

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Ontogenetic onset of immune-relevant genes in the common sole (*Solea solea*)

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Ferraresso, S., Bonaldo, A., Parma, L., Buonocore, F., Scapigliati, G., Gatta, P.P., et al. (2016). Ontogenetic onset of immune-relevant genes in the common sole (*Solea solea*). FISH AND SHELLFISH IMMUNOLOGY, 57, 278-292 [10.1016/j.fsi.2016.08.044].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/573134> since: 2020-02-28

*Published:*

DOI: <http://doi.org/10.1016/j.fsi.2016.08.044>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Serena Ferraresso, Alessio Bonaldo, Luca Parma, Francesco Buonocore, Giuseppe Scapigliati, Pier Paolo Gatta, Luca Bargelloni, *Ontogenetic onset of immune-relevant genes in the common sole (Solea solea)*, Fish & Shellfish Immunology, Volume 57, 2016, Pages 278-292,

The final published version is available online at:

<https://doi.org/10.1016/j.fsi.2016.08.044>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

**When citing, please refer to the published version.**

**Ontogenetic onset of immune-relevant genes in the common sole**  
**(*Solea solea*)**

*Serena Ferraresso*<sup>1§</sup>, *Alessio Bonaldo*<sup>2</sup>, *Luca Parma*<sup>2</sup>, *Francesco Buonocore*<sup>3</sup>,  
*Giuseppe Scapigliati*<sup>3</sup>, *Pier Paolo Gatta*<sup>2</sup>, *Luca Bargelloni*<sup>1</sup>.

<sup>1</sup> Department of Comparative Biomedicine and Food Science, University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy;

<sup>2</sup> Department of Veterinary Medical Sciences, Alma Mater Studiorum, University of Bologna, via Tolara di Sopra 50, 40064 Ozzano Emilia (BO), Italy;

<sup>3</sup> Department for Innovation in Biological, Agro-food and Forest Systems, Tuscia University, Via San Camillo de Lellis s.n.c., 01100 Viterbo, Italy.

§ **Corresponding author:** [serena.ferraresso@unipd.it](mailto:serena.ferraresso@unipd.it) Tel +39 049 8272506, Fax +39 049 8272973

**Email addresses:** [serena.ferraresso@unipd.it](mailto:serena.ferraresso@unipd.it) (S. Ferraresso), [alessio.bonaldo@unibo.it](mailto:alessio.bonaldo@unibo.it) (A. Bonaldo), [luca.parma@unibo.it](mailto:luca.parma@unibo.it) (L. Parma), [fbuono@unitus.it](mailto:fbuono@unitus.it) (F. Buonocore), [scapigg@unitus.it](mailto:scapigg@unitus.it) (G. Scapigliati), [pierpaolo.gatta@unibo.it](mailto:pierpaolo.gatta@unibo.it) (P. P. Gatta), [luca.bargelloni@unipd.it](mailto:luca.bargelloni@unipd.it) (L. Bargelloni).

27 **ABSTRACT**

28 Fish are free-living organisms since initial stages of development and are exposed to numerous  
29 pathogens before their lymphoid organs have matured and adaptive immunity has developed.  
30 Susceptibility to diseases and juvenile mortality represent key critical factors for aquaculture. In this  
31 context, the characterization of the appearance kinetics of the immune system key members will be  
32 useful in understanding the ability of a particular species in generating immune protection against  
33 invading pathogens at different developmental stages.

34 The present study characterized, for the first time, the transcriptional onset of un-explored relevant  
35 genes of both innate and adaptive immune system during the *Solea solea* ontogenesis.

36 Gene expression profiles of immune relevant genes was investigated, by means of DNA microarray,  
37 in ten developmental stages, from hatching (1 day post-hatching, dph) to accomplishment of the  
38 juvenile form (33 dph).

39 The obtained results revealed that transcripts encoding relevant members of innate immune repertoire,  
40 such as lysozyme, AMPs (hepcidin,  $\beta$ -defensin), PPRs and complement components are generally  
41 characterized by high expression levels at first stages (i.e. hatch and first feeding) indicating  
42 protection from environmental pathogens even at early development.

43 Transcription of adaptive immune genes (i.e. Class I and class II MHC, TCRs) differs from that of  
44 the innate immune system. Their onset coincides with metamorphosis and larvae-to-juvenile  
45 transition, and likely overlaps with the appearance and maturation of the main lymphoid organs.  
46 Finally, data collected suggest that at the end of metamorphosis *S. solea* cell-mediated immune  
47 system hasn't still undergone full maturation.

48

49

50

51

52 **Keywords:** *Solea solea*, ontogeny, innate immunity, adaptive immunity, transcriptome, larval  
53 development, gene expression.

## 1. INTRODUCTION

Teleost represent the largest class of vertebrates with more than 40,000 species whose diversity allowed their adaptation to every aquatic environment. This heterogeneous group of organisms is the earliest vertebrate group that possess an immune system comparable to higher vertebrates, with an acquired immune system consisting of B- and T-lymphocytes and memory formation (Fischer et al. 2005, Rauta et al. 2012). However, in contrast to mammals, fish are free-living organisms since initial stages of development and are exposed to numerous pathogens before their lymphoid organs have matured and adaptive immunity has developed.

Immunocompetence is a feature acquired by the organism during ontogeny. For survival, newly hatched fish rely on their innate immune repertoire (Rombout et al. 2005) that acts, in a non-specific manner, as the earliest mechanism that defends the host from infection. However, the development of long lasting immunological memory depends on the ability to activate the adaptive immune system (Øvergård et al. 2011).

A better understanding of the mechanisms related to innate immune responses is crucial for the development of effective approaches for disease management. In the same way, a proper knowledge of adaptive immune system is important since memory responses are possible targets for prevention strategies and it has been proven that immunostimulation before the immune system is fully developed induce tolerance resulting in the lack of response to later stimulation (Patel et al. 2009).

In recent years, considerable progress in the understanding of the development of the immune system in fish has been made. Extensive numbers of studies have been reported on the ontogeny of lymphomyeloid organs of marine fish and freshwater species (Patel et al. 2009, Cunha et al. 2008, Mulero et al. 2007, Falk-Petersen 2005 and references herein). In addition, studies on relevant genes of both innate and adaptive immunity have been conducted in many commercial fish species, such as rainbow trout (*Oncorhynchus mykiss*), Atlantic salmon (*Salmo salar*), sea bass (*Dicentrarchus labrax*), gilthead seabream (*Sparus aurata*), Japanese flounder (*Paralichthys olivaceus*), channel catfish (*Ictalurus punctatus*), turbot (*Scophthalmus maximus*), Atlantic halibut (*Hippoglossus hippoglossus*).

Although the literature in this field is getting extremely wide, few are the studies focused in analyzing, in a comprehensive way, the ontogeny of immune-relevant genes. Due to the importance of disease resistance in the fish industry, the characterization of the appearance kinetics of the immune system key members will be useful in understanding the ability of a particular species in generating immune protection against invading pathogens at different developmental stages.

The common sole (*Solea solea*) is a promising candidate for the European aquaculture, due to its high flesh quality, high market value and increasing demand by consumers (Parma et al. 2013). As for

88 other flatfish species, however, several critical bottlenecks remain to be solved in order to establish  
89 large scale sole farming production: susceptibility to diseases and juvenile mortality represent key  
90 critical factors for sole aquaculture.

91 Although increasing attention has been devoted on characterizing the transcriptome of this species  
92 (Ferraresso et al. 2013, Mazurais et al. 2014, Benzekri et al. 2014), still few information is available  
93 about its immune repertoire.

94 To overcome this limitation, the present study characterized and assessed the transcriptional onset of  
95 un-explored relevant genes of both innate and adaptive immune system during the *S. solea*  
96 ontogenesis, from hatching to accomplishment of the juvenile form. The obtained findings shed light  
97 on the development of common sole defense machinery and demonstrated that the main players of  
98 innate immunity are present during early developmental stages.

99     **2. MATERIAL AND METHODS**

100    **2.1. *S. solea* transcriptome and in silico identification of immune related transcripts**

101    Similarity analyses were performed on a total of 225,944 *S. solea* transcripts (22,252 isotigs/contigs  
102    and 203,692 singletons) obtained through Roche 454 GS FLX pyrosequencing, starting from a  
103    normalized cDNA library constructed by pooling larval stages (1, 4, 6, 8, 11, 13, 20, 33 dph) and  
104    adult tissues as described by Ferraresso et al. (2013). All sequences are stored in the public database  
105    Transcriptome Shotgun Assembly Sequence Database (TSA) under accession number  
106    GAAQ000000000.

107    The Basic Local Alignment Search Tool (BLAST) was used to perform annotation of *S. solea*  
108    isotigs/contigs and singletons (Ferraresso et al. 2013). Briefly, blastx analyses (cut off e-value of <  
109    1.0 E-5) against NCBI amino acid nr database, SWISSPROT database and high quality draft protein  
110    databases of *Danio rerio*, *Gasterosteus aculeatus*, *Oryzias latipes*, *Fugu rubripes*, *Tetraodon*  
111    *nigroviridis*, *Homo sapiens*, and *Mus musculus*, available on Ensembl Genome Browser (release 56),  
112    were performed. In addition, dedicated searches through TBLASTN, using known teleost immune  
113    genes families as queries, were performed. Default algorithm parameters were modified using a less  
114    stringent e-value in order to score sequences with low homology. Identity of all contigs identified as  
115    potentially encoding immune-related genes were further screened for confirmation by FASTA  
116    searches (Pearson W 2004). *S. solea* sequences showing an e-value lower than 10 E-15 were manually  
117    checked and, whenever possible, prolonged at 5' and 3' ends by overlapping.

118    **2.2. *Solea solea* larvae rearing and sampling**

119    *S. solea* larvae employed in the present study are part of an experiment already described by  
120    Ferraresso et al. 2013. Briefly, all animals came from one batch of fertilized eggs obtained from  
121    spontaneous spawning of a broodstock maintained at the Laboratory of Aquaculture, Department of  
122    Medical Veterinary Sciences, University of Bologna, Italy.

123    Newly hatched larvae were maintained in a single incubator until mouth opening at 4 days post  
124    hatching (dph) and then allocated in three flat bottom 280-liter square tanks (2000 larvae tank<sup>-1</sup>).  
125    Animals were reared until 33 dph as described by Ferraresso et al. (2013).

126    Larvae were randomly collected at 1, 4, 6, 8, 11, 13, 15, 18, 24 and 33 dph. The onset of  
127    metamorphosis occurred at 13-14 dph (start of left eye migration) and ended at 24-25 dph (completion  
128    of left eye migration and visibility of left orbital arch on the dorsal side). Larvae were sacrificed by  
129    anesthetic overdose and sampled as described by Ferraresso et al. (2013). Euthanasia method and all  
130    experimental procedures were evaluated and approved by the Ethical-Scientific Committee for  
131    Animal Experimentation of the University of Bologna, in accordance with the European Community  
132    Council directive (86/609/ECC).

### 2.3. RNA extraction

Ten (10) developmental stages, as listed above, from hatching until completed metamorphosis, were used in the present work to characterize the larval transcriptome of *S. solea* by means of DNA microarray (as described later). Eight stages (i.e. 1, 4, 6, 11, 13, 18, 24, 33 dph) were previously employed (Ferraresso et al. 2013) while two (2) additional time points, 8 and 15 dph, were added in the present study. Total RNA was extracted from pools of *S. solea* larvae using the RNAeasy Mini Kit (Qiagen, Hilden, Germany) according to manufacturer's specifications.

Each pool of larvae from 1 to 15 days post hatching (dph) contained approximately 10 individuals while from 18 to 33 dph pools were constituted by 5 larvae each. For each sample, RNA concentration was determined using a UV-Vis spectrophotometer NanoDrop® ND-1000 (NanoDrop Technologies, Wilmington, USA). RNA integrity and quality was then estimated on Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and RNA integrity number (RIN) index was calculated for each sample, only RNAs with RIN number > 8.5 were further processed.

### 2.4. *Solea solea* oligonucleotide microarray

Gene expression analyses were performed using the Agilent-036353 *Solea solea* oligo microarray (GEO accession: GPL16124). This platform represents 15,385 unique transcripts, mainly represented by one probe (60 nt) *in situ* synthesized onto the array (Ferraresso et al. 2013), using Agilent non-contact ink-jet technology (4 x 18K format, including default positive and negative controls). A single dye (Cy3) labeling scheme was implemented, a mixture of 10 different viral poly-adenilated RNAs (Agilent Spike-In Mix) was also added to each RNA sample to monitor labeling and hybridization quality as well as microarray analysis work-flow. Sample labelling and hybridization were performed as reported in Ferraresso et al. (2013). Processed slides were scanned at 5 µm resolution using an Agilent G2565BA DNA microarray scanner. Default settings were modified to scan the same slide twice at two different sensitivity levels (XDR Hi 100% and XDR Lo 10%). The two linked images generated were analyzed together, data were extracted and background subtracted using the standard procedures contained in the Agilent Feature Extraction (FE) Software, version 9.5.1.

New microarray experiments were conducted on two developmental stages, 8 and 15 dph, by processing four (4) pools per stage. Newly produced microarray data were normalized together with the previous from Ferraresso et al. (2013) (i.e. 1, 4, 6, 11, 13, 18, 24 and 33 dph, 4 pools/stage). The normalization procedure was performed using R statistical software, microarray data were cyclic lowess normalized across all arrays. All raw and normalized gene expression data are available at the Gene Expression Omnibus (GEO) online microarray data repository under accession GSE81944.



## 167 **2.5. Phylogenetic analyses**

168 Phylogenetic trees were inferred using Bayesian inference (BI) and Maximum Likelihood (ML)  
169 methods (Felsenstein 2004). The BI tree was obtained with MrBayes 3.2.6 (Ronquist et al. 2012).  
170 Two simultaneous runs, each of four chains, were performed in each dataset analysis. Each run  
171 consisted of 2,000,000 generations, and trees were sampled every 200 generations (trees generated =  
172  $2 \times 10^4$ ). Stationarity was considered to be reached when the average standard deviation of split  
173 frequencies was less than 0.005. Burn-in was very stringent and only the last 2,000 generated trees  
174 were used to compute the majority-rule posterior consensus trees. The ML analyses were performed  
175 with Phyml 3.0 program (Guindon and Gascuel 2003), non-parametric bootstrap resampling (1000  
176 replicates) was also performed to evaluate the robustness of tree topology.  
177 For each amino acid alignment, the most appropriate model of protein evolution was selected using  
178 the Akaike information criterion as implemented in ProtTest v1.4 (Abascal et al. 2005).

179

### 180 3. RESULTS

181 *In silico* analysis of *S. solea* larval transcriptome led to the identification of 149 immune-related  
182 transcripts (listed in Supplementary file 1), many of which are represented by complete cDNA  
183 sequences that were identified for the first time in this species.

184 Herein we report on the characterization of those transcripts encoding some of the major proteins  
185 involved in both innate and adaptive immune system. Their transcriptional onset was investigated by  
186 means of DNA microarray in *S. solea* larvae from hatching to accomplishment of the juvenile form.  
187 In order to facilitate the results discussion, these proteins have been subdivided into functional  
188 categories, as follows.

#### 189 3.1. INNATE IMMUNE SYSTEM

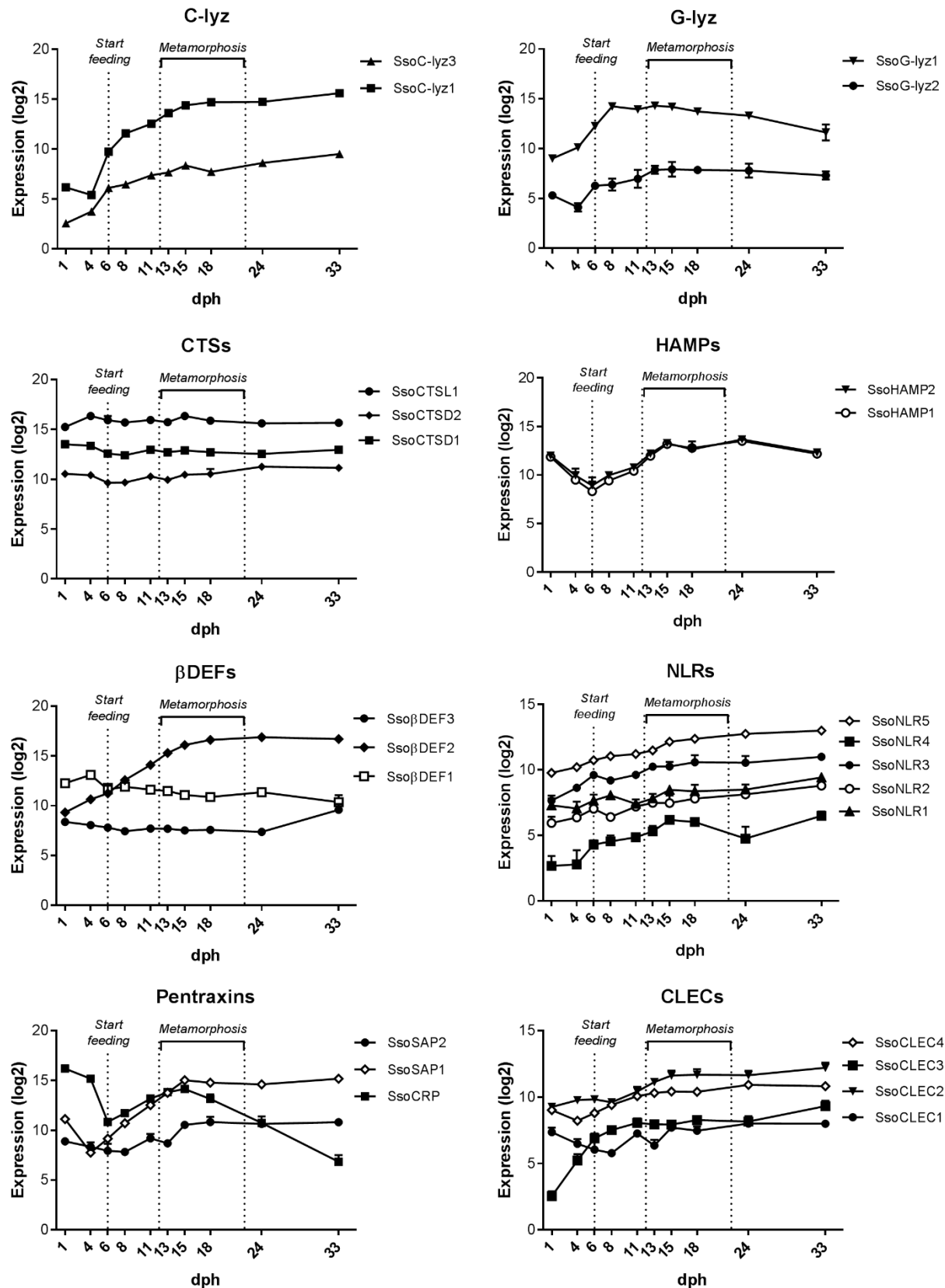
##### 190 3.1.1. Lytic enzymes

191 Lytic enzymes are important defence elements especially against bacteria (Magnadóttir, 2006), in the  
192 present study specific blast searches were conducted in order to identify transcripts coding different  
193 classes of lytic enzymes and their expression during larval development has been assessed.

194 In total, three transcripts coding C-type lysozyme (c-lyz) family members herein referred to as SsoC-  
195 lyz1(isotig00798), SsoC-lyz2 (isotig00800) and SsoC-lyz3 (isotig06704), and two related to G-type  
196 lysozyme (g-lyz) family members, SsoG-lyz1 (isotig19917) and SsoG-lyz2 (isotig13329), were  
197 identified in *S. solea*.

198 Gene expression during larval development of c-type (isotig00798 and isotig06704) and g-type  
199 (isotig19917 and isotig13329) *S. solea* lysozymes were also assessed by DNA microarray. Marked  
200 differences in expression profiles were observed between c-lyz and g-lyz members (see Figure 1).  
201 Both c-lyz transcripts showed a strong increase in expression levels over time, showing at 33 dph a  
202 fold-change of 121-fold and 638-fold for SsoC-lyz3 and SsoC-lyz1, respectively.

203 Slightly different expression profiles were appreciated for g-lyz members, SsoG-lyz1 showed a rapid  
204 increase of mRNA levels soon after start feeding (4-8 dph, 37-fold) followed by a decrease during  
205 and after metamorphosis, while SsoG-lyz2 is characterized by a gradual increase of expression until  
206 13 dph (6.5-fold compared to 1 dph), a steady state during metamorphosis, and a decrease at 33 dph.  
207 A second class of lytic enzymes are cathepsins (CTSs), a complex superfamily of lysosomal proteases  
208 that participate in many physiologic and pathophysiologic cellular processes. Among the different  
209 classes of CTSs, cathepsin L (CTSL) has a documented role in immune response in both vertebrate  
210 and invertebrate.



211

212 **Figure 1. Temporal expression of innate immune genes.** Gene expression levels (log2), from 1 to 33 dph, of *S. solea*  
 213 transcripts coding innate immune genes. For each developmental stage, Standard deviation (SD) across biological  
 214 replicates was calculated and represented as error bar, however in many points the error bar would be shorter than the  
 215 height of the symbol and was not represented.

216 One cathepsin L gene was identified from the transcriptome of *S. solea* using *tblastn* searches, from  
217 now referred to as SsoCTSL1(isotig05294). SsoCTSL1 deduced protein sequence exhibited a  
218 conserved mature peptide with the catalytic triad of amino acids (Cys, His and Asn) usually present  
219 in its active site (see Supplementary file 2). Pairwise similarities between SsoCTSL1 and CTSL genes  
220 from other teleost species revealed 94% and 91% identity with *Cynoglossus semilaevis* (AEL31666)  
221 and *Epinephelus coioides* (AEN28617) CTSL, respectively. SsoCSTL1 gene expression pattern  
222 during early development revealed a high and constitutive expression from hatching to  
223 metamorphosis (see Figure 1).

224 Within cathepsin superfamily, also cathepsin D (CTSD) has a suggested, albeit still uncharacterized,  
225 role in immune response. In total two transcripts coding cathepsin D were identified in *S. solea*,  
226 namely SsoCTSD1 (isotig00536) and SsoCTSD2 (isotig07059). SsoCTSD1 and SsoCTD2 are not  
227 full length transcripts, since they encode 131 and 155 amino acids, respectively. However they both  
228 contain conserved Asp residue (D281) known to be involved in the CTSDs catalytic site (data not  
229 shown) and show 96.2% (SsoCTSD1) and 92.9% (SsoCTSD2) sequence identity with *Lates calcifer*  
230 CTSD (ABV59077) and *H. hippoglossus* CTSD (ABI85390), respectively. Gene expression analysis  
231 revealed a marked difference between SsoCTSD1 and SsoCTSD2 during larval development.  
232 SsoCTSD1 shows a pick of expression at 4 dph and then it maintains high and stable mRNA levels  
233 over time, while SsoCTSD2 shows a drop in expression prior to start feeding (1.9-fold at 6 dph  
234 compared to 1 dph) and then an increase (3-fold) during metamorphosis (see Figure 1).

### 235 **3.1.2. Antimicrobial peptides (AMPs)**

236 Sequence analysis of *S. solea* contigs allowed the identification of several antimicrobial peptides  
237 (AMPs) belonging to different classes.

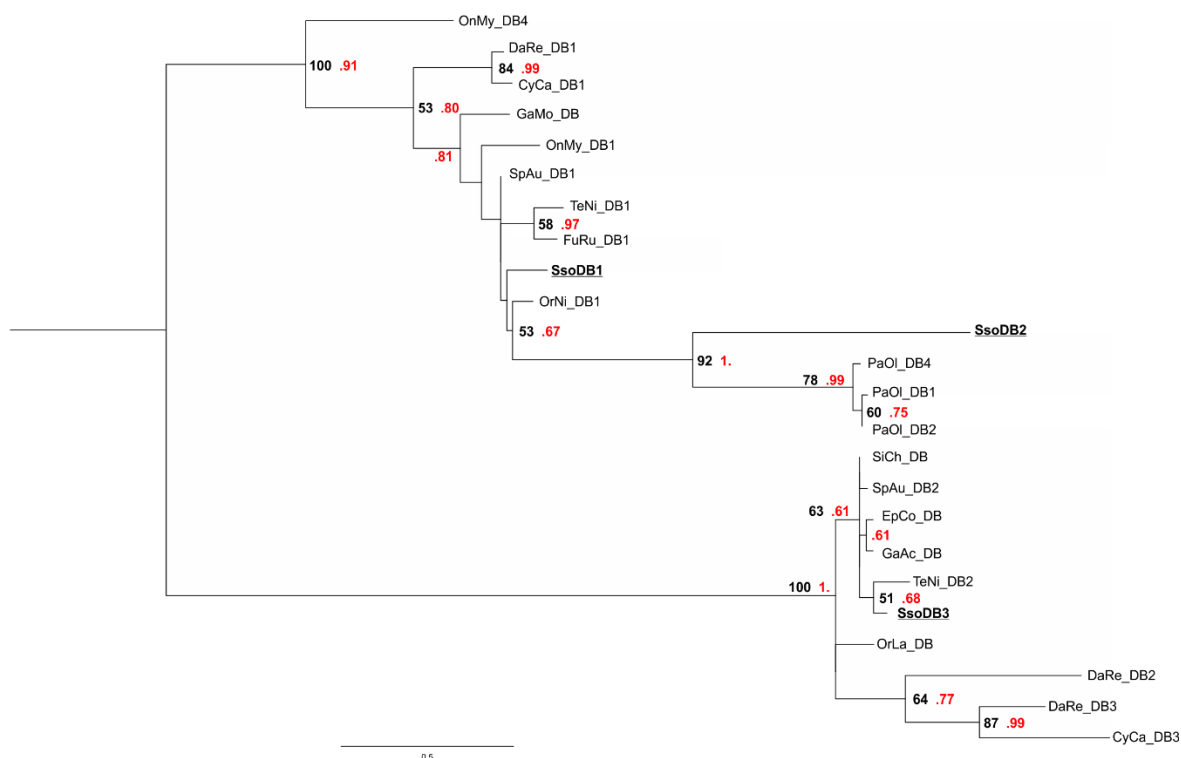
238 Two contigs encoding partially hepcidin antimicrobial peptides (hamp), namely SsoHAMP1  
239 (isotig22007) and SsoHAMP2 (isotig03373), were identified showing significant match with *S.*  
240 *aurata* hepcidin 1 (ABV01929.1, 81% identity) and *S. maximus* hepcidin-2 (Q5CAJ5, 57.8%  
241 identity), respectively. Both transcripts show the same expression profile, mRNA levels are higher  
242 at hatching (1 dph) followed by a decrease at 4 and 6 dph and, after that period, transcription patterns  
243 show a marked increase prior to (15 dph, >20-fold compared to 6 dph) and during metamorphosis  
244 (24 dph, ~26-fold), and, finally, a new decline at 33 dph (see Figure 1).

245 It's worth of notice that *S. solea* contigs encode two not overlapping portions of hamp peptides and,  
246 even in light of their identical transcription profiles, it cannot be excluded they are sequences related  
247 to the same gene.

A.



B.



**Figure 2. Phylogenetic analysis of *S. solea*  $\beta$ -defensin genes.** A) Multiple alignment of the deduced amino acid sequences of *S. solea* (Sso) and other teleost  $\beta$ -defensin (DB) genes. Published DB sequences described from *Danio rerio* (DaRe\_DB1, CAJ57442.1; DaRe\_DB2, NP\_001075023.1; DaRe\_DB3, NP\_001075024.1), *Oncorhynchus mykiss* (OnMy\_DB1, CAK54950.1; OnMy\_DB4, CAR82092.1), *Oryzias latipes* (OrLa\_DB, ACG55699.1), *Paralichthys olivaceus* (PaOl\_DB1, ADA84138.1; PaOl\_DB2, ADA84139.1; PaOl\_DB4, ADA84141.1), *Cyprinus carpio* (CyCa\_DB1, AEZ00740.1; CyCa\_DB3, AGZ03658.1), *Epinephelus coioides* (EpCo\_DB, AAN02164.1), *Gasterosteus aculeatus* (GaAc\_DB, ENSGACP00000027373), *Tetraodon nigroviridis* (TeNi\_DB1, CAJ57644.1; TeNi\_DB2, CAG00590.1), *Fugu rubripes* (FuRu\_DB1, CAJ57646.1), *Gadus morhua* (GaMo\_DB, AEB69787.1), *Oreochromis niloticus* (OrNi\_DB1, AGW83444.1), *Siniperca chuatsi* (SiCh\_DB, ACO88907.1), *Sparus aurata* (SpAu, unpublished) were included in the alignment. Cysteines (C) involved in disulphide bridge formation are highlighted with a connection showing which pairs form bonds. B) Phylogenetic tree showing the relationships between teleost  $\beta$ -defensin. Branch lengths correspond to the number of amino acid substitutions, with the scale indicated at the bottom of the figure. Black numbers represent bootstrap values (values less than 50 are not shown) while red numbers refer to Bayesian Inference posterior probabilities. These latter values are provided in a compressed way (e.g. 1. for 1.00 ; .99 for 0.99).

265 Three full length transcripts encoding putative  $\beta$ -defensin were found, namely Sso $\beta$ DEF1  
266 (isotig03407), Sso $\beta$ DEF2 (isotig19582) and Sso $\beta$ DEF3 (isotig18247). They all share the common  
267 features of vertebrate defensins, including a small size with a precursor molecule ranging from 59 to  
268 65 aa, six conserved cysteines in the mature region and the presence of a negatively charged glutamic  
269 acid residue in the middle of the mature peptide, a feature peculiar of the fish  $\beta$ -defensin (Figure 2A).  
270 MatGAT software was employed to determine the homology across *S. solea*  $\beta$ -defensin: the highest  
271 identity/similarity was appreciated between Sso $\beta$ DEF1 and Sso $\beta$ DEF2 (38.8% identity and 56.9%  
272 similarity), while lower identities/similarities were appreciated when comparing Sso $\beta$ DEF3 with  
273 Sso $\beta$ DEF1 and Sso $\beta$ DEF2 (29.9%/47.7% and 26.2%/47.6%, respectively).

274 In order to analyze the evolutionary relationship of *S. solea*  $\beta$ -defensin, phylogenetic analyses were  
275 conducted (by applying the JTT+I+G evolutionary model) considering all publicly available teleost  
276  $\beta$ -defensin proteins. Trees produced by ML and BI methods share the same topology (Figure 2B). As  
277 for other species, evolutionary analysis divided *S. solea* defensins into two major clusters; the first  
278 cluster groups Sso $\beta$ DEF1 and Sso $\beta$ DEF2, while Sso $\beta$ DEF3 belongs to the second one. A strong  
279 relationship can be noticed between Sso $\beta$ DEF2 and *P. olivaceus*  $\beta$ -defensins (bootstrap support  
280 >90%, posterior probability=1.0), the only other flatfish species represented in the dataset.

281 Transcriptional onset of all Sso $\beta$ DEFs during ontogeny was investigated. *S. solea*  $\beta$ -defensin  
282 transcripts showed patterns of expression markedly different across developmental stages (Figure 1).  
283 Sso $\beta$ DEF1 showed the highest expression levels soon after hatching (1-4 dph), followed by a decrease  
284 over time; Sso $\beta$ DEF3 mRNA levels are generally low from 1 to 24 dph, followed by peak of  
285 expression (4.9-fold) at 33 dph, while Sso $\beta$ DEF2 shows a considerable increase of expression over  
286 time (~187-fold between 1 and 24 dph).

### 287 **3.1.3. Pattern Recognition Receptors (PPRs)**

288 Nod-like receptors (NLRs) and Toll-like receptors (TLRs) are two major forms of innate immune  
289 sensors, which provide immediate responses against pathogenic invasion or tissue injury.

290 *In silico* analysis of *S. solea* transcriptome identified at least 5 different NLRs herein named SsoNLR1  
291 (isotig00221), SsoNLR2 (isotig11055), SsoNLR3 (isotig18107), SsoNLR4 (isotig08270) and  
292 SsoNLR5 (isotig07671). Sequence analysis of SsoNLRs highlighted a high sequence conservation  
293 with *Larimichthys crocea* NLRC3 (KKF31313) with identity percentages ranging from 65 to 91%  
294 (78 to 99% similarity).

295 Analysis of SsoNLRs gene expression through DNA microarray highlighted very similar patterns of  
296 expression during larval development for all five transcripts showing a gradual increase of expression  
297 over time with fold-changes (1 dph vs 33 dph) ranging from 4- to 13-fold for SsoNLR1 and SsoNLR4,  
298 respectively (see Figure 1). Even if the trend of expression along time is similar for all SsoNLRs, the

299 same is not true for their mRNA levels since a difference up to 100-fold can be appreciated, with  
300 SsoNLR4 exhibiting the lowest level of expression and SsoNLR5 showing the highest.

301 All similarity analyses failed to identify TLRs in *S. solea* transcriptome. Specific tblastn searches by  
302 using all teleost TLRs as query found significant matches with several sequences coding leucine-rich  
303 repeats but no TIR domains were detected even if using only teleost TIRs as query. Neither MYD88  
304 (Myeloid Differentiation factor 88) nor TICAM1 (Toll-Like Receptor Adaptor Molecule 1) genes,  
305 two key members of the TLR signaling pathway were found in *S. solea*, while other effectors such as  
306 TIRAP (TIR domain-containing adapter protein), IRAK4 (Interleukin-1 receptor-associated kinase  
307 4) and PI3K (Phosphoinositide-3-kinase) were identified. Those genes didn't show big variation in  
308 expression during larval development except for IRAK4 that exhibited a gradual increase of mRNA  
309 levels over time reaching its peak of expression at 33 dph (FC 6-fold compared to 1 dph, data not  
310 shown).

311 It's worth noting that similarity analyses conducted in both *S. solea* isotigs and singletons identified  
312 9 transcripts coding Novel-immune type receptors (NITRs), a family of natural killer receptors  
313 recently identified in teleost. These transcripts were all singletons thus only incomplete sequence  
314 information were available and no data upon their gene expression have been collected.

#### 315 **3.1.4. Agglutinins and precipitins**

316 Important components of the innate immunity are agglutinins and pentraxins that play a pivotal role  
317 in opsonization of invading pathogens and activation of the complement cascade.

318 Two contigs (isotig14311 and isotig19418) were identified in *S. solea* transcriptome encoding one C-  
319 reactive protein (CRP). SsoCRP deduced protein sequence is 221 aa long and exhibits the Pentraxin  
320 family signature H-x-C-x-[ST]-W-x-[ST], in which the conserved cysteine is involved in an inter-  
321 chain disulfide bond (see Supplementary file 2). Pairwise similarities between SsoCRP and CRP  
322 proteins from other teleost species display a high level of sequence conservation with highest identity  
323 shared with *L. calcifer* CRP (73.5% identity, 89.2% similarity) and *Siniperca chuatsi* CRP (70.3%  
324 identity, 87.4% similarity).

325 SsoCRP pattern of expression during development shows highest transcript levels at 1 dph, a  
326 pronounced drop in expression before start feeding (~40-fold, 1 dph vs 6 dph), followed by an increase  
327 before and during metamorphosis, with a peak at 15 dph (9.7-fold compared to 6 dph), and a  
328 successive decrease of expression thereafter (see Figure 1).

329 Similarity searches identified also two transcripts coding *S. solea* Serum amyloid P-component  
330 (SAP), herein referred to as SsoSAP1 (isotig15744) and SsoSAP2 (isotig13411). As for SsoCRP,  
331 these transcripts exhibit the conserved Pentraxin family signature and show 58% (SsoSAP1) and 68%  
332 (SsoSAP2) sequence identity with *Perca flavescens* SAP (ADX97225) and *Anoplopoma fimbria*

333 SAP (ACQ58218), respectively. During larval development, both SsoSAPS exhibit similar  
334 expression patterns, with a decrease of expression soon after hatching followed by a great increase of  
335 expression until 15 dph (~150-fold for SsoSAP1) or 18 dph (7-fold for SsoSAP2).

336 C-type lectins (CLEC) functions as an endocytic receptors and are involved in antigen uptake at the  
337 site of infection. This gene family counts many members that share a common protein fold and  
338 signature. In *S. solea* transcriptome four transcripts coding CLEC were identified, namely SsoCLEC1  
339 (isotig10440), SsoCLEC2 (isotig04630), SsoCLEC3 (isotig01907) and SsoCLEC4 (isotig05634).  
340 SsoCLECs deduced amino acid sequences share a conserved C-type lectin domain signature, despite  
341 pairwise alignment showed a high level of sequence variation across SsoCLECs (data not shown).  
342 Gene expression analysis revealed similar trends for SsoCLEC2, SsoCLEC3 and SsoCLEC4, that  
343 showed a gradual increase of mRNA levels over time with fold-changes (1 dph vs 33 dph) ranging  
344 from 3.5- to 109-fold for SsoCLEC4 and SsoCLEC3, respectively (see Figure 1). Different is the  
345 expression profile of SsoCLE1 that showed a 3-fold decrease of expression from 1 to 8 dph, followed  
346 by a gradual increase thereafter (4.7-fold, 33 dph vs 8 dph).

347 Member of the CLEC family is also the Macrophage mannose receptor 1(MRC1), a protein that acts  
348 as phagocytic receptor for bacteria, fungi and other pathogens and mediates the endocytosis of  
349 glycoproteins by macrophages. *In silico* analysis identified one transcript (isotig03928) coding *S.*  
350 *solea* MRC1 (SsoMRC1) that showed consistent sequence identity (63%) with *Oreochromis niloticus*  
351 MRC1 (XP\_005476589). SsoMRC1 pattern of expression during larval development is similar to that  
352 of SsoCLE1 with a 2-fold decrease of mRNA levels until 8 dph, followed by an increase until 24 dph  
353 (data not shown).

### 354 **3.1.5. Complement cascade**

355 The complement is a major humoral system of innate immunity. Members of both classical, lectin  
356 and alternative pathways have been identified in several teleost species and almost all the homologues  
357 of the mammalian complement components have been identified in fish model species (i.e. zebrafish,  
358 trout) (Nakao et al. 2011). *In silico* analysis of *S. solea* sequences let to the identification of at least  
359 14 members of the complement system, all listed in Table 1, that show a high level of conservation  
360 across teleost species.

361 Transcription profiles during development highlighted different patterns of expression across  
362 complement components (Figure 3). C3 component, key molecule of the alternative pathway, shows  
363 a gradual increase of expression over time with a fold increase up to 43-fold between 4 and 33 dph.  
364 Similar evidence was found also for B/C2 with both transcripts showing a strong increase in  
365 expression during and after metamorphosis, and a peak at 33 dph (56- and 58-fold compared to 1  
366 dph), while all transcripts coding C4 and MASP-1 show no difference in expression during



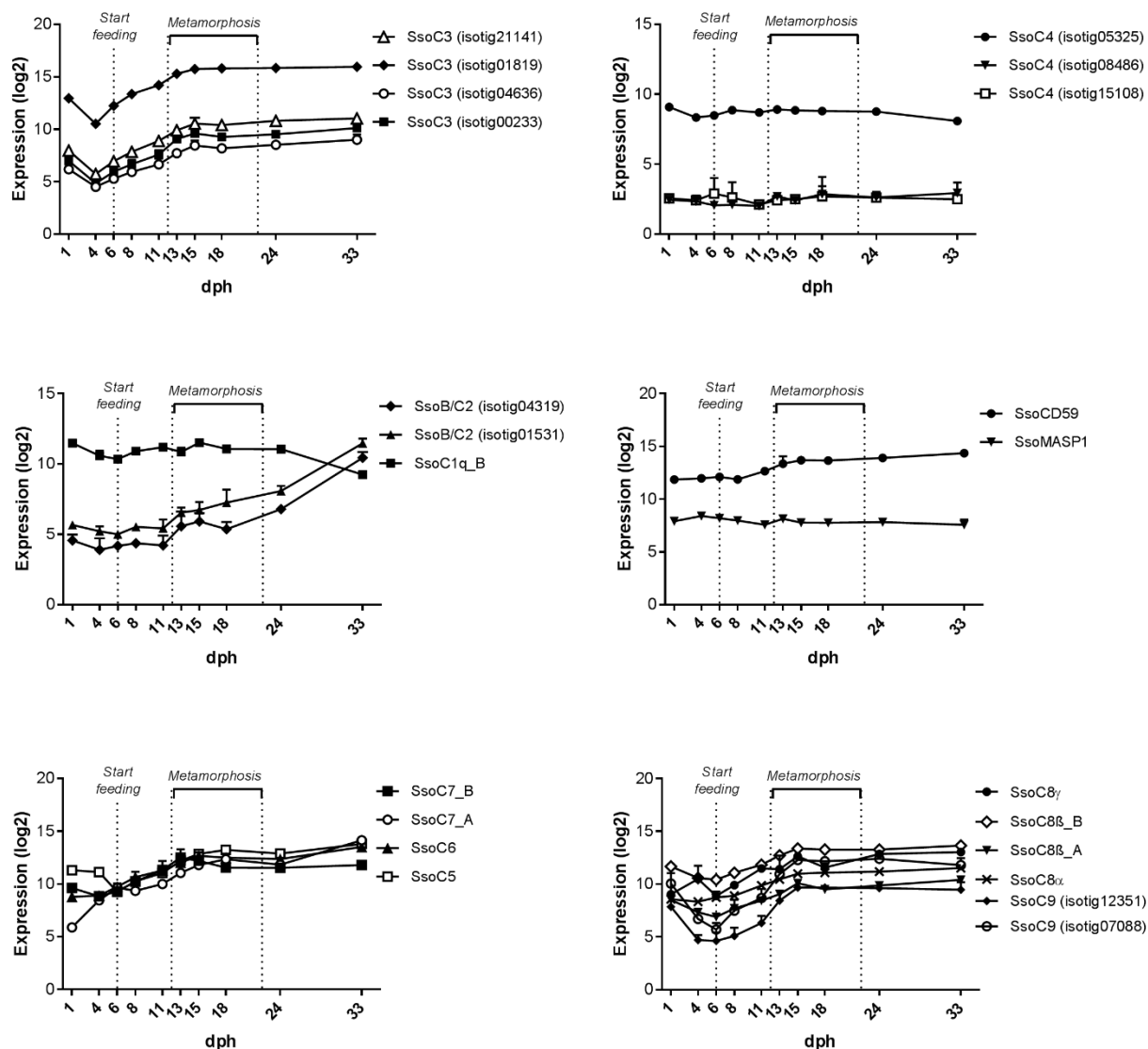
367 development. Worth of notice is the very high level of expression of C1q-B transcript, while no data  
 368 are available for the C1q-A isoform. Terminal complement components C5, C6, C7, C8 and C9 are  
 369 all characterized by a similar expression profile with an increase of expression until the onset of  
 370 metamorphosis and a second increase after metamorphosis completion (Figure 3).

371

372 **Table 1:** Complement component members identified in *S. solea*  
 373

Component	Putative isoform	<i>S. solea</i> Seq ID	% Identity (% Similarity) with other teleost species
<b>C1q</b>	<b>A</b>	Isotig00144	39.5% (72.2%) <i>Salmo salar</i>
	<b>B</b>	Isotig02657	41.3% (77.9%) <i>Danio rerio</i>
<b>B/C2</b>		Isotig04319	72% (84%) <i>Larimichthys crocea</i> ACS83542
		Isotig01531	77% (86%) <i>Larimichthys crocea</i> ACS83542
<b>C3</b>		Isotig00233	77% (88%) <i>Sparus aurata</i>
		Isotig04636	78% (87%) <i>Sparus aurata</i>
		Isotig01819	73% (80%) <i>Sparus aurata</i>
		Isotig21141	73% (90%) <i>Sparus aurata</i>
<b>C4</b>		Isotig15108	57% (75%) <i>Oreochromis niloticus</i> XP_005472530
		Isotig08486	61% (73%) <i>Oreochromis niloticus</i> XP_005472530
		Isotig05325	65% (83%) <i>Oreochromis niloticus</i> XP_005472530
<b>C5</b>		Isotig10341	67% (82%) <i>Fugu rubripes</i> XP_003974669
<b>C6</b>		Isotig18274	80% (86%) <i>Epinephelus bruneus</i> AEM37769
<b>C7</b>	<b>A</b>	Isotig01050	74% (83%) <i>Oreochromis niloticus</i> XP_003446192
	<b>B</b>	Isotig10343	73% (79%) <i>Oplegnatus fasciatus</i> AFZ93893
<b>C8<math>\alpha</math></b>		Isotig01719	77% (86%) <i>Oplegnatus fasciatus</i> AFZ93888
<b>C8<math>\beta</math></b>	<b>A</b>	Isotig16508	86% (92%) <i>Pseudopleuronectes americanus</i> AAT01914
	<b>B</b>	Isotig03362	81% (91%) <i>Paralichthys olivaceus</i> Q9PVW7
<b>C8<math>\gamma</math></b>		Isotig07766	76% (85%) <i>Oryzias latipes</i> AGE44247
<b>C9</b>		Isotig12351	80% (86%) <i>Oplegnatus fasciatus</i> AFU81223
		Isotig07088	76% (86%) <i>Oplegnatus fasciatus</i> AFU81223
<b>MASP-1</b>		isotig13433	82% (92%) <i>Oreochromis niloticus</i> XP_005462795
<b>CD59</b>		Isotig15198	75% (82%) <i>Oryzias latipes</i> XP_005457336
<b>Complement factor H</b>		isotig00434	64% (76%) <i>Perca flavescens</i>
		Isotig13606	66% (80%) <i>Perca flavescens</i>

374

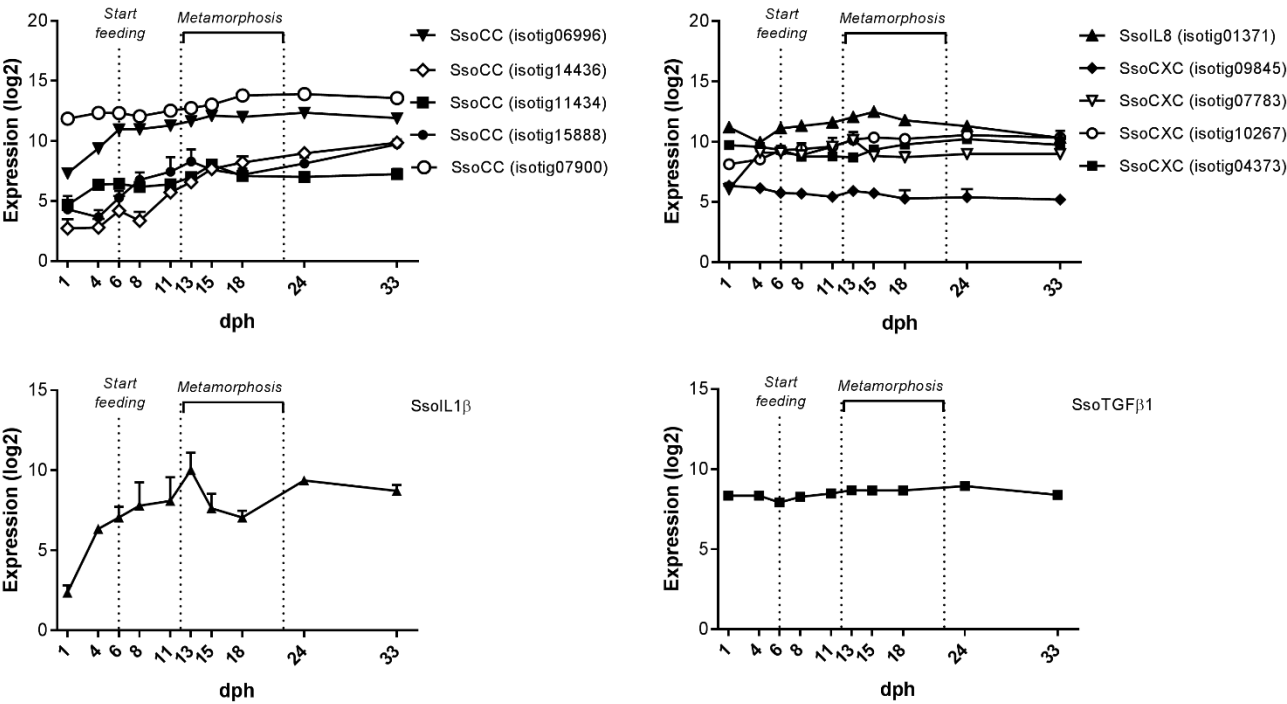


**Figure 3. Temporal expression of complement components.** Gene expression levels (log2), from 1 to 33 dph, of *S. solea* transcripts coding complement component members. For each developmental stage, Standard Deviation (SD) across biological replicates was calculated and represented as error bar, however in many points the error bar would be shorter than the height of the symbol and was not represented.

### 3.1.6. Cytokines cascade

Functional and genetic studies in many teleost species have proven that fish possess a network of signalling molecules, cytokines and chemokines, that function as important effectors of inflammatory responses. The *in silico* analysis of *S. solea* transcriptome identified a total of ten contigs encoding both CC-motif and CXC-motif chemokines showing different expression profiles during larval development (see Figure 4). All putative CC chemokines (isotig07900, isotig15888, isotig11434, isotig14436 and isotig06996) exhibit increasing expression levels over time with fold-changes up to 124-fold (33 vs 1 dph) while considerably different profiles were observed for CXC chemokines (isotig04373, isotig07783, isotig09845, isotig01371 and isotig10267).

390 When looking at inflammatory cytokines, similarity analyses identified one transcript (isotig14919)  
 391 showing 75% identity with *C. semilaevis* Interleukin-1 $\beta$  (IL1 $\beta$ ; XP\_008333362) and one contig  
 392 (isotig10739) exhibiting 76% identity with *E. coioides* Transforming growth factor  $\beta$ 1 (TGF $\beta$ 1;  
 393 ACV96791), from now referred as SsoIL1 $\beta$ , and SsoTGF $\beta$ 1, respectively.  
 394 SsoIL1 $\beta$  gene expression is almost undetectable soon after hatching (1dph) and shows a dramatic  
 395 increase of expression at the onset of metamorphosis (237-fold at 13dph), while SsoTGF $\beta$ 1 maintains  
 396 moderate and stable mRNA levels from 1 to 33 dph (see Figure 4).  
 397



398  
 399 **Figure 4. Temporal expression of *S. solea* cytokines.** Gene expression levels (log2), from 1 to 33 dph, of *S. solea*  
 400 transcripts coding cytokines. For each developmental stage, Standard Deviation (SD) across biological replicates was  
 401 calculated and represented as error bar, however in many points the error bar would be shorter than the height of the  
 402 symbol and was not represented.  
 403

## 404 3.2. ACQUIRED IMMUNE SYSTEM

### 405 3.2.1. Antigen presentation

406 Major histocompatibility complex (MHC) genes encode key molecules in the immune response and  
 407 members of MHC complex can be distinguished in Class I or Class II MHC proteins.  
 408 Class I MHC is a non-covalently bound heterodimer composed of an MHC class I $\alpha$  chain and  $\beta$ 2-  
 409 microglobulin. MHC class I $\alpha$  is expressed on all the nucleated cells and generally presents  
 410 endogenous antigens to cytotoxic T lymphocytes (Cuesta et al. 2007). The MHC class I genes often  
 411 exist as multiple copies, analysis of *S. solea* sequences lead to the identification of 3 different  
 412 transcripts for MHC I $\alpha$ , from now named SsoMHC I $\alpha$ \_01 (isotig00026), SsoMHC I $\alpha$ \_02 (isotig00049)

413 and SsoMHCI $\alpha$ \_03 (isotig00066). Isotig00026 encodes the complete SsoMHCI $\alpha$ \_01 protein  
414 sequence with leader peptide region,  $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 3 domains, connecting peptide, transmembrane  
415 domain and cytoplasmic domain (See Supplementary file 3), while isotig00049 and isotig00066 are  
416 not full-length transcripts and they encode putative proteins lacking the leader peptide and a portion  
417 of the  $\alpha$ 1 domain.

418 The three deduced MHC I $\alpha$  protein sequences were aligned to other vertebrate MHC class I alpha  
419 proteins and showed all the features conserved across vertebrates, such as the presence of cysteines  
420 involved in the  $\alpha$ 2 and  $\alpha$ 3 intra-domain disulphide bridge (C125, C189, C225 and C285), a  
421 potential *N*-glycosylation site (NQT, position 111–113 in alignment) and conserved residues  
422 important for peptide and  $\beta$ 2-microglobulin interaction (See Supplementary file 3).

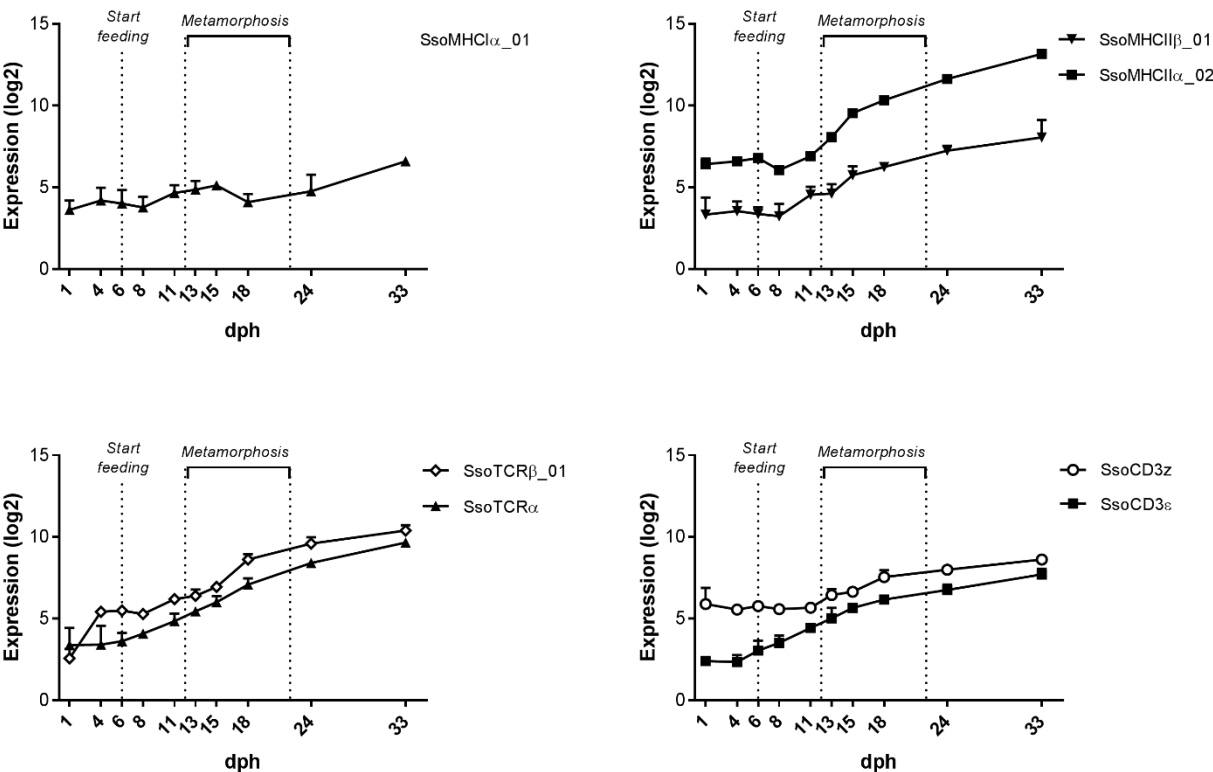
423 Phylogenetic analyses were conducted (by applying the JTT+I+G evolutionary model) to determine  
424 the evolutionary relationship of sole MHC I $\alpha$  genes (Supplementary file 3), and all *S. solea* alleles  
425 cluster together with bootstrap support and posterior probability and are in close relationship with the  
426 Japanese flounder and the turbot sequences, the other two teleost species belonging to the order  
427 Pleuronectiformes.

428 Class II MHC is a heterodimer composed of  $\alpha$  and  $\beta$  chains, expressed in the surface of only a few  
429 cell types and generally presents exogenous antigens to T helper lymphocytes. *In silico* analysis of *S.*  
430 *solea* transcriptome identified two and three different isoforms for MHC II $\alpha$  and MHC II $\beta$ ,  
431 respectively. SsoMHCII $\alpha$ \_01 (isotig01270) and SsoMHCII $\alpha$ \_02 (isotig01271) encoded full-length  
432 transcripts and both present Leader peptide,  $\alpha$ 1 and  $\alpha$ 2 domains, connecting peptide, transmembrane  
433 region and cytoplasmic domain (see Supplementary file 3).

434 Three different amino acid sequences were identified for *S. solea* MHC II $\beta$ , namely SsoMHCII $\beta$ \_01  
435 (isotig00164), SsoMHCII $\beta$ \_02 (isotig00165) and SsoMHCII $\beta$ \_03 (isotig00169), sharing 91-94%  
436 amino acid identity. SsoMHCII $\beta$ \_01 and SsoMHCII $\beta$ \_02 encoded full-length transcripts and possess  
437 Leader peptide,  $\beta$ 1 and  $\beta$ 2 domains, connecting peptide, transmembrane region and cytoplasmic  
438 domain, while SsoMHCII $\beta$ \_03 sequence lack the Leader peptide.

439 The deduced *S. solea* MHC II $\alpha$  protein sequences share 48.9-78.6% amino acid identity with other  
440 teleost species and similar results were obtained from the alignment of the deduced amino acid  
441 sequence of the sole MHC II $\beta$  (46.8-73.6%). To evaluate the evolutionary relationships of the  
442 common sole MHC II $\alpha$  and MHC II $\beta$ , two phylogenetic trees (by applying the WAG+G+F  
443 evolutionary model) were conducted based on the protein sequences of teleost MHC II  
444 (Supplementary file 3). Both phylogenetic trees evidenced a close relationship with teleost belonging  
445 to the order Pleuronectiformes and *S. solea* MHCs grouped in a unique cluster with high bootstrap  
446 support and posterior probability.

447 Gene expression levels of both class I and class II MHC (each represented onto the microarray  
 448 platform by one member) have been assessed from hatching to the end of the metamorphosis. MHC  
 449 I $\alpha$ , MHC II $\alpha$  and MHC II $\beta$  showed totally comparable patterns of expression, with mRNA levels  
 450 barely detectable soon after hatching followed by a gradual increase during and after metamorphosis  
 451 and a peak of expression at 33 dph, with fold-changes ranging from 7.5 to 106-fold (see Figure 5).  
 452



453  
 454 **Figure 5. Temporal expression of *S. solea* adaptive immune genes.** Gene expression levels (log2), from 1 to 33 dph,  
 455 of *S. solea* transcripts coding adaptive immune genes. For each developmental stage, Standard Deviation (SD) across  
 456 biological replicates was calculated and represented as error bar, however in many points the error bar would be shorter  
 457 than the height of the symbol and was not represented.  
 458

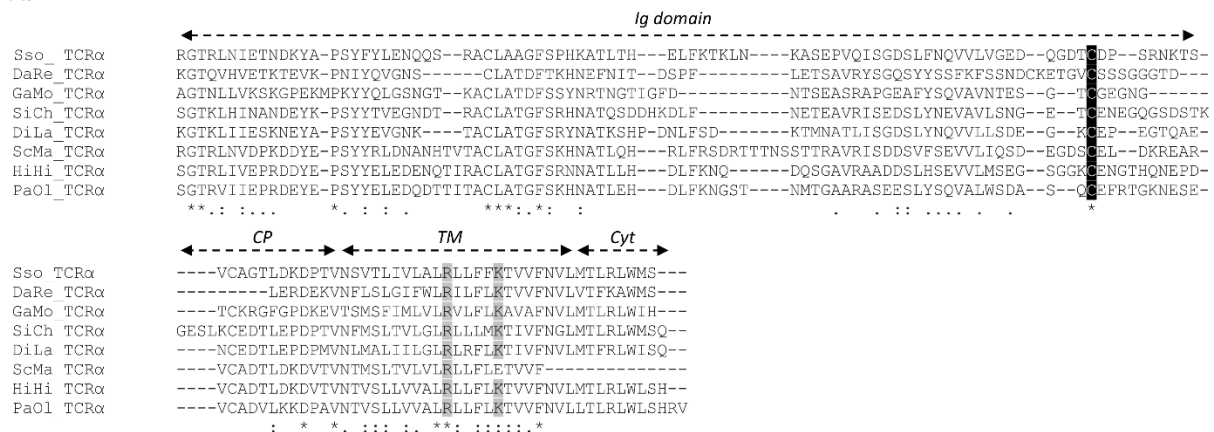
### 459 3.2.2. Antigen recognition

460 Antigen-specific T-cell responses are initiated through the interaction of a T cell antigen receptor  
 461 (TCR) and the corresponding peptide-MHC protein complex expressed on cell surface. TCR $\alpha$  and  
 462 TCR $\beta$  genes have been characterized in both teleost and cartilaginous fishes (Castro et al. 2011 for a  
 463 review). In the present study, one transcript (Isotig02715) was identified coding TCR $\alpha$  Constant  
 464 region (C $\alpha$ ) and it exhibits 62% identity at the protein level with the orthologous sequence in *H.*  
 465 *hippoglossus* (ACY54771.1). *S. solea* C $\alpha$  possess several conserved structurally important features,  
 466 such as a Cys residue in the CP domain, thought to form an inter-chain disulphide bond with the  
 467 TCR $\beta$  chain in mammals, and the positively charged residues arginine (Arg) and lysine (Lys) in the

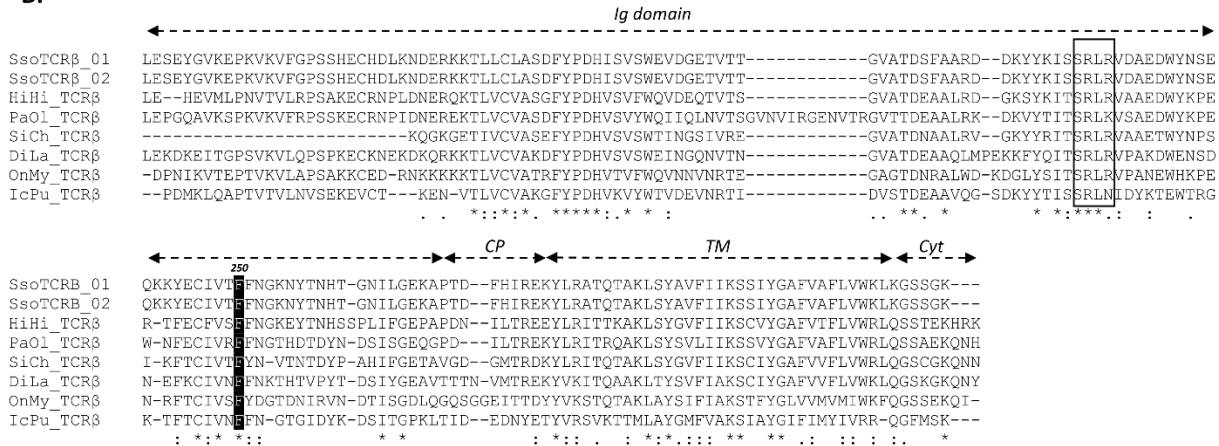
468 TM region, involved in the assembly and the surface expression of the TCR-CD3 complex (see Figure  
469 6A).

470 Similarity analyses of *S. solea* transcriptome identified also two different TCR $\beta$  chains, SsoTCR $\beta$ \_01  
471 and SsoTCR $\beta$ \_02 (Isotig00319 and Isotig00320, respectively). The two cDNAs, that codify the TCR  
472 Constant region (C $\beta$ ), share 81% identity at the protein level and are closely related to *H. hippoglossus*  
473 TCR $\beta$  (ACY54772.1, 63-61% aa identity). *S. solea* TCR $\beta$  Ig domains possess the conserved  
474 aminoacid motif S-R-L-R, which is required to form the TCR  $\alpha\beta$  heterodimer, and the phenylalanine  
475 (Phe<sub>250</sub>) residue, which is considered important for TCR recognition (see Figure 6B).  
476

#### A.



#### B.



477

478

479 **Figure 6. Multiple alignment of teleost TCR constant regions. A)** Alignment of the deduced amino acid sequences of  
480 *S. solea* (Sso) and other teleost TCR  $\alpha$  regions. Published TCR $\alpha$  sequences described from *Danio rerio* (DaRe,  
481 AAG31714.1), *Paralichthys olivaceus* (PaOl, BAB82531.1), *Gadus morhua* (GaMo, CAD28810.1), *Siniperca chuatsi*  
482 (SiCh, ABV60622.1), *Dicentrarchus labrax* (DiLa, AAV88084.1), *Scophthalmus maximus* (ScMa, AAP76193.1),  
483 *Hippoglossus hippoglossus* (HiHi, ACY54771.1) were included in the alignment. Highlighted in grey are residues  
484 involved in the assembly of the TCR-CD3 complex while conserved Cys residue is black shadowed. **B)** Alignment of the  
485 deduced amino acid sequences of *S. solea* (Sso) and other teleost TCR  $\beta$  regions. Published TCR $\beta$  sequences described  
486 from *Hippoglossus hippoglossus* (HiHi, ACY54772.1), *Paralichthys olivaceus* (PaOl, BAB82607.1), *Siniperca chuatsi*  
487 (SiCh, ABV60623.1), *Dicentrarchus labrax* (DiLa, CBK52290.1), *Oncorhynchus mykiss* (OnMy, AAA92558.1),  
488 *Ictalurus punctatus* (IcPu, AAL07367.1) were included in the alignment. The core motif S-R-L-R is boxed while  
489 conserved Phe residue is highlighted in black. In both alignments, sequences are defined by the four-letter species  
490 abbreviation. At the bottom of each alignment, identical residues are indicated by " \* ", conservative and less conservative  
491 residues are indicated by " : " and " . ", respectively.  
492

493 T cell signaling complex in fish seems to include counterparts of mammalian CD3 $\epsilon$  and CD3Z  
494 (mainly known as CD247). This complex plays an important role in coupling antigen recognition to  
495 several intracellular signal-transduction pathways. In the present study two full-length transcripts,  
496 one (isotig06954) coding SsoCD3 $\epsilon$  and one (isotig14459) coding SsoCD3Z, were identified showing  
497 52% identity at protein level with *H. hippoglossus* CD3 $\epsilon$  (ACY54760.1) and 68% identity with *L.*  
498 *crocea* CD3Z (XP\_010753352), respectively.

499 Changes in expression of *S. solea* TCR $\alpha$ , TCR $\beta$  CD3 $\epsilon$ , and CD3Z during development were also  
500 assessed (see Figure 5), and all transcripts were proven to exhibit the same trend, with mRNA levels  
501 almost undetectable soon after hatching followed by a gradual increase over time with fold-changes  
502 at 33 dph ranging from 5.4- (CD3Z) to 217.8-fold (TCR $\beta$ ).

503

504

#### 505 4. DISCUSSION

506 Since the early stages, fish life occurs in the open environment. In particular, during the first critical  
507 transitions of larval development (i.e. hatching, mouth opening, first feeding) interactions between  
508 the developing teleost immune system and the environment intensify, increasing dramatically the  
509 need for efficient defence mechanisms against pathogens.

510 In the last decade, extensive numbers of studies have been conducted in order to improve the  
511 understanding of teleost immune system. Transcriptome sequencing through NGS strategies allowed  
512 a rapid increase of sequence information also in non-model species (Huang et al. 2015, Ali et al. 2014,  
513 Tong et al. 2015, Xia et al. 2010), however studies have been mainly limited to identify immune  
514 relevant genes and/or characterize their response to pathogen invasions, while little is known about  
515 their temporal appearance during development as a whole.

516 To date, information on the ontogeny of the teleost immune system is largely restricted to a few fish  
517 species and usually to a limited number of genes (Cecchini et al. 2013, Seppola et al. 2009, Rise et  
518 al. 2012, Nayak et al. 2011, Picchiatti et al. 2009, Broekman et al. 2011), making it difficult a  
519 comparison between different species.

520 Herein, similarity searches led to the identification of 149 immune-related isotigs/contigs in a total of  
521 225,944 *S. solea* sequences. Many of these transcripts are represented by complete cDNA sequences  
522 that were identified for the first time in this species. The identification of such a repertoire in common  
523 sole provided the basis for studying the transcriptional profiles of many immune components across  
524 ten (10) larval stages, from hatching to accomplishment of the juvenile form, and to address if genes  
525 that have an important role in immune defense are present early during development.

#### 526 *Innate immune elements*

527 During evolution, marine eukaryotes have developed a wide array of anti-infective molecules and  
528 strategies by which they protect themselves against prokaryotic and viral attack (Smith et al. 2010).  
529 Innate immune system is the earliest immune mechanism that acts as first line of defense against all  
530 pathogens, and protect the host from invasion by other organisms in a nonspecific manner. It is known  
531 that teleost immunological capacity is limited in early embryos. The immune system is not completely  
532 developed during embryonic and larval stages, thus conferring to innate mechanisms a key role for  
533 survival.

534 Here, the temporal appearance during larval development of relevant members of innate immune  
535 repertoire, such as lysozyme, AMPs, PPRs and complement has been assessed.

536 Five lysozyme transcripts including three c-type and two g-type lysozymes have been identified in *S.*  
537 *solea*. Only a few studies assessed the ontogeny of lyz genes in teleost fish (Seppola et al. 2009,  
538 Nayak et al. 2011) reporting, in particular for g-type lysozyme, its transcription from embryos



539 onwards. A confirmation was found in the present study with SsoLYZs mRNA levels detectable from  
540 hatching and, with gene-specific trends, generally increasing after the start of the exogenous feeding.  
541 AMPs are host produced low molecular weight peptides (or small proteins), typically exerting a  
542 potent broad-spectrum activity against pathogenic bacteria, fungi, parasites, and viruses (Cole et al.  
543 2000, Dezfuli 2010). Extensively studied in mammals, amphibians and invertebrates, these molecules  
544 have recently received attention in teleost, leading to the identification of many different families of  
545 AMPs in the last two decades (Smith et al. 2010 for a review). In the present study two classes of  
546 AMPs, hepcidin and  $\beta$ -defensin, have been identified in *S. solea*.

547 Hepcidin (HAMP) is a small cysteine-rich protein with not only antimicrobial activity but also pivotal  
548 function in iron homeostasis (Wang et al. 2012). It is known that HAMP genes can be distinguished  
549 into two isoforms with specific activities (Neves et al. 2015, Hilton et al. 2008). Due to the incomplete  
550 sequence information, no phylogenetic analyses were conducted on SsoHAMP transcripts, preventing  
551 us from defining whether this/these peptide/s belong to hamp1- or hamp2-type isoforms, the latter  
552 exerting primarily an antimicrobial function.

553 SsoHAMPs gene expression patterns during larval development show higher expression levels at  
554 hatching and a new increase in expression from 8 dph. This totally reflect what already observed in  
555 the Atlantic cod (i.e. *Gadus morhua*, Seppola et al. 2009) and in the half-smooth tongue sole (i.e. *C.*  
556 *semilaevis*, Wang et al. 2012), while is markedly different to what reported for the Medium Carp  
557 homologue (i.e. *Puntius sarana*, Das et al. 2015), where a stable expression from 6 h post fertilization  
558 onwards was observed. This might be due to isoform-specific differences in expression profiles,  
559 already described in teleost healthy and challenged tissues (Neves et al. 2015), as well as during  
560 ontogenesis (Martin-Antonio et al. 2009). In both Atlantic cod (Seppola et al. 2009), half-smooth  
561 tongue sole (Wang et al. 2012) and redbanded seabream (Martin-Antonio et al. 2009), some HAMP  
562 isoforms are expressed in embryos and at hatching followed by a decrease in expression in later  
563 stages, leading to the hypothesis that those genes might play a role in embryogenesis and early larval  
564 development, likely through their known function in iron homeostasis (Martin-Antonio et al. 2009,  
565 Shi et al. 2006).

566 Under this scenario, even for *S. solea* HAMPs, a non immunological function might explain their  
567 higher expression at hatching compared to immediately following stages, while the second increase  
568 of transcript levels is ascribable to the growing needs for antimicrobial defense at the commencement  
569 of exogenous feeding.

570 Beta-defensins are antimicrobial peptides that have an important role in innate immune responses at  
571 epithelial barriers. Herein, three full-length transcripts coding  $\beta$ -defensin have been identified, and  
572 like their counterpart in higher vertebrates and other teleost species (Zou et al. 2007, Casadei et al.

2009, Nam et al. 2010), they fall into two subgroups based on phylogenetic analysis. Noteworthy, *in silico* analysis of Sso $\beta$ DEFs protein sequences highlighted that Sso $\beta$ DEF2 possesses an anionic net charge.

In general, defensins are known as cysteine-rich cationic antimicrobial peptides. However, non-cationic defensins have been found in insects (Lai et al. 2004) and, recently, in *P. olivaceus* (Nam et al. 2010), where they have been reported to exert antimicrobial activity against Gram-positive bacteria, even though their mechanism of action is still unknown. In the present study, phylogenetic analysis reported Sso $\beta$ DEF2 as closely related to non-cationic *P. olivaceus*  $\beta$ -defensins, but further investigations are required to verify if these sequence features contribute to their antimicrobial activities.

$\beta$ -defensin gene expression in teleost has been assessed in tissues from healthy individuals (Zou et al. 2007) and/or following bacterial challenge (Casadei et al. 2009) reporting that, in the same species and under the same conditions, members of  $\beta$ -defensin family show different tissue distribution and patterns of expression (Zuo et al. 2007, Casadei et al. 2009). The same evidence was observed in the present study with Sso $\beta$ DEF transcripts showing markedly different patterns of expression during development, suggesting a distinct role of these peptides in host defense.

Although different trends of expression were observed, all Sso $\beta$ DEF transcripts show high transcription levels across all larval stages.  $\beta$ -defensin constitutive expression during early developmental stages was already reported in another teleost species (*i.e.* *P. olivaceus*, Nam et al. 2010), in insects (Wen et al. 2009) and higher vertebrates (Meyerholz et al. 2004, Meada et al. 2009). In human,  $\beta$ -defensins have also been proven *in vitro* to participate in keratinocytes growth/differentiation and in maturation of skin DCs (Frye et al. 2001, Ferris et al. 2013). In this context, expression patterns observed herein might suggest that these peptides, in addition to a primary function in protecting newly hatched fish from pathogenic assault, might play a role even in cellular growth/differentiation in developing fish larvae.

Vertebrates have evolved a vast array of both extracellular and intracellular pathogen recognition receptors (PRRs) for detecting and responding to pathogen-associated molecular patterns (PAMPs) (Hansen et al. 2011), and their role is critical to the initiation of both innate and adaptive immune responses.

Nod-like receptors (NLRs) constitute a multigene family of cytoplasmic PRRs in which three subfamilies can be distinguished; i) NLR-A, putative orthologous of mammalian NODs; ii) NLR-B, putative orthologous of mammalian NALPs, and iii) NLR-C that appears to be unique of teleost fish and is currently still not well defined (Zhu et al. 2013, Laing et al. 2008).

606 Herein, five different *S. solea* NLRs were identified, all showing consistent similarities to teleost  
607 NLR-C family members (data not shown). In human, Nod-like receptors have been proposed playing  
608 a crucial role not only in innate immune responses, but also in apoptosis and early development (Kim  
609 et al. 2016). All SsoNLRs are expressed early during development and their expression levels  
610 constantly increase over time. To our knowledge this is the first time that NLRs ontogeny is  
611 investigated in both teleost and higher vertebrates, and the present findings suggest that these  
612 receptors might take part in immune repertoire maturation and larval development.

613 Within PRRs, Toll-like receptors play a pivotal role in early recognition of pathogens as well as in  
614 the initiation of a robust and specific adaptive immune response. At least 10 TLRs have been  
615 identified in fish (Palti 2011, Pietretti et al. 2014) from several different orders including  
616 Pleuronectiformes (Yu et al. 2009, Hwang et al. 2010), and have been proven to be transcriptionally  
617 active even during embryogenesis in both teleost and higher vertebrates (Yu et al. 2009, Peterson et  
618 al. 2005, Kannaki et al. 2015).

619 Despite targeted Blast searches were carried out using entire TLR and TIR domain sequences from  
620 several teleost species, in the present study no putative TLR transcripts have been detected in *S. solea*  
621 larval transcriptome. Nevertheless some effectors molecules, such as TIRAP, IRAK4 and PI3K, that  
622 have an essential role in the TLR signaling cascade (Kawagoe et al. 2007), were found. Even though  
623 IRAK4 and PI3K are involved also in other biological processes such as IL1 signaling pathway  
624 (Suzuki et al. 2002, Sizemore et al. 1999), TIRAP seems to be specifically involved in signaling  
625 through the members of the TLR family and it could provide an indirect evidence of the  
626 transcriptional activation of TLR repertoire in common sole even if we could not detect any related  
627 sequence in our search.

628 A possible explanation of the lack of TLRs transcripts in our data could be given by the low  
629 sequencing depth of the obtained sole transcriptome. At the time of *S. solea* transcriptome sequencing  
630 (2011) the Roche 454 GS FLX technology was chosen mainly because of the average read length  
631 (300-600 nt), significantly longer compared to other platforms (Illumina and Solid) and more suitable  
632 for *de novo* sequencing/assembly. However, 454 technology produces 0.4-0.6 Gb per run that can be  
633 converted into an extremely low number of reads/run, explaining the difficulty of  
634 detecting/sequencing low represented transcripts, even if normalized cDNA libraries have been  
635 employed as template. If ever formulated, the hypothesis of the absence of TLR repertoire in *S. solea*  
636 was definitely ruled out by the study from Benzekri et al (2014) in which at least 7 TLRs were  
637 identified.

638 In recent years, a new multigene family encoding non-rearranging receptors belonging to  
639 immunoglobulin superfamily, called Novel immune-type receptors (NITRs), has been discovered and

640 their occurrence seems to be restricted to teleost species (Ferraresso et al. 2009, Yoder et al. 2010,  
641 Meng et al. 2014). Herein, NITR transcripts have been identified in *S. solea* larvae, unfortunately  
642 they were represented by only partial sequences and, so far, no further analyses have been conducted.  
643 However, this represents the first evidence of the presence of multiple NITR genes on Soleidae  
644 family. Recognition of potential pathogens can be mediated also by some humoral lectins, such as  
645 pentraxins, that function as effector factors by promoting their phagocytosis and activating the  
646 complement system. In vertebrates, pentraxin family is generally composed by two members, C-  
647 reactive protein (CRP) and Serum amyloid P (SAP).

648 CRP and SAP share extensive sequence homology, indicating that they evolved from a common  
649 ancestor and often, in teleost, these proteins are generally named as "pentraxin" or "pentraxin-like".  
650 However, CRPs and SAPs vary considerably in terms of ligand-binding specificity and in some  
651 species they were characterized based on their substrate selectivity (Huong Ciang et al. 2010). In the  
652 present study, three transcripts coding pentraxin family members were identified, having best match  
653 with either teleost CRP or SAP. To properly classify *S. solea* pentraxins, a phylogenetic analysis was  
654 conducted by using sequences of both CRP and SAP from several mammals and teleost. Phylogenetic  
655 tree showed that all pentraxin sequences grouped in two clusters, corresponding to mammalian SAPs  
656 and CRPs, with high bootstrap support, providing high confidence in the *S. solea* gene identity as  
657 initially revealed by BLAST searches (data not shown). To our knowledge, this is the first time that  
658 the transcriptional onset of both CRP and SAP transcripts have been investigated in teleost during  
659 larval development. The only available information is by Seppola and colleagues (2009), that  
660 investigated the gene expression profile of a pentraxin transcript during cod development. This study  
661 reported expression profiles comparable to *S. solea* homologues, with increased mRNA levels during  
662 embryo development, a drop in expression soon after hatching and a second increase during  
663 metamorphosis. The meaning of this trend of expression is unclear. In mammals, it has been reported  
664 that some pentraxin members are involved also in synapse formation and remodeling (Boulanger  
665 2009). During metamorphosis, flatfishes dramatically transform from a symmetrical to an  
666 asymmetrical body shape, accompanied by eye migration, cranium deformation and tissue  
667 remodeling. In this context it is not unlikely that *S. solea* pentraxin play a role in neurological  
668 processes that take place during both early embryogenesis and sole metamorphosis.

669 The complement system modulates several physiological processes, from inflammation and pathogen  
670 opsonization to hematopoiesis, tissue regeneration and lipid metabolism (Forn-Cuni et al. 2014). In  
671 particular, it plays a central role in early pathogen defense and bridges the innate with the adaptive  
672 immune response (Sunyer, et al. 2003). There are three distinct pathways through which complement  
673 can be activated on pathogen surfaces; i) the Classical pathway, ii) the Alternative pathway and the

674 ii) Mannan-binding (MB)-lectin pathway (Janeway et al. 2001). These pathways depend on different  
675 molecules for their initiation, but they converge to generate the same set of effector molecules  
676 (terminal complement components). Multiple forms of complement component genes have been  
677 previously identified in fish, suggesting a key role for this system in aquatic organisms (Forn-Cuni et  
678 al. 2014).

679 In the present study, the main components of both classical, alternative and MB-lectin pathways have  
680 been identified, clearly indicating that all three pathways of the complement cascade are represented  
681 in the common sole. As expected, almost all members of the complement system, terminal  
682 components included, are expressed early during development and show an increase in expression  
683 over time, in particular after the start of the exogenous feeding, in a way totally compatible with the  
684 growing needs for efficient defense mechanisms against pathogens that accompany larval  
685 development. The only exception is represented by C1q, C4 and MASP-1 transcripts, main  
686 components of the classical and MB-lectin pathways, that exhibit no difference in expression,  
687 however they have a high/moderate constitutive expression in all larval stages analyzed, proving  
688 evidence that these pathways are already "active" during early development.

689 The third complement component (C3) is a central protein in all three overlapping pathways of  
690 complement activation (Wang et al. 2015). Herein, we report the identification of four different  
691 complement C3 transcripts in the common sole transcriptome. As already observed in other teleost  
692 species (Forn-Cuni et al 2014, Lange et al. 2004, Lovoll et al. 2007), gene expression analysis of  
693 SsoC3 genes during ontogeny showed for all transcripts similar patterns of expression with mRNA  
694 levels increasing soon after mouth opening (4 dph). Of particular interest, in both *S. solea* and  
695 zebrafish (Forn-Cuni et al 2014), a decreased expression of C3 transcripts soon after hatching can be  
696 appreciated, suggesting for this molecule a non-immunological function during embryo development  
697 as already demonstrated in mammalian species (Carmona-Fontaine et al. 2011, Jeanes et al. 2015).

698 Cytokines are a broad group of small proteins that act as intercellular mediators to regulate the  
699 immune response and are known acting as a bridge between innate and adaptive response. Release of  
700 pro-inflammatory cytokines is essential for the activation of an effective innate host defense, and  
701 subsequently for the modulation of adaptive immune responses. Among these, interleukin-1 $\beta$  is one  
702 of the most studied pro-inflammatory cytokine, mainly produced by activated macrophages and blood  
703 monocytes (Lu et al. 2008). Once secreted, it initiates and enhances the inflammatory response by  
704 inducing the expression of pro-inflammatory molecules and adhesion molecules in diverse  
705 stromal/inflammatory cells (Apte et al. 2006). In addition to its function in inflammation, IL1 $\beta$  plays  
706 a pivotal role in the proliferation and differentiation of cells of both innate and adaptive immunity  
707 (Lu et al. 2008). Ontogenesis of IL1 $\beta$  was recently investigated in zebrafish by Dios et al. (2010) that

708 reported its transcriptional onset at 8 dpf. This is in agreement with our findings (mRNA levels barely  
709 detectable at hatching and increasing thereafter) and it is consistent with the primary role of IL1 $\beta$  in  
710 maturation of the immune system.

711 Contrary to IL1 $\beta$ , in the present study other crucial pro-inflammatory cytokines such as TNF $\alpha$  and  
712 IFNs were not identified. This can be linked to the biological function of these molecules. IFN $\gamma$  is  
713 known to be involved in both innate and adaptive immunity (Robertsen 2006), nevertheless it is  
714 mainly secreted by activated NK cells and T-helper lymphocytes, thus it is likely that in both  
715 developmental stages and (not inflammatory) conditions employed for cDNA library construction  
716 this transcript was expressed at low levels and therefore not detectable. We could argue the same for  
717 TNF $\alpha$ , produced by a wide array of cells but, usually, mainly in response to inflammatory stimuli.  
718 Numerous chemokine genes have been identified in teleost species and they exert a vast array of  
719 functions yet not fully elucidated. In addition to their immunological function, chemokines have a  
720 role in organogenesis and have been demonstrated to regulate homing, maturation and even micro-  
721 environmental segregation within lymphoid organs (Alejo and Tafalla 2011 for a review). Herein, at  
722 least ten chemokines were found showing different trends in expression, however they are all  
723 characterized by moderate/high transcript levels even at early stages thus supporting the hypothesis  
724 that at least some of these molecules may have an additional role in larval development.

#### 725 *Acquired immune elements*

726 All jawed vertebrates possess the genetic elements essential for the functioning of adaptive immune  
727 responses (Rauta et al. 2012). The hallmark of the adaptive immunity is the ability to mount a highly  
728 specific response against virtually any invading organism. This specificity is mainly determined by  
729 B or T lymphocytes and requires a number of key molecules expressed on effector leukocytes and  
730 target cells, such as the T cell receptor (TCR), its accessory molecules CD4 and CD8, and Class I or  
731 Class II major histocompatibility complex (MHC) molecules.

732 Class I and class II MHC play a pivotal role in mediating antigen recognition by lymphocytes through  
733 the interaction with TCR molecules. Until now, MHC genes have been isolated and characterized in  
734 various fish species, including zebrafish (*D. rerio*), rainbow trout, channel catfish, turbot, Nile tilapia  
735 (*O. niloticus*), sea bass and halibut (*C. semilaevis*) (Pang et al. 2013).

736 In the present study, we identified and characterized three transcripts coding Class I MHC  $\alpha$  chain,  
737 as well as two transcripts coding MHCII $\alpha$  and three MHCII $\beta$  molecules. SsoMHCs all share the main  
738 features conserved across vertebrates and phylogenetic analysis confirmed their close relationships  
739 with those of teleost belonging to the order Pleuronectiformes.

740 Antigens bound to either class I or class II MHC proteins are then recognized by lymphocytes through  
741 the TCR. Little is known about the development of T-cells and the expression of T-cell markers

742 during development in Solenidae, and the information is restricted to a few species in teleost  
743 (Picchietti et al. 2009, Øvergård et al. 2011). In the present study three genes important in T-cell  
744 differentiation, TCR $\alpha$ , TCR $\beta$  and CD3 $\epsilon$ , were identified and their expression assessed during *S. solea*  
745 development.

746 SsoMHC and SsoTCR transcripts differ in their expression from those of the innate immune system  
747 being barely detectable soon after hatching and showing a marked increase at later stages. This is not  
748 unexpected since it was already reported in other teleost species that the major immune events leading  
749 to immunocompetence coincide with metamorphosis and with the larvae to juvenile transitory phase  
750 (Dios et al. 2010, Picchietti et al. 2015). This is consistent with the timing of the main lymphoid  
751 organs appearance. To date, no reports can be found in literature upon the ontogenetic development  
752 of *S. solea* lymphoid organs, while a few information upon the appearance of hematopoietic cells in  
753 immune organs is available for the closely related species Senegalese sole (*Solea senegalensis*, Cunha  
754 et al. 2008). The timing at which the haematopoietic elements could be seen coincided with the  
755 development of the thymus and the start of exogenous feeding, suggesting that the same can  
756 reasonably happens in the common sole.

757 There is increasing evidence demonstrating a cross-talk between neuroendocrine and immune system.  
758 In teleost fish numerous studies already revealed an important immunomodulatory role of GH/IGFI  
759 axis (Welniak et al. 2002, Yada 2007, Franz et al. 2016), in particular during thymus and thymocytes  
760 development, through both endocrine and paracrine mechanism. In higher vertebrates it has been also  
761 demonstrated that thyroid hormones (THs) modulate specific immune responses, including cell-  
762 mediated immunity, lymphocyte differentiations and proliferation, and natural killer cell activity  
763 (Hodkinson et al. 2009), while thyrotropin (TSH) acts as a growth factor for developing T cells (van  
764 der Weerd et al. 2014). Although there is scarce information about a possible role of THs in fish  
765 immune system, first evidence appear to prove that immune organs and cells are responsive to THs  
766 (Quesada-Garcia et al. 2014) and that a relationship between thyroid and thymus development and  
767 lymphopoiesis is conserved even in teleost (Lam et al. 2005).

768 In *S. solea* the ontogeny during larval development of both GH/IGFI axis and TH cascade was  
769 previously reported (Ferraresso et al. 2013) and their expression was proven to be concurrent with  
770 larval organogenesis and the onset of metamorphosis. In addition, temporal regulation of their  
771 transcript levels is coherent with a stimulatory role of thymocytes as deduced by looking at SsoTCR  
772 and SsoCD3 $\epsilon$  gene expression profiles. Taken together these results might provide indirect evidence  
773 of a hormonal modulation of adaptive immune system in *S. solea*.

774 It is worth of notice that SsoTCR $\beta$  expression is detectable earlier than TCR $\alpha$  proving that *S. solea*  
775 TCR $\beta$  is the first to be rearranged and expressed on the surface of thymocytes, as observed in other  
776 teleost species (Øvergård et al. 2011).

777 In the present study, neither RAG1 (Recombination-activating 1), that plays an essential role in TCR  
778 locus rearrangement, nor CD4 and CD8, key markers of T-cell differentiation, were found.

779 The onset of CD4 and/or CD8 positive cells is correlated to the maturation of the thymus, and appear  
780 notably later compared to the appearance of double negative cells. In addition, it was already reported  
781 that CD4 and CD8 expression is delayed in comparison with RAG1 and TCR genes (e.g. ~25 days in  
782 *D. labrax*, Picchietti et al. 2009). In this context, the failure to identify/detect CD4 and CD8 transcripts  
783 might suggest that the developmental stages analyzed were too early for T-cells having undergone a  
784 full maturation. In addition, it should be taken into account that whole larvae were employed for  
785 transcriptome sequencing, thus making more difficult the detection of tissue- or cell-specific genes.

786

## 787 5. CONCLUSIONS

788 Young fish use innate mechanisms during the first weeks of their development. In this respect,  
789 ontogenetic studies are important to understand pathways appearance and to defend farmed fish  
790 against pathogens at early age. The present study provides a temporal comparative overview on the  
791 transcriptional onset of most major proteins of innate and adaptive immunity during early  
792 development of *S. solea*, a commercially important species.

793 We demonstrated that the main players of the innate immunity are conserved and already expressed  
794 at first larval stages providing protection from environmental pathogens even at early development.  
795 Many of these genes show higher transcript levels at hatching, leading to the hypothesis that they  
796 might play a role during embryogenesis, and their expression increase dramatically as the needs for  
797 antimicrobial defense grow (e.g. start of exogenous feeding).

798 Adaptive immune transcripts differ in their expression from those of the innate immune system, their  
799 onset coincides with metamorphosis and larvae to juvenile transition, and likely overlaps with the  
800 appearance and maturation of the main lymphoid organs. However, data collected suggest that at the  
801 end of metamorphosis *S. solea* cell-mediated immune system hasn't still undergone full maturation.

802



803 **REFERENCES**

- 804 Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution.  
805 Bioinformatics. 2005;21(9):2104–2105.
- 806 Alejo A, Tafalla C. Chemokines in teleost fish species. Dev Comp Immunol. 2011;35(12):1215-  
807 1222.
- 808 Ali A, Rexroad CE, Thorgaard GH, Yao J, Salem M. Characterization of the rainbow trout spleen  
809 transcriptome and identification of immune-related genes. Front Genet. 2014;5:348.
- 810 Apte RN, Dotan S, Elkabets M, White MR, Reich E, Carmi Y, Song X, Dvozkin T, Krelin Y,  
811 Voronov E. The involvement of IL-1 in tumorigenesis, tumor invasiveness, metastasis and tumor-  
812 host interactions. Cancer Metastasis Rev. 2006;25(3):387-408.
- 813 Benzekri H, Armesto P, Cousin X, Rovira M, Crespo D, Merlo MA, Mazurais D, Bautista R,  
814 Guerrero-Fernández D, Fernandez-Pozo N, Ponce M, Infante C, Zambonino JL, Nidelet S, Gut  
815 M, Rebordinos L, Planas JV, Bégout ML, Claros MG, Manchado M. De novo assembly,  
816 characterization and functional annotation of Senegalese sole (*Solea senegalensis*) and common  
817 sole (*Solea solea*) transcriptomes: integration in a database and design of a microarray. BMC  
818 Genomics. 2014;15:952.
- 819 Boulanger LM. Immune proteins in brain development and synaptic plasticity. Neuron.  
820 2009;64(1):93-109.
- 821 Broekman DC, Frei DM, Gylfason GA, Steinarsson A, Jörnvall H, Agerberth B, Gudmundsson  
822 GH, Maier VH. Cod cathelicidin: isolation of the mature peptide, cleavage site characterisation  
823 and developmental expression. Dev Comp Immunol. 2011;35(3):296-303.
- 824 Carmona-Fontaine C, Theveneau E, Tzekou A, Tada M, Woods M, Page KM, Parsons M,  
825 Lambris JD, Mayor R. Complement fragment C3a controls mutual cell attraction during collective  
826 cell migration. Dev Cell. 2011;21(6):1026-1037.
- 827 Casadei E, Wang T, Zou J, González Vecino JL, Wadsworth S, Secombes CJ. Characterization  
828 of three novel beta-defensin antimicrobial peptides in rainbow trout (*Oncorhynchus mykiss*). Mol  
829 Immunol. 2009;46(16):3358-3366.
- 830 Castro R, Bernard D, Lefranc MP, Six A, Benmansour A, Boudinot P. T cell diversity and TcR  
831 repertoires in teleost fish. Fish Shellfish Immunol. 2011;31(5):644-654.

- Cecchini S, Paciolla M, Biffali E, Borra M, Ursini MV, Lioi MB. Ontogenetic profile of innate immune related genes and their tissue-specific expression in brown trout, *Salmo trutta* (Linnaeus, 1758). *Fish Shellfish Immunol.* 2013;35(3):988-92.
- Cole AM, Darouiche RO, Legarda D, Connell N, Diamond G. Characterization of a fish antimicrobial peptide: gene expression, subcellular localization, and spectrum of activity. *Antimicrob Agents Chemother.* 2000;44(8):2039-2045.
- Cuesta A, Meseguer J, Esteban MA. Cloning and regulation of the major histocompatibility class I alpha gene in the teleost fish gilthead seabream. *Fish Shellfish Immunol.* 2007;22(6):718-726.
- Cunha MC, Makridis P, Soares F, Rodrigues P, Dinis MT. Timing of appearance of lymphoid cells during early development of senegalese sole, *Solea senegalensis* kaup. *Journal of the world aquaculture society.* 2008; 39(3):436-439.
- Das A, Mohapatra A, Sahoo PK. Cloning and Characterization of Antimicrobial Peptide, Hepcidin in Medium Carp, *Puntius sarana*. *Int J Pept Res Ther.* 2015;21:139–147.
- Dezfuli BS, Pironi F, Giari L, Noga EJ. Immunocytochemical localization of piscidin in mast cells of infected seabass gill. *Fish Shellfish Immunol.* 2010;28(3):476-482.
- Dios S, Romero A, Chamorro R, Figueras A, Novoa B. Effect of the temperature during antiviral immune response ontogeny in teleosts. *Fish Shellfish Immunol.* 2010;29(6):1019-1027.
- Falk-Petersen IB. Comparative organ differentiation during early life stages of marine fish. *Fish & Shellfish Immunology.* 2005; 19(5):397-412.
- Felsenstein J. *Inferring phylogenies.* 2004 Sinauer, Sunderland
- Ferraresso S, Bonaldo A, Parma L, Cinotti S, Massi P, Bargelloni L, Gatta PP. Exploring the larval transcriptome of the common sole (*Solea solea* L.). *BMC Genomics.* 2013;14:315.
- Ferraresso S, Kuhl H, Milan M, Ritchie DW, Secombes CJ, Reinhardt R, Bargelloni L. Identification and characterisation of a novel immune-type receptor (NITR) gene cluster in the European sea bass, *Dicentrarchus labrax*, reveals recurrent gene expansion and diversification by positive selection. *Immunogenetics.* 2009;61(11-12):773-788.
- Ferris LK, Mburu YK, Mathers AR, Fluharty ER, Larregina AT, Ferris RL, Falo LD Jr. Human beta-defensin 3 induces maturation of human langerhans cell-like dendritic cells: an antimicrobial peptide that functions as an endogenous adjuvant. *J Invest Dermatol.* 2013;133(2):460-468.

861 Fischer U, Dijkstra JM, Köllner B, Kiryu I, Koppang EO, Hordvik I, Sawamoto Y, Ototake M.  
862 The ontogeny of MHC class I expression in rainbow trout (*Oncorhynchus mykiss*). Fish Shellfish  
863 Immunol. 2005;18(1):49-60.

864 Forn-Cuní G, Reis ES, Dios S, Posada D, Lambris JD, Figueras A, Novoa B. The evolution and  
865 appearance of C3 duplications in fish originate an exclusive teleost c3 gene form with anti-  
866 inflammatory activity. PLoS One. 2014;9(6):e99673.

867 Franz AC, Faass O, Köllner B, Shved N, Link K, Casanova A, Wenger M, D'Cotta H, Baroiller  
868 JF, Ullrich O, Reinecke M, Eppler E. Endocrine and Local IGF-I in the Bony Fish Immune  
869 System. Biology (Basel). 2016;5(1):E9.

870 Frye M, Bargon J, Gropp R. Expression of human beta-defensin-1 promotes differentiation of  
871 keratinocytes. J Mol Med. 2001;79(5-6):275-282.

872 Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by  
873 maximum likelihood. Syst Biol. 2003;52:696–704.

874 Hansen JD, Vojtech LN, Laing KJ. