MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2012) doi: 10.1111/j.1755-0998.2012.03126.x

Exploitation of a turbot (*Scophthalmus maximus* L.) immune-related expressed sequence tag (EST) database for microsatellite screening and validation

R. NAVAJAS-PÉREZ,* F. ROBLES,* M. J. MOLINA-LUZÓN,* R. DE LA HERRÁN,* J. A. ÁLVAREZ-DIOS,† B. G. PARDO,‡ M. VERA,‡ C. BOUZA‡ and P. MARTÍNEZ‡
*Departamento de Genética, Facultad de Ciencias, Campus de Fuentenueva s/n, Universidad de Granada, 18071 Granada, Spain, †Departamento de Matemática Aplicada, Facultad de Matemáticas, Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain, ‡Departamento de Genética, Facultad de Veterinaria, Universidad de Santiago de Compostela, Campus de Lugo, 27002 Lugo, Spain

Abstract

In this study, we identified and characterized 160 microsatellite loci from an expressed sequence tag (EST) database generated from immune-related organs of turbot (*Scophthalmus maximus*). A final set of 83 new polymorphic microsatellites were validated after the analysis of 40 individuals of Atlantic origin including both wild and farmed individuals. The allele number and the expected heterozygosity ranged from 2 to 18 and from 0.021 to 0.951, respectively. Evidences of null alleles at moderate–high frequencies were detected at six loci using population data. None of the analysed loci showed deviations from Mendelian segregation after the analysis of five full-sib families including approximately 92 individuals/family. The markers are used to consolidate the turbot genetic map, and because they are mostly EST-derived, they will be very useful for comparative genomic studies within flatfishes and with model fish species. Using an *in silico* approach, we detected significant homologies of microsatellite sequences with the EST databases of the flatfish species with highest genomic resources (Senegalese sole, Atlantic halibut, bastard halibut) in 31% of these turbot markers. The conservation of these microsatellites within Pleuronectiformes will pave the way for anchoring genetic maps of different species and identifying genomic regions related to productive traits.

Keywords: cross-species analysis, EST database, microsatellites, Pleuronectiformes, *Scophthalmus maximus*, turbot

Received 24 November 2011; revision received 13 January 2012; accepted 25 January 2012

Introduction

Expressed sequence tags (ESTs) resources have increased during the last decade in aquaculture species including fish (Pardo *et al.* 2008; Sha *et al.* 2010; Bowman *et al.* 2011), shellfish (Hedgecock *et al.* 2005; Zhang & Guo 2010) and crustaceans (Du *et al.* 2010; Gorbach *et al.* 2010; Leu *et al.* 2011). These resources are essential to develop molecular markers [microsatellites and single nucleotide polymorphism (SNPs)] useful for constructing genetic maps and QTL identification and for population screening and parentage analysis (Canario *et al.* 2008).

The turbot (*Scophthalmus maximus*; Scophthalmidae; Pleuronectiformes) is a flatfish species for which human consumption has sharply risen in the last decade. In fact, a burgeoning industry is being developed mainly in Europe and lately in China, currently making turbot a promising aquaculture species. Genetic studies are being conducted to improve turbot production by optimizing broodstock organization and through genetic breeding programmes. The development of highly polymorphic, codominant and easily assayed molecular markers, such as microsatellites, is necessary to support breeding programmes.

Microsatellite loci have been characterized in turbot to evaluate the genetic resources (Coughlan *et al.* 1996; Estoup *et al.* 1998; Iyengar *et al.* 2000; Bouza *et al.* 2002) and to construct genetic linkage maps (Bouza *et al.* 2007, 2008; Pardo *et al.* 2007; Martínez *et al.* 2008; Ruan *et al.* 2010). Drawing medium- to high-density maps is a requirement for the screening of QTL related to productive traits and for their subsequent application in marker-assisted selection (MAS) programmes. However, most microsatellites characterized to date have been isolated from genomic DNA libraries, and thus, anonymous microsatellites (type II markers) have been mainly used for genetic map construction in turbot (Bouza *et al.* 2007; Martínez *et al.* 2008; Ruan *et al.* 2010). As these markers are mostly associated with nonannotated sequences, they

Correspondence: Paulino Martínez, Fax: +34 982822428;
E-mail: paulino.martinez@usc.es

are less conserved than EST linked markers, and so they show lower cross-species amplification with related organisms hampering comparative genomic analyses. Accordingly, microsatellites linked to genes (type I markers) are useful to identify the syntenies by comparative mapping with model species or with other flatfish species where cross-amplification has proven to be feasible (Cerdá *et al.* 2010). A few EST-derived microsatellites have been described in turbot (Chen *et al.* 2007; Bouza *et al.* 2008), and to date, only 31 of them have been mapped in this species (Bouza *et al.* 2008). It is therefore necessary to characterize new EST-derived markers in order to consolidate the turbot map and to facilitate comparative genomic studies. Additionally, enlarging EST-linked microsatellite resources will be useful to search for adaptive variation in turbot populations (Vilas *et al.* 2010).

As in other economically important flatfish species—Senegalese sole (*Solea senegalensis*; Cerdà *et al.* 2008), bastard halibut (*Paralichthys olivaceus*; Aoki *et al.* 1999; Arma *et al.* 2005), winter flounder (*Pseudopleuronectes americanus*, Douglas *et al.* 1999), European flounder (*Platichthys flesus*; Williams *et al.* 2006) and Atlantic halibut (*Hippoglossus hippoglossus*; Douglas *et al.* 2007)—an EST database was developed in turbot from cDNA libraries of immune-related organs (Pardo *et al.* 2008). This database, which contains 6170 unique sequences, was updated with new ESTs from a nodavirus-infected head kidney library (Park *et al.* 2009), and from cDNA libraries of new immune organs (pyloric caeca and thymus) and pathogens (*Enteromyxum scophthalmi*) (Vera *et al.* 2011). Additionally, 4339 sequences from previously exploited microsatellite-enriched genomic libraries were included in the turbot database (Pardo *et al.* 2006).

In this work, we developed and technically validated a large set of mostly EST-associated microsatellite markers using the updated turbot EST database. The aim of our study was to increase EST-linked microsatellite resources for (i) consolidating the turbot genetic map; (ii) enhancing comparative genomic strategies; and (iii) searching for adaptive variation in natural populations. Additionally, these microsatellite-containing sequences were used to screen EST databases of other flatfish (Senegalese sole, Atlantic halibut and bastard halibut) in order to carry out a preliminary evaluation on their conservation and utility for comparative genomics within Pleuronectiformes.

