Characterization of the European sea bass (*Dicentrarchus labrax*) gonadal transcriptome during sexual development

Ribas L¹⁺, Crespo B²,³⁺, Sánchez-Baizán N¹, Xavier D⁴, Kuhl H⁵,⁶, Rodríguez JM⁷,⁸, Díaz N¹,⁹, Boltañá S¹⁰,¹¹, MacKenzie S¹⁰,¹², Morán F⁴, Zanuy S², Gómez A²*, Piferrer F¹*  

¹ Institute of Marine Sciences (ICM), Spanish National Research Council (CSIC), Barcelona, Spain  
² Institute of Aquaculture of Torre de la Sal (IATS-CSIC), Castellón, Spain  
³ Current address: UCL GOS Institute of Child Health, University College London, London, UK  
⁴ Dept. of Biochemistry and Molecular Biology I, Complutense University, Madrid, Spain  
⁵ Max Planck Institute for Molecular Genetics, Berlin, Germany  
⁶ Current address: Department of Ecophysiology and Aquaculture, Leibniz Institute for Freshwater Ecology and Inland Fisheries, Berlin, Germany.  
⁷ Spanish National Bioinformatics Institute, Madrid, Spain  
⁸ Current address: Cardiovascular Proteomics Laboratory, Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain  
⁹ Current address: Max Planck Institute for Molecular Biomedicine, Muenster, Germany  
¹⁰ Autonomous University of Barcelona, Barcelona, Spain.  
¹¹ Current address: Interdisciplinary Center for Aquaculture Research (INCAR), Dept. of Oceanography, Biotechnology Center, University of Concepción, Concepción, Chile  
¹² Current address: Institute of Aquaculture, University of Stirling, Scotland, UK  

⁺These authors contributed equally  
*Correspondence: Dr. Francesc Piferrer, Institute of Marine Sciences, Spanish Council for Scientific Research (CSIC), Barcelona, Spain. Tel. +34-932309567; fax, +34-932309555; e-mail: piferrer@icm.csic.es. Dr. Ana Gómez. Institute of Aquaculture Torre la Sal (CSIC). Ribera de Cabanes s/n. 12595 Torre la Sal. Castellón. Spain. Tel. +34 964 319500; fax, +34 964 319509; e-mail: a.gomez@csic.es
Abstract

The European sea bass is one of the most important cultured fish in Europe and has a marked sexual growth dimorphism in favor of females. It is a gonochoristic species with polygenic sex determination, where a combination between still undifferentiated genetic factors and environmental temperature determine sex ratios. The molecular mechanisms responsible for gonadal sex differentiation are still unknown. Here, we sampled fish during the gonadal developmental period (110 to 350 days post fertilization, dpf), and performed a comprehensive transcriptomic study by using a species-specific microarray. This analysis uncovered sex-specific gonadal transcriptomic profiles at each stage of development, identifying larger number of differentially expressed genes in ovaries when compared to testis. The expression patterns of 54 reproduction-related genes were analyzed. We found that *hsd17β10* is a reliable marker of early ovarian differentiation. Further, three genes, *pdgfb, snx1* and *nfy*, not previously related to fish sex differentiation, were tightly associated with testis development in the sea bass. Regarding signaling pathways, lysine degradation, bladder cancer and NOD-like receptor signaling were enriched for ovarian development while eight pathways including basal transcription factors and steroid biosynthesis were enriched for testis development. Analysis of the transcription factor abundance showed an earlier increase in females than in males. Our results show that, although many players in the sex differentiation pathways are conserved among species, there are peculiarities in gene expression worth exploring. The genes identified in this study illustrate the diversity of players involved in fish sex differentiation and can become potential biomarkers for the management of sex ratios in the European sea bass and perhaps other cultured species.

Keywords: genomics, transcriptomics, reproduction, sex differentiation, gonads, aquaculture
Introduction

Many fish species exhibit sexual dimorphic growth where one sex, in many cases the females, grow more than the other. Mixed sex rearing constitutes a problem for the aquaculture industry because it means producing at suboptimal capacity when compared to monosex culture. In addition, farming conditions results in many cases in masculinization, which further aggravates the problem if the desired sex is the females. Thus, the control of sex ratios is of major importance for many farmed species (Piferrer 2001; Budd et al. 2015; Wang et al. 2019). Deciphering the molecular mechanisms involved in gonadal development and the establishment of population sex ratios has then not only basic interest but also is of practical importance for modern fish farming.

Transcriptomic studies on gonadal tissues have been conducted in a relatively large number of fish species, both cultured and not. For example, in the channel catfish (Ictalurus punctatus), the genes differentially expressed (DEG) were described during testis development (Zeng et al. 2016). In fugu (Takifugu rubripes) and in the spotted knifejaw (Oplegnathus punctatus), a larger number of DEG were found in adult testes when compared to ovaries (Du et al. 2017; Wang et al. 2017). However, the number of studies focusing on sex-related differences precisely during sexual development and with a genomics approach are much more limited, and include those carried out in Nile tilapia (Oreochromis niloticus) (Tao et al. 2013), turbot (Scophthalmus maximus) (Ribas et al. 2016), zebrafish (Danio rerio) (Ribas et al. 2017) and yellow river carp (Cyprinus carpio) (Jia et al. 2018).

The European sea bass (Dicentrarchus labrax) stands as the third most important cultured species of marine fish in Europe with a production of ~180 thousand tons per year (Food and Agriculture Organization of the United Nations 2016). Its production has benefited from advances in reproduction and the implementation of breeding programs (reviewed in Felip and Piferrer, 2018; Wang et al. 2019). In this species, females grow about 30% more than males (Saillant et al. 2001). However, temperature during early development can affect sex ratios, favoring a higher number of males in the populations (reviewed in Vandeputte and Piferrer, 2019). This problem can be even more aggravated if males undergo precocious maturation as it slows down their growth (Carrillo et al. 2015). Thus, there is interest in producing monosex female stocks.
The European sea bass is one of the richest species in terms of genomic resources among cultured fish, which include the availability of the genome and Single Nucleotide Polymorphism (SNP) markers, among others. The European sea bass has a small genome size (675 Mb) with a total of 26,719 annotated genes (Tine et al. 2014; Chaves-Pozo et al. 2017) and has 24 haploid chromosome pairs (Aref’yev 1989). It is a gonochoristic species with a polygenic sex determination system with both genetic and environmental influences (Piferrer et al. 2005; Vandeputte et al. 2007; Vandeputte and Piferrer 2019). Although the genetic factors are still not known, efforts have been done towards identifying sex-determinant markers to aid in genetic selection programs. Studies using ~6,700 SNP markers showed that there are at least three loci linked to sex (Palaiokostas et al. 2015) but these may be family-specific. Studies on the effects of food supply during early development evidenced changes in energy balance during testicular development (Díaz et al. 2014). In a similar manner, transcriptomic profiles of differentiating gonads subjected to different temperatures or estrogen treatments identified the involvement of steroidogenic- and epigenetic-related genes (Díaz and Piferrer 2015; Díaz and Piferrer 2017). Recently, the epigenetic regulation of key sex-related genes has been reported and a method based on the analysis of epimutations to predict sex in the European sea bass devised (Anastasiadi et al. 2018). Nevertheless, a study of the transcriptomic changes that occur during sexual development had never been carried out.

The objective of this study was to fill this gap. To do so, we transcriptomically analyzed sexually undifferentiated, differentiating and differentiated gonads. Gene expression levels were evaluated by using a homologous custom-made microarray enriched with reproduction-related genes. Studying transcriptomes by using a microarray platform provides accuracy and reproducibility of the performed analysis and allows studying a broad range, if not all, of the transcripts of the genome (Shi et al. 2006). In particular, our custom microarray contained 78.5% of the annotated coding genes of the European sea bass genome, thus providing a powerful molecular tool to study gene expression patterns of this species.
Materials and Methods

European sea bass gonad sampling

In order to obtain the widest possible range of expressed transcripts, gonads were dissected from fish (5.2 ± 0.5 cm of standard length, SD) at 110 days post fertilization (dpf), when they are still morphologically undifferentiated but can be sexed by analyzing the expression of sex-markers (i.e., cyp19a1a (Blázquez et al. 2008), when differentiating at 250 dpf (12.7 ± 5.7 cm SD and 11.2 ± 0.6 cm SD in females and males, respectively) and when differentiated at 350 dpf (16 ± 1.3 cm SD and 14.8 ± 1.1 cm SD, in females and males, respectively). The set of samples consisted of eleven gonads at 110 dpf, twelve gonads (six testes and six ovaries) at 250 dpf and eighteen gonads (nine testes and nine ovaries) at 350 dpf. When possible, gonads were fully isolated from fish at 250 and 350 dpf and thus gonadal tissue was devoid of any other tissue. However, a clean gonad isolation was not always feasible in younger fish (110 dpf) due to their extremely small size, and thus some epithelial contamination could not be ruled out.

Microarray platform

The microarray platform used in the present study consists of 1,417 Agilent control probes and a total of 43,803 transcript probes that represent 20,978 genes of which 20,028 have two probe copies each while the rest have between 1 to 6 copies per gene (Supplementary Table S1). Genes with known reproduction-related functions had at least four copies. Microarray was based on sequences obtained from two 454 FLX Titanium runs on European sea bass gonad tissues at different ages (from 40 dpf up to 6 years), a former custom European sea bass microarray platform (GPL13443) available in our laboratory (Díaz et al. 2014; Díaz and Piferrer 2015), a previously published European sea bass microarray (Ferraresso et al. 2010) and a battery of selected reproduction canonical genes. Only non-redundant and annotated sequences were selected. This microarray was submitted to Gene Expression Omnibus (GEO) database (Edgar et al. 2002) with the platform number GPL16767 and its functionality was reported in a previous work using European sea bass larvae (Schaeck et al. 2017). In the present work, this microarray has been re-annotated by using the European sea bass genome (Tine et al. 2014) and used to study the gonadal development in the European sea bass. Microarray data of the present study were submitted to GEO and are
accessible through GEO Series accession number GSE115841. For a complete list of
gene names and abbreviations, see Dataset 1.

RNA isolation and microarray hybridization

Total RNAs was extracted from 41 gonad samples using RNeasy Mini Kit (Qiagen, Germany) following the manufacturer's instructions. Quantity was determined by a Nanodrop spectrophotometer (Nanodrop Technologies, USA) and quality (RNA integrity number, RIN) measured with a Bioanalyzer (Agilent Technologies, USA). Only RNA samples with a mean RIN ≥ 8.4 were further processed for microarray analysis. RNA was labelled using the Low Input Quick Amp Labelling Kit, One-Color (Cy3; Agilent Technologies) and cRNA was hybridized overnight with the corresponding buffers during 17 h at 65°C and washed on the following day. Samples were hybridized individually in the European sea bass custom 4x44K Agilent microarray described above at the Parc d’Investigació Biomédica de Barcelona (PRBB) and slides were scanned using an Agilent G2565B microarray scanner (Agilent Technologies, USA). Agilent software was used to avoid saturation and the extraction feature generated the raw data for further pre-processing.

Data analysis

Statistical analyses were performed with R software (2.13.1 version; www.R-project.org). Array normalization was implemented using the Quantile method in the Linear Models for Microarray Analysis (Limma) R package (Wettenhall and Smyth 2004; Ritchie et al. 2015). Potential batch effects were removed by ComBat correction, a bioinformatic tool based on Empirical Bayes algorithms (Chen et al. 2011). Data visualization, Principal Component Analysis of the variance and identification of clusters and outliers (two samples at 350 dpf were detected as outliers and excluded from further analysis) were performed using R. TIGR Multiexperiment Viewer version 4.9 (TMeV) software (Saeed et al. 2003) was used to determine the number of differentially expressed genes (DEG) between sexes at a given stage of development or between stages of development within the same sex. Significance was assessed by Significant Analysis of Microarrays (SAM) statistical test with a False Discovery Rate (FDR) and adjusted $P$ values < 0.01 and < 0.001 were applied to identify genes with statistically significant differences in expression. The above-mentioned analysis generated lists of DEG at each stage (110, 250 and 350 dpf) in the same sex or in
comparison to the other sex, including the log\textsubscript{2} transformation of fluorescence intensity measured for each gene.

\textbf{Gene ontology terms and KEGG pathway analysis}

The over-represented gene ontology (GO) functional categories of the DEG between females and males at each stage were obtained by GO-terms enrichment analysis using GO.db and topGO packages from the Bioconductor Project (Gentleman et al. 2004; Alexa and Rahnenfuhrer 2016; Carlson 2017) in R software (R Core 2017). The graphs and heatmaps were produced using gplots and ggplot2 packages (Wickham 2009; Warnes et al. 2016).

We used the Gene Set Variation Analysis (GSVA) from Pathway Processor 2.0 to study the signalling pathways involved in gonadal development. GSVA transforms the gene expression values into a normalized expression matrix with enrichment scores of differentially regulated pathways (DRP) with the corrected \(P\) value between males and females at each developmental stage (Beltrame et al. 2013).

Four pathways involved in sex differentiation: fanconi anemia and wnt signaling pathways, associated with female differentiation (Rodríguez-Marí and Postlethwait 2011; Sreenivasan et al. 2014), p53 signaling and cytokine-cytokine interaction receptor pathways, associated with male differentiation (Yasuda et al. 2012; Ribas et al. 2017) were also studied. The lists of genes making-up these pathways were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG), using zebrafish as background. The numbers of DEG from these pathways as well as expression values (in log\textsubscript{2} Fold Change, FC) were plotted together; upregulated in male \textit{vs.} upregulated in female over time.

\textbf{Transcription factors analysis}

The transcription factors (TFs) present in the microarray were identified with the aid of the TF checkpoint database, a list of TFs compiled from nine databases (Chawla et al. 2013). Using this list as background, the percentage of TFs present in the DEG between sexes and differentially expressed at each age was calculated.

\textbf{Validation of the microarray}
Microarray results were validated by quantitative real time polymerase chain reaction (qPCR) analyzing the expression of twelve genes selected with a wide range of FC values and equal amount of upregulated and downregulated genes when ovaries and testes were compared. Two house-keeping genes were chosen as reference: Elongation factor-1 alpha (ef-1α) and 40S ribosomal protein (f40) that were previously validated in the European sea bass (Mitter et al. 2009). One hundred nanograms of total RNA were reverse transcribed into cDNA using Superscript III (Invitrogen) and 100 ng of random hexamer primers (Sigma) following the manufacturer’s instructions. The reaction was carried out with SYBR Green chemistry (Power SYBR Green PCR Master Mix; Applied Biosystems). qPCR reactions contained 1X SYBR green master mix (Applied Biosystems), 10 pmol of each primer and 1 µl of the RT reaction. Samples were run individually and in triplicate in optically clear 384-well plates in Applied Biosystems 7900 machine. Cycling parameters were: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Finally, a temperature-determining dissociation step was performed at 95°C for 15 s, 60°C for 15 s and 95°C for 15 s at the end of the amplification phase. qPCR data were collected by SDS 2.3 and RQ Manager 1.2 software and relative quantity (RQ) values for each reaction replicate were calculated by the 2ΔΔCT method (Schmittgen and Livak 2008). Primer sequences used for gene expression study are shown in Supplementary Table S2.

**Results**

**Microarray platform and validation**

Hybridization repeatability and consistency of results was verified in 55 genes related to reproduction and previously selected in turbot (Ribas et al. 2016) and zebrafish (Ribas et al. 2017) to study sex differentiation and reproduction in fish (Supplementary Table S3). Probe copy tendency for 52 of these genes were the same for all the copies and only three of them (representing ~5% of the 52 tested probes) showed different probe tendency in at least one of the copies. Some examples of upregulated (Supplementary Fig. S1 a, c, e, g), downregulated (Supplementary Fig. S1 b, d, f, h) or variable (i) gene expression are shown. Since most probes had two or four copies (Supplementary Fig. 1j) to further evaluate the hybridization accuracy, we determined the magnitude of variation between technical replicates. The mean of the standard deviations for all 54 probes ranged between 0.205 and 0.347 (Supplementary Fig. 1j). Thus, given the low
standard error among probe copies of the same gene, the average FC value of all probe copies was used for each gene.

Microarray validation by qPCR for 12 DEGs showed a good correlation between the results obtained either by microarray and qPCR techniques ($R^2 = 0.748, P = 0.0003$; Supplementary Fig. S2a). Additionally, gene expression values of aromatase (cyp19a1a) from samples at different stages of gonadal development (110, 250 and 350 dpf) determined either by using microarray and by qPCR, further validated our results (Supplementary Fig. S2b).

Gonadal transcriptomes overview

Samples clustered in two distinctive groups according to sex and within each group samples tended to group according to age, except for males at 350 dpf, which showed more variation in the PCA (Fig. 1a). The component 1 of the PCA alone explained 87.14% of the variance, while components 2 and 3 contributed to 8.35% and 1.98% of the total variance, respectively. Thus, the first three components together explained 97.47% of variance.

A total of seven comparisons were analyzed with the SAM statistical test with an adjusted $P$ value ≤ 0.01: three between sexes, two within males and two within females at the three different ages (Fig. 1b). Among the 20,978 genes included in the microarray, 64.93% were differentially expressed at one or several of these comparisons. The number of DEGs between ovaries and testes was 708; 7,639 and 6,926, at 110, 250 and 350 dpf, respectively (Fig. 1b, Dataset 1). A larger number of genes were upregulated in females when compared to M: 685 vs. 23 at 110 dpf; 3,870 vs. 3,769 at 250 dpf, and 6,097 vs. 829 at 350 dpf. Between 110-250 and 250-350 dpf, the number of upregulated genes was 3,564 and 1,100 in females while 451 and 309 in males, respectively. The number of downregulated genes between 110-250 and 250-350 dpf was 3,060 and 671 in females while 1,737 and 1,179 in males, respectively. Among the upregulated genes in the developing ovaries, higher FC values were found in 110 dpf and 350 dpf while at 250 dpf FC values were higher in testes when compared to ovaries (Supplementary Table S4). The highest FC values were observed at 250 dpf ($P < 0.001$) which was 7.5 and 6.7 Log2 in males and females, respectively (Dataset 1),
indicating important sex-related differences in expression levels (Supplementary Table S4).

Gene ontology and gene pathway enrichments along gonadal development

The GO term enrichment analysis of the DEGs between F and M revealed several categories related to biological processes (BP), molecular function (MF) and cellular component (CC) throughout development (Dataset 2). A total of 39 GO terms in the three categories were enriched during ovarian formation. The 15 GO terms significantly enriched in BP common at 110, 250 and 350 dpf ($P < 0.01$) are shown in Fig. 2a. The three developmental stages were enriched in GO terms related to metabolic processes (GO:0008152, GO:0071704, GO:0044237), catalytic activity (GO:0003824), oxidoreductase activity (GO:0016491), coenzyme and cofactor binding (GO:0050662; GO:0048037), and biosynthetic processes (GO:0009058, GO:1901576). Among the GO terms enriched for testis formation, there were a total of 52 significantly enriched terms that were common at 250 and 350 dpf, but none at 110 dpf which were related to catabolic processes (GO:0000956, GO:0006402, GO:0006401), regulation of ion transmembrane activity (GO:1904427, GO:0032414, GO:0034767), regulation of calcium ion (GO:0010524, GO:0050850, GO:0051281, GO:0060316, GO:1901021, GO:1904427) and positive regulation of growth (GO:0045927). Fig. 2b shows the enriched GO terms found for testis formation in BP category ($P < 0.02$).

Then, we determined the significantly Differentially Regulated Cellular Pathways (DRP) between males and females along gonadal development. A total of 41, 151 and 106 DRP were found between males and females at 110, 250 and 350, respectively (Dataset 3, $P < 0.05$). Some of these DRP were related to sex differentiation, for example, at early gonadal development (110 dpf): p53 signaling pathway, steroid hormone biosynthesis or erbβ signaling pathway; at 250 dpf: wnt signaling pathway, oocyte meiosis or steroid biosynthesis and at 350 dpf: MAPK signaling pathway or cytokine-cytokine receptor interaction among others. A total of 16 DPR were consistently differentially regulated at the three developmental stages in both sexes (Table 1). The lysine degradation, bladder cancer and the nucleotide-binding oligomerization (NOD)-like receptor signaling pathways were upregulated in females at the three gonadal developmental stages when compared to males. The dorso-ventral axis formation pathway was significantly downregulated in females at 110 dpf but
Expression of canonical genes and pathways related to sex differentiation

Of the 54 canonical genes known to be relevant for reproduction and sex differentiation in fish according to the primary literature, 49 of them had sex-related significant differences in at least one of the three ages studied. The majority (80.5%) were DEG at 250 dpf (Supplementary Table S3). Of the 49 just mentioned above, 25 were pro-female and 24 pro-male genes. Hierarchical clustering analysis and the corresponding heatmaps of the 25 pro-female genes (Fig. 3a) and the 24 pro-male genes (Fig. 3b) showed that gene expression results mostly matched according to their phenotypic gender. The expression profiles of twelve key genes is shown in Fig. 4. Six of them are related to the steroidogenic pathway (cyp19a1a, hsd17β10, hsd3β, cyp11β, ara, fshr, Fig. 4a-f) while the other six genes are TFs related to sex differentiation (foxl2, sox3, figlα, nr5a1a, sox9b and dmrt1, Fig. 4g-l). The genes cyp11β1, ara, hsd3β and fshr were upregulated in males when compared to females at 250 dpf and onwards while hsd17β10 was upregulated in females already at 110 dpf. In all these genes, sex-specific significant differences in expression were observed at least in one of the three sampling ages. All genes except hsd3β had maximal sex-related expression differences at 250 dpf. Regarding the six canonical TFs, they were upregulated as expected according to sex: foxl2, sox3 and figlα in females while nr5a1a, sox9b and dmrt1 in males. All of them were differentially expressed between sexes at least in 250 dpf and some also at 350 dpf (sox3, figlα, sf1α and sox9b).

Next, we looked specifically at the four signaling pathways known to be associated with sex differentiation from previous studies (see Materials and methods). Among genes upregulated at 250 and 350 dpf. The p53 signaling pathway and the Chagas disease (American trypanosomiasis) pathways showed significance, being upregulated at 110 and 350 dpf in females but downregulated at 250 dpf when compared to males. There were two pathways, phosphatidylinositol signaling system and the ErbB signaling pathway that were upregulated in females at 110 dpf but at 250 and 350 dpf in males. Finally, there were up to eight pathways upregulated in males when compared to females throughout the studied period: butirosin and neomycin biosynthesis, basal transcription factors, amino sugar and nucleotide sugar metabolism, type II diabetes mellitus, glycine, serine and threonine metabolism, steroid biosynthesis and ribosome and folate biosynthesis pathways.
that constitute these signaling pathways, we looked at the number of DEG and the magnitude of the gene expression values. The two selected pathways related to ovarian development, fanconi anemia (Fig. 5a, b) and wnt (Fig. 5c, d) signaling pathways, had a larger number of DEG and a higher gene expression (FC) values in females. Similarly, when looking among pathways related to testis development, p53 signaling pathway (Fig. 5e, f) and cytokine-cytokine interaction receptor pathway (Fig. 5g, h), they had a larger number of DEG and a higher gene expression (FC) values in males.

The role of transcription factors during gonadal development

Our microarray included 2,822 TFs in total, i.e., 13.5% of all the probes. The proportion of differentially expressed TFs was variable between sexes and across time (Fig. 6). At 110 dpf 8.0% of the DEG were upregulated TFs in ovaries when compared to testes, while no TFs were upregulated in testes. At 250 dpf, the percentage of differentially upregulated TFs increased up to 13.0% in ovaries and 9.6% in testis. Up to 25.6% of the DEG at 350 dpf were identified as TFs; 14.5% upregulated in ovaries and 11.1% were upregulated in testis.

Discussion

Robustness of the microarray

This study provides a comprehensive transcriptomic analysis of gonad differentiation in the European sea bass using a custom species-specific microarray (Schaeck et al. 2017) that has been here further validated. First, the microarray was completely re-annotated and includes almost 80% of the genes identified in the European sea bass genome. Quality control showed that it had high reproducibility and accuracy. Transcript expression values were very robust as the standard deviations of probe replicates was very low (average 0.276 for 20,029 the duplicated probes), confirming the high reproducibility of RNA analysis using the Agilent oligo-array (Shi et al. 2006). To date, microarray analyses have been very useful in the study of fish transcriptomes, e.g., (Millan et al. 2010; Jantzen et al. 2011; Tingaud-Sequeira et al. 2013; Schaeck et al. 2017). Although in the last years RNA sequencing have gained favor over array platforms, analysis of the same samples with the two different techniques gives similar results (Zhao et al. 2014). Here we provide a validated, fast and cost-effective tool for
aquaculture research to study the expression patterns of genes, including all major reproduction-related genes, in the European sea bass (Schaeck et al. 2017).

Transcriptomic differences between females and males during gonadal development

PCA classified individuals in well-defined and separated clusters according to sex and stage. At 110 dpf, when gonads were still histologically undifferentiated, transcriptomic analysis was already capable of classifying samples according to phenotypic sex.

Statistical analysis showed that the highest number of DEG were found in differentiated females when compared to differentiated males, in particular at 250 dpf, probably explained by the fact that sex differentiation in the European sea bass starts earlier in females (Piferrer et al. 2005) as in many other fish species (Piferrer 2001; Devlin and Nagahama 2002; Wang et al., 2019). This contrasts with results found in other fish species such as zebrafish (Small et al. 2009), tilapia (Tao et al. 2013) or turbot (Ribas et al. 2016), where male-related genes were enriched with respect to female-related genes.

Testis development implied downregulation of genes in a certain stage when compared to the previous developmental stages. This tendency was also observed in turbot (Ribas et al. 2016), supporting the importance of active gene repression for testis development.

This is in accordance with the current view stating that positive and negative regulatory loops are required for sex differentiation in vertebrates (Munger et al. 2013; Capel 2017).

There were 15 enriched GO terms in upregulated genes during ovarian development (from 110 dpf to 350 dpf), all of them related to metabolic functions. In this regard, the metabolic process category was the most enriched one found in the differentiating ovary of the protogynous ricefield eel (*Monopterus albus*) (Cai et al. 2017) and was described as well for ovarian development in turbot (Ribas et al. 2016). In contrast, we did not find any common GO term in the three developmental stages studied in testis.

Nevertheless, 30 common GO terms were found enriched between 250 and 350 dpf in developing males and were related to catabolic processes, regulation and positive regulation of growth, among others.

Three pathways (i.e., lysine degradation, bladder cancer and NOD-like receptor signaling) showed a consistent upregulation in the ovaries in all stages when compared to testes. Lysine is an essential amino acid required for protein synthesis, enzyme...
catalysis and L-carnitine biosynthesis and thus essential for energy metabolism in all body tissues, including ovaries (Ramseyer and Garling 1994; Hallen et al. 2013). The bladder cancer pathway includes genes involved in gonadal development such as the tumor protein tp53 (Mitra et al. 2006), a gene present in germ cells and that induces apoptosis and atresia in oocytes (Rodriguez-Mari et al. 2010; Sayed et al. 2018). The NOD-like receptor signaling pathways is activated in response to host defense and inflammatory disease response (Caruso et al. 2014) and it is upregulated in human polycystic ovaries (Wang et al. 2014). On the other hand, during testicular development eight pathways were identified, including pathways related to metabolism (e.g., butirosin and neomycin biosynthesis, amino sugar and nucleotide sugar metabolism), to genetic information processing (e.g., basal transcription factors, type II diabetes mellitus ribosome) and to steroid biosynthesis. A sex-biased representation of these molecular pathways might also be species-specific. Thus, for example, in Japanese flounder (Paralichthys olivaceus) the upregulation of metabolic-related pathways was found in ovaries rather than in testes (Fan et al. 2014). Also, in the tilapia, steroidogenic pathways were more expressed in females than in males, particularly at early stages of development (Tao et al. 2013).

Sex-biased expression of canonical genes during sex differentiation

We selected 54 pro-female or pro-male genes to study their expression along gonadal development. We found 49 DEGs whose expression matched the expected sex bias described in previous studies of other fish species (see Materials and methods). However, six of the genes upregulated in male European sea bass had been previously described as upregulated in female zebrafish (Ribas et al. 2017), namely: *hsd3β*, *cyp19b1*, *tradd*, *er1*, *fshr* and *er2b*.

In females, the steroidogenic gene *hsd17β10* was upregulated at 110 dpf. The expression of *hsd17β*, *cyp19a1a* and *star* was downregulated in European sea bass at 170 dpf in fish previously exposed to high temperature, due to the masculinization of the ovary (Díaz and Piferrer 2017). The *cyp19a1a* gene, a key enzyme responsible for converting androgens into estrogens (Guiguen et al. 2009) is considered an early marker of ovarian differentiation in several fish species, including sea bass (Blázquez et al. 2008) and also Atlantic halibut (*Hippoglossus hippoglossus*) (Matsuoka et al. 2006) and turbot (Ribas et al. 2016), among others. In the present study, *cyp19a1a* showed...
differentially expression at 110 dpf but differences were not significant until 250 dpf. 

*Hsd17β10* is a mitochondrial enzyme involved in multiple cellular functions, which 
include fatty acid oxidation, amino acid degradation and steroid metabolism (Yang et al. 
2007; Zschocke 2012). In humans, *hsd17β10* is related to neurodegenerative diseases 
such as Parkinson or Alzheimer and has been fully documented (Zschocke 2012; Yang 
et al. 2014). In fish, there is a lack of information about this isoform as it has been only 
described in amphioxus (*Branchiostoma belcheri*) (Zhang et al. 2008) and zebrafish (He 
and Yang 2009) but no biological functions have been yet ascribed. However, 
information does exist for other genes of the same family such as *hsd17β1* and *hsd17β3*, 
which are involved in sex steroid biosynthesis: *hsd17β1* is responsible to convert 
inactive estrone to active estradiol and leads to female sex differentiation while *hsd17β3* 
is required for 11-ketotestosterone synthesis (Tokarz et al. 2015). *Hsd17β1* has been 
identified in some fish species such as Nile tilapia (Zhou et al. 2005), Atlantic cod 
(*Gadus morhua*) (Breton and Berlinsky 2014) and olive flounder (Fan et al. 2014) while 
*hsd17β3* in zebrafish and in medaka (*Oryzias latipes*) only (Mindnich et al. 2004; Kim 
et al. 2014). *Hsd17β1* was already detected at early stages of development in pre-
differentiated fathead minnow (*Pimephales promelas*) embryos, although its expression 
was not correlated to any sex in particular (Wood et al. 2015). Recently, *hsd17β1* has 
been suggested as the sex determining gene in the California yellowtail (*Seriola 
dorsalis*), which a putative ZW sex determination system (Purcell et al. 2018). In our 
data, we did not find any differential expression in *hsd17β1* and *hsd17β3* genes but we 
did it in *hsd17β10*. To our knowledge, this is the first time that the *hsd17β10* is 
described in European sea bass and it is detected early in the ovaries. Therefore, it is a 
candidate to be considered as a novel early ovarian marker in this species, although 
research on its functional role during ovarian differentiation needs further attention.

In males, the first signs of sex-biased expression of canonical reproduction-related 
genes were detected at 250 dpf onwards and not earlier, indicating that their expression 
starts somewhere between 110 to 250 dpf as previously showed in fish subjected to high 
temperature treatments at 170 dpf (Díaz and Piferrer 2015). In the present study, genes 
involved in the steroidogenic pathway such as *hsd3β* or *cyp11β1*, in androgen action 
such as *ara*, or in gonadotropin signaling, *fshr* and *lhr*, were differentially expressed 
during testis development, as previously reported (Blázquez and Piferrer 2005; Mazón
et al. 2014). In contrast, the gonadotropin subunits fshb and lhb, detected in the gonads of the sea bass in this study, as also described in other fish species (Wong and Zohar 2004; von Schalburg et al. 2005; Levavi-Sivan et al. 2010) were not differentially expressed.

When looking for DEG ($P < 0.01$) in testes compared to ovaries at 110 dpf, a total of 15 genes were found, although none could be considered as canonical reproduction-related genes. Among them, we identified three genes that were previously described to be expressed in the reproductive system: platelet-derived growth factor beta polypeptide (pdgfb), sorting nexin 1 (snx1) and nuclear transcription factor Y beta (nfy), although there are few data on the role of these genes in testis, not only in fish, but also in mammals. For example, pdgfb is involved in the regulation of many biological processes including embryonic development and sexual phenotype, since alteration of this gene generated male and female infertility in several species, including humans (Donnem et al. 2010). Snx is involved in cellular endocytosis functions and its role in oogenesis was described in the gibel carp (Carassius gibelio) (Wen et al. 2003) while nfy is a pleiotropic transcription factor that participates in multiple processes such as cell proliferation and development (Li et al. 2018). For example, it has been detected in breast cancer cells (Lagadec et al. 2014) and it is involved in spermatogenesis (Vanwert et al. 2008) but its specific role in transcriptional regulation is not fully understood although several functional studies have been published so far. It is known that it can bind to the piwil1 promoter in the germ cells (Chang et al. 2015), but also binds together with the orphan nuclear receptor steroidogenic factor-1 (Nr5a1), to the promoter of fshb gene (Jacobs et al. 2003).

Next, we looked for canonical KEGG pathways involved in sex differentiation and previously described in some but few fish species. This was the case of two pathways required for ovarian development: the fanconi anemia pathway, identified in zebrafish (Rodriguez-Marí and Postlethwait 2011) and in common carp (Cyprinus carpio) (Jia et al. 2018), and the wnt signaling pathway, identified in zebrafish (Sreenivasan et al. 2014) and in rainbow trout (Oncorhynchus mykiss) (Nicol and Guiguen 2011). In the European sea bass, we found that the number of genes differentially expressed ascribed to these two pathways increased at 250 dpf and onwards in the developing ovaries. In testes, we studied the apoptotic pathway p53 previously described in zebrafish...
(Rodríguez-Marí et al. 2010), medaka (Yasuda et al. 2012) and spotted knifejaw testes (Oplegnathus punctatus) (Du et al. 2017), and the cytokine-cytokine interaction pathway identified in Japanese flounder (Zhang et al. 2015) and in zebrafish (Ribas et al. 2017) gonads. In the European sea bass, we found an increase in the number of genes differentially expressed and associated with these pathways at 250 dpf that then decreased. Thus, these results confirm that, as occurs in other fish species, these four pathways are also involved in gonad development in the European sea bass.

Transcription factors during gonadal development

TFs tightly control gene expression in a large number of processes including gonadal development (Migeon and Wisniewski 2000) and so, in the last years, many studies have revealed their importance in fish sex differentiation (Herpin and Schartl 2011; Nakamura et al. 2011; Shen and Wang 2014; Tanaka 2016). Consequently, with the aim of deciphering the involvement of TFs in gonadal development in the European sea bass, we studied the expression of TFs already known to be sexually dimorphic. These included foxl2 (Yamaguchi et al. 2007) and figlα (Kanamori et al. 2008), related to ovarian development, and sox9b (Bagheri-Fam et al. 2010), nr5a1a (Crespo et al. 2013) and dmrt1 (Deloffre et al. 2009) related to testis development, the expression of some being in accordance to what was previously described in the European sea bass gonads fish at 170 dpf and subjected to high temperatures (Díaz and Piferrer 2015). Sox3 was considered as a male-determining gene in ricefish (Oryzias dancena) (Takehana et al. 2014) although its expression was related to both oocyte and testis development in other fish species. This is the case found in grouper (Epinephelus coioides) (Yao et al. 2007) and in Japanese flounder (Jeng et al. 2018) with an expression bias towards female development. In the present study, sox3 clearly showed a female bias in the European sea bass gonads.

Then, we explored the presence of TFs at each specific stage of gonadal development. The number of DE TFs increased as the gonadal development progressed. Thus, at 350 dpf, when gonads were fully differentiated, the largest number of DE TFs were detected. In all stages, there was a larger number of DE TFs in females than in males that is in concordance with the larger number of DEGs found in females in this study. This skewed number towards females was evident already at 110 dpf. DE TFs were only detected in ovaries, probably due to sex differentiation starting earlier in females.
(Piferrer et al. 2005) and to the increased activity of the tissue by ovary formation and meiotic division actions (D’Cotta et al. 2001).

Conclusions

A species-specific microarray enriched for reproduction-related genes was used to study gene expression during European sea bass gonadal development. In contrast to what had been described in other species, a larger number of DEG and DE TFs were observed in ovaries when compared to testis. The expression profiles of 54 genes previously associated to sex differentiation in other species were examined and the steroidogenic gene \textit{hsd17}\beta\textsubscript{10} is described as a promising ovarian marker capable of identifying females as early as 110 dpf. Also, three genes: \textit{pdgfb}, \textit{snx1} and \textit{nfy}, were identified as potential markers for male development. Further, three and eight pathways that are consistently enriched along gonadal development in ovary or testis, respectively, were also identified. Taken together, these results contribute to our understanding of gene expression during sexual development in an economically important species in particular and in non-mammalian vertebrates in general, and emphasize the great diversity, also at the molecular level, of fish sexual development.

Acknowledgments

We would like to thank Dr. Richard Reinhardt for his work on the 454 sequencing of the European sea bass gonads performed at Max Planck Institute for Molecular Genetics, Berlin (Germany). Work was supported by Spanish Government Consolider-Ingenio 2010 grant “Aquagenomics” (ref. CDS2007-0002) to SMK, FM, SZ and FP. BC was supported by an Aquagenomics contract; LR was supported by Aquagenomics postdoctoral contract; ND was supported by Aquagenomics predoctoral contract. NSB was supported by AGL2016-78710-R.

Conflict of Interest

The authors declare that they have no competing interests.
References


34 DOI: 10.1371/journal.pone.0063604.


**Figure 1.** Overview of transcriptomic changes during European sea bass sex differentiation. a) Principal component analysis of microarray results at three developmental stages: 110, 250 and 350 days post fertilization (dpf). Samples cluster together by gender: females (pink ellipse), males (blue ellipse). b) Number of differentially expressed genes found along gonadal development (110, 250 and 350 dpf) for the seven comparisons analyzed using the SAM test.

**Figure 2.** Common Biological Processes Gene Ontology terms at 110, 250 and 350 days post fertilization (dpf) of differentially expressed genes during European sea bass sex differentiation a) Female-related genes ($P < 0.01$) b) Male-related genes ($P < 0.02$).

**Figure 3.** Heatmap of the microarray expression data for 49 out of 5 reproduction-related canonical genes: a) 25 up- and b) 24 downregulated genes in females. Each row represents a gene and each column represents a group of fish by age and sex: (M110 = 8 males at 110 days post fertilization (dpf), M250 = 6 males at 250 dpf and M350 = 9 males at 350 dpf; F110 = 4 females at 110 dpf); F250 = 6 females at 250 dpf and F350 = 9 females at 350 dpf. The key color represents the level of expression scaled by gene (yellow: high expression and blue: low expression). The dendrograms inform of the similarity between genes and between the different samples. Notice that all genes were grouped as pro-female and pro-male as expected from studies in other species. See Dataset 1 for a complete list of gene names and abbreviations.

**Figure 4.** Changes in expression of a set of canonical genes related to sex differentiation and reproduction in female and male gonads during European sea bass sex differentiation. a-f) Canonical genes of the steroidogenic pathway. g-l) Canonical transcription factors. Abbreviations: dpf, days post fertilization; * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$).

**Figure 5.** Number of differentially expressed genes (DEG; left panels) and fold change of reproduction-related pathways (right panels) during European sea bass sex differentiation (pink, females; blue, males): a-b) Fanconi anemia signaling pathway, c-d) Wnt signaling pathway, e-f) p53 signaling pathway, g-h) Cytokine-cytokine...
interaction receptor signaling pathway. In the left panels data is expressed as the total number of genes (absolute values) differentially expressed at each time of development, in the right panels data is expressed as fold change using male values at 110 dpf as control group set at 0.

Figure 6. Diagram showing at three ages during European sea bass sex differentiation the percentage of pro-male and pro-female transcription factors (TFs) differentially expressed at each developmental stage of female-related genes, in pink or male-related genes, in blue.
Table 1. Sixteen common differentially regulated pathways identified from the microarrays data along gonadal development \((P < 0.05)\). Pink color means the pathways that are upregulated in females whereas blue are the upregulated pathways in males

<table>
<thead>
<tr>
<th>Pathway Description</th>
<th>110 dpf</th>
<th>250 dpf</th>
<th>350 dpf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOD-like receptor signaling pathway</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dorso-ventral axis formation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p53 signaling pathway</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chagas disease (American trypanosomiasis)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphatidylinositol signaling system</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ErbB signaling pathway</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butirosin and neomycin biosynthesis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basal transcription factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amino sugar and nucleotide sugar metabolism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type II diabetes mellitus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycine, serine and threonine metabolism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steroid biosynthesis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribosome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folate biosynthesis</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1

(a) Scatter plot showing the distribution of gene expression levels at different developmental stages (350 dpf, 250 dpf, 110 dpf) for females (F) and males (M). The plot highlights the differences in gene expression between sexes.

(b) Diagram illustrating the comparison of gene expression levels between different developmental stages (110, 250, 350 dpf) and sexes (F, M). The table shows the total number of upregulated and downregulated genes for each comparison, along with the statistical significance (P < 0.01 or P < 0.001).
Figure 3
Figure 4

**Males**

**Females**

a. cyp19a1a

b. hsd17β10

c. cyp11βc1

d. ara

e. hsd3β

f. fshr

g. foxl2

h. sox3

i. figlα

j. nr5a1a

k. sox9b

l. dmrt1

Click here to access/download:Figure;Figure 4.pdf
Figure 6

Pro-male TF  Pro-female TF  DEG (no TF)
Click here to access/download

Supplementary Material

Supplementary Information_Ribas et al_FINAL.docx
Click here to access/download **Supplementary Material**
Dataset 1. DEG_new.xlsx
Click here to access/download
Supplementary Material
Dataset 2. GO_TERMS.xlsx
Click here to access/download Supplementary Material
Dataset 3. Pathways_GSVA.xls