Candidate gene variation in gilthead sea bream reveals complex spatiotemporal selection patterns between marine and lagoon habitats


To cite this version:


HAL Id: hal-03046445
https://hal.umontpellier.fr/hal-03046445
Submitted on 8 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Candidate gene variation in gilthead sea bream reveals complex spatiotemporal selection patterns between marine and lagoon habitats

B. Guinand1,2,*, C. Chauvel2,3, M. Lechene2,3, J. Tournois4, C. S. Tsigenopoulos5, A. M. Darnaude4, D. J. McKenzie4, P. A. Gagnaire2,3

1Département Biologie-Ecologie, Université de Montpellier, Place E. Bataillon, 34095 Montpellier, France
2Unité Mixte de Recherche Institut des Sciences de l’Evolution de Montpellier (UMR ISEM), Université de Montpellier, Place E. Bataillon, 34095 Montpellier, France
3Station Biologique Marine de Sète, 2 rue des Chantiers, 34200 Sète, France
4Unité Mixte de Recherche Marine Biodiversity, Exploitation and Conservation (UMR MARBEC), Université de Montpellier, Place E. Bataillon, 34095 Montpellier, France
5Institute of Marine Biology and Genetics, Hellenic Center for Marine Research, PO Box 2214, Gournes Pediados, 71500 Heraklion, Crete, Greece

ABSTRACT: In marine fishes, the extent to which spatial patterns induced by selection remain stable across generations remains largely unknown. In the gilthead sea bream Sparus aurata, polymorphisms in the growth hormone (GH) and prolactin (Prl) genes can display high levels of differentiation between marine and lagoon habitats. These genotype–environment associations have been attributed to differential selection following larval settlement, but it remains unclear whether selective mortality during later juvenile stages further shapes genetic differences among habitats. We addressed this question by analysing differentiation patterns at GH and Prl markers together with a set of 21 putatively neutral microsatellite loci. We compared genetic variation of spring juveniles that had just settled in 3 ecologically different lagoons against older juveniles sampled from the same sites in autumn, at the onset of winter outmigration. In spring, genetic differentiation among lagoons was greater than expected from neutrality for both candidate gene markers. Surprisingly, this signal disappeared completely in the older juveniles, with no significant differentiation for either locus a few months later in autumn. We searched for signals of haplotype structure within GH and Prl genes using next-generation amplicon deep sequencing. Both genes contained 2 groups of haplotypes, but high similarities among groups indicated that signatures of selection, if any, had largely been erased by recombination. Our results are consistent with the view that differential selection operates during early juvenile life in sea bream and highlight the importance of temporal replication in studies of post-settlement selection in marine fish.

KEY WORDS: Candidate gene · Growth hormone · Prolactin · Genetic differentiation · Amplicon sequencing · Local selection

INTRODUCTION

The demography of coastal organisms has a profound influence on the genetic structure of their populations (Hellberg 2006, 2009, Gagnaire et al. 2015). During early life stages, genetic variation can be shaped by a combination of stochastic, dispersive and selective events which affect pelagic larval dispersal but also settlement and colonisation of nursery habitats by juveniles (e.g. Larson & Julian 1999, Broquet et al. 2013). Disentangling these various sources of variation throughout the larval and juvenile stages remains extremely challenging.

© The authors, CNRS and TOTAL Foundation 2016. Open Access under Creative Commons by Attribution Licence. Use, distribution and reproduction are unrestricted. Authors and original publication must be credited.

Publisher: Inter-Research · www.int-res.com
In species with high fecundity, discrete broadcast spawning events followed by very large mortality rates of pelagic larvae (>99%, type III survivorship curve) can produce successive waves of settlers that differ in their genetic makeup. Because of the marked variance in reproductive success, each group of recruits may comprise a different subset of individuals, representing only a fraction of the parental genetic pool (Hedgecock 1986, 1994). Hence, even if selection occurs during the planktonic stage (Johnson & Black 1984), genetic drift is thought to be the main force shaping genetic structuring among groups of settlers, sometimes promoting increased relatedness among individuals (e.g. Larson & Julian 1999, Planes & Lenfant 2002, Selkoe et al. 2006, Lachemi et al. 2013, Aglieri et al. 2014, review in Hauser & Carvalho 2008). This phenomenon, called sweepstake recruitment, has been widely documented (Hellberg 2006, 2009, Selkoe et al. 2008, Hedgecock & Pudovkin 2011). As the process is random and depends on factors such as the strength of the cohorts or variation in local dispersal and connectivity, it generally produces unpatterned spatial (or temporal) genetic differentiation among habitat patches, resulting in chaotic genetic patchiness, detectable over the whole genome (Hedgecock & Pudovkin 2011).

Conversely, if the variance in reproductive success is small enough or if pools of related larvae are sufficiently mixed by currents before settlement, panmixia might be observed among groups of recruits from distinct locations (Domingues et al. 2011, Guinand et al. 2011). In such cases, patterns of locus-specific genetic differentiation may indicate selection operating on juveniles that settled in different nursery habitats (Koehn et al. 1980, Lemaire et al. 2000, Planes & Roman 2004, Guinand et al. 2011). Moreover, patterns of genetic differentiation generated by selection are generally associated with spatial habitat variation (mosaics of different habitats or environmental gradients), and the temporal stability of such patterns provides further support for local selection (Schmidt & Rand 2001, Vézil et al. 2004, Gagnaire et al. 2012).

Studies focusing on post-settlement local adaptation in marine organisms usually lack the level of replication of sweepstake recruitment studies (e.g. David et al. 1997, Moberg & Burton 2000). Population studies in marine organisms are often only performed once, with observed genetic patterns then assumed to illustrate species population dynamics (Hedgecock et al. 2007). This may be misleading because within- or among-generational differences in population structure and patterns of connectivity among subpopulations are poorly captured by one single snapshot observation (Carson et al. 2010, Bert et al. 2014). The impact on observed genetic structure of demographic and ecological factors such as cohort size, larval mixing, predictability of habitat patches, local currents driving recruitment or the local strength of selection most likely varies from one generation to the next (Lotterhos & Markel 2012, Therkildsen et al. 2013). Ideally, a model of fluctuating selection that varies in space and/or time would offer a better representation of reality (Hellberg 2006, Hedgecock & Pudovkin 2011, Moody et al. 2015). This makes the eco-evolutionary dynamics of such metapopulations potentially idiosyncratic, with little scope for the systematic directional changes that are typically expected under habitat-based selection (Hanski 2011).

The gilthead sea bream *Sparus aurata* is a euryhaline coastal marine fish in which spatially varying selection between marine and lagoon habitats has been reported in the northwestern Mediterranean (Chaoui et al. 2012). After hatching at sea from December to February, a large proportion of the juveniles colonise coastal lagoons in March and April for foraging (Mercier et al. 2011, Isnard et al. 2015) before returning to the open sea in October and November to overwinter, when lagoon water temperatures drop (Audouin 1962, Lasserre 1974, Mercier et al. 2012). In the Gulf of Lions, 2 lagoons, located about 30 km apart and offering very different nursery habitats, have been studied: (1) the Mauguio (MA) lagoon (also known as Etang de l’Or), which is shallow and highly productive, with warm summer temperatures and wide salinity variations; and (2) the Thau (TH) lagoon, which is larger and deeper and ecologically more similar to the coastal marine environment due to major connections with the Mediterranean (e.g. Mercier et al. 2012, Tournois et al. 2013 and references therein). Using candidate gene markers located in the proximal promoter of the growth hormone (GH) and prolactin (Prl) genes, Chaoui et al. (2012) found non-neutral allele frequency shifts between young juveniles caught in marine and lagoon habitats and interpreted these patterns as a footprint of post-settlement selection. Because cis-regulatory polymorphisms have been described for both GH (Astola et al. 2003) and Prl (Almuly et al. 2008) genes in sea bream, Chaoui et al. (2012) hypothesised that functional variants affecting gene expression might be selected differently among habitats due to their phenotypic effects on growth and/or osmoregulation (e.g. Streelman & Kocher 2002, Blel et al. 2010, Shimada et al. 2011).

In this study, we first specifically assessed the temporal stability of habitat-based genetic differentia-
tion at the GH and Prl loci, extending our spatial coverage to another large coastal lagoon, Salses-Leucate (SL), which is a marinelike habitat ecologically similar to the TH lagoon (Bec et al. 2011, Mercier et al. 2012, Tournois et al. 2013). Twenty-one putatively neutral microsatellite loci were analysed together with GH and Prl markers to partition the relative influence of drift and selection in patterns of genetic variation. We also used next-generation amplicon sequencing to reconstruct haplotype of the GH and Prl genes in 24 individuals from the TH and MA lagoons to search for additional haplotype-based signals of selection.

**MATERIALS AND METHODS**

**Sampling**

The northwestern Mediterranean coast of the Gulf of Lions (Fig. 1) is characterised by a series of contiguous lagoons that offer seasonal nurseries for the juveniles of numerous highly prized fish species, including the gilthead sea bream (Quignard et al. 1984). The SL, TH and MA lagoons are located along a 150 km stretch of coastline interspersed by various other lagoons. Sea bream juveniles were captured by local fishermen using fyke nets in spring (from late April to May 2011) as they entered the lagoons and in autumn (from September to November 2011) as they started to migrate out to sea, as described in Isnard et al. (2015). Sampling locations were located within each lagoon (i.e. not at the inlets/outlets, where the strong marine influence makes this transition zone more similar to the marine environment). Autumn samples considered in this study were also analysed by Isnard et al. (2015) for their differences in growth rates and condition. Using back calculation from otolith readings, Isnard et al. (2015) showed that autumn individuals have recruited at different periods in the 3 lagoons (Fig. S1 in Supplement 1 at www.int-res.com/articles/
Estimation of the recruitment date in each lagoon closely matched the dates of sampling for our entering individuals in SL and TH. In MA, however, individuals were sampled after the estimated recruitment date (Fig. S1). Fork lengths (mean ± SD) of entering individuals were 31.3 ± 2.5 mm in SL, 34.7 ± 3.7 mm in TH and 41.3 ± 4.4 mm in MA, with population means being significantly different from each other (Tukey’s post hoc test; all p < 0.05). Larger sizes observed at MA should thus reflect longer time for growth between estimated recruitment and fish sampling dates. Hence, it is likely that entering and migrating sea bream individuals belong to the same demographic unit within each lagoon (i.e. individuals sampled in autumn are likely to have recruited in the same lagoon in spring and do not belong to distinct arrival waves except in MA). Total lengths (mean ± SD) of autumn (migrating) individuals were found to be significantly larger in MA (184 ± 11 mm) than in other lagoons (TH: 172 ± 11 mm; SL: 167 ± 7 mm) that were not different from each other (Isnard et al. 2012, Isnard et al. 2015). On the day of capture, fish were transported on ice to the laboratory, and tissues were sampled and kept in ethanol for DNA analysis (n = 276), as summarised in Table 1.

### Microsatellite genotyping

A small piece of tissue was incubated overnight in a lysis buffer with 5 µl Proteinase K (Qiagen). DNA was isolated using the protocol described in Aljanabi & Martinez (1997), and the concentration of each DNA sample was evaluated using NanoDrop 8000 (ThermoScientific) and standardised to individual DNA sample was evaluated using Nano & Martinez (1997), and the concentration of each was isolated using the protocol described in Aljanabi lysis buffer with 5 µl Proteinase K (Qiagen). DNA analysis (n = 276), as summarised in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>f</th>
<th>p</th>
<th>pcorr</th>
<th>rxy</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spring</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL-E</td>
<td>56</td>
<td>0.0279</td>
<td>&lt;0.05</td>
<td>0.0115</td>
<td>–</td>
<td>–0.0184</td>
</tr>
<tr>
<td>TH-E</td>
<td>55</td>
<td>0.1038</td>
<td>&lt;0.001</td>
<td>0.0861</td>
<td>&lt;0.001</td>
<td>–0.0207</td>
</tr>
<tr>
<td>MA-E</td>
<td>54</td>
<td>0.0710</td>
<td>&lt;0.001</td>
<td>0.0472</td>
<td>&lt;0.01</td>
<td>–0.0198</td>
</tr>
<tr>
<td><strong>Autumn</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL-M</td>
<td>32</td>
<td>0.0238</td>
<td>–</td>
<td>0.0075</td>
<td>–</td>
<td>–0.0401</td>
</tr>
<tr>
<td>TH-M</td>
<td>53</td>
<td>0.0487</td>
<td>&lt;0.01</td>
<td>0.0234</td>
<td>–</td>
<td>–0.0296</td>
</tr>
<tr>
<td>MA-M</td>
<td>26</td>
<td>0.0673</td>
<td>&lt;0.001</td>
<td>0.0219</td>
<td>–</td>
<td>–0.0776</td>
</tr>
</tbody>
</table>

Table 1. *Sparus aurata*. Spring and autumn 2011 samples used in this study, from southwest to northeast. Spring samples are composed of newly recruited entering (E) juveniles, while autumn samples are composed of late juveniles migrating (M) out of the lagoons to the open sea. \(f\) and \(f_{corr}\) represent, respectively, estimates of deviations from Hardy-Weinberg expectations for the original data set (23 loci) and after removing the C67b and Prl loci from the analysis (see ‘Results’ section). Levels of significance are reported. For each sample, the coefficients of relatedness \(r_{xy}\) (Queller & Goodnight 1989) estimated using the IDENTIX computer package (Belkhir et al. 2002) are also reported, together with their associated \(p\)-values. Corrections for multiple tests were applied when necessary. SL: Salses-Leucate; TH: Thau; MA: Mauguio (Fig. 1); n: sample size; –: not significant.

Population genetic analyses

Preliminary to genetic analyses, data consistency was checked with Micro-Checker 2.2.3 using default settings (van Oosterhout et al. 2004) to evaluate the presence of null alleles, large-allele dropout and scoring errors. Deviations from the Hardy-Weinberg expectations (HWE) within samples were investigated using GENETIX v4.05 (http://mbb.univ-montp 2.fr/MBB/subsection/downloads.php?section=2) by testing the null hypothesis of no significant departure from HWE \((f = 0)\) through 5000 random permutations. Among-population differentiation was measured using \(\theta\) (Weir & Cockerham 1984), which estimates Wright’s (1951) \(F_{ST}\). To deal with unequal sample
sizes between spring (non-limiting) and autumn migrating samples (Table 1), we tested for consistency of population differentiation estimates by randomly sampling 26 individuals in each spring sample. This made the sample sizes of spring and autumn samples similar (Table 1) and allowed us to perform unbiased comparisons. Five hundred subsamples of \( n = 26 \) were randomly drawn for each spring sample, and these bootstrapped samples were randomly associated in triads containing 1 sample of each location: MA, TH and SL. Each triad was then associated with the real (i.e. non-bootstrapped) autumn samples, and population differentiation indices were estimated as described in this paragraph.

Patterns of linkage disequilibrium in each sample were investigated as proposed by Weir (1979). Corrections for multiple testing were performed according to Narum (2006), when necessary, for both real and bootstrapped data to maintain the significance level at \( \alpha = 0.05 \). We estimated genetic relatedness among individuals within each sample (i.e. kin aggregation) by estimating the coefficient of relatedness \( r_{ST} \) (Queller & Goodnight 1989) using IDENTIX (Belkhir et al. 2002). Briefly, to test whether individuals within each sample were genetically more related than expected by chance, the mean identity index of all sea bream pairs was compared with the null distribution assuming no relatedness. The null distribution was obtained by calculating identity indices for randomly generated samples of similar size (Table 1).

Finally, we searched for putatively selected loci using the \( F_{ST} \) outlier detection method by Beaumont & Nichols (1996) implemented in the software LOSITAN (Antao et al. 2008). Outlier loci were searched using 3 different levels of comparison that were performed among (1) all 6 samples from the MA, TH and SL lagoons, (2) 3 entering (E) juvenile samples and (3) 3 migrating (M) juvenile samples. In all 3 tests, locus-specific \( F_{ST} \) values were compared with the null distribution of \( F_{ST} \) generated with 100,000 simulated loci using the estimated neutral mean \( F_{ST} \) to determine the average differentiation level targeted in simulations.

Next-generation sequencing of GH and Prl long-range amplicons

Long-range amplicon sequencing of the 2 candidate genes was performed to investigate haplotype structure. We also tested whether estimates of genetic differentiation reached higher levels at each candidate microsatellite locus or if nearby gene regions provided even larger estimates, potentially indicating the causative mutations experiencing selection. We hypothesised that under habitat-based selection, genetic differentiation would be higher in older juveniles caught in autumn than in younger ones caught soon after entering coastal lagoons. Therefore, we randomly selected 12 autumn individuals from MA and 12 autumn individuals from TH to evaluate genetic differentiation between the 2 most extreme habitats (Chaoui et al. 2012). New primers were designed to amplify the full gene sequences of GH and Prl genes using long-range PCR (LR-PCR) (see LR-PCR primers in Table S1 in Supplement 1). We targeted a 3685 bp region extending from the first exon (E1) to the last exon (E6) for the GH gene and a 4277 bp region spanning the full gene sequence, including both 5' and 3' untranslated regions (UTRs), for the Prl gene. LR-PCR was performed using the PCR Extender System (5 Prime). The reaction mix contained 2 µl of 10× LR-PCR buffer (15 mM MgCl₂), 1 µl of each primer (20 µM), 0.5 µl of enzyme mix and 2 µl of DNA (25 ng µl⁻¹) in a 20 µl final reaction volume. LR-PCR amplification parameters were as follows: initial denaturation at 94°C for 3 min; 10 cycles of denaturation at 94°C for 20 s, annealing at 59°C for 30 s, and elongation at 68°C for 4 min; 25 cycles of denaturation at 94°C for 20 s, annealing at 59°C for 30 s, and elongation at 68°C for 4 min plus 2 s per cycle; and final elongation at 68°C for 10 min. LR-PCR products obtained for the 24 autumn juveniles from MA and TH were quantified for each gene using a Qubit dsDNA BR Assay Kit (Invitrogen) and pooled in equimolar proportions for each individual. The 24 individual pools were then submitted to tagmentation and individual indexing using the Nextera XT DNA Sample Preparation Kit (24 samples) (Illumina), following the library preparation guide. The 24 individual libraries were pooled using 1 ng DNA per individual and quantified on a Bioanalyzer 2100 (Agilent) prior to sequencing on the Illumina MiSeq platform, with 250 paired-end reads.

Reconstruction and analysis of individual haplotypes

Individual raw sequencing data were quality filtered using Trimomatic (Bolger et al. 2014) to remove adapters and low-quality reads. We used GH and Prl gene sequences retrieved from GenBank to perform a reference assembly for each individual using Geneious 7.1.5 (Kearse et al. 2012). Five iterations were used for alignment, and after each itera-
tion, the reference sequence was corrected using the previous consensus alignment. Heterozygote positions were identified for each individual and encoded using the International Union of Pure and Applied Chemistry ambiguity code, and individual consensus sequences were aligned for each gene using ClustalW (Thompson et al. 1994). Alignments were edited by hand to remove ambiguities within regions of low complexities (mini- and microsatellites). Therefore, variation in repeat number was not scored, and we only focused on single nucleotide polymorphism (SNP) variation in subsequent analyses.

Individual haplotype reconstruction was performed using the coalescent Bayesian inference method implemented in PHASE v2.1 (Stephens et al. 2001) using default settings. We then evaluated the extent of haplotype structure to search for signatures of partial selective sweeps. To do this, phased individual haplotypes were then used to compute haplotype diversity ($H$), nucleotide diversity ($\pi$) and Tajima’s $D$ (Tajima 1989) for the whole gene sequences as well as in a 100 bp sliding window using DnaSP (Librado & Rozas 2009). Synonymous and nonsynonymous polymorphisms were searched with the coding regions of each gene. Finally, the similarity relationships among haplotypes were represented with a NeighborNet network for each gene using SplitsTree4 (Huson & Bryant 2006).

Microsatellite and haplotype data produced during this work are provided in Supplement 2 (available at www.int-res.com/articles/suppl/m558p115_supp.xls).

RESULTS

Hardy-Weinberg equilibrium and relatedness

No microsatellite marker displayed evidence of null alleles, allelic dropout or allele scoring problems. No consistent linkage disequilibrium was found among loci in any sample. Thirty-six significant linkage disequilibrium values were detected over 1516 tests (~3.37%), a proportion that would be expected by chance for $\alpha = 0.05$. Six loci displayed significant departures from HWE (deficits in heterozygotes: Dd16, 172EP, C67b, Fd92H, GH, Prl), but only 2 of them (C67b, Prl) presented significant $f > 0$ values in 4 samples (spring and autumn samples of MA and TH: MA-E, MA-M, TH-E and TH-M; Table 1). Deficits in heterozygotes were detected in all samples except SL-M and remained significant in only 2 samples when the C67b and Prl loci were removed from the data set (Table 1). This reduced dataset indicated that deviations from HWE were most likely due to a few loci and were not generated by a genome-wide effect due to the mixing of larvae from discrete reproductive units. No significant relatedness was detected within each individual sample, with all values of $r_{xy}$ being slightly negative (Table 1). This did not mean that related individuals were not present but simply that their number was too low to influence genetic structure.

Microsatellite spatiotemporal variation patterns

Real datasets

The level of genetic differentiation estimated using all 23 markers over the whole data set was not significant ($\theta = 0.0020 \pm 0.0160$), nor was genetic differentiation among E and M samples ($\theta = 0.0049 \pm 0.0108$ and $\theta = -0.0035 \pm 0.0094$, respectively). However, these multilocus average values did not reflect the variance among loci (Fig. 2). The GH and Prl loci showed a significant global genetic differentiation (Fig. 2a; $\theta_{GH} = 0.0192, \theta_{Pr} = 0.0188; p\text{-values} < 0.001$). Surprisingly, genetic differentiation at GH and Prl markers was significant among the E samples but not among the M samples (Fig. 2b,c). All pairwise differentiation values calculated between spring E samples were significant for both GH and Prl loci (GH: $\theta_{SL-E/TH-E} = 0.0285, \theta_{SL-E/MA-E} = 0.0331, \theta_{TH-E/MA-E} = 0.0199, \text{all } p < 0.01; \text{Prl: } \theta_{SL-E/TH-E} = 0.0110, \theta_{SL-E/MA-E} = 0.0223, \theta_{TH-E/MA-E} = 0.0162, \text{all } p < 0.01$). Significant genetic differentiation at these loci in spring is moreover related with observed phenotypic differentiation for size recorded among E samples (see ‘Materials and methods: Sampling’). The only marker showing significant genetic differentiation among autumn M samples was locus C67b ($\theta_{C67b} = 0.0181; p < 0.01$). Significant genetic differentiation at this locus was mostly explained by the SL-M sample that differed from both TH-M and MA-M ($\theta_{SL-M/TH-M} = 0.0224, p = 0.013; \theta_{SL-M/MA-M} = 0.0286, p < 0.001$), while TH-M and MA-M samples were not differentiated ($\theta_{MA-M/TH-M} = 0.0055$, non-significant [ns]). This result did not match phenotypic differentiation for size, as the size of M individuals was found significantly larger in MA compared to the other lagoons but similar between SL and TH (see ‘Materials and methods: Sampling’). Outlier detection tests corroborated the outlying genetic differentiation level of GH and Prl loci among the 3 E samples as well as for locus C67b among the 3 M samples (Fig. S2 in Supplement 1).
Genetic differentiation estimates between E and M samples within each lagoon revealed no significant temporal genetic differentiation in MA (Prl: \( \theta = 0.0094, \) ns; GH: \( \theta = 0.0143, \) ns) but significant allele frequency changes in TH (Prl: \( \theta = 0.0233, p < 0.01; \) GH: \( \theta = 0.0193, p < 0.01 \)) and in SL (Prl: \( \theta = 0.0254, p < 0.01; \) GH: \( \theta = 0.0289, p < 0.01 \)). Moreover, autumn juveniles migrating out of TH and SL were genetically undifferentiated (Prl: \( \theta = -0.0045, \) ns; GH: \( \theta = 0.0004, \) ns). Thus, the loss of genetic differentiation among lagoons in migrating samples was largely due to similar temporal changes in allele frequencies that occurred in the TH and SL lagoons between spring and autumn. Fish sampled in TH and SL were shown to likely belong to a demographic unit that had undergone a similar life history throughout the year (see ‘Materials and methods: Sampling’ and Fig. S1), suggesting that these demographic units inhabiting ecologically similar lagoons had undergone similar allele frequency variation between spring and autumn.

Bootstrapped datasets

Because spring and autumn sample sizes differed, we performed bootstrap resampling to investigate the consistency of patterns of genetic differentiation reported in the previous subsection. Sample sizes of E samples were standardised to \( n = 26 \) individuals to match the sample size of M samples (Table 1). The level of overall genetic differentiation (i.e. among the 6 samples) estimated using all 23 markers was found significant in only 7 (1.4\%) of 500 replicates. Triads of bootstrapped E samples were also compared, and significant multilocus genetic differentiation was detected in only 13 (2.6\%) of 500 bootstrap resamplings. Such patterns are expected by chance alone; therefore, multilocus genetic homogeneity detected with the real dataset among all 6 samples or among the E samples was likely not affected by unequal sampling sizes between spring and autumn juveniles.

When looking specifically at patterns of genetic differentiation at the GH and Prl loci in bootstrapped samples, results reported significant overall genetic differentiation in 448 (89.6\%) and 463 (93.0\%) of 500 bootstrap resamplings for Prl and GH, respectively. These numbers reached 463 (92.6\%) and 484 (96.8\%) for Prl and GH, respectively, when considering only the triads of bootstrapped E samples. Differentiation patterns observed at the GH and Prl loci were thus not dependent on sample size. By contrast,
we found a lack of genetic differentiation among E samples in 7.4 and 3.2% of resampled data sets for Prl and GH, respectively. Hence, the lack of observed genetic differentiation among the real M samples was likely not affected by sample sizes (Fig. 2).

Similarly, when standardising E samples to n = 26, locus C67b displayed significant genetic differentiation in only 21 (4.2%) of 500 cases. This suggested that the absence of genetic differentiation among real E samples was true for this locus.

Overall, our bootstrap procedure supported that the observed genetic differentiation patterns at the GH, Prl and C67b loci were robust to unequal sample sizes, with error rates generally <0.05. We hence considered that estimates of pairwise genetic differentiation and outlier detection tests reported in the previous subsection for the real data set were relevant.

**Haplotype structure in sea bream for GH and Prl genes**

Microsatellite GH and Prl markers did not display genetic differentiation in autumn samples, so we additionally investigated whether genetic differentiation could be higher at other genomic locations within or close to each candidate gene. Among 14 407 916 filtered paired-end reads, 5 067 129 and 5 753 776 were successfully aligned to the GH and Prl gene reference sequences, respectively, providing extremely profound sequencing depth for each gene in each individual (Table S2 in Supplement 1). The GH and Prl gene sequences analysed in this study have lengths of 3633 and 4079 bp, respectively (Table 2). Haplotype diversity (H) was close to 1 at each locus, indicating that almost every individual had 2 unique haplotypes (Table 2). Nucleotide diversity was almost twice as high for Prl (π = 0.0072) as for GH (π = 0.0039). Albeit non-significant, Tajima’s D showed a slight excess of intermediate-frequency sites at locus Prl and an excess of rare variants at locus GH (positive and negative estimates of D, respectively; Table 2). Nucleotide diversity (π) profiles calculated in overlapping 100 bp windows revealed heterogeneous diversity along each gene (Fig. 3). These profiles largely reflected the intron/exon structure within each gene, with exonic regions showing reduced diversity. Only a few non-synonymous mutations segregating at low frequencies were found in each gene (not reported), making them unlikely to be related to the functional variants under spatially varying selection. The highest diversity levels were in the 5’ UTR as well as in introns 1 and 3 of the Prl

<table>
<thead>
<tr>
<th>Gene</th>
<th>Length (bp)</th>
<th>No. of ind.</th>
<th>No. Haplotypes</th>
<th>Nucleotide diversity (H)</th>
<th>Nucleotide diversity (π)</th>
<th>Tajima’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prl</td>
<td>4079</td>
<td>24</td>
<td>46</td>
<td>0.997</td>
<td>0.0072</td>
<td>0.0707 (ns)</td>
</tr>
<tr>
<td>GH</td>
<td>3633</td>
<td>24</td>
<td>47</td>
<td>0.999</td>
<td>0.0039</td>
<td>-0.5711 (ns)</td>
</tr>
</tbody>
</table>

Table 2. *Sparus aurata*. Genetic diversity indices for prolactin (Prl) and growth hormone (GH) genes after haplotype reconstruction. Tajima’s D values are both non-significant (ns)

Fig. 3. *Sparus aurata*. Profiles of nucleotide diversity averaged in 100 bp sliding windows along (a) growth hormone (GH) and (b) prolactin (Prl) genes in the Mauguio and Thau lagoons. E: exon; I: intron; UTR: untranslated region
Gene and in introns 1 and 3 of the GH gene. Little linkage disequilibrium was found among the most variable regions within each gene, indicative of low haplotype structure. However, we could distinguish for each gene 2 haplotype groups that were evenly distributed among TH-M and MA-M juveniles (Fig. 4).

**DISCUSSION**

The aim of this study was to re-investigate patterns of genetic differentiation among sea bream populations collected in coastal lagoons of the Gulf of Lions using a larger array of microsatellite markers than in previous studies in this species (Alarcón et al. 2004, De Innocentiis et al. 2004, Chaoui et al. 2009; but see Coscia et al. 2012). A particular focus was made on microsatellite markers located close to 2 candidate genes, GH and Prl, that are functionally related to growth and osmoregulatory capacities in fish (Streelman & Kocher 2002, Almuly et al. 2008, Quéré et al. 2010) and have been previously showed to be involved in habitat-based genetic differentiation in sea bream (Chaoui et al. 2012). This latter study suggested that GH and Prl genotypes with varying selective advantage among lagoons impact mortality in juvenile sea breams that spent their first year of life in ecologically contrasted lagoons (i.e. with different salinity, temperature, food availability conditions). Here, we re-evaluated this result by analysing genetic differentiation patterns at GH and Prl genes together with a new set of neutral loci in another cohort of sea bream sampled at 2 different periods of their first year of life.

The set of 23 microsatellite markers used in this study did not show any significant multilocus genetic differentiation among lagoons in the newly recruited juveniles, confirming that the sea bream population in the Gulf of Lions broadly behaves as a panmictic unit, as reported in Chaoui et al. (2009), with fewer markers. Nevertheless, sea bream larvae from discrete reproductive events have been reported to colonise the 3 lagoons at different dates, separated by up to 3 wk (Isnard et al. 2015; Fig. S1 in Supplement 1). Sweepstake recruitment may therefore occur, although it did not translate into significant multilocus genetic differentiation in our dataset. The absence of significant kinship structure was also suggestive of a large local population size, as previously reported for the sea bass *Dicentrarchus labrax* in the Gulf of Lions (Lemaire et al. 2000, Guinand et al. 2008).

The candidate GH and Prl markers were found genetically differentiated in our study. Bootstrap resampling showed that genetic differentiation patterns observed at these loci were robust to variation in sampling size among samples. If observed genetic differentiation was due to neutral demographic factors, other loci would have displayed non-null differentiation values. These results thus confirm that habitat-based selection operates at the GH and Prl genes, as previously reported in Chaoui et al. (2012). Besides, the only other marker showing significant genetic differentiation and a possible footprint of selection was locus C67b. In this case, genetic differentiation was observed among M samples and was
mainly explained by the distinctiveness of the SL-M sample. This locus was associated with the ubiquitin ligase nedd4-2 (neural precursor cell-expressed, developmentally down-regulated protein 4-2) gene, a gene involved in regulation of the epithelial sodium channel in several organs, including kidney in mouse (Gormley et al. 2003, Bens et al. 2006). It may therefore be associated with osmoregulation in sea bream, but this function has never been investigated in fish. This unexpected result for locus C67b should be re-evaluated in an independent study, and it will not be considered further here.

Considering candidate GH and Prl markers, only E samples were found genetically differentiated. Together with observed differences in body size among E individuals from each lagoon, these results suggest that early selection may have occurred at or just after settlement in response to specific selective pressures in each lagoon. A sequential temporal genetic response to a sea–lagoon selective gradient was already suggested for young sea bass recruits in the MA lagoon, with a possible signature of early selection appearing very soon after colonisation, while further selective differences were only observed after some months (Guinand et al. 2015) (see also Planes & Roman 2004 for the white sea bream Diplodus sargus). In addition, Chaoui et al. (2012) showed that a mechanism based on genotype-dependent habitat choice with no selective mortality was very unlikely to produce the observed genetic structure at the GH and Prl loci.

Conversely, no genetic differentiation at candidate microsatellite gene markers was detected among M samples. This result was confirmed by long-range amplicon sequencing, which could not detect any signal of haplotype structure. These results challenge the traditional view that genetic differentiation at candidate gene markers accumulates with the time spent within the habitats where selection is at play, as previously recorded in the flounder Platichthys flesus (Laroche et al. 2002) or the common sole Solea vulgaris (Guinand et al. 2011). Temporal genetic differentiation has been tested previously in sea bream, although using a distinct sampling scheme (Chaoui et al. 2012). Interestingly, a trend toward a reduction in genetic differentiation was already observed at locus Prl in the MA lagoon. Thus, the genetic homogenisation observed in this study between spring and autumn juveniles is not contradictory with previous observations based on a different year cohort. However, it affected both GH and Prl markers and resulted in a loss of genetic differentiation at candidate gene markers in autumn.

Hypotheses explaining the genetic homogeneity among M samples of sea bream first involve the potential mixing of individuals from the different habitats, prior to migration of individuals to the sea. Otolith microchemistry analyses have revealed that individuals from MA, TH and SL can be distinguished by specific chemical signatures associated with each nursery habitat (Mercier et al. 2011, 2012, Tournois et al. 2013). These signatures are relatively stable across multiple annual cohorts despite variations in meteorological conditions across the years. This suggests that there are very low levels of migration among lagoons at the juvenile stage during the first year, even at these relatively short distances. While lagoons such as TH are connected by canals to several adjacent lagoons that are ecologically similar to MA, the mixing of juveniles among the different habitats is unlikely.

A second hypothesis would be that similar environmental conditions later in the year make allele frequencies at candidate genes converge to similar equilibrium frequencies. In the present study, allele frequency changes between E and M juveniles of the 2011 sea bream cohort mainly occurred in TH and SL for both genes. Since these marinelike nurseries have almost identical fish species composition (Pérez-Ruzafa et al. 2011) and similar physico-chemical characteristics (e.g. Bec et al. 2011), temporal changes in allele frequency observed between TH and SL may suggest convergence to equilibrium frequencies while fish were genetically distinct at the 2 candidate gene markers in spring. TH and SL individuals were also different for the average size in spring but were not in autumn, indicating that both genotypic and phenotypic convergence occurred in TH and SL. Isnard et al. (2015) reported that compared to sea bream in the MA lagoon, sea bream inhabiting the TH and SL lagoons have (1) similar characteristics in body condition in autumn and (2) similar growth rates during their life history. However, these observations do not explain why similar allele frequencies are also observed in the MA sample in autumn, since this lagoon is ecologically distinct from SL and TH (e.g. Bec et al. 2011, Pérez-Ruzafa et al. 2011) and is also phe-notypically differentiated (Isnard et al. 2015). Results observed at candidate genes for the MA lagoon are also puzzling, as no genetic differentiation was observed between E and M individuals at the Prl and GH loci in this study, while substantial genetic differentiation was observed at each locus among young and older sea bream in Chaoui et al. (2012). Therefore, the relative roles of migration and post-settlement selection in the autumnal dedifferentiation remain difficult to disentangle in sea bream.
Our results hence suggest that patterns of genetic differentiation estimated for markers associated with candidate genes reflect spring selection in the 2011 cohort of sea bream. However, the temporal analysis of genetic differentiation patterns at each candidate gene in each lagoon suggests a lack of stability for the patterns established in spring. This may be due to highly dynamic and heterogeneous environmental conditions of coastal lagoons (Pérez-Ruzafa et al. 2005, 2007, Tournois et al. 2013) that may generate within-year variations in the strength and direction of selection. This emphasises the need for increased spatiotemporal resolution to monitor the potentially complex responses to selection during early development of marine organisms (Broquet et al. 2013, Hernández-Garcia et al. 2015).

As a supplement to the study of microsatellite markers located in the proximal promoter region of the GH and Prl genes, the study of molecular variation by amplicon resequencing allowed investigation of the footprint produced by spatially varying selection within these genes. A detailed map of genetic variation was obtained along GH and Prl genes using 250 bp paired-end resequencing. This approach was efficient for reconstructing individual haplotypes defined by combinations of neighbouring SNPs but could not determine the exact number of repeats associated with micro- and mini-satellite repeats. However, we detected an average number of 35 repeats for the 17 bp motif (GAC CTG TCT GTC TCT CT) of the regulatory mini-satellite located in the first intron of the GH gene (Almuly et al. 2008). Interestingly, intron 1 was the most variable region within the GH gene (due to variation within and among repeats), and the proximal promoter region involved in the modulation of Prl gene expression (Astola et al. 2003) was also among the most variable regions of the Prl gene. These results are therefore consistent with a form of balancing selection maintaining high levels of polymorphism in regulatory regions, whereas, initially, the allele frequency shifts observed at candidate microsatellite loci have been attributed to direct selection acting on cis-regulatory polymorphisms (see Chaoufi et al. 2012). Although our analysis revealed haplotype groups for each gene, close similarities among groups indicated that any possible signature of selection had largely been erased by recombination. This is consistent with the prediction that the chromosomal signature left by local (i.e. habitat-based) selection in panmixia typically affect narrow genomic regions, especially if spatially varying selection has been maintaining polymorphism for a long time (Gagnaire et al. 2015).

Acknowledgements. This research was funded by IFREMER (project SpartaGen of the GDR MarCo), the CNRS (project EC2CO METASPAR) and the TOTAL Foundation (project LAGUNEX). The authors are grateful to local fishermen for fish collections, to E. Isnard and F. Ferraton for fish dissection, to L. Velez for helping with the figures and to M. T. Augé for assistance during lab work. Our studies on population genetics of sea bream were first initiated by F. Bonhomme.

LITERATURE CITED

- Blel H, Panfilj I, Guinand B, Berrebi P, Sait K, Durand JD


Imsar E (2012) Comparison of condition and life history traits of gilthead seabream * Sparus aurata* L. juveniles after their first summer in various lagoons on the Gulf of Lions (Mediterranean Sea). MSc thesis, Université de Montpellier II


Imsar E (2012) Comparison of condition and life history traits of gilthead seabream *Sparus aurata* L. juveniles after their first summer in various lagoons on the Gulf of Lions (Mediterranean Sea). MSc thesis, Université de Montpellier II


Selkoe KA, Hentzler CM, Gaines SD (2008) Seascape genetics and the spatial ecology of marine populations. Fish Fish 9:363–377


Editorial responsibility: Philippe Borsa, Montpellier, France

Submitted: December 14, 2015; Accepted: July 27, 2016

Proofs received from author(s): September 28, 2016