HAL
open science

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

B Guinand, C Chauvel, M Lechene, J Tournois, Costas S Tsigenopoulos, Audrey M. Darnaude, David J. Mckenzie, P.-A Gagnaire

## To cite this version:

B Guinand, C Chauvel, M Lechene, J Tournois, Costas S Tsigenopoulos, et al.. Candidate gene variation in gilthead sea bream reveals complex spatiotemporal selection patterns between marine and lagoon habitats. 2016. hal-01370327

**HAL Id: hal-01370327**

**https://hal.science/hal-01370327v1**

Preprint submitted on 22 Sep 2016

**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.

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

B. Guinand[1, 2 *], C. Chauvel[2, 3], M. Lechene[2, 3], J. Tournois[4], C.S. Tsigenopoulos[5], A.M. Darnaude[4], D.J. McKenzie[4], P.-A. Gagnaire[2, 3]

**Accepted Manuscript: doi 103354/meps1185**

[1] : Université de Montpellier, Département Biologie-Ecologie, Place E. Bataillon, 34095 Montpellier, France

[2] : UMR ISEM, Institut des Sciences de l'Evolution de Montpellier, Université de Montpellier, 34095 Montpellier, France

[3] : Station Biologique Marine de Sète, 2 rue des Chantiers, 34200 Sète, France

[4] : UMR MARBEC, Marine Biodiversity, Exploitation and Conservation, Université de Montpellier, Place E. Bataillon, 34095 Montpellier, France

[5] : Institute of Marine Biology and Genetics, Hellenic Center for Marine Research, PO Box 2214, Gournes Pediados, 71500 Heraklion, Crete, Greece

*: corresponding author

Running head: Complex sea-lagoon selection patterns in sea bream

**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 twenty-one putatively neutral microsatellite loci. We compared genetic variation of spring juveniles that had just settled in three 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 two 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 colonization 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.

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; Iacchei 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 upon factors such as the strength of the cohorts or variation in local dispersal and connectivity, it generally produces un-patterned 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; Véliz 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 (Lottherhos & 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 North Western Mediterranean (Chaoui et al. 2012). After hatching at sea in December-February, a large proportion of the juveniles colonize coastal lagoons in March-April for foraging (Mercier et al. 2011, Isnard et al. 2015), before returning to the open sea in October-November to overwinter when lagoon water temperatures drop (Audoin 1962, Lasserre 1974, Mercier et al., 2012). In the Gulf of Lions, two lagoons located about 30km apart have been studied, which offer very different nursery habitats: (*i*) the Mauguio lagoon, which is shallow and highly productive, with warm summer temperatures and wide salinity variations, and (*ii*) the Thau 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* and *Prl* genes in sea bream (Astola et al. 2003, Almuly et al. 2008; respectively), Chaoui et al. (2012) hypothesized 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. 2010).

In this study, we first specifically assessed the temporal stability of habitat-based genetic differentiation at the *GH* and *Prl* loci, extending our spatial coverage to another

large coastal lagoon, Salses-Leucate, which is a marine-like habitat ecologically similar to Thau 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, in order to partition the relative influence of drift and selection in patterns of genetic variation. We also used next-generation amplicon sequencing to reconstruct haplotypes of the *GH* and *Prl* genes in 24 individuals from Thau and Mauguio lagoons to search for additional haplotype-based signals of selection.

## Materials and Methods

### *Sampling*

The north-western Mediterranean coast of the Gulf of Lions (Fig. 1) is characterized 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 Salses-Leucate, Thau, and Mauguio (also known as 'Etang de l'Or') 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 three lagoons (Electronic Suppl. Fig S1). Estimation of the recruitment date in each lagoon closely matched the dates of sampling for our entering individuals in Salses-Leucate and Thau. In Mauguio, however, individuals were sampled after the estimated recruitment date (see Electronic Suppl. Fig S1). Fork length (mean ± SD) of entering individuals were 31.3 ± 2.5 mm in Salses-Leucate, 34.7 ± 3.7 mm in Thau, and 41.3 ± 4.4 mm in Mauguio, with populations means being significantly different from each other (Tukey post-hoc test; all $P$'s < 0.05). Larger sizes observed at Mauguio 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 Mauguio). Total lengths (mean ± SD) of autumn (migrating) individuals were found significantly larger in Mauguio (184 ± 11mm) than in other lagoons (Thau: 172 ± 11mm; Salses-Leucate: 167 ± 7mm) that were not different of each other (Isnard 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 ($N$ = 256), as summarized in Table 1.

*Microsatellite genotyping*

A small piece of tissue was incubated overnight in a lysis buffer with 5µl proteinase K (Qiagen). DNA was isolated using the protocol described in Aljanabi & Martinez (1997), and the concentration of each individual DNA sample was evaluated using NanoDrop 8000 (ThermoScientific) and standardised to 20ng.µl$^{-1}$ of genomic DNA. Twenty-three microsatellite loci were analyzed (Electronic Suppl. Table S1). Two of them were candidate *cis*-regulatory microsatellites described for *GH* (Almuly et al. 2005) and *Prl*

(Astola et al. 2003) loci, respectively. Protocols for PCR amplification and genotyping at these two loci are reported in Chaoui et al. (2009). The remaining twenty-one microsatellite loci were developed by Franch et al. (2006) and Coscia et al. (2012) (Electronic Suppl. Table S1), and were considered as putatively neutral. These loci were grouped in two multiplexes of ten and eleven loci, respectively. Multiplexed PCR was optimised using the QIAGEN multiplex PCR kit in a 10µl final volume with 4.5µl Qiagen multiplex PCR master mix, 1µl Q-solution, 1µl of genomic DNA and 1.5µl of each primer. PCR conditions were as follow: initial denaturation 5 min (95°C), 30 cycles at 94°C for 30s, annealing at 58°C for 30s, elongation at 72°C for 60s, then a final elongation of 5 min. Amplifications were performed on a Mastercycler Gradient (Eppendorf) or a Gradient Cycler PTC200 (Bio-Rad) according to the above-mentioned PCR conditions. Genotyping of individuals was performed on an ABI PRISM® 3130XL DNA Analyzer (Life Technologies), using 5′-labelled primers and the GeneScanTM-500 LIZ® Size Standard (Life Technologies) as internal size standard (1µl of multiplex-PCR product; 12µl formamide, 0.2µl internal size standard). Allele scoring was performed using the GeneMapper software v.4.0 (Life Technologies).

*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 drop-out and scoring errors. Deviations from the Hardy–Weinberg expectations (HWE) within samples were investigated using GENETIX v4.05 (http://mbb.univ-montp2.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's 1984) which estimates Wright's (1951) $F_{ST}$. In order 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 twenty-six 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 one sample of each location: MA, TH and SL. Each triad was then associated to the real (i.e. non bootstrapped) autumn samples, and population differentiation indices were estimated as described above.

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 $r_{xy}$ (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 three different levels of comparison that were performed among (*i*) all six samples from Mauguio (MA), Thau (TH) and Salses-Leucate (SL) lagoons, (*ii*) three entering (E) and (*iii*) three migrating (M) juvenile samples. In all

three 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 two candidate genes was performed in order 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 hypothesized 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 individuals from TH to evaluate genetic differentiation between the two 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 (LR) PCR (see LR- PCR primers Electronic Suppl. Table S1). We targeted a 3685 base pair (bp) region extending from the first (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'UTRs (UnTranslated Regions) for the *Prl* gene. LR-PCR was performed using the PCR Extender System (5 Prime). Reaction mix comprised 2µl of 10X LR-PCR buffer (15 mM $MgCl_2$), 1µl of each primer (20 µM), 0.5µl of enzyme mix and 2µl of DNA (25ng.$\mu l^{-1}$) in a 20µl 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 20s, annealing at 59°C for 30s, elongation at 68°C for 4 min, 25 cycles of denaturation at 94°C for 20s, annealing at 59°C for 30s, elongation at 68°C for 4 min + 2s per cycle, and ended by 10

min of final elongation at 68°C. LR-PCR products obtained for 24 autumn juveniles from MA and TH were quantified for each gene using 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 1ng 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 TRIMMOMATIC (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, after each of which the reference sequence was corrected using the previous consensus alignment. Heterozygote positions were identified for each individual and encoded using the IUPAC 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 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 (π), and Tajima's *D* (Tajima 1989) for the whole gene sequences, as well as in a 100bp 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).

## Results

### *Hardy-Weinberg equilibrium and relatedness*

No microsatellite marker displayed evidence of null alleles, allelic drop-out 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 two of them (*C67b, Prl*) presented significant $f > 0$ values in four samples (spring and autumn samples of Mauguio and Thau: 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 two 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 ± 0.0160), nor was genetic differentiation among E and M samples ($\theta$ = 0.0049 ± 0.0108, and $\theta$ = -0.0035 ± 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, $\theta_{Prl}$ = 0.0188; *P*'s<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*: $\theta_{SL\text{-}E/TH\text{-}E}$ = 0.0285, $\theta_{SL\text{-}E/MA\text{-}E}$ = 0.0331, $\theta_{TH\text{-}E/MA\text{-}E}$ = 0.0199, all *P*'s<0.01; *Prl*: $\theta_{SL\text{-}E/TH\text{-}E}$ = 0.0110, $\theta_{SL\text{-}E/MA\text{-}E}$ = 0.0223, $\theta_{TH\text{-}E/MA\text{-}E}$ = 0.0162, all *P*'s<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 ($\theta_{SL\text{-}M/TH\text{-}M}$ = 0.0224, $P$ = 0.013; $\theta_{SL\text{-}M/MA\text{-}M}$ = 0.0286, $P$ < 0.001), while TH-M and MA-M samples were not differentiated ($\theta_{MA\text{-}M/TH\text{-}M}$ = 0.0055, 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 three E samples, as well as for locus *C67b* among the three M samples (Electronic Suppl. Fig. S2).

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 TH and SL lagoons between spring and autumn. Fish sampled in TH and SL were showed to likely belong to a demographic unit that undergone a similar life-history throughout the year (see "Materials and Methods – *Sampling*", and Electronic Suppl. Fig. S1), suggesting that these demographic units inhabiting ecologically similar lagoons 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 standardized 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%) out of 500 replicates. Triads of bootstrapped E samples were also compared and significant multilocus genetic differentiation was detected in only 13 cases (2.6%) out of 500 bootstrap resamplings. Such patterns are expected by chance alone, and therefore multilocus genetic homogeneity detected with

the real dataset among all 6 samples or among the E samples were 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 465 (93.0%) cases out 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 standardizing E samples to $N$ = 26, locus *C67b* displayed significant genetic differentiation in only 21 (4.2%) cases out of 500. 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 test reported in the previous subsection for the real data set were relevant.