Materials and methods

EST database and microsatellite screening

Microsatellite-bearing sequences were selected from the updated turbot database (Pardo *et al.* 2006; Vera *et al.* 2011). Mining of microsatellites was carried out using the SPUTNIK program (http://espressosoftware.com/sputnik/index.html) while looking for dinucleotide motifs with more than five repeats, trinucleotide motifs with more than three repeats and tetranucleotide motifs with more than two repeats. Sequences from SPUTNIK output were selected according to the following criteria: (i) high-quality and enough flanking region for primer design; (ii) annotated ESTs preferentially; and (iii) appropriate technical parameters (product size between 100 and 300 pb; primer Tm 54–65 °C; primer %GC up to 50; Max self-complementarity = 5.00; Max 3′ self-complementarity = 3.00; Max Poly-X = 5) for primer pair design using the PRIMER3 program (Rozen & Skaletsky 2000).

Microsatellite genotyping

Standard phenol–chloroform protocols (Sambrook *et al.* 1989) were used to extract the DNA from the caudal fin in a sample of turbot from natural and cultured populations. Each selected microsatellite was amplified at a range of annealing temperatures and $MgCl_2$ concentrations in four individuals and checked in 2% agarose gels. The microsatellites that showed appropriate amplification (discrete bands of expected size) were genotyped on an ABI 3100 DNA sequencer using the forward primer labelled for fluorescent detection. Each amplification reaction was carried out in a 15-μL reaction mixture containing 30 ng of DNA sample; 1.5–2 mM of $MgCl_2$; 10 mM Tris, pH = 8.3; 5 mM NH_4Cl; 50 mM KCl; 0.2 mM of each deoxyribonucleotide triphosphate; 5 pmol of both forward and reverse PCR primers; and 0.5 U Taq DNA polymerase for a initial denaturation at 94 °C for 10 min, 35 cycles of denaturation at 94 °C for 1 min, variable annealing temperature for 45 s, and extension at 72 °C for 45 s (for full details, see Table 1). A final extension step was performed at 72 °C for 10 min. Multiplex PCR from two to four microsatellites was carried out when temperature and $MgCl_2$ concentrations were similar and amplification size and/or label colour were compatible. Results were analysed using GeneMapper 3.7 software (Applied Biosystems).

Gene diversity and population analysis

Genetic diversity was evaluated on 40 turbot individuals, all of them of Atlantic origin: 22 coming from a natural population of NW Spain and 18 parents or grandparents from seven unrelated families (F1–F7) currently used for genetic mapping and QTL identification (Bouza *et al.* 2008; Martínez *et al.* 2009; Rodríguez-Ramilo *et al.* 2011; Sánchez-Molano *et al.* 2011). These families came from the 2nd generation of the genetic breeding programme of Stolt Sea Farm SA (SSF), and pedigree information was

Table 1 Summary of the 83 polymorphic microsatellite loci validated from the turbot (*Scophthalmus maximus*) expressed sequence tag database

Locus	Primers 5′–3′	Annealing temp [MgCl2]	No. of alleles	Repeated motif	Size	Expected heterozygosity	Hardy–Weinberg *P*-value
Sma-E50	F: CACATCGTTGGGACAATCAG R: AGCAGTGAGCCTCTTTGGAC	60 °C 1.5 mM	8	(GT)9(GA)16	242–274	0.562	0.190
Sma-E51	F: ATTGCTTCACGGATTGTTCC R: TAGAGGAGGCCCACACAGAC	60 °C 1.5 mM	2	(AC)9	254–256	0.344	1.000
Sma-E52	F: CTCGATGATGTGCCAGAAGA R: GAGAGAGAAGCGGAACAGGA	56 °C 1.5 mM	2	(TAT)4	305–307	0.021	1.000
Sma-E61	F: ATGGGTAATGAAATGGTCCG R: GGATTCCCTGGCTTTACTCC	54 °C 1.5 mM	6	(TA)7	640–657	0.628	0.000
Sma-E71	F: CAGATCGTCTTCTCGCTCCT R: TGAGAGGAGTCACTTGTCCG	60 °C 1.5 mM	4	(GCT)8	139–152	0.672	0.157
Sma-E72	F: GGAGACACACAGTGCCGAC R: CGTTCTCCTAAGTTGCAGCG	62 °C 1.5 mM	11	(ACA)13	203–244	0.768	0.718
Sma-E74	F: ACCGGCTGTGTCTCTTGC R: CGGGTGTTCGAGAAGTACG	54 °C 1.5 mM	4	(CCG)13	403–418	0.599	0.916
Sma-E78	F: AAGAACTGCATCGACCGACT R: CGTGTGTTTCCATCAACTGG	58 °C 1 mM	2	(CA)5	175–177	0.229	1.000
Sma-E79	F: GCAGCGACTTGCTTCTTTCT R: GTCAGTTTGTGGTGTGTGGG	58 °C 1.5 mM	15	(GT)6-(AT)14-(GT)9-(TA)7	274–318	0.794	0.766
Sma-E82	F: TTGAACGGAACTTCTACACTCG R: GCGGTTTCGTCGTTAGTGTT	58 °C 1.5 mM	3	(CCT)4	113–119	0.650	0.835
Sma-E84	F: TGCATCTATTCCTGTTGGTGA R: TGTTGGTTCATAACTGAGCGAC	47 °C 1.5 mM	5	(GT)19	98–108	0.758	0.190
Sma-E86	F: CAGGAGGACTTCTCTGCCAA R: TTTACTCTCCACAGGCAGCA	51 °C 1.5 mM	9	(TA)18	292–325	0.852	1.000
Sma-E91	F: GACGGACGATACCTGCTGAT R: ACACTCGCCTCGTTTCTCAT	59 °C 1.5 mM	2	(TG)5	186–188	0.461	0.060
Sma-E96	F: TCTGCTGGCTCACCTTAACA R: ATAGGGTCTGCACTCATGGC	54 °C 1.5 mM	9	(ACC)8	646–668	0.676	0.000
Sma-E97	F: CTAACAGACGCAAATGCACC R: CCATGCAAACACTCACCTGT	56 °C 1.5 mM	11	(CT)(GT)25	289–331	0.668	0.930
Sma-E99	F: AACGACTTCTCCAGAGCCAA R: TACAGACAGATGACGGCTCG	54 °C 1.5 mM	3	(CAG)4	328–342	0.351	0.694
Sma-E100	F: CCGAGCTAACCACTGACCTT R: CGAGCACGCAGTAATGGATA	54 °C 1.5 mM	3	(TG)5	303–311	0.509	0.817
Sma-E105	F: TTCACAAACCACATCCAAGG R: TGGCACAAGCTCAAACTGAC	60 °C 1.5 mM	3	(GT)6	286–302	0.200	1.000
Sma-E112	F: GGTGCAGGCCATAGTCATTT R: TGTGAGTGATTCGGCAACAG	59 °C 1.5 mM	6	(TA)12	275–294	0.721	0.373
Sma-E113	F: CACACATCCACAGACTCGCT R: AAACATTCCTCTCAGTGCCG	52 °C 1.5 mM	8	(TA)25	318–330	0.710	0.247
Sma-E117	F: GCACAAACAGACAAACACGC R: TCAAATGCAACCATGACGTT	57 °C 1.5 mM	3	(CA)9	321–333	0.520	0.000
Sma-E118	F: TATTATGGAGGGATCGGCTG R: TCAACGTGATGTTTGCCTTC	55 °C 1.5 mM	7	(TG)21	216–244	0.485	0.067
Sma-E120	F: TACTGGGTCTACTGGGTGCC R: CCGTCCGTTTCCTTCAAATA	52 °C 1.5 mM	4	(AGG)4	206–216	0.376	0.001
Sma-E127	F: TGAGATTTGCATGGATGTGG R: GACTCCTGGCTCCTCCTTCT	52 °C 1.5 mM	6	(GT)10	78–102	0.791	0.304
Sma-E128	F: CTTCATCGCCATCTCCATTT R: GGCCGAATACTCCGATAACA	55 °C 1.5 mM	18	(ATC)8	270–312	0.917	0.156
Sma-E132	F: GGTCGGTCATCTCGTAGCAT R: AAGCCCTGCACATGGAAGTA	59 °C 1.5 mM	7	(GCC)6	350–370	0.669	0.000
Sma-E134	F: CGGCTTTCTCTCCTCCTGTT R: AGCTCACGGCCAGATTAGAA	48 °C 1.5 mM	2	(TG)7	115–119	0.142	1.000

Table 1 (Continued)

Locus	Primers 5′–3′	Annealing temp [MgCl2]	No. of alleles	Repeated motif	Size	Expected heterozygosity	Hardy–Weinberg *P*-value
Sma-E136	F: ATGGAGACTCACACGGAGGT R: AAAGAACTTCGGCACTGAGG	51 °C 1.5 mM	3	(TGT)6	198–204	0.279	0.572
Sma-E137	F: CTGTGTCCCTTGGAGATGGT R: AAAGGGTCGTGCAGAAGCTA	57 °C 1.5 mM	6	(CGC)5	614–624	0.778	0.005
Sma-E139	F: GAACAATGACTTGCTGCTGG R: AGCTGAAGACGCTATGGGAG	50 °C 1.5 mM	4	(CTG)10	400–408	0.672	0.500
Sma-E142	F: TCCATCGCAATATCACAGGA R: TCAAACAAAGCTGCACAAGC	58 °C 1.5 mM	9	(AGA)12	266–294	0.793	0.002
Sma-E144	F: CTTCTACAGCCAAACGAGGG R: CATTGATGCGCCTTTCCTAT	48 °C 1.5 mM	7	(TA)10	292–306	0.793	0.020
Sma-E145	F: CTGTCTCCCTGTCCGTCTGT R: GAGAAGCTCGGGATGATGAC	58 °C 1.5 mM	3	(TC)6	430–440	0.395	1.000
Sma-E154	F: CTCTTCTCTGCGTTTCTGCC R: GAGTCTCGTGAACCTGGAGC	52 °C 1.5 mM	14	(TCC)4	560–588	0.919	0.000
Sma-E156	F: GTGATGAGGGTGATGAGGGT R: CCAGCCTCTCTTTGTTGCTC	54 °C 1.5 mM	4	(CTG)10	337–346	0.647	0.002
Sma-E158	F: GTCTCGCACTTCCTGTCTCC R: TGGAATCTGTCCGTCTGTTG	56 °C 1.5 mM	4	(TCT)9	309–336	0.452	1.000
Sma-E159	F: GATCAATGTGGTCCTCCACC R: CTCCTTCTCCAAGTCCACCA	52 °C 1.5 mM	3	(TC)7	156–161	0.357	0.072
Sma-E164	F: ATTCTCAGCCATCTGGAACC R: AGTGATGACCACGACCACAA	54 °C 1.5 mM	3	(TA)10	289–293	0.294	0.300
Sma-E167	F: TTACGTTTGTGAGTCGTCTGG R: CATCAGTCCACATCCGTCTG	63 °C 1.5 mM	2	(CA)6	86–88	0.221	0.540
Sma-E168	F: CGTCTTTGTACGCGAAGCTC R: GATTTCAAAGTCAAGGCCCA	63 °C 1.5 mM	3	(AC)9	117–129	0.451	0.842
Sma-E170	F: TTCACCATGAAGCCATGAAA R: TGACGTAACAAGACGGAGGA	54 °C 1.5 mM	9	(AC)18	328–346	0.796	0.812
Sma-E174	F: CCCAGATGAGACATGGACAA R: ACAGTATGTGGGCCTTTCAG	58 °C 1.5 mM	3	(CA)12	82–86	0.563	0.055
Sma-E180	F: AGAGCAATGTAAGCGCCTTT R: CTTGGTACAGCATTCACGATG	58 °C 1.5 mM	3	(AC)12	218–222	0.485	0.403
Sma-E183	F: GAAACAGGAAGGGAACAGCA R: CTTTGGTCCTTGCCAACACT	58 °C 1.5 mM	2	(TTG)5	289–292	0.145	1.000
Sma-E184	F: AGGACGACACAACCATCACA R: AACCTCCTCTCTCTGGAGCC	58 °C 1.5 mM	6	(GGA)9	241–266	0.754	0.001
Sma-E187	F: GTTCGTGTCGCTGAAGATGA R: ACAGACGGAACAGCAGTGAG	50 °C 1.5 mM	2	(CTC)9	279–288	0.099	1.000
Sma-E189	F: CGACTGACCTCTGCATCGTA R: GCCTCCTGAAGACGCTATTG	56 °C 1.5 mM	4	(TG)8	271–279	0.552	0.390
Sma-E191	F: GGAGGGCGAAGAAGAAGAAG R: GCTGCTCCAGTCTGCGTT	58 °C 1.5 mM	6	(CGA)4	267–282	0.670	0.247
Sma-E194	F: CCACACGTTGCTATACACGG R: ACGGTAAGAGAGGAGACGCA	52 °C 1.5 mM	6	(TC)10	116–126	0.669	0.905
Sma-E195	F: CGCCTGAGAGTTTCTCTTCC R: AACAAACAAAGCTCCGCAGT	52 °C 1.5 mM	11	(TG)23	313–344	0.820	0.000
Sma-E197	F: AGCTCTGTTGGAGGAACACG R: GTAGCAGAGGAGCTGGATGG	58 °C 1.5 mM	2	(GAC)4	378–380	0.082	1.000
Sma-E205	F: GTCCCGGTGAGGAGTACAGA R: TCAGCCGGATAGGGAAGATA	54 °C 1.5 mM	2	(AGG)6	235–237	0.498	0.400
Sma-E215	F: TGTTGCATTCCGAGAAACTG R: GACCATGCCCTTGATTTGTT	57 °C 1.5 mM	2	(AGA)11	410–419	0.569	0.444

Table 1 (Continued)

Locus	Primers 5′–3′	Annealing temp [MgCl2]	No. of alleles	Repeated motif	Size	Expected heterozygosity	Hardy–Weinberg *P*-value
Sma-E218	F: GGATTGGCTTCTGAAATGGA R: GAGGCTGGACACCAAGACTG	49 °C 1.5 mM	3	(AAAC)6	169–177	0.413	0.900
Sma-E220	F: CAGGATTGAGGAGGAGCTTG R: ACCACAGACTTGGACCTTGG	55 °C 1.5 mM	2	(AC)5	212–214	0.175	1.000
Sma-E224	F: GCTCAGAGAGAAGAGAGCGG R: CACCACAGCAAGTATGCCAC	59 °C 1.5 mM	4	(AACA)7	248–267	0.746	0.078
Sma-E225	F: GCCAAAGGAATGTCGGTAAA R: CACACACACACACTCACCCA	56 °C 1.5 mM	3	(GT)12	284–288	0.604	0.620
Sma-E227	F: GAAGGCGGTAATCATCCAGA R: GCTTCACACCTGCTGTTTCA	58 °C 1.5 mM	2	(CT)12	315–319	0.474	0.216
Sma-E231	F: CAGTTGTGGGTGTGAGGTTG R: GGTCACGAGAGAAATGAGGC	57 °C 1.5 mM	5	(TCC)6	302–308	0.490	0.512
Sma-E244	F: TCCATGCAAAGCAGACACAT R: CACACCGTGCATTCAAGTTC	54 °C 1.5 mM	3	(CA)9	303–307	0.417	0.140
Sma-E248	F: TGGGACTTAATGGGACAAGG R: GAATACCCACCCAAATGCAC	58 °C 1.5 mM	3	(TG)18	271–287	0.764	0.786
Sma-E254	F: ATGCCGTCCCACTACAGTTC R: CCACATTCTACTGGCGAGGT	54 °C 1.5 mM	3	(GCT)8	188–241	0.648	0.797
Sma-E255	F: TCTATGGAGCCCACAAGTCC R: TCAACCTGGTGAAGAAAGGC	56 °C 1.5 mM	3	(CA)6	314–324	0.509	0.000
Sma-E261	F: CTGGAAGGAGGAAAGAACCC R: GCTGAGCGGAGAGAGAGAGA	51 °C 1.5 mM	8	(CT)7	90–95	0.532	0.786
Sma-E270	F: TGACACCATTTCTGGGAACA R: GAGGCACGCGACTACTTCAC	57 °C 1.5 mM	2	(TTGA)3	319–332	0.474	0.444
Sma-E272	F: TGCAACTAGCCGATTTAACCA R: GTTGAGGACAAAGCCGAGAG	47 °C 1.5 mM	3	(CTT)4	207–212	0.192	0.068
Sma-E276	F: CTCAATCACGCTCTCACACG R: CCGAGGGACGGAGATACATA	63 °C 2 mM	4	(CA)12	115–145	0.475	0.236
Sma-E277	F: AGACACAAGCGCACACAGAC R: TCCAGAGCTGAACATCACCA	58 °C 2 mM	2	(TC)9	322–330	0.743	0.276
Sma-E279	F: TGTTATAGCCGACAGCAGCA R: TCACTCCCGGTCTGATGTTT	54 °C 2 mM	3	(TCC)9	297–313	0.687	0.756
Sma-E283	F: TCACAGCTTGGGCCTTATTT R: AGTTACAGCAGCAGGCAACA	54 °C 2 mM	8	(AC)15	283–293	0.672	0.222
Sma-E284	F: ACTTCATCCGCTTTGACTGC R: GGGCGAAGGAGTTGTGTTTA	57 °C 2 mM	5	(GT)9	319–323	0.444	0.457
Sma-E286	F: ACGACAGCGACACACACACT R: TACATTCGGTGACGATGCTG	54 °C 0.5 mM	7	(GT)12	243–257	0.756	0.756
Sma-E289	F: CAATGAGGACTGATGCTTCG R: GTTCAGCGACAGGAAGTGCT	54 °C 0.5 mM	7	(GT)30	240–366	0.951	0.324
Sma-E290	F: GAGACCCACAGACCTCGTGTA R: TGTTCTTTGGTCCCTTGCTC	53 °C 2 mM	4	(TG)16	336–354	0.762	0.548
Sma-E293	F: CTGTAGCAGCCTCCTCCCT R: GGAGAACAAAGTCCGTCCAC	59 °C 1.5 mM	7	(CGT)7	215–221	0.506	0.901
Sma-E294	F: GCATCGTGAAACACTGGAGA R: GAACGAACCAAACCACGACT	58 °C 2 mM	2	(TAA)6	218–224	0.221	0.067
Sma-E302	F: TCTTTGTCCAGAACAGTCGG R: CATGTGAAATTGGCAGCATC	57 °C 2 mM	2	(AGC)7	313–331	0.270	0.149
Sma-E305	F: GATTTGTGTGGAAACTGCCAT R: CCATGCAAACACTCACCTGT	56 °C 2 mM	3	(TG)16	211–215	0.302	0.006
Sma-E310	F: CGCTCCTGCACATCTACACT R: GGCTCCCTCAACACACAAAT	55 °C 2 mM	5	(AT)10	215–225	0.547	0.545
Sma-E315	F: CCGTTCAGAATACCTGCTCC R: TGTCGCTCTCTGCTGGTCTA	50 °C 2 mM	2	(TTA)8	187–190	0.379	0.720

Table 1 (Continued)

Locus	Primers 5′–3′	Annealing temp [MgCl2]	No. of alleles	Repeated motif	Size	Expected heterozygosity	Hardy–Weinberg *P*-value
Sma-E316	F: GAATGGAAATGGATGCAGTGT R: TGTAACAACCGTGTGTCTGTC	56 °C 1.5 mM	5	(AC)9	282–292	0.697	0.836
Sma-E317	F: GTGACCCTCTGACCTTTGCT R: ACACACCTCAGTGCAGAACG	58 °C 1.5 mM	3	(GT)8	80–84	0.557	0.440
Sma-E318	F: CCTGAACACTGGAACCTTCA R: AATAACTCACCTAGCACTCACG	56 °C 1.5 mM	3	(GT)14	247–257	0.579	0.332

available for all of them. However, because some parents or grandparents had died when offspring were collected, DNA was only available for 18 of them. Genotyping information from these families is highly valuable for further mapping analysis. The GENEPOP 4.0 program (Raymond & Rousset 1995; Rousset 2008) was used to estimate the genetic diversity [expected heterozygosity (He) and number of alleles (A)] and to check for conformance to Hardy–Weinberg (HW) expectations. To manage an appropriate sample size for these analyses, we decided to include the 18 individuals from SSF families because of their Atlantic origin and the low genetic structure reported for turbot populations from this area (Vilas *et al.* 2010). However, caution was taken especially when analysing genetic disequilibria by reason of divergence after two generations of selection. Micro-Checker 2.2.3 was used to investigate the causes of HW deviations (Van Oosterhout *et al.* 2004) and CERVUS 3.0.3 (Kalinowski *et al.* 2007) to estimate the frequency of null alleles and combined exclusion probabilities for parentage assignment.

Conformance to Mendelian segregation

Progenies ranging between 85 and 96 offspring in five of the aforementioned families (F1–F5) were finally used to evaluate the conformance to Mendelian segregation of polymorphic loci. Chi-square tests were applied to check the null hypothesis, using the progeny from at least one of the families. Segregation distortion was tested at each locus, adjusting the significance level for multiple tests within each type of segregation (1:1, 1:1:1:1 or 1:2:1) using Bonferroni correction (Rice 1989).

Microsatellite conservation analysis within Pleuronectiformes

We were interested in analysing the degree of evolutionary conservation of microsatellites, but especially within Pleuronectiformes, given the low cross-species amplification previously reported in this order (Bouza *et al.* 2002; Castro *et al.* 2006). The *in silico* information obtained would also be of practical interest to develop new microsatellites in other species and to obtain shared microsatellites useful for comparative genomics between flatfish species. The increase in EST resources in several Pleuronectiformes pertaining to different families made this analysis possible. Flanking regions of all polymorphic and monomorphic turbot microsatellites were BLASTed using BLASTn (Altschul *et al.* 1990) against EST resources of the three flatfish species with highest genomic information: bastard halibut (13 869 sequences), Atlantic halibut (20 886 sequences) and Senegalese sole (5208 sequences). All sequences were retrieved from the NCBI-EST database (http://www.ncbi.nlm.nih.gov/dbEST/) as they were of December 2010, except for *S. senegalensis* sequences, which were extracted from the Pleurogene project database (Cerdà *et al.* 2008). Best hits at $<10^{-5}$ *E*-value cut-off and >80 bp were considered. Alignments of significant hits were performed using ClustalX (Thompson *et al.* 1997) followed by manual adjustment. As the homologies found between sequences were based on flanking regions, we considered the microsatellite motif, the number of repeats and presence/absence status of the microsatellite to evaluate the degree of conservation. Accordingly, we established three different categories: (i) monomorphic microsatellites in turbot, with either identical or different repeat number in other flatfish species; (ii) polymorphic microsatellites in turbot with either identical or different repeat number in other species; and (iii) microsatellite-containing sequences in turbot lacking microsatellite in other species.

Annotation

For putative function determination and annotation, EST-bearing microsatellite sequences were BLASTed against several model fish genomes (http://genome.ucsc.edu/index.html) and the GenBank nucleotide collection (http://www.ncbi.nlm.nih.gov/genbank/) by means of BLASTx (*E*-value cut-off: $<10^{-3}$) or BLASTn (*E*-value cut-off: $<10^{-5}$) when no BLASTx hits were found.

Results

Microsatellite selection and amplification conditions

Among the 16 255 high-quality ESTs (>100 bp; PHRED >20) sequences included in the turbot database, we identified 298 microsatellite-containing sequences according to the established criteria and with long enough flanking regions to design primers. In 228 of them, a band with the expected size was found when the PCR products were run in agarose gels. In another 13 cases, the amplicons exceeded the expected size (with bands from 400 to 600 bp), which may indicate the presence of intron(s) in the genomic DNA. All the 241 primer pairs were tested in the automatic sequencer.

Gene diversity and population analysis

From the 241 technically selected microsatellites, 81 (33.6%) showed poor resolution, while the remaining 160 yielded unambiguous genotyping in the automatic sequencer. These 160 sequences were deposited in GenBank and given a corresponding accession number (Table S1). Among them, 77 (48.1%) corresponded to monomorphic loci and 83 (51.9%) resulted polymorphic. Genetic diversity was estimated in a sample of 40 individuals of Atlantic origin for all polymorphic loci. The number of alleles per locus ranged from 2 to 18 (mean = 4.9 ± 0.354) and expected heterozygosity between 0.021 and 0.951 (mean = 0.540 ± 0.039) (Table 1). Among the 83 polymorphic loci, six (Sma-E145, Sma-E225, Sma-E227, Sma-E231, Sma-E276 and Sma-E277) belonged to sequences reanalysed from partial genomic DNA libraries (Pardo *et al*. 2006) incorporated to the EST-enriched database. Seventy-six of the 83 polymorphic loci conformed to HW expectations after Bonferroni correction. The remaining 7 loci showed significant deviation owing to heterozygote excess (Sma-E154) or deficit (Sma-E61, Sma-E96, Sma-E117, Sma-E132, Sma-E195 and Sma-E255; Table 1), suggesting the presence of null alleles in the last group. Null allele frequencies were estimated using CERVUS 3.0.3: Sma-E61 (0.482), Sma-E96 (0.367), Sma-E117 (0.786), Sma-E132 (0.438), Sma-E195 (0.182) and Sma-E225 (0.149). The presence of null alleles was confirmed by using the Micro-Checker 2.2.3 software, which also suggested genotyping errors owing to stuttering at Sma-E61, Sma-E96, Sma-E117 and Sma-E132. No evidence for allele dropout was found at any locus ($P > 0.05$). Eight of 13 loci with amplicons longer than expected were polymorphic and three of them (Sma-E61, Sma-E96 and Sma-E154) deviated from HW expectations (Table 1). The probabilities to exclude a false parent for paternity inference ranged from 0.005 (Sma-E187) to 0.787 (Sma-E289) for the first parent (Excl1) and from 0.069 (Sma-E187) to 0.881 (Sma-E289) for the second parent (Excl2). Combined exclusion probabilities for parentage assignment using the 83 polymorphic loci were virtually 1 for both Excl1 and Excl2.

Family analysis

All but three microsatellites (Sma-E134, Sma-E145 and Sma-E293) could be tested for Mendelian segregation in the families. Eight of the 80 tests performed deviated from Mendelian expectations (Table S2), but none of them after Bonferroni correction.

Microsatellite conservation within flatfishes

In order to analyse the conservation of turbot microsatellite sequences in other flatfish species, we explored EST databases of three representative Pleuronectiformes: bastard halibut, Atlantic halibut and Senegalese sole, each belonging to different families (Paralichthyidae, Pleuronectidae and Soleidae, respectively). Of 160 sequences tested, including both polymorphic and monomorphic microsatellites, 49 (approximately 31% of the total) showed a significant match (*E*-value $<10^{-5}$) with any of the EST databases explored (Table 2). Of these, 10 (approximately 6%) showed a match with all species analysed, while 24 (approximately 15%) and 15 (approximately 9%) showed match with one or two species, respectively. We found a higher number of matches with the bastard halibut EST database (39) than with the other species: halibut (26) and sole (22). The

Table 2 Significant hits of the 160 turbot microsatellite-containing sequences with other flatfish species: Senegalese sole (*Solea senegalensis*), bastard halibut (*Paralichthys olivaceus*) and Atlantic halibut (*Hippoglossus hippoglossus*)

	Hits (%)	B. halibut	S. sole	Halibut	B. halibut/S. sole	B. halibut/A. halibut	S. sole/A. halibut
Turbot-specific microsatellites	111 (69.4)	–	–	–	–	–	–
Match with one species	24 (15)	15	3	6	–	–	–
Match with two species	15 (9.4)	–	–	–	6	7	2
Match with all species	10 (6.2)	–	–	–	–	–	–

S. sole, Senegalese sole; A. halibut, Atlantic halibut; B. halibut, bastard halibut.

remaining 111 sequences (approximately 69%) were turbot-specific and displayed no match with the EST databases explored (Table 2). On the other hand, variable sequence conservation was observed when comparing turbot microsatellite-containing sequences against the same flatfish databases (Table 3). Of the 49 homologous microsatellites found in the other flatfish databases, 26 were monomorphic and 23 polymorphic in turbot. All of them showed the same microsatellite motif in the other species where a significant match was detected, although with imperfections in some cases. Among monomorphic loci, nine showed the same repeat number in the other species, 13 showed different repeat numbers, and three lacked microsatellite. Among polymorphic ones, five microsatellites showed the same repeat number as a certain turbot allele, six differed in repeat number, and 10 lacked microsatellite. Approximately in half of the cases the homologous repetitive motif was imperfect. Despite the conservation of one of the flanking regions, no data could be obtained for three microsatellites because the corresponding EST sequences were incomplete in flatfish databases, and no information existed for the repetitive region (Table 3).

EST annotations

After BLASTing polymorphic microsatellite-containing sequences against different protein and nucleotide databases, most polymorphic loci showed significant sequence similarity with annotated genes and, in three cases (Sma-E254, Sma-E283 and Sma-E317), with genomic DNA or mRNA clones of fish model species (Table S3). Among the sequences belonging to partial genomic libraries, there were even three that showed significant homology with gene-related sequences (Sma-E145, Sma-E225 and Sma-E231). The other three could also be considered informative markers because they showed homology with specific unidentified cDNA sequences or with clones of genomic DNA regions of other fish species (Table S3).

Discussion

Genomic resources have greatly increased in aquaculture species especially after the arrival of new generation sequencing (NGS) technologies, and several transcriptomes and whole-genome sequencing projects are underway in different fish species (Davidson *et al.* 2010; Kuhl *et al.* 2010). Exploitation of these resources requires their organization in databases with appropriate bioinformatic tools for sequence edition, clustering, annotation and functional classification among others. EST databases have proven to be a valuable source of molecular markers such as microsatellites (Bouza *et al.* 2008; Cerdá *et al.* 2010) and SNPs (Pardo *et al.* 2008; Vera *et al.* 2011). Within economic important flatfish species such as *H. hippoglossus* (Douglas *et al.* 2007), *P. olivaceus* (Liu *et al.* 2006) and *Cynoglossus semilaevis* (Liu *et al.* 2007), EST sequences have been used to identify microsatellite loci mostly to be used for linkage map construction. In turbot, a screening of the EST database v1.0 (Pardo *et al.* 2008) yielded a set of 31 type I markers useful for genetic mapping and population genomics (Bouza *et al.* 2008; Vilas *et al.* 2010).

In this study, we characterized 83 new polymorphic microsatellites from the updated turbot EST database (Pardo *et al.* 2008; Vera *et al.* 2011). Genetic diversity of these 83 loci showed allele number and expected heterozygosity in the range previously observed in other EST microsatellites characterized in flatfish (Liu *et al.* 2006, 2007; Chen *et al.* 2007; Douglas *et al.* 2007), including turbot (Bouza *et al.* 2008). Deviation from HW expectations was found in seven loci (8.4%), mostly attributable to the presence of null alleles. The 83 microsatellites characterized in this study, together with the previous ones and the 77 SNPs characterized by Vera *et al.* (2011), constitute a suitable set of EST-linked markers to consolidate the turbot genetic map mostly based on anonymous markers (Bouza *et al.* 2007, 2008; Ruan *et al.* 2010). Also, these markers will be essential for comparative genomics strategies and to extend analysis on adaptive variation in turbot populations (Vilas *et al.* 2010).

Table 3 Classification of microsatellite sequence showing significant hits according to their motif, repeat number and presence/absence status

	Same motif				
Turbot loci	Same repeat number	Different repeat number	Lacking microsatellite	No data	Total
Monomorphic	9 (3)	13 (6)	4	1	26
Polymorphic	5 (2)	6 (3)	12	2	23
Total	14 (5)	19 (9)	16	3	49

All turbot microsatellites considered in the analysis were perfect. In parentheses, the number of imperfect repeats is given. When comparing polymorphic microsatellites, we considered the same repeat number if a match with any of turbot alleles was observed at that locus.

There are several flatfish species of great commercial value, and relevant genomic advances have been made in four of them (bastard halibut, Senegalese sole, Atlantic halibut and turbot). Cross-checking this information is then crucial to obtain insights on flatfish evolution and to identify the genomic regions and/or candidate genes related to productive traits. To achieve this goal, it is essential to construct a flatfish database, which would be additionally useful to identify the common genetic markers to anchor genetic and physical maps between species. In our study, we carried out a first approach by comparing the conservation of turbot microsatellites with the existing EST flatfish databases. This comparison provided data on microsatellite evolution and also useful information to identify common sets of microsatellites between different flatfish species. We could detect significant homology of 31% of turbot microsatellite flanking sequences in the other EST flatfish databases, which demonstrates a certain conservation of turbot microsatellites in the genomes of related species, as reported for other fish groups (Rico *et al*. 1996; DeWoody & Avise 2000). Higher similarities were detected with the bastard halibut EST database (39 matches) than with the other species [Atlantic halibut (26 matches); Senegalese sole (22 matches)]. This agrees with phylogenetic relationships of Pleuronectiformes, which place the family Scophthalmidae closer to Paralichthyidae and Pleuronectidae than to Soleidae (Pardo *et al*. 2005). Also, the lower similarity observed with Senegalese sole is in accordance with the low cross-species amplification previously reported (Castro *et al*. 2006). In our comparison, a notable proportion of perfect turbot microsatellites (43%) turned out to be imperfect in the other flatfish species analysed, and all data indicate that microsatellite conservation was higher at monomorphic loci, suggesting a higher evolutionary rate for polymorphic ones. Moreover, dramatic changes, involving the complete loss of the microsatellite, were observed at 27% of loci despite the fact that flanking regions were mostly conserved. Our study is preliminary and we cannot rule out a certain bias in these figures because of the incompleteness of the databases explored. In fact, the Atlantic halibut database is the best represented one with one and a half times more sequences than that of bastard halibut, both databases showing a similar average sequence size (around 600 bp). On the other hand, the Senegalese sole database contains only 25% sequences of the halibut database, but with higher average size (715 bp).

Microsatellite markers characterized in this study are being integrated together with a SNP panel (Vera *et al*. 2011) in the updated version of turbot genetic map (C. Bouza , M. Hermida, B. G. Pardo, M. Vera, C. Fernández, R. de la Herrán, R. Navajas, J.A. Álvarez-Dios, A. Gómez-Tato and P. Martínez, submitted) and in a new Senegalese sole map (M. J. Molina-Luzón, M. Hermida, J. I. Navas, F. Robles, R. Navajas-Pérez, P. Martínez, C. Ruiz-Rejón and R. de la Herrán, unpublished data). They are mostly annotated and have exhibited significant homologies with other flatfish species highly valuable for comparative mapping to look for candidate genes related to adaptive variation or productive characters.

Acknowledgements

This work was supported by the Consolider-Ingenio AQUAGENOMICS Project (CSD2007-00002) and two Nacional Plan (AGL2006-13158-C03-03/ACU and AGL2009-11872) Spanish Projects and by the Xunta de Galicia local Government Project (09MMA011261PR). Navajas-Pérez R is a Ramón y Cajal researcher funded by Spanish MICINN at University of Granada (Spain). Pardo B.G. was supported by an Isidro Parga Pondal research fellowship from Xunta de Galicia (Spain).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Aoki T, Nam BH, Hirono I, Yamamota E (1999) Sequences of 596 cDNA clones (565,977 bp) of Japanese flounder (*Paralichthys olivaceus*) leukocytes infected with hirame rhabdovirus. *Marine Biotechnology*, **1**, 477–488.

Arma NR, Hirono I, Aoki T (2005) Identification of genes expressed in the liver of Japanese flounder *Paralichthys olivaceus* by expressed sequence tags. *Fisheries Science*, **71**, 504–518.

Bouza C, Presa P, Castro J, Sánchez L, Martínez P (2002) Allozyme and microsatellite diversity in natural and domestic populations of turbot (*Scophthalmus maximus*) in comparison with other Pleuronectiformes. *Canadian Journal of Fisheries and Aquatic Sciences*, **59**, 1460–1473.

Bouza C, Hermida M, Pardo BG *et al*. (2007) A microsatellite genetic map of the turbot (*Scophthalmus maximus*). *Genetics*, **177**, 2457–2467.

Bouza C, Hermida M, Millán A *et al*. (2008) Characterization of EST-derived microsatellites for gene mapping and evolutionary genomics in turbot. *Animal Genetics*, **39**, 666–670.

Bowman S, Hubert S, Higgins B *et al*. (2011) An integrated approach to gene discovery and marker development in Atlantic cod (*Gadus morhua*). *Marine Biotechnology*, **13**, 242–255.

Canario A, Bargelloni L, Volckaert F, Houston RD, Massault C, Guiguen Y (2008) Genomics toolbox for farmed fish. *Reviews in Fisheries Science*, **16**, 1–13.

Castro J, Pino A, Hermida M *et al*. (2006) A microsatellite marker tool for parentage assessment in Senegal sole (*Solea senegalensis*): genotyping errors, null alleles and conformance to theoretical assumptions. *Aquaculture*, **261**, 1194–1203.

Cerdá J, Douglas S, Reith M (2010) Genomic resource for flatfish research and their applications. *Journal of Fish Biology*, **77**, 1045–1070.

Cerdà J, Mercadé J, Lozano JJ *et al*. (2008) Genomic resources for a commercial flatfish, the Senegalese sole (*Solea senegalensis*): EST sequencing, oligo microarray design, and development of the Soleamold bioinformatic platform. *BMC Genomics*, **9**, 508.

Chen SL, Ma HY, Jiang Y, Liao XL, Meng L (2007) Isolation and characterization of polymorphic microsatellite loci from an EST library of turbot (*Scophthalmus maximus*) and cross-species amplification. *Molecular Ecology*, **7**, 848–850.

Coughlan JP, McCarthy E, McGregor D *et al*. (1996) Four polymorphic microsatellites in turbot *Scophthalmus maximus*. *Animal Genetics*, **27**, 433–442.

Davidson WS, Koop BF, Jones SJM *et al*. (2010) Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biology*, **11**, 403.

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 46–473.

Douglas SE, Gallant JW, Bullerwell CE, Wolff C, Munholland J, Reith ME (1999) Winter flounder expressed sequence tags: establishment of an EST database and identification of novel fish genes. *Marine Biotechnology*, **1**, 458–464.

Douglas SE, Knickle LC, Kimball J, Reith ME (2007) Comprehensive EST analysis of Atlantic halibut (*Hippoglossus hippoglossus*), a commercially relevant aquaculture species. *BMC Genomics*, **8**, 144–155.

Du ZQ, Ciobanu DC, Onteru SK *et al.* (2010) A gene-based SNP linkage map for pacific white shrimp, *Litopenaeus vannamei*. *Animal Genetics*, **41**, 286–294.

Estoup A, Gharbi K, San Cristobal M, Chevalet C, Haffray P, Guyomard R (1998) Parentage assignment using microsatellites in turbot (*Scophthalmus maximus*) and rainbow trout (*Oncorhynchus mykiss*) hatchery populations. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 715–725.

Gorbach DM, Hu ZL, Du ZQ, Rothschild MF (2010) Mining ESTs to determine the usefulness of SNPs across shrimp species. *Animal Biotechnology*, **21**, 100–103.

Hedgecock D, Gaffney PM, Goulletquer P, Guo X, Reece K, Warr GW (2005) The case for sequencing the Pacific oyster genome. *Journal of Shellfish Research*, **24**, 429–441.

Iyengar A, Piyapattanakorn S, Heipel DA *et al.* (2000) A suite of highly polymorphic microsatellite markers in turbot (*Scophthalmus maximus* L.) with potential for use across several flatfish species. *Molecular Ecology*, **9**, 368–371.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.

Kuhl H, Beck A, Wozniak G, Canario AVM, Volckaert FAM, Reinhardt R (2010) The European sea bass *Dicentrarchus labrax* genome puzzle: comparative BAC-mapping and low coverage shotgun sequencing. *BMC Genomics*, **11**, 68.

Leu JH, Chen SH, Wang YB *et al.* (2011) A review of the major penaeid shrimp EST studies and the construction of a shrimp transcriptome database based on the ESTs from four penaeid shrimp. *Marine Biotechnology*, **13**, 608–621.

Liu YG, Zheng MG, Liu LX, Lin H, Wang DF (2006) Five new microsatellite loci for olive flounder (*Paralichthys olivaceus*) from an expressed sequence tag (EST) library and cross-species amplification. *Molecular Ecology Notes*, **6**, 371–373.

Liu YG, Sun XQ, Gao H, Liu LX (2007) Microsatellite markers from an expressed sequence tag library of half-smooth tongue sole (*Cynoglossus semilaevis*) and their application in other related fish species. *Molecular Ecology Notes*, **7**, 1242–1244.

Martínez P, Hermida M, Pardo BG *et al.* (2008) Centromere-linkage in the turbot (*Scophthalmus maximus*) through half-tetrad analysis in diploid meiogynogenetics. *Aquaculture*, **280**, 81–88.

Martínez P, Bouza C, Hermida M *et al.* (2009) Identification of the major sex-determining region of turbot (*Scophthalmus maximus*). *Genetics*, **183**, 1443–1452.

Pardo BG, Machordom A, Foresti F *et al.* (2005) Phylogenetic analysis of flatfish (order Pleuronectiformes) based on mitochondrial 16S rDNA sequences. *Scientia Marina*, **69**, 531–543.

Pardo BG, Hermida M, Fernández C *et al.* (2006) A set of highly polymorphic microsatellites useful for kinship and population analysis in turbot (*Scophthalmus maximus* L.). *Aquaculture Research*, **37**, 1578–1582.

Pardo BG, Fernández C, Hermida M *et al.* (2007) Development and characterization of 248 novel microsatellite markers in turbot (*Scophthalmus maximus*). *Genome*, **50**, 329–332.

Pardo BG, Fernández C, Millán A *et al.* (2008) Expressed sequence tags (ESTs) from immune tissues of turbot (*Scophthalmus maximus*) challenged with pathogens. *BMC Veterinary Research*, **4**, 37.

Park KC, Osborne JA, Montes A *et al.* (2009) Immunological responses of turbot (*Psetta maxima*) to nodavirus infection or polyriboinosinic polyribocytidylic acid (pIC) stimulation, using expressed sequence tags (ESTs) analysis and cDNA microarrays. *Fish and Shellfish Immunology*, **26**, 91–108.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2) Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rico C, Rico I, Hewitt G (1996) 470 million years of conservation of microsatellite loci among fish species. *Proceeding of Royal Society of London, Series B, Biological Sciences*, **263**, 549–557.

Rodríguez-Ramilo ST, Toro MA, Bouza C *et al.* (2011) QTL detection for *Aeromonas salmonicida* resistance related traits in turbot (*Scophthalmus maximus*). *BMC Genomics*, **12**, 541.

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

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener SE), pp. 365–386. Humana Press, Totowa, NJ.

Ruan X, Wang W, Kong J, Yu F, Huang X (2010) Genetic linkage mapping of turbot (*Scophthalmus maximus* L.) using microsatellite markers and its application in QTL analysis. *Aquaculture*, **308**, 89–100.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sánchez-Molano E, Cerna A, Toro MA *et al.* (2011) QTL detection for *Aeromonas salmonicida* resistance related traits in turbot (*Scophthalmus maximus*). *BMC Genomics*, **12**, 473.

Sha Z, Wang S, Zhuang Z *et al.* (2010) Generation and analysis of 10 000 ESTs from the half-smooth tongue sole *Cynoglossus semilaevis* and identification of microsatellite and SNP markers. *Journal of Fish Biology*, **76**, 1190–1204.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876–4882.

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

Vera M, Alvarez-Dios JA, Millán A *et al.* (2011) Validation of single nucleotide polymorphism (SNP) markers from an immune Expressed Sequence Tag (EST) turbot, *Scophthalmus maximus*, database. *Aquaculture*, **313**, 31–41.

Vilas R, Bouza C, Vera M, Millán A, Martínez P (2010) Variation in anonymous and EST-microsatellites suggests adaptive population divergence in turbot. *Marine Ecology Progress Series*, **420**, 231–239.

Williams TD, Diab AM, George SG *et al.* (2006) Development of the GENIPOL European flounder (*Platichthys flesus*) microarray and determination of temporal transcriptional responses to cadmium at low dose. *Environmental Science and Technology*, **40**, 6479–6488.

Zhang L, Guo X (2010) Development and validation of single nucleotide polymorphism markers in the eastern oyster *Crassostrea virginica* Gmelin by mining ESTs and resequencing. *Aquaculture*, **302**, 124–129.

Two research teams from Granada (UGR) and Santiago de Compostela (USC) Universities collaborated to carry out this work. F.R. and M.J.M.-L. (UGR) and B.G.P. and M.V. (USC) carried out the technical work to validate different sets of microsatellites in the laboratory. R.N.-P. (UGR) and C.B. (USC) mined EST sequences from the turbot EST database to look for microsatellites and performed the comparison with EST flatfish databases. J.A.A.-D. was the bioinformatics responsible of developing the tools for microsatellite screening in the EST databases. R.H. (UGR) and P.M. (USC) coordinated the work and wrote the manuscript.

Data Accessibility

DNA sequences: GenBank accessions (see Table S1 or Dryad repository).

Data deposited in the Dryad repository: doi:10.5061/dryad.q2c86hb2.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Comparative analysis of turbot microsatellites with Senegalese sole (*Solea senegalensis*), bastard halibut (*Paralichthys olivaceus*) and Atlantic halibut (*Hippoglossus hippoglossus*).

Table S2 Segregation analyses of 83 EST-derived microsatellites in five turbot families.

Table S3 Annotation of EST-derived microsatellites using BLASTx and BLASTn against several model fish genomes and the nucleotide collection of GenBank.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.