O'Leary, D. B., Coughlan, J., McCarthy, T. V. et al. 2006. Application of a rapid method of SNP analysis (glycosylase mediated polymorphism detection) to mtDNA and nuclear DNA of cod *Gadus morhua*. – J. Fish Biol. 69(Suppl. A): 145–153.

O'Leary, D. B., Coughlan, J., Dillane, E. et al. 2007. Microsatellite variation in cod *Gadus morhua* throughout its geographic range. – J. Fish Biol. 70(Suppl. C): 310–335.

Pampoulie, C., Stefánsson, M. Ö., Jörundsdóttir, T. D. et al. 2008. Recolonization history and large-scale dispersal in the open sea: the case study of the North Atlantic cod, *Gadus morhua* L. – Biol. J. Linn. Soc. 94: 315–329.

Pogson, G. H., Mesa, K. A. and Boutilier, R. G. 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. – Genetics 139: 375–385.

Pogson, G. H., Taggart, C. T., Mesa, K. A. et al. 2001. Isolation by distance in the Atlantic cod, *Gadus morhua*, at large and small geographic scales. – Evolution 55: 131–146.

Primmer, C. R., Møller, A. P. and Ellegren, H. 1996. A wide-range survey of cross-species microsatellite amplification in birds. – Mol. Ecol. 5: 365–378.

Primmer, C. R., Koskinen, M. T. and Piironen, J. 2000. The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. – Proc. Biol. Sci. 267: 1699–1704.

Raymond, M. and Rousset, F. 1995. GENEPOP (ver. 1.2): population genetics software for exact tests and ecumenicism. – J. Hered. 86: 248–249.

Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. – Mol. Ecol. Resour. 8: 103–106.

Rozen, S. and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. – Meth. Mol. Biol. 132: 365–386.

Rubin, B. E., Makarewich, C. A., Talaba, A. L. et al. 2009. Isolation and characterization of microsatellite markers from the acacia-ant *Crematogaster mimosae*. – Mol. Ecol. Resour. 9: 1212–1214.

Ryman, N., Palm, S., André, C. et al. 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. – Mol. Ecol. 15: 2031–2045.

Schlötterer, C., Amos, B. and Tautz, D. 1991. Conservation of polymorphic simple sequence loci in cetacean species. – Nature 354: 63–65.

Schoebel, C. N., Brodbeck, S., Buehler, D. et al. 2013. Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. – J. Evol. Biol. 26: 600–611.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. – Nat. Biotechnol. 18: 233–234.

Selkoe, K. A. and Toonen, R. J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. – Ecol. Lett. 9: 615–629.

Sell, J., Mioduchowska, M., Kaczmarczyk, A. et al. 2013. Identification and characterization of the first microsatellite loci for the thick-shelled river mussel *Unio crassus* (Bivalvia: Unionidae). – J. Exp. Zool. A Ecol. Genet. Physiol. 319: 113–116.

Skirnisdottir, S., Olafsdottir, G., Olafsson, K. et al. 2013. Twenty-two novel microsatellite loci for lumpfish (*Cyclopterus lumpus*). – Conserv. Genet. Resour. 5: 177–179.

Squirrell, J., Hollingsworth, P. M., Woodhead, M. et al. 2003. How much effort is required to isolate nuclear microsatellites from plants? – Mol. Ecol. 12: 1339–1348.

Star, B., Nederbragt, A. J., Jentoft, S. et al. 2011. The genome sequence of Atlantic cod reveals a unique immune system. – Nature 477: 207–210.

Steffens, D. L., Sutter, S. L. and Roemer, S. C. 1993. An alternate universal forward primer for improved automated DNA sequencing of M13. – BioTechniques 15: 580–582.

Taguchi, M., Shigenobu, Y., Ohkubo, M. et al. 2013. Characterization of 12 polymorphic microsatellite DNA loci in the blue shark, *Prionace glauca*, isolated by next generation sequencing approach. – Conserv. Genet. Resour. 5: 117–119.

Tapper, S-L., Byrne, M. and Gardner, M. G. 2013. Characterisation of microsatellite markers for the granite endemic *Kunzea pulchella* (Lindl.) A. S. George (Myrtaceae) identified using next generation sequencing. – Conserv. Genet. Resour. 5: 129–131.

Tozaki, T., Mashima, S., Hirota, K. et al. 2001. Characterization of equine microsatellites and microsatellite-linked repetitive elements (eMLREs) by efficient cloning and genotyping methods. – DNA Res. 8: 33–45.

Treangen, T. J. and Salzberg, S. L. 2012. Repetitive DNA and next-generation sequencing: computational challenges and solutions. – Nat. Rev. Genet. 13: 36–46.

Untergasser, A., Nijveen, H., Rao, X. et al. 2007. Primer3Plus, an enhanced web interface to Primer3. – Nucleic Acids Res. 35: W71–74.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. et al. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. – Mol. Ecol. Notes. 4: 535–538.

Xu, D., Zhang, Y. and Peng, Z. 2013. High-throughput microsatellite marker development in Amur catfish (*Silurus asotus*) using next-generation sequencing. – Conserv. Genet. Resour. 5: 487–490.

Zalapa, J. E., Cuevas, H., Zhu, H. et al. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. – Am. J. Bot. 99: 193–208.

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