*Haplotype structure in the 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 a total of 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 (Electronic Suppl. Table S2). The *GH* and *Prl* gene sequences analysed in this study have lengths of 3,633bp and 4,079bp, respectively (Table 2). Haplotype diversity (*H*) was close to one at each locus, indicating that almost every individual had two unique haplotypes (Table 2). Nucleotide diversity was almost twice as high for *Prl* ($\pi$ = 0.0072) as for *GH* ($\pi$ = 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 ($\pi$) profiles calculated in overlapping 100pb windows revealed very 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* 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 two 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 (De Innocentiis et al.

2004; Alarcon et al. 2004; Chaoui et al. 2009; but see Coscia et al. 2012). A particular focus was made on microsatellite markers located close to two candidate genes, *GH* and *Prl*, that are functionally related to growth and osmoregulatory capacities in fish (Streelman and 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 two different periods of their first year of life.

The set of twenty three 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 three lagoons at different dates, separated by up to three weeks (Isnard et al. 2015; Electronic Suppl. Fig. S1). Sweepstake recruitment may therefore occur, although it did not translate into significant multilocus genetic differentiation in our data set. 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 his 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 marker 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 homogenization 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 to converge to similar equilibrium frequencies. In the present study, allele frequency changes between the E and M juveniles of the 2011 sea bream cohort mainly occurred in TH and SL for both genes. Since these marine-like 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 an equilibrium frequencies while fish were genetically distinct at the two candidate gene markers in spring. TH and SL individuals were also different for the average size in spring while they were not anymore in autumn, indicating that both genotypic and phenotypic convergence occurred in TH and SL. Isnard et al. (2015) reported that sea bream inhabiting these two lagoons (*i*) have similar characteristics in body condition in autumn compared to MA, and (*ii*) that their growth rates were similar during their life-history, also compared to MA. However, these observations do not explain why similar allele frequencies are also observed in the MA sample in autumn, while this lagoon is ecologically distinct from SL and TH (e.g., Bec et al. 2011; Pérez-Ruzafa et al. 2011) and is also phenotypically differentiated (Isnard et al. 2015). Results observed at candidate genes for the MA lagoon are also puzzling, as no genetic differentiation was observed between the 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 to 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 emphasizes the need for increased spatiotemporal resolution, to monitor the potentially complex responses to selection during early development of marine organisms (Broquet et al. 2013; Hernandez-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 the 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 (GACCTGTCTGTCTCTCT) 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 Chaoui 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), by 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.

*Archiving of data:* Data will be deposited on a public database upon final acceptance of the current submission.

## References


Aglieri G, Papetti C, Zane L, Milisenda G, Boero F, Piraino S (2014) First evidence of inbreeding, relatedness and chaotic genetic patchiness in the holoplanktonic jellyfish *Pelagia noctiluca* (Scyphozoa, Cnidaria). PLoS ONE 9: e99647

Alarcón JA, Magoulas A, Georgakopoulos T, Zouros E, Alvarez MC (2004) Genetic comparison of wild and cultivated European populations of the gilthead sea bream (*Sparus aurata*), Aquaculture 230: 65-80.

Almuly R, Poleg-Danin Y, Gorshkov S, Gorshkova G, Rapoport B, Soller M, Kashi Y, Funkenstein B (2005) Characterization of the 5′ flanking region of the growth hormone gene of the marine teleost, gilthead sea bream *Sparus aurata*: analysis of a polymorphic microsatellite in the proximal promoter. Fish Sci 71: 479–490

Almuly R, Skopal T, Funkenstein B (2008) Regulatory regions in the promoter and first intron of *Sparus aurata* growth hormone gene: repression of gene activity by a polymorphic minisatellite. Comp Biochem Physiol D: Genomics Proteomics 3: 43–50

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

Astola A, Ortiz M, Calduch-Giner JA, Perez-Sanchez J, Valdivia MM (2003) Isolation of *Sparus auratus* prolactin gene and activity of the *cis*-acting regulatory elements. Gen Comp Endocrinol 134: 57–61

Audouin J (1962) La daurade de l'étang de Thau *Chrysophrys aurata* (Linné). Revue des Travaux l'Institut des Pêches Maritimes 26

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

Bec B, Collos Y, Souchu P, Vaquer A, Lautier J, Fiandrino A, Benau L, V. Orsoni V, Laugier T (2011) Distribution of picophytoplankton and nanophytoplankton along an anthropogenic eutrophication gradient in French Mediterranean coastal lagoons. Aquat Microb Ecol 63: 29−45

Belkhir K, Castric V, Bonhomme F (2002) IDENTIX, a software to test for relatedness in a population using permutation methods. Mol Ecol Notes 2: 611–614

Bens M, Chassin C, Vandewalle A (2006) Regulation of NaCl transport in the renal collecting duct: lessons from cultured cells. Pflugers Arch - Eur J Physiol 453: 133-146

Bert TM, Arnold WS, Wilbur AE, Seyoum S, McMillen-Jackson AL, Stephenson SP, Weisberg RH, Yarbro LA 2014. Florida Gulf bay scallop (*Argopecten irradians concentricus*) population genetic structure: form, variation, and influential factors. J Shellfish Res 33: 199-236

Blel H, Panfili J, Guinand B, Berrebi P, Said K, Durand JD (2010) Selection footprint at the first intron of the *Prl* gene in natural populations of the flathead mullet (*Mugil cephalus*, L. 1758). J Exp Mar Biol Ecol 210: 60-67

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics btu170

Broquet T, Viard F, Yearsley JM (2013) Genetic drift and collective dispersal can result in chaotic genetic patchiness. Evolution 67: 1760-1775

Chaoui L, Kara MH, Quignard JP, Faure E, Bonhomme F (2009) Strong genetic differentiation of the gilthead sea bream *Sparus aurata* (L., 1758) between the two western banks of the Mediterranean. C R Biol 332: 329–335

Chaoui L, Gagnaire PA, Guinand B, Quignard JP, Tsigenopoulos C, Kara MH, Bonhomme F (2012) Microsatellite length variation in candidate genes correlates with habitat in the gilthead sea bream *Sparus aurata*. Mol Ecol 21: 5497–5511.

Carson HS, López-Duarte PC, Rasmussen L, Wang D, Levin LA (2010) Reproductive timing alters population connectivity in marine metapopulations. Curr Biol 20: 1926-1931

Coscia I, Vogiatzi E, Kotoulas G, Tsigenopoulos CS, Mariani S (2012) Exploring neutral and adaptive processes in expanding populations of gilthead sea bream, *Sparus aurata* L., in the North-East Atlantic. Heredity 108: 537-546

De Innocentiis S, Lesti A, Livi S, Rossi AR, Crosetti D, Sola L (2004) Microsatellite markers reveal population structure in gilthead sea bream *Sparus auratus* from the Atlantic Ocean and Mediterranean Sea. Fish Sci 70: 852-859

David P, Perdieu MA, Pernot AF, Jarne P (1997) Fine-grained spatial and temporal population genetic structure in the marine bivalve *Spisula ovalis* L. Evolution 51: 1318–1322

Domingues CP, Creer S, Taylor MI, Queiroga H, Carvalho GR (2011) Temporal genetic homogeneity among shore crab (*Carcinus maenas*) larval events supplied to an estuarine system on the Portuguese northwest coast. Heredity 106: 832–840

Franch R, Louro B, Tsalavouta M, Chatziplis, Tsigenopoulos CS, Sarropoulou E, Antonello J, Magoulas A, Mylonas CC, Babbucci M, Patarnello T, Power DM, Kotoulas G, Bargelloni L (2006) A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. Genetics 174: 851-861

Gagnaire P-A, Normandeau E, Côté C, Hansen MM, Bernatchez L (2012) The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). Genetics 190: 725-736

Gagnaire P-A, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F, Arnaud-Haond S, Bierne N (2015) Using neutral, selected, and hitchhicker loci to assess connectivity of marine populations in the genomic era. Evol Applic 8: 769-786

Gormley K, Dong Y, Sagnella GA (2003) Regulation of the epithelial sodium channel by accessory proteins. Biochem J 371: 1-14

Guinand B, Dujardin E, Dufour V, Tsigenopoulos C (2008) Characterisation of genetic structure of *Dicentrarchus labrax* larvae in two nurseries of the Gulf of Lions (NW Mediterranean). Aquat Living Resour 21: 81–87

Guinand B, Durieux EDH, Dupuy C, Cerqueira F, Bégout ML (2011) Phenotypic and genetic differentiation in young-of-the-year common sole (*Solea solea*) at differentially contaminated nursery grounds. Mar Environm Res 71: 195-206

Guinand B, Quéré N, Desmarais E, Lagnel J, Tsigenopoulos CS, Bonhomme F. (2015) From the laboratory to the wild: salinity-based genetic differentiation of the European sea bass (*Dicentrarchus labrax*) using gene-associated and gene-independent microsatellite markers. Mar Biol 162: 515-538

Hanski IA (2011) Eco-evolutionary spatial dynamics in the Glanville fritillary butterfly. Proc Natl Acad Sci USA 108: 14397-14404

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish 9: 333–362

Hedgecock D (1986) Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates. Bull Mar Sci 39: 550-564

Hedgecock D (1994) Does variance in reproductive success limit effective population size of marine organisms? In: AR Beaumont (Ed.) Genetics and evolution of aquatic organisms. Chapman & Hall, London, p 122-134

Hedgecock D, Launey S, Pudovkin AI, Naciri Y, Lapègue S, Bonhomme F (2007) Small effective number of parents ($N_b$) inferred for a naturally spawned cohort of juvenile European flat oysters *Ostrea edulis*. Mar Biol 150: 1173–1182

Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. Bull Mar Sci 87: 971-1002

Hellberg ME (2006) Genetic approaches to understanding marine metapopulation dynamics. In: JP Kritzer, Sale PF (eds) Marine metapopulations. Elsevier, Amsterdam, p 431–455

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. Ann Rev Ecol Evol Syst 40: 310–491

Hernández-Garcia R, Muñoz I, López-Capel A, Marcos C, Pérez-Ruzafa Á (2015) The influence of environmental variability of a coastal lagoon ecosystem on genetic diversity and structure of white seabream [*Diplodus sargus* (Linnaeus 1758)] populations. Mar Ecol 36: 1144-1154

Huson DH, Bryant D (2006) Applications of phylogenetic networks in evolutionary studies. Mol Biol Evol 23: 254-267

Iacchei M, Ben-Horin T, Selkoe KA, Bird CE, García-Rodríguez FJ, Toonen RJ (2013) Combined analyses of kinship and $F_{ST}$ suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Mol Ecol 22: 3476-3494

Isnard 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. Univ. Montpellier 2. 41p

Isnard E, Tournois J, McKenzie DJ, Ferraton F, Bodin N, Aliaume C, Darnaude AM (2015) Getting a good start in life? A comparative analysis of the quality of lagoons as juvenile habitats for the gilthead seabream *Sparus aurata* in the Gulf of Lions. Estuar Coasts 38: 1937-1950

Johnson MS, Black R (1984) Patterns beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. Evolution 38: 1371-1383

Kearse M, Moir R, Wilson A, Stone-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647-1649

Koehn RK, Newell RI, Immermann F (1980) Maintenance of an aminopeptidase allele frequency cline by natural selection. Proc Natl Acad Sci USA 77: 5385–5389

Laroche JQuiniou L, Juhel G, Auffret M, Moraga D (2002) Genetic and physiological responses of flounder (*Platichthys flesus*) populations to chemical contamination in estuaries. Environm Toxicol Chem 21: 2705-2712.

Larson RJ, Julian RM (1999) Spatial and temporal genetic patchiness in marine populations and their implications for fisheries management. CalCOFI Rep 40: 94-99

Lasserre G (1974) Stock-number, growth, production and migration of giltheads *Sparus auratus* (L. 1758) of group 0+ from étang de thau. Cah Biol Mar 15: 89-111

Lemaire C, Allegrucci G, Naciri M, Bahri-Sfar L, Kara H, Bonhomme F (2000) Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*)? Mol Ecol 9: 457–467

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452

Lottherhos KE, Markel RW (2012) Oceanographic drivers of offspring abundance may increase or decrease reproductive variance in a temperate marine fish. Mol Ecol 21: 5009-5026

Mercier L, Panfili J, Paillon C, N'diaye A, Mouillot D, Darnaude AM (2011) Otolith reading and multi-model inference for improved estimation of age and growth in the gilthead seabream *Sparus aurata* (L.). Estuar Coast Shelf Sci 92: 534-545

Mercier L, Mouillot D, Bruguier O, Vigliola L, Darnaude AM (2012) Multi-element otolith fingerprints unravel sea-lagoon lifetime migrations of gilthead sea bream *Sparus aurata*. Mar Ecol Progr Ser 444: 175–194

Moberg PE, Burton RS (2000) Genetic heterogeneity among adult and recruit red sea urchins, *Strongylocentrotus franciscanus*. Mar Biol 136: 773–784.

Moody KN, Hunter SN, Childress MJ, Blob RW, Schoenfuss HL, Blum MJ, Ptacek MB (2015) Local adaptation despite high gene flow in the waterfall-climbing Hawaiian goby, *Sicyopterus stimpsoni*. Mol Ecol 24: 545-563

Narum SR (2006) Beyond Bonferroni: less conservative analyses for conservation genetics. Conserv Genet 7: 783–787

Pérez-Ruzafa A, Fernández AI, Marcos C, Gilabert J, Quispe JI, García-Charton JA (2005) Spatial and temporal variations of hydrological conditions, nutrients and chlorophyll in a Mediterranean coastal lagoon (Mar Menor, Spain). Hydrobiologia 550: 11–27

Pérez-Ruzafa A, Marcos C, Pérez-Ruzafa IM, Barcala E, Hegazi MI, Quispe J (2007) Detecting changes resulting from human pressure in a naturally quick-changing and heterogeneous environment: spatial and temporal scales of variability in coastal lagoons. Estuar Coast Shelf Sci 75: 175-188

Pérez-Ruzafa A, Marcos C, Pérez-Ruzafa I, Pérez-Marcos M (2011) Coastal lagoons: “transitional ecosystems” between transitional and coastal waters. J Coast Conserv 15: 369-392

Planes S, Lenfant P (2002) Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. Mol Ecol 11: 1515-1524

Planes S, Roman P (2004) Evidence of genetic selection for growth in new recruits of a marine fish. Mol Ecol 13: 2049-2060

Queller D, Goodnight K (1989) Estimating relatedness using genetic markers. Evolution 43: 258-275

Quéré N, Guinand B, Kuhl H, Reinhardt R, Bonhomme F, Desmarais E. 2010. Genomic sequences and genetic differentiation at associated tandem repeat markers in growth hormone, somatolactin and insulin-like growth factor-1 genes of the sea bass, *Dicentrarchus labrax*. Aquat Living Resour 23: 285-296.

Quignard JP, Man-Wai R, Vianet R (1984) Les poissons de l’Etang de Mauguio (Hérault, France) inventaire, structure du peuplement, croissance et polymorphisme des tailles. Vie Milieu 34: 173–183

Schmidt PS, Rand DM (2001) Adaptive maintenance of genetic polymorphism in an intertidal barnacle: habitat- and life-stage-specific survivorship of *Mpi* genotypes. Evolution 55: 1336-1344

Selkoe KA, Gaines SD, Caselle JE, Warner RR (2006) Current shifts and kin aggregation explain genetic patchiness in fish recruits. Ecology 87: 3082–3094

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

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68: 978-989

Streelman JT, Kocher TD (2002) Microsatellite variation associated with prolactin expression and growth of salt-challenged tilapia. Physiol Genom 9:1-4

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595

Therkildsen NO, Hememr-Hansen J, Als TD, Swain DP, Morgan MJ, Trippel EA, Palumbi SR, Meldrup D, Nielsen EE (2013) Microevolution in time and space: SNP analysis of historical DNA reveals dynamic signatures of selection in Atlantic cod. Mol Ecol 22: 2424-2440

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22: 4673-4680

Tournois J, Ferraton F, Velez L, McKenzie DJ, Aliaume C, Mercier L, Darnaude AM (2013) Temporal stability of otolith elemental fingerprints discriminates among lagoon nursery habitats. Estuar Coast Shelf Sci 131: 182-193

van Oosterhout C, Hutchinson WF, Willis DPM, Shipley PF (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4: 535–538

Véliz D, Bourget E, Bernatchez L (2004) Regional variation in the spatial scale of selection at *MPI** and *GPI** in the acorn barnacle *Semibalanus balanoides* (Crustacea). J Evol Biol 17: 953-966

Weir BS (1979) Inference about linkage disequilibrium. Biometrics 31: 235–254

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

Wright S (1951) The genetical structure of populations. Ann Eugen 15: 323–354

**Legends of tables:**

**Table 1:** *Sparus aurata*. Spring and autumn 2011 samples used in this study, from south-west to north-east: SL: Salses-Leucate, T: Thau, and M: Mauguio (Fig. 1). Spring samples are composed of newly recruited 'entering' (E) juveniles, while autumn samples are made of late juveniles 'migrating' (M) out of the lagoons to the open sea. $f$ and $f_{corr}$ represents, 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 and Goodnight 1989) estimated using the Identix computer package (Belkhir et al. 2002) is also reported, together with their associated *p*-values. Corrections for multiple tests were applied when necessary. *N*: sample size.

| | | ***N*** | *f* | *P* | $f_{corr}$ | *P* | $r_{xy}$ | *p* |
|---|---|---|---|---|---|---|---|---|
| Spring | SL-E | 56 | 0.0279 | <0.05 | 0.0115 | - | -0.0184 | 0.354 |
| | TH-E | 55 | 0.1038 | <0.001 | 0.0861 | <0.001 | -0.0207 | 0.394 |
| | MA-E | 54 | 0.0710 | <0.001 | 0.0472 | <0.01 | -0.0198 | 0.140 |
| Autumn | SL-M | 36 | 0.0238 | - | 0.0075 | - | -0.0401 | 0.723 |
| | TH-M | 27 | 0.0487 | <0.01 | 0.0234 | - | -0.0296 | 0.902 |
| | MA-M | 26 | 0.0673 | <0.001 | 0.0219 | - | -0.0776 | 0.845 |

- : not significant

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

| Gene | Length (bp) | No. of individuals | No. of haplotypes | Haplotype diversity (*H*) | Nucleotide diversity ($\pi$) | Tajima's D |
|---|---|---|---|---|---|---|
| *Prl* | 4079 | 24 | 46 | 0.997 | 0.0072 | 0.0707 (ns) |
| *GH* | 3633 | 24 | 47 | 0.999 | 0.0039 | -0.5711 (ns) |

## Legends of figures

**Fig. 1:** *Sparus aurata*. Map of the study area showing the three studied lagoons (Mauguio, Thau and Salses-Leucate) in bold.

**Fig. 2:** *Sparus aurata*. Observed genetic differentiation as measured by $F_{ST}$ at each of the twenty-three loci considered in this study. Loci that showed significant genetic differentiation are indicated by filled circles, while absence of genetic differentiation is indicated by open squares (bars represent one standard deviation). The order of loci along the horizontal axis is the same in each panel; this order corresponds to the one adopted in Supplementary Table S1. Only labels of markers showing significant genetic differentiation are indicated in the figure. Left panel: genetic differentiation calculated among all six samples. Central panel: genetic differentiation among the three samples of young recruits entering the lagoons in spring. Right panel: genetic differentiation among the three samples of individuals migrating out of the lagoons in autumn (Table 1).

**Fig. 3:** *Sparus aurata*. Profiles of nucleotide diversity averaged in 100 base pair sliding windows along *GH* and *Prl* genes in the Mauguio and Thau lagoons. E: exons; I: introns.

**Fig 4:** *Sparus aurata*. NEIGHBORNET networks representing similarities among haplotypes for *GH* and *Prl* genes. Shaded areas indicate groups of related haplotypes (i.e. located further apart in the network) in each gene. Haplotypes found in the Mauguio and Thau migrating juveniles are represented by green and blue dots, respectively.

Fig. 1

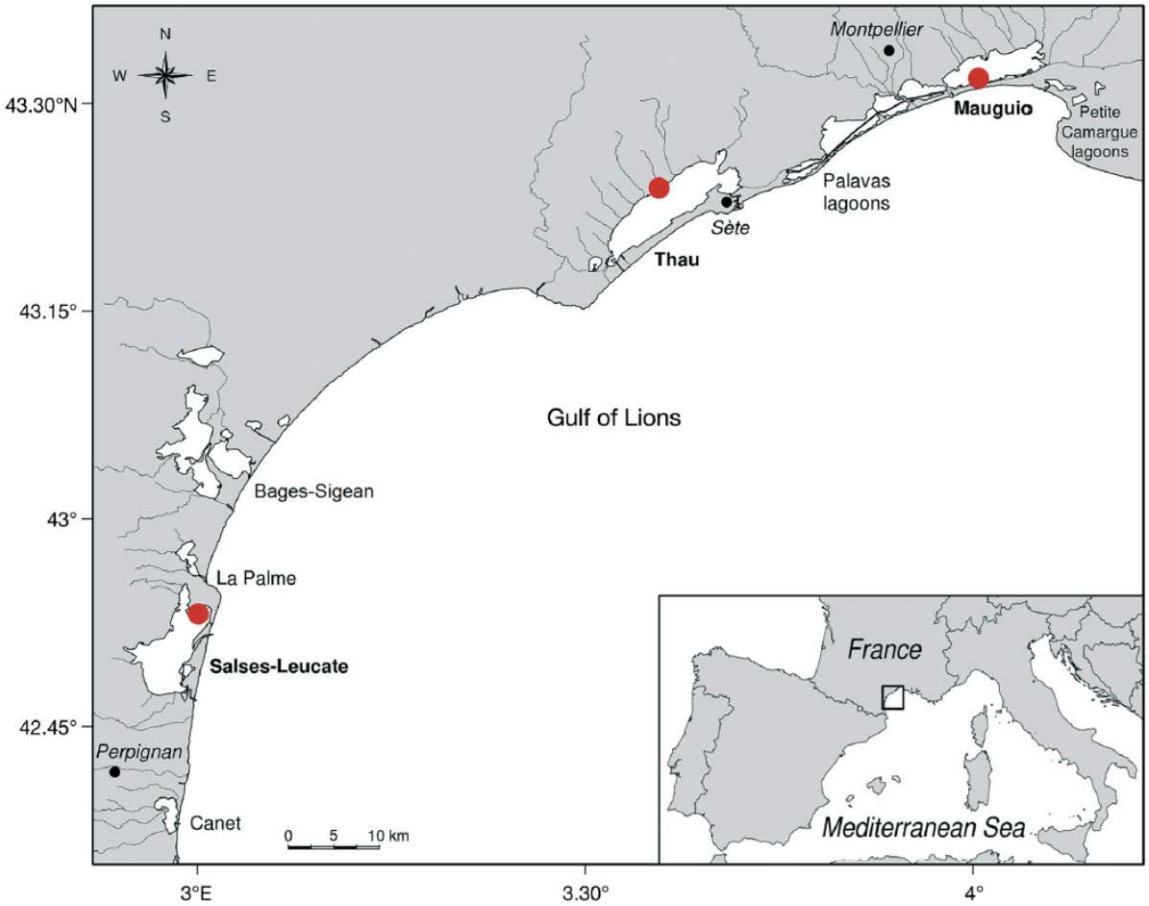

N
W
E
S
43.30°N
43.15°
43°
42.45°
3°E
3.30°
4°
Montpellier
Mauguio
Petite Camargue lagoons
Palavas lagoons
Sète
Thau
Gulf of Lions
Bages-Sigean
La Palme
Salses-Leucate
Perpignan
Canet
0 5 10 km
France
Mediterranean Sea

Fig. 2

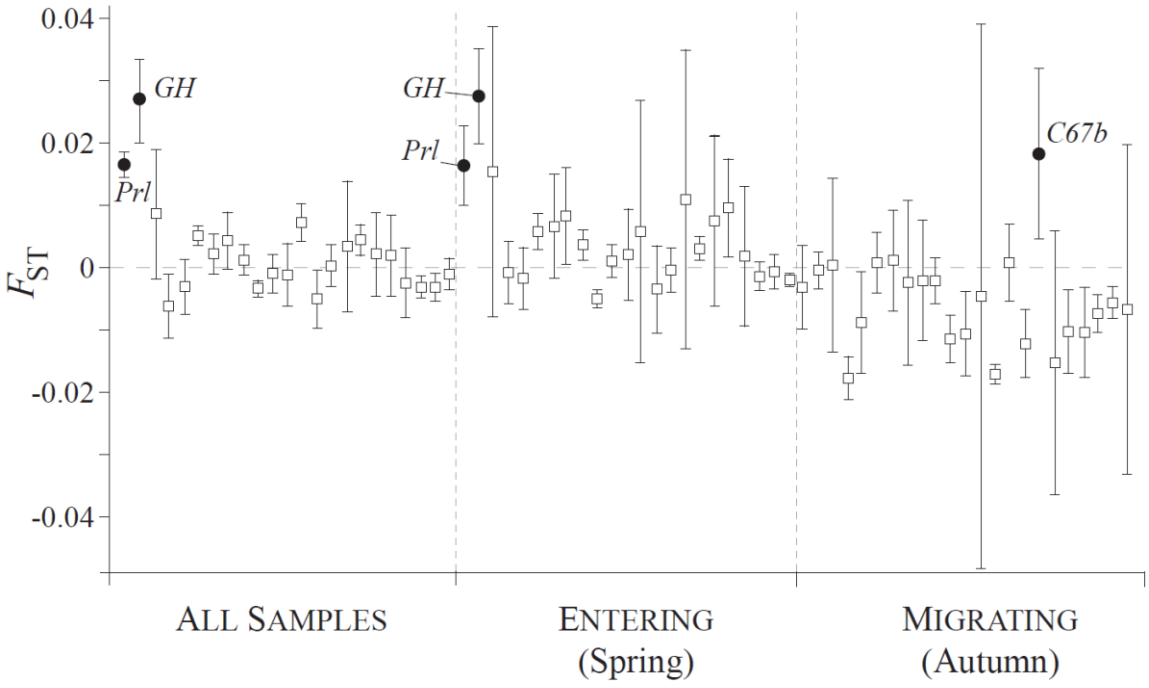
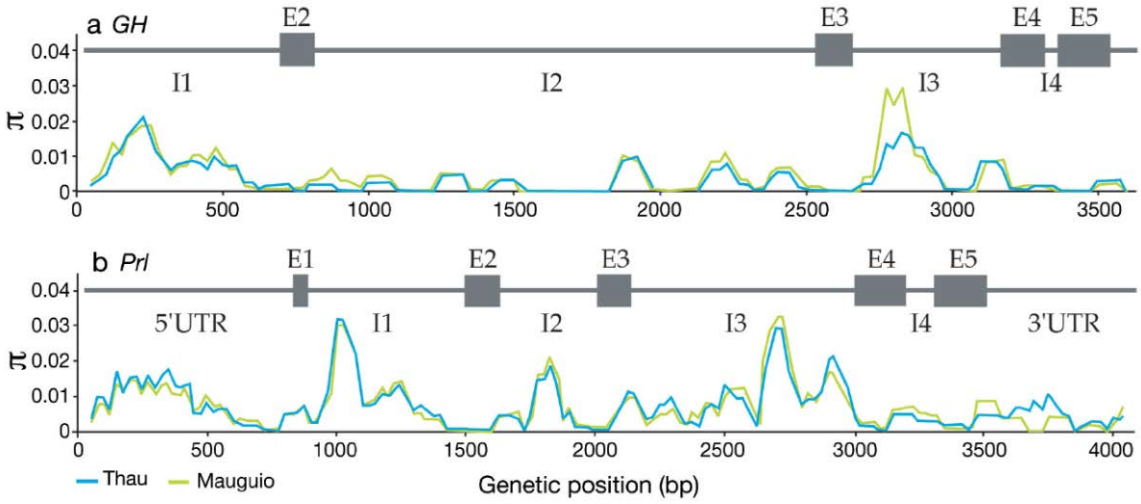

0.04
0.02
0
-0.02
-0.04
$F_{ST}$
GH
Prl
GH
Prl
C67b
ALL SAMPLES
ENTERING
(Spring)
MIGRATING
(Autumn)

Fig.


a GH
E2
E3
E4
E5
I1
I2
I3
I4
π
0.04
0.03
0.02
0.01
0
0
500
1000
1500
2000
2500
3000
3500
b Prl
E1
E2
E3
E4
E5
5'UTR
I1
I2
I3
I4
3'UTR
4000
Thau
Mauguio
Genetic position (bp)

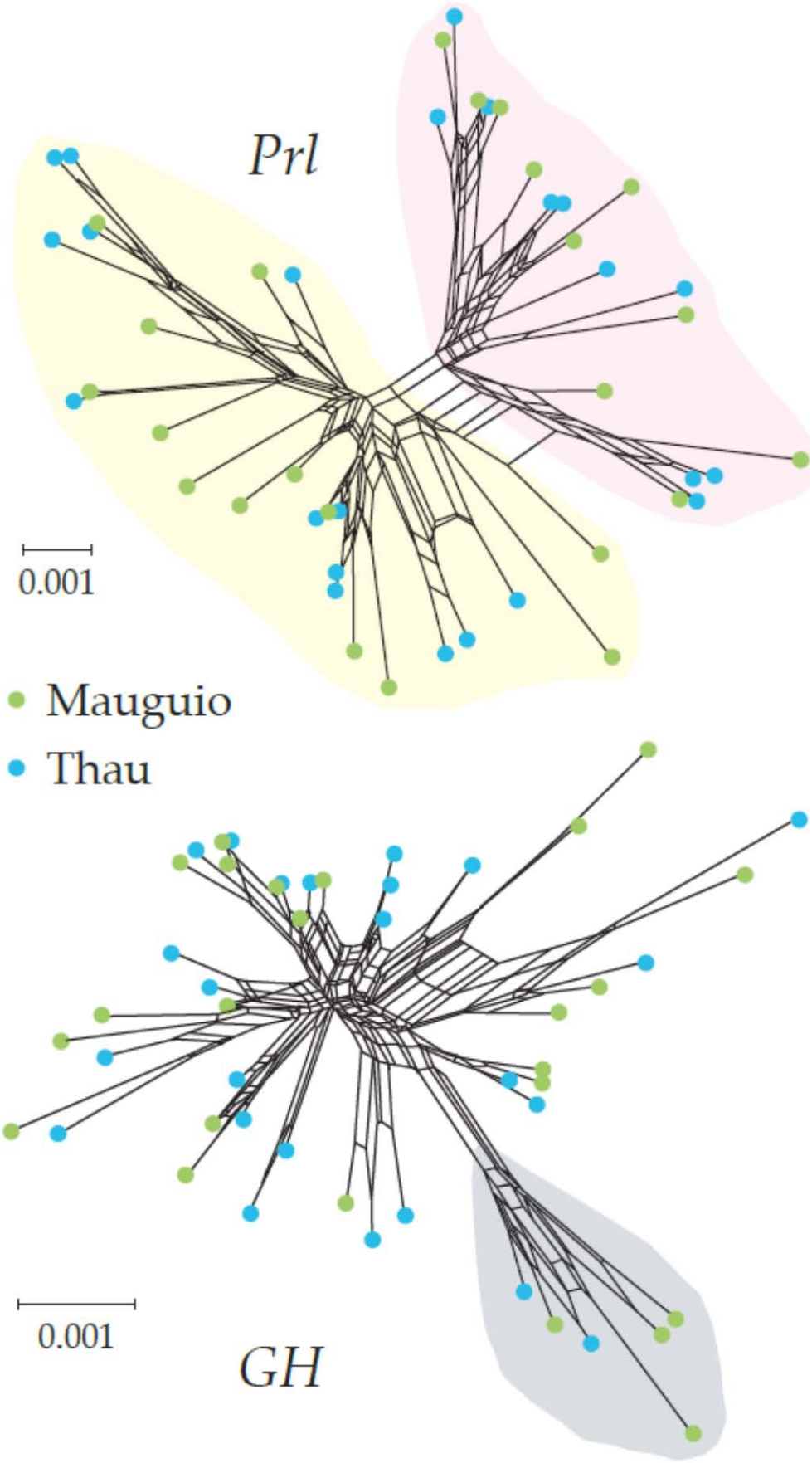



Fig. 4