Vol. 420: 231–239, 2010
doi: 10.3354/meps08874

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published December 16

# Variation in anonymous and EST-microsatellites suggests adaptive population divergence in turbot

**Román Vilas*, Carmen Bouza, Manuel Vera, Adrián Millán, Paulino Martínez**

**Departamento de Xenética, Universidade de Santiago de Compostela, Facultade de Veterinaria, Lugo 27002, Spain**

ABSTRACT: We studied the variation at 30 anonymous and 30 expressed sequence tag (EST)-associated microsatellites in 4 natural populations of turbot *Scophthalmus maximus* living in habitats with different salinity and temperature conditions. We identified putative divergent selection effects on 3 genes: the fibroblast growth factor receptor, the β microglobuline, and the trap alpha gene for translocon associate protein. The markers closely linked to these genes showed significant deviations from the neutral expectations using 2 different statistical methods in several pairwise population comparisons involving samples from salty and brackish environments. Our results confirmed the weak genetic structure among populations from the northeast Atlantic and the low but significant genetic differentiation of turbot from the Baltic Sea. These results suggest that populations from the Baltic–Atlantic transition area could be accumulating adaptive polymorphisms in the face of high gene flow.

KEY WORDS: Population genetic structure · Divergent selection · Local adaptation · Microsatellites · Candidate genes · *Scophthalmus maximus*



## INTRODUCTION

The turbot *Scophthalmus maximus* is a species with a high dispersal potential due to pelagic stages of the life cycle, high fecundity, and large population size. It is widely distributed in the Northeast Atlantic, including both sides of 2 biogeographical barriers, the Baltic–Atlantic transition zone (Johannesson & André 2006) and the Orán–Almería front (Blanquer et al. 1992, Borsa et al. 1997). These areas show strong differences in salinity and temperature, respectively. Although several studies have confirmed that populations of flatfish are characterized by spatial genetic homogeneity (Blanquer et al. 1992, Bouza et al. 1997, 2002, Hoarau et al. 2002, Nielsen et al. 2004), significant genetic discontinuities in populations of flatfish including turbot have been observed in the Baltic–Atlantic area (Nielsen et al. 2004, Hemmer-Hansen et al. 2007a,b, Florin & Höglund 2008).

Turbot generally inhabit shallow areas on both sandy and rocky ground. This coastal species is likely more exposed to differences in temperature and salinity than other species living in more homogeneous habitats at greater depths. The survival of turbot in environments with very different salinity and temperature can be explained by at least 2 different hypotheses. It could be the result of phenotypic plasticity rather than natural selection if local adaptation is critically hampered by gene flow (Florin & Höglund 2007). On the other hand, large panmictic populations showing weak structure at neutral markers suggest a low genetic drift scenario where even relatively small selection effects could shape the spatial patterns of genetic variability (Nielsen et al. 2009). Under this scenario, migration could be restricted by local adaptation at particular genes because immigrant alleles of lower fitness would be efficiently purged from the population. Therefore, adaptive differentiation may be accumulated in the face of high gene flow retarding the divergence of unlinked loci (Hemmer-Hansen et al. 2007a), although divergent selection, gene flow, and phenotypic plasticity may interact in a number of ways (Crispo 2008).

To test the local adaptation hypothesis, we assume that genes evolve independently and that mechanisms like gene flow and genetic drift influence all loci in a similar

*Email: roman.vilas@usc.es

way. These assumptions are important because the detection of molecular signatures of positive selection is based on the distinction between the locus-specific effects, such as those caused by selection, and genome-wide effects resulting from demographic processes (Luikart et al. 2003). It is possible to make such distinction by analyzing neutral variation, because neutral variants that are closely linked to the beneficial alleles are also affected by natural selection (Schlötterer 2003). These variants should be identified as outliers when they are compared against patterns of variation derived from a random sample of neutral markers across the genome, which provides a control for the effects of historical population subdivision or genetic drift. However, population structure and natural selection can render similar patterns of neutral variation (Charlesworth et al. 2003). For instance, under complex demographic scenarios, genetic drift can increase the variance of the divergence among loci, hindering the identification of potential effects of divergent selection. In contrast, species with a relatively simple demographic history would facilitate the identification of signatures of divergent selection by analyzing the distribution of $F_{ST}$ estimates (Beaumont 2005), a statistic that reflects inbreeding due to the subdivision of the population.

Turbot from the Northeast Atlantic provide a good model to test local adaptation in marine fishes because they show relatively weak population genetic structure at neutral markers along temperature and salinity gradients. The high genetic variability and the low population structure suggest large population sizes and random mating, so it appears unlikely that a postulated signature of selection inferred from population genetic data of this fish is the product of complicated demographics. The turbot is also a good model because important genetic resources including a large number of anonymous and expressed sequence tag (EST)-linked markers placed in a genetic map are now available for this species (Bouza et al. 2007, 2008, Martínez et al. 2008, 2009). Given the relatively small size of its genome (Cuñado et al. 2001), a reasonable number of markers can be selected on the basis of their linkage relationships to detect molecular signatures of natural selection.

In this study, we analyzed the variation at 60 microsatellites, covering the 26 linkage groups that make up the current map of the species (Bouza et al. 2007), in 4 populations living in environments that substantially differ in salinity and temperature. Particularly, the studied area shows a salinity gradient spanning from 30 to 35‰ in the North Sea and the northwest of the Iberian Peninsula to 8‰ in the Baltic proper, with levels varying from 1 to 2‰ in the northernmost Bothnian Bay to 20‰ in the Kattegat. Mean differences in temperature range from 16°C in the Iberian Peninsula to 8°C in the south part of the Baltic Sea (HELCOM 2003, Álvarez et al. 2005). To detect adaptive variation, we followed 2 different methodological approaches based on the prediction that markers closely linked to genes affected by divergent selection should reveal an unusually high degree of differentiation among populations from different environments. We report 3 loci showing significant deviations from neutral expectations which can be treated as candidate genes involved in differential adaptation.

## MATERIALS AND METHODS

**Sampling and DNA analysis.** Fin clip samples were obtained from 4 geographic populations from the North Atlantic: in the Baltic Sea at the Island of Bornholm, Denmark (BS); in the North Sea at the west coast of Denmark (NS); in the Cantabric Sea at Xove, Spain (CS); and on the Atlantic Galician coast at Vigo, Spain (AG). Whole genomic DNA was extracted by using standard phenol–chloroform procedures. In total, 190 individuals (48 site$^{-1}$ except at CS, n = 46) were analyzed and 60 microsatellite loci used: 30 obtained by screening enriched genomic libraries (Pardo et al. 2006, 2007) and 30 obtained by screening ESTs (Bouza et al. 2008). Loci were selected on the basis of their linkage relationships, so they belong to the 26 linkage groups previously identified in turbot (Bouza et al. 2007, 2008). Genotypic equilibrium between markers can be assumed because those assigned to the same linkage group were generally separated by large distances on the genetic map. The only major exceptions were *Sma146* and *SmaE28*, which were spaced by less than 1 cM (Bouza et al. 2007, 2008). PCR products were resolved by using an ABI PRISM® 3730 automatic sequencer (Applied Biosystems). Allele scoring was performed with GENEMAPPER 3.7 software (Applied Biosystems).

**Genetic variation within and among populations.** Allele frequencies and estimates of genetic variation within populations (average number of alleles per locus and heterozygosity, Nei 1978) and among populations (estimator of Wright's $F_{ST}$, Weir & Cockerham 1984) were calculated using FSTAT 2.9.3 (Goudet 2001). We calculated both global and pairwise $F_{ST}$ values. FSTAT was also used to perform a significance test of 10 000 permutations of the obtained estimates and to calculate the allelic richness per population based on the minimum sample size of 46 individuals. We explored the possibility of population structure regardless of the geographic origin by using the Bayesian clustering analysis implemented in BAPS 3.1 (Corander et al. 2003). The analysis determines the existence of clusters of populations by minimizing Hardy-

Weinberg and linkage disequilibrium within clusters. The difference between BAPS clusters was estimated with the Kullback-Leibler (K-L) divergence, which can be used as a measure of distance in a genetic context (Anderson & Thompson 2002, Corander et al. 2003). Conformity to Hardy-Weinberg proportions was tested using exact tests as implemented in GENEPOP 3.4 (Raymond & Rousset 1995). Data analyses were done separately for anonymous and EST-linked markers. Critical significance levels were adjusted for multiple tests using the Bonferroni correction.

**Outlier tests for selection.** To search for possible signatures of selection, we applied 2 different statistical methods. The Beaumont & Nichols (1996) method derives the expected neutral distribution of $F_{ST}$ values conditional on heterozygosity using coalescent simulations under a symmetrical island model and assuming migration-drift equilibrium. This method provides evidence for divergent selection by looking for outliers with higher $F_{ST}$ values than expected under neutrality, accounting for variation in allele frequency. The tests were performed using the software Fdist (Beaumont & Nichols 1996) as implemented through LOSITAN (Antao et al. 2008). Simulations were run for 10 000 replications using the options for neutral and forced mean $F_{ST}$. Analyses were performed at 95 and 99% confidence levels. A stepwise mutation model was assumed. The method performs particularly well under adaptive divergence of large populations under simple demographic scenarios (Beaumont & Nichols 1996). This is expected for species such as turbot that show weak population genetic structure.

The second method used to detect loci under selection is based on the approach of Beaumont & Balding (2004). The method estimates $F_{ST}$ values that are specific for each population and locus using a hierarchical Bayesian approach that models locus-specific and population effects using a logistic-regression model. The posterior probability of including a locus-specific effect is estimated using a reversible jump Markov chain Monte Carlo approach in BayeScan (Foll & Gaggiotti 2008). This probability cannot be compared directly to the p-value in the Fdist software. For each method, we analyzed both anonymous and EST-linked loci collectively in a first screen for outliers with the aim of revealing loci with a major overall effect. We also conducted pairwise comparisons because results are more reliable compared to approaches across several populations (Robertson 1975, Tsakas & Krimbas 1976, Vitalis et al. 2001).

**Gene function.** EST-linked markers that were outliers for any of the 2 tests of selection applied were investigated further to determine their predicted function. Blast searches against GenBank using tBlastn (National Center for Biotechnology Information, NCBI) were used to find significant hits (E-value below $10^{-5}$; Altschul et al. 1990). These genes were categorized according to known biological processes, molecular functions, and cellular components using gene ontology (GO) functional terms (Ashburner et al. 2000).

## RESULTS

### Genetic variation within and among populations

Only 2 loci showed consistent deviations from Hardy-Weinberg equilibrium (*Sma142* and *SmaE43*; $p < 0.0008$; Table 1). Deviations were always toward a deficit of heterozygotes, suggesting the presence of null alleles in the sample. Gene diversity and allelic richness for EST-linked and anonymous microsatellites were similar among all populations studied. However, anonymous microsatellites derived from short repeat-enriched libraries showed higher variation and lower population differentiation than EST-derived microsatellites (mean within-population gene diversity = 0.741 and 0.566, respectively; mean number of alleles across populations = 14.93 and 8.33, respectively; Mann-Whitney *U*-test, $p < 0.05$; global $F_{ST}$ = 0.017 and 0.024, respectively; Fig. 1). The difference between EST-based and anonymous loci in global $F_{ST}$ was not significant. Global $F_{ST}$ for all populations was significant when the analysis included the 60 loci ($F_{ST}$ = 0.020, $p < 0.001$). The 4 geographic samples were clustered by BAPS into a single group after analyzing the variation at anonymous markers (data not shown). However, the best partition at EST-linked markers separated the Baltic population from the other 3 (KL-divergence = 0.318). When the outliers *SmaE4*, *SmaE7*, and *SmaE12* (see below) were removed from the data set, the 4 samples were clustered by BAPS into a single group. This result suggests that the discrepancy between genetic structure as estimated with anonymous and EST-linked markers is explained by the strong influence of a few outlier loci among the EST-linked ones. All pairwise $F_{ST}$ values involving the BS population were statistically significant ($p < 0.0083$ in all cases; Table 2). Although low, anonymous markers also revealed significant genetic structure among North Sea and any of the 2 Iberian samples ($F_{ST}$ = 0.011 and 0.012; $p < 0.0083$). Only the 2 Iberian populations showed no significant differentiation between them (Table 2). Despite the small number of samples, which does not allow for multiple independent comparisons between populations living in different environments, the North Sea, Baltic Sea, and Iberian samples can be treated as 3 different populations on the basis of the genetic differentiation at presumably neutral markers.

Table 1. *Scophthalmus maximus.* Gene diversity ($H_e$), Hardy-Weinberg equilibrium departures ($F_{IS}$), and test of conformity to expected values by locus and population. BS, NS, CS, and AG are samples from the Baltic Sea, North Sea, Cantabric Sea, and Atlantic Galician coast, respectively. *p < 0.0008

| Locus | BS | | NS | | CS | | AG | |
|---|---|---|---|---|---|---|---|---|
| | $H_e$ | $F_{IS}$ | $H_e$ | $F_{IS}$ | $H_e$ | $F_{IS}$ | $H_e$ | $F_{IS}$ |
| *Sma284* | 0.823 | +0.038 | 0.863 | −0.062 | 0.850 | +0.011 | 0.855 | +0.079 |
| *Sma42* | 0.924 | +0.053 | 0.911 | +0.062 | 0.929 | −0.027 | 0.925 | −0.035 |
| *Sma168* | 0.545 | −0.032 | 0.711 | −0.084 | 0.684 | +0.142 | 0.635 | +0.082 |
| *Sma247* | 0.751 | −0.026 | 0.722 | +0.134 | 0.753 | +0.170 | 0.697 | −0.046 |
| *2/5TG14* | 0.845 | −0.060 | 0.840 | −0.140 | 0.880 | +0.044 | 0.830 | +0.051 |
| *Sma135* | 0.785 | +0.045 | 0.892 | +0.066 | 0.859 | −0.058 | 0.882 | −0.039 |
| *Sma22* | 0.847 | +0.041 | 0.904 | −0.014 | 0.899 | −0.006 | 0.908 | +0.082 |
| *Sma147* | 0.729 | +0.028 | 0.737 | −0.032 | 0.722 | −0.007 | 0.670 | +0.099 |
| *Sma14* | 0.854 | −0.025 | 0.888 | −0.055 | 0.886 | +0.003 | 0.878 | +0.003 |
| *Sma137* | 0.141 | −0.035 | 0.250 | +0.001 | 0.301 | +0.094 | 0.342 | −0.081 |
| *Sma142* | 0.906 | +0.563* | 0.921 | +0.593* | 0.825 | +0.449* | 0.850 | +0.474* |
| *Sma149* | 0.852 | −0.052 | 0.815 | −0.048 | 0.816 | −0.059 | 0.886 | +0.072 |
| *F8I11/8/17* | 0.694 | +0.130 | 0.752 | +0.335* | 0.793 | +0.260 | 0.850 | +0.143 |
| *Sma113* | 0.686 | −0.003 | 0.686 | +0.240 | 0.714 | +0.100 | 0.722 | −0.039 |
| *Sma117* | 0.792 | +0.027 | 0.820 | +0.111 | 0.817 | +0.103 | 0.815 | −0.045 |
| *Sma34* | 0.733 | +0.021 | 0.863 | −0.011 | 0.893 | +0.059 | 0.851 | −0.078 |
| *Sma18* | 0.897 | −0.015 | 0.904 | +0.009 | 0.904 | −0.031 | 0.902 | +0.030 |
| *Sma184* | 0.342 | −0.145 | 0.392 | −0.010 | 0.388 | +0.003 | 0.479 | +0.003 |
| *Sma185* | 0.516 | +0.116 | 0.539 | −0.083 | 0.551 | +0.113 | 0.563 | −0.148 |
| *Sma38* | 0.843 | −0.009 | 0.806 | −0.112 | 0.857 | −0.015 | 0.809 | +0.098 |
| *Sma100* | 0.300 | +0.029 | 0.498 | +0.146 | 0.446 | +0.103 | 0.516 | −0.051 |
| *Sma144* | 0.951 | −0.029 | 0.969 | +0.121 | 0.964 | +0.104 | 0.962 | +0.025 |
| *Sma175* | 0.539 | −0.106 | 0.588 | +0.256 | 0.416 | +0.004 | 0.472 | −0.103 |
| *Sma205* | 0.395 | +0.084 | 0.375 | +0.262 | 0.471 | +0.197 | 0.478 | +0.110 |
| *Sma19* | 0.882 | −0.040 | 0.933 | +0.084 | 0.906 | +0.044 | 0.900 | −0.041 |
| *Sma146* | 0.807 | +0.251* | 0.791 | +0.026 | 0.804 | +0.115 | 0.856 | +0.030 |
| *Sma278* | 0.757 | +0.119 | 0.733 | +0.119 | 0.737 | −0.055 | 0.782 | −0.039 |
| *Sma282* | 0.593 | +0.052 | 0.746 | −0.061 | 0.760 | +0.006 | 0.723 | +0.117 |
| *Sma77* | 0.731 | +0.088 | 0.757 | +0.016 | 0.763 | +0.031 | 0.796 | +0.084 |
| *Sma21* | 0.799 | −0.017 | 0.828 | +0.044 | 0.882 | +0.098 | 0.875 | +0.047 |
| *SmaE1* | 0.705 | +0.261* | 0.823 | +0.164 | 0.820 | +0.099 | 0.806 | +0.121 |
| *SmaE2* | 0.000 | – | 0.021 | +0.000 | 0.044 | −0.011 | 0.021 | +0.000 |
| *SmaE3* | 0.550 | +0.204 | 0.599 | +0.041 | 0.574 | −0.276 | 0.536 | −0.127 |
| *SmaE7* | 0.573 | −0.127 | 0.505 | +0.174 | 0.591 | +0.001 | 0.563 | −0.109 |
| *SmaE4* | 0.259 | +0.015 | 0.513 | −0.218 | 0.517 | +0.243 | 0.447 | +0.000 |
| *SmaE29* | 0.154 | −0.080 | 0.357 | −0.051 | 0.437 | −0.352 | 0.448 | −0.209 |
| *SmaE8* | 0.557 | +0.027 | 0.681 | +0.082 | 0.666 | +0.213 | 0.666 | +0.156 |
| *SmaE12* | 0.537 | −0.010 | 0.766 | +0.075 | 0.794 | −0.068 | 0.833 | +0.100 |
| *SmaE14* | 0.820 | −0.038 | 0.837 | +0.229 | 0.845 | +0.059 | 0.822 | +0.146* |
| *SmaE16* | 0.081 | −0.025 | 0.137 | −0.068 | 0.085 | −0.020 | 0.101 | −0.035 |
| *SmaE25* | 0.812 | −0.026 | 0.744 | +0.048 | 0.768 | +0.179 | 0.780 | −0.041 |
| *SmaE28* | 0.280 | −0.115 | 0.438 | +0.030 | 0.302 | +0.152 | 0.385 | −0.082 |
| *SmaE38* | 0.861 | −0.016 | 0.879 | −0.042 | 0.868 | +0.148 | 0.854 | −0.049 |
| *SmaE20* | 0.525 | +0.069 | 0.509 | +0.289 | 0.510 | −0.090 | 0.536 | +0.008 |
| *SmaE35* | 0.158 | +0.207 | 0.122 | −0.047 | 0.066 | −0.015 | 0.062 | −0.022 |
| *SmaE40* | 0.576 | +0.023 | 0.827 | −0.058 | 0.818 | +0.103 | 0.795 | +0.056 |
| *SmaE41* | 0.686 | +0.162 | 0.577 | +0.057 | 0.540 | +0.178 | 0.544 | −0.111 |
| *SmaE42* | 0.466 | +0.042 | 0.554 | +0.058 | 0.530 | −0.006 | 0.574 | −0.039 |
| *SmaE19* | 0.738 | +0.097 | 0.661 | +0.022 | 0.741 | +0.135 | 0.735 | −0.072 |
| *SmaE22* | 0.722 | +0.106 | 0.576 | −0.121 | 0.627 | +0.220 | 0.778 | +0.304* |
| *SmaE26* | 0.491 | −0.104 | 0.506 | +0.218 | 0.502 | −0.168 | 0.491 | −0.104 |
| *SmaE32* | 0.696 | −0.047 | 0.623 | +0.063 | 0.601 | −0.096 | 0.588 | +0.059 |
| *SmaE30* | 0.425 | −0.127 | 0.522 | −0.158 | 0.393 | +0.076 | 0.525 | −0.071 |
| *SmaE43* | 0.327 | +0.618* | 0.369 | +0.718* | 0.411 | +0.717* | 0.332 | +0.423 |
| *SmaE10* | 0.439 | +0.052 | 0.318 | +0.332 | 0.511 | +0.086 | 0.525 | +0.130 |
| *SmaE13* | 0.570 | +0.013 | 0.855 | −0.024 | 0.815 | +0.086 | 0.817 | −0.016 |
| *SmaE31* | 0.697 | +0.013 | 0.766 | −0.033 | 0.753 | +0.223 | 0.654 | +0.089 |
| *SmaE33* | 0.760 | −0.070 | 0.621 | +0.127 | 0.741 | +0.031 | 0.700 | −0.094 |
| *SmaE36* | 0.690 | −0.026 | 0.683 | +0.115 | 0.689 | +0.011 | 0.656 | −0.079 |
| *SmaE39* | 0.827 | −0.032 | 0.812 | +0.051 | 0.806 | +0.048 | 0.825 | +0.020 |

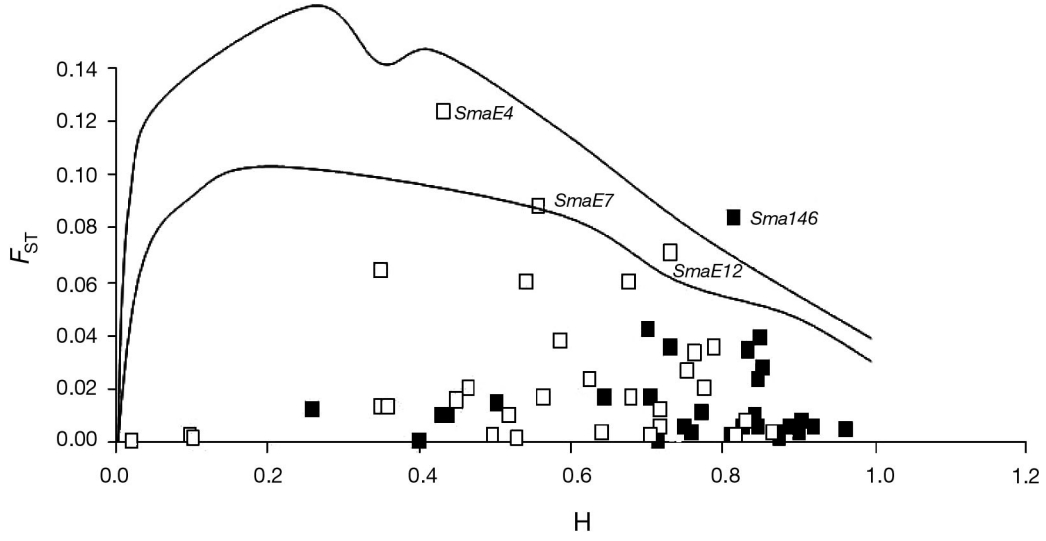


Fig. 1. *Scophthalmus maximus.* Plot of $F_{ST}$ versus heterozygosity values for anonymous (black squares) and expressed sequence tag (EST)-linked (white squares) markers. Top and bottom lines indicate the 99 and 95% confidence levels for the global analysis performed with Fdist. Outliers are indicated by name

Table 2. *Scophthalmus maximus.* Pairwise $F_{ST}$ values from expressed sequence tag (EST)-linked (above diagonal) and anonymous (below diagonal) markers. BS, NS, CS, and AG are samples from the Baltic Sea, North Sea, Cantabric Sea, and Atlantic Galician coast, respectively. $^{\star}p < 0.0083$

| | BS | NS | CS | AG |
|---|---|---|---|---|
| BS | – | 0.038$^{\star}$ | 0.044$^{\star}$ | 0.038$^{\star}$ |
| NS | 0.021$^{\star}$ | – | 0.002 | 0.016$^{\star}$ |
| CS | 0.028$^{\star}$ | 0.011$^{\star}$ | – | 0.006 |
| AG | 0.025$^{\star}$ | 0.012$^{\star}$ | 0.002 | – |

### Tests for divergent selection

The 2 different approaches for testing whether differentiation patterns at anonymous and EST-linked microsatellite loci could depart from neutral expectations were highly consistent in identifying a set of outlier loci that could represent signatures of divergent selection among populations (Table 3). Following the Fdist approach in a global analysis across populations, a total of 4 markers, 1 anonymous (*Sma146*) and 3 EST-linked microsatellites (*SmaE4*, *SmaE7*, and *SmaE12*), showed significantly higher $F_{ST}$ values than expected ($p < 0.05$; Fig. 1). These markers were also identified as highly likely to be outliers in a global analysis with BayeScan. Thus, the posterior probability that these loci are not selectively neutral was higher than 0.95 in all cases (Table 3). In contrast, the Fdist and the Bayesian methods were not concordant in the detection of low differentiation outliers, which could represent signatures of stabilizing selection among populations. The lack of concordance was observed in both the analysis across populations and the pairwise analysis. No outliers with low $F_{ST}$ values were identified as outliers simultaneously with the 2 tests of selection (Table 4).

In total, 4 EST-linked and 2 anonymous microsatellites were identified as outliers with the 2 statistical approaches when signatures of divergent selection were addressed by performing pairwise population analyses. Only 3 markers (*SmaE7*, *SmaE4*, and *SmaE12*) were statistically significant with Fdist and BayeScan in at least 2 of the 3 comparisons involving samples from the Baltic (Table 3). The anonymous marker *Sma146* was identified as an outlier with both methods in the comparisons between the North Sea (NS) and the 2 Iberian samples, Cantabric (CS) and Atlantic (AG). Despite the high statistical significance, we cannot exclude that *Sma146* represents a false positive because the analysis of the closely linked loci *SmaE28*, located less than 1 cM away did not confirm its outlier status. The loci *2/5TG14* and *SmaE22* also showed a significantly higher $F_{ST}$ value than the neutral expectation with Fdist and BayeScan for the NS/AG comparison and *SmaE22* for the CS/AG comparison (Table 3).

Outliers were particularly frequent in comparisons between samples involving the Baltic. The highest $F_{ST}$ values were more frequent in EST-linked markers than in anonymous markers. However, all markers with the lowest differentiation were anonymous. While the Fdist and the Bayesian methods revealed a similar number of

Table 3. *Scophthalmus maximus.* Anonymous (A) and expressed sequence tag (EST)-linked (E) markers that presented higher $F_{ST}$ values than expected; $p$ ($\alpha \neq 0$) is the posterior probability that each locus is affected by divergent selection. Results in pairwise population comparisons and across populations are shown. BS, NS, CS, and AG are samples from the Baltic Sea, North Sea, Cantabric Sea, and Atlantic Galician coast, respectively

| Comparison | Marker | Type | $F_{ST}$ | p-value Fdist | $p$ ($\alpha \neq 0$) BayeScan | Annotation |
|---|---|---|---|---|---|---|
| BS/NS | *SmaE7* | E | 0.168 | 0.033 | 0.994 | Fibroblast growth factor receptor |
| | *SmaE4* | E | 0.269 | 0.005 | 0.93 | β microglobuline |
| BS/CS | *Sma146* | A | 0.125 | 0.021 | 0.753 | – |
| | *SmaE4* | E | 0.257 | 0.012 | 0.845 | β microglobuline |
| | *SmaE12* | E | 0.164 | 0.023 | 0.989 | Trap alpha-translocon protein |
| BS/AG | *SmaE7* | E | 0.168 | 0.042 | 0.991 | Fibroblast growth factor receptor |
| | *SmaE12* | E | 0.148 | 0.028 | 0.997 | Trap alpha-translocon protein |
| | *SmaE29* | E | 0.166 | 0.049 | 0.572 | – |
| NS/CS | *Sma146* | A | 0.147 | 0 | 1.000 | – |
| NS/AG | *Sma146* | A | 0.114 | 0 | 0.999 | – |
| | *2/5TG14* | A | 0.072 | 0.036 | 0.991 | – |
| | *SmaE22* | E | 0.133 | 0.008 | 1.000 | – |
| CS/AG | *Sma149* | A | 0.031 | 0.036 | 0.457 | – |
| | *SmaE22* | E | 0.088 | 0.001 | 0.949 | – |
| Global | *Sma146* | A | 0.084 | 0 | 1.000 | – |
| | *SmaE7* | E | 0.088 | 0.031 | 0.995 | Fibroblast growth factor receptor |
| | *SmaE4* | E | 0.123 | 0.017 | 0.963 | β microglobuline |
| | *SmaE12* | E | 0.071 | 0.02 | 1.000 | Trap alpha-translocon protein |

Table 4. *Scophthalmus maximus.* Markers that presented lower $F_{ST}$ values than expected (they are all anonymous); $p$ ($\alpha \neq 0$) is the posterior probability that each locus is affected by stabilizing selection. Results in pairwise population comparisons and across populations are shown. BS, NS, CS, and AG are samples from the Baltic Sea, North Sea, Cantabric Sea, and Atlantic Galician coast, respectively. Outliers were not detected in the CS/AG comparison

| Comparison | Marker | $F_{ST}$ | p-value Fdist | $p$ ($\alpha \neq 0$) BayeScan |
|---|---|---|---|---|
| BS/NS | *Sma137* | 0.007 | 0.49 | 0.999 |
| | *Sma144* | 0.003 | 0.304 | 1.000 |
| BS/CS | *Sma42* | 0.006 | 0.367 | 0.999 |
| | *Sma19* | 0 | 0.086 | 0.999 |
| | *Sma144* | 0.01 | 0.519 | 1.000 |
| | *Sma38* | 0 | 0.034 | 0.878 |
| BS/AG | *Sma22* | 0.007 | 0.201 | 1.000 |
| | *Sma142* | 0.007 | 0.292 | 0.993 |
| | *Sma144* | 0.009 | 0.486 | 1.000 |
| NS/CS | *Sma144* | 0 | 0.258 | 0.999 |
| NS/AG | *Sma144* | 0 | 0.144 | 0.999 |
| Global | *Sma19* | 0.005 | 0.17 | 0.996 |
| | *Sma144* | 0.006 | 0.223 | 0.999 |
| | *Sma14* | 0.001 | 0.035 | 0.683 |

outlier patterns consistent with divergent selection, BayeScan identified a relatively high proportion of loci exhibiting significant departures from the neutral expectations consistent with stabilizing selection. Outliers were not associated with any particular linkage group.

### Gene function

The 3 outliers (*SmaE4*, *SmaE7*, and *SmaE12*) consistently identified by both methods are linked to ESTs with the following homologies based on the highest Blast hits: the β microglobuline (E-value = $4 \times 10^{-79}$), the fibroblast growth factor receptor (E-value = $2 \times 10^{-14}$) and the trap alpha-translocon protein (E-value = $3 \times 10^{-39}$), respectively. These 3 genes presumably linked to outlier loci can be treated as candidate genes involved in local adaptation to differences in salinity. Based on GO terms associated with these genes, they are related with membrane processes.

## DISCUSSION

The 4 populations of turbot living in habitats with very different salinity and temperature showed similar levels of genetic variability at anonymous and EST-linked microsatellite loci. However, anonymous markers showed higher variation and lower population differentiation than EST-linked markers. Results confirmed both the weak genetic structure previously observed in turbot populations (Blanquer et al. 1992, Bouza et al. 1997, 2002) and a higher significant differentiation of samples from the Baltic Sea, which might be related to adaptation to differences in salinity (Nielsen et al. 2004). According to the local adaptation hypothesis, the study of the variation of 60 microsatel-

lites in populations from contrasting habitats allowed the identification of putative divergent selection effects on 3 genes. These genes showed evidence for divergent selection on the basis of 2 statistical tests that differ in their approaches and assumptions. Thus, 1 method uses frequentist-like tests to infer outlier patterns and assumes an island model of migration, while the other follows a Bayesian approach and does not assume equal migration rates and population sizes. This difference is relevant because the violation of that assumption may be an important source of false positives (Excoffier et al. 2009). However, it is expected that the Fdist model is robust for species that have a simple genetic structure (Beaumont 2005). The North Sea versus Baltic Sea and the Iberian samples versus Baltic Sea cannot be properly considered as independent comparisons between populations living in habitats that differ markedly in salinity because the same low saline sample was used in all comparisons. However, the identification of the same loci as outliers by using 2 different tests of selection in several comparisons involving the Baltic Sea sample indicates that this sample is driving the outlier pattern of such loci (*SmaE4*, *SmaE7*, and *SmaE12*). This suggests adaptation to the Baltic Sea environment. However, the observed negative association between $F_{ST}$ and gene diversity ($H_e$) suggests that outliers at low $H_e$ might be the result of low mutation rates and demographic influences alone if the range of mutation rates is large or there is frequent homoplasy. Thus, the identification of *SmaE4* as an outlier must be taken with some caution due to its relatively low variation (average $H_e$ = 0.434). This marker also showed a lower posterior probability than *SmaE7* and *SmaE12* (p = 0.963; Table 3)

The 2 tests of selection were highly consistent in identifying signs of divergent selection among populations. The congruence in the identification of outliers could be explained by the weak population genetic structure of turbot. It is expected that the efficiency of selective tests used in the detection of divergent selection is increased in large panmictic populations with high gene flow because the selection signatures are distinguished from a lower background neutral signal caused by drift (Beaumont 2005). The identification of outliers in a scenario of simple demographics characterized by large effective population sizes and life cycle forms leading to high gene flow should produce a relatively low number of false positives because both violations of the model assumptions are less likely and population subdivision increases the variance of $F_{ST}$. However, the low mean $F_{ST}$ estimated among populations reduces the statistical power for detecting outliers with low $F_{ST}$. Accordingly, it is expected that the inconsistency of results between methods is particularly frequent in the detection of deviations from neutrality towards low $F_{ST}$ values. We found that none of the loci presenting lower $F_{ST}$ values than expected was simultaneously supported by the 2 methods applied, suggesting that many, if not all, such outliers could be the result of a type I error. In contrast, it is unlikely that outliers indicating divergent selection with the 2 neutrality tests represent false positives, particularly if they were identified in several comparisons. When we restrict the analysis to divergent selection, the outliers were more frequent in EST-linked markers than in anonymous markers, confirming the expectation that adaptive variation is more likely in coding sequences.

In the present study, only 3 markers were identified as outliers simultaneously with the 2 neutrality tests in at least 2 of the pairwise population comparisons involving samples from contrasting habitats. They are 3 EST-associated microsatellites showing an unusually high genetic divergence between the Baltic and any other population. This result supports the hypothesis that such markers are in strong linkage disequilibrium with genes involved in adaptation to differences in salinity. However, it is possible that other environmental factors with which salinity could be associated are responsible for the observed patterns. Candidate genes showed homology to genes that exhibit a wide range of putative functions: the fibroblast growth factor receptor, the β microglobuline, and the trap alpha gene for translocon associated protein. These genes are involved in membrane-related processes which are part of the immune system (Pardo et al. 2008, Peatman et al. 2008). It is known that different environmental factors including salinity and temperature exert direct effects on the immune system of fish (Boutet et al. 2006, Bowden 2008), so that adaptation to salinity differences could involve changes in immune-related genes. Further, osmotic regulation processes are used by the immune system as a defense strategy, and it is likely that they are influenced by the environment. An alternative explanation of the outlier behavior of such loci is that they could be reflecting differences associated with the hybridization of 2 divergent gene pools at the Baltic–Atlantic transition zone (Nielsen et al. 2004), not necessarily related to adaptation to different environments.

Only 1 anonymous marker (*Sma146*) was identified as an outlier simultaneously by both methods in several comparisons. However, differentiation of this marker is not associated with salinity and temperature. Further, *Sma146* is closely linked to *SmaE28*, which did not confirm its outlier status. However, *Sma146* may be a candidate locus if (1) recombination rapidly obscured the selective footprint on *SmaE28*, which is expected in large panmictic populations, (2) it is tightly linked to a second gene directly affected by selection,

and (3) this anonymous locus is under selection in a scenario of high recombination. Given the high gene density of turbot, this hypothesis seems plausible. Indeed, the genome of this species is among the smallest of the vertebrates (800 Mb, Cuñado et al. 2001) and shows a low relationship between physical and genetic distances (0.53 Mb $cM^{-1}$, Bouza et al. 2007).

The proportion of outlier loci identified in the present study (restricting the analysis to divergent selection) was higher among gene-associated loci (10%) than among anonymous loci (3%). These percentages are slightly lower than those obtained in previous studies using microsatellites as genetic markers. For instance, a recent genome scan in contrasting populations of the marine angiosperm *Zostera marina* identified 2 of 14 (14.2%) and 1 of 11 (9.0%) outliers among gene-associated and anonymous loci, respectively (Oetjen & Reusch 2007). Other genome scans using EST-linked microsatellites obtained similar values (Vasemägi et al. 2005, Kane & Rieseberg 2007, Mäkinen et al. 2008). The higher percentage of outliers in EST-linked markers is concordant with the differences obtained with the Bayesian clustering of groups of populations by using anonymous and EST-linked markers. Assuming Hardy-Weinberg and linkage equilibrium, the analysis of anonymous markers yielded a single group, while the analysis of EST-linked markers clearly discriminated between the Baltic Sea and the remaining populations. However, pairwise neutral $F_{ST}$ between the Baltic Sea sample and other samples were significant, which is concordant with previous data (Nielsen et al. 2004). This result suggests that the neutral structure detected between the North Sea and the Baltic Sea populations could reflect barriers to gene flow caused by adaptation to diverse environments.

The results of the present study suggest that divergent selection between turbot populations from both sides of the Baltic–Atlantic barrier seems strong enough to dominate over random drift and migration. They are consistent with recent studies that suggest adaptation of flatfish populations from the study area to differences in salinity (Hemmer-Hansen et al. 2007a, Larsen et al. 2007, 2008). If differential adaptation is actually the cause of the observed population structure, our work shows that microsatellites linked to coding sequences are useful to reveal selection signatures, even in a presumed high recombination scenario. However, we cannot rule out that factors other than adaptation to differences in salinity and temperature are causing a decrease in gene flow at specific chromosome locations. The 3 candidate EST-loci identified in this study may be used in a further sequence analysis to validate an explanation (in terms of adaptation) of their atypical behavior.

*Acknowledgements*. We thank E. Nielsen and M. Hansen for providing samples from the North Sea and the Baltic Sea, M. Foll and A. Pérez-Figueroa for help with BayeScan, and L. Insua for technical assistance. This work was funded by the grant PGIDIT06PXIB261178PR from Xunta de Galicia, a Consolider Ingenio Aquagenomics project (CSD2007-00002) from Spanish Ministry of Education and Science, and an Isidro Parga Pondal Fellowship to R.V.

## LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Álvarez I, de Castro M, Gómez-Gesteira M, Prego R (2005) Inter- and intra-annual analysis of the salinity and temperature evolution in the Galician Rías Baixas-ocean boundary (northwest Spain). J Geophys Res 110: C04008 doi: 10.1029/2004TC002S04

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a $F_{ST}$-outlier method. BMC Bioinformatics 9:323

Ashburner M, Ball CA, Blake JA, Botstein D and others (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29

Beaumont MA (2005) Adaptation and speciation: What can $F_{ST}$ tell us? Trends Ecol Evol 20:435–440

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol 13:969–980

Beaumont MA, Nichols RA (1996) Evaluating loci for the use in the genetic analysis of population structure. Proc Biol Sci 263:1619–1626

Blanquer A, Alayse JP, Berrada-Rkhami O, Berrebi P (1992) Allozyme variation in turbot (*Psetta maxima*) and brill (*Scophthalmus rhombus*) (Osteichtyes, Pleuronectiformes, Scophthalmidae) throughout their range in Europe. J Fish Biol 41:725–736

Borsa P, Blanquer A, Berrebi P (1997) Genetic structure of the flounders *Platichthys flesus* and *P. stellatus* at different geographic scales. Mar Biol 129:233–246

Boutet I, Long Ky CL, Bonhomme F (2006) A transcriptomic approach of salinity response in the euryhaline teleost, *Dicentrarchus labrax*. Gene 379:40–50

Bouza C, Sánchez L, Martínez P (1997) Gene diversity analysis in natural populations and cultured stocks of turbot (*Scophthalmus maximus* L.). Anim Genet 28:28–36

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. Can J Fish Aquat Sci 59:1460–1473

Bouza C, Hermida M, Pardo BG, Fernández C and others (2007) A microsatellite genetic map of turbot (*Scophthalmus maximus*). Genetics 177:2457–2467

Bouza C, Hermida M, Millán A, Vilas R and others (2008) Characterization of EST-derived microsatellites for gene mapping and evolutionary genomics in turbot. Anim Genet 39:666–670

Bowden TJ (2008) Modulation of the immune system of fish by their environment. Fish Shellfish Immunol 25:373–383

Charlesworth B, Charlesworth D, Barton NH (2003) The effects of genetic and geographic structure on neutral variation. Annu Rev Ecol Evol Syst 34:99–125

Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. Genetics 163:367–374

Crispo E (2008) Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. J Evol Biol 21:1460–1469

Cuñado N, Terrones J, Sánchez L, Martínez P, Santos JL (2001) Synaptonemal complex analysis in spermatocytes and oocytes of turbot, *Scophthalmus maximus* (Pisces, Scophthalmidae). Genome 44:1143–1147

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103:285–298

Florin AB, Höglund J (2007) Absence of population structure of turbot (*Psetta maxima*) in the Baltic Sea. Mol Ecol 16: 115–126

Florin AB, Höglund J (2008) Population structure of flounder (*Platichthys flesus*) in the Baltic Sea: differences among demersal and pelagic spawners. Heredity 101:27–38

Foll M, Gaggiotti O (2008) Estimating selection with different markers and varying demographic scenarios: a Bayesian perspective. Genetics 180:977–993

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available at www2.unil.ch/pogen/softwares/fstat.htm

HELCOM (2003) The Baltic marine environment 1999–2002. Baltic Sea Environ Proc 87:1–46

Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007a) Adaptive divergence in a high gene flow environment: *Hsc70* variation in the European flounder (*Platichthys flesus* L.). Heredity 99:592–600

Hemmer-Hansen J, Nielsen EE, Grønjkær P, Loeschcke V (2007b) Evolutionary factors shaping the genetic population structure of marine fishes; lessons from the European flounder (*Platichtys flesus* L.). Mol Ecol 16:3104–3118

Hoarau G, Rijnsdorp AD, Van der Veer HV, Stam WT, Olsen JL (2002) Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: microsatellites revealed large scale spatial and temporal homogeneity. Mol Ecol 11:1165–1176

Johannesson K, André C (2006) Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. Mol Ecol 15:2013–2029

Kane NC, Rieseberg LH (2007) Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annuus.* Genetics 175: 1823–1834

Larsen PF, Nielsen EE, Williams TD, Hemmer-Hansen J and others (2007) Adaptive differences in gene expression in European flounder (*Platichthys flesus*). Mol Ecol 16: 4674–4683

Larsen PF, Nielsen EE, Williams TD, Loeschke V (2008) Intraspecific variation in expression of candidate genes for osmoregulation, heme biosynthesis and stress resistance suggests local adaptation in European flounder (*Platichthys flesus*). Heredity 101:247–259

Luikart G, England PR, Tallmon D, Jordon S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4: 981–994

Mäkinen HS, Cano JM, Merilä J (2008) Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. Mol Ecol 17:3565–3582

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

Martínez P, Bouza C, Hermida F, Fernández J and others (2009) Identification of the major sex-determining region of turbot (*Scophthalmus maximus*). Genetics 183: 1443–1452

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Nielsen EE, Nielsen PH, Meldrup D, Hansen MM (2004) Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. Mol Ecol 13:585–595

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18: 3128–3150

Oetjen K, Reusch TBH (2007) Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina.* Mol Ecol 16:5156–5167

Pardo BG, Hermida M, Fernández C, Bouza C and others (2006) A set of highly polymorphic microsatellites useful for kinship and population analysis in turbot (*Scophthalmus maximus* L.). Aquac Res 37:1578–1582

Pardo BG, Fernández C, Hermida M, Vázquez-López A and others (2007) Development and characterization of 248 novel microsatellites markers in turbot (*Scophthalmus maximus*). Genome 50:329–332

Pardo BG, Fernández C, Millán A, Bouza C and others (2008) Expressed sequence tags (ESTs) from immune tissues of turbot (*Scophthalmus maximus*) challenged with pathogens. BMC Vet Res 4:37

Peatman E, Terhune J, Baoprasertkul P, Xu P and others (2008) Microarray analysis of gene expression in the blue catfish liver reveals early activation of the MHC class I pathway after infection with *Edwardsiella ictaluri.* Mol Immunol 45:553–566

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

Robertson A (1975) Gene frequency distributions as a test of selective neutrality. Genetics 81:775–785

Schlötterer C (2003) Hitchhiking mapping—functional genomics from the population genetics perspective. Trends Genet 19:32–38

Tsakas S, Krimbas CB (1976) Testing the heterogeneity of *F* values: a suggestion and a correction. Genetics 84: 399–401

Vasemägi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphism for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). Mol Biol Evol 22:1067–1076

Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. Genetics 158:1811–1823

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370

*Editorial responsibility: Hans Heinrich Janssen, Oldendorf/Luhe, Germany*

*Submitted: August 27, 2010; Accepted: October 13, 2010*
*Proofs received from author(s): December 10, 2010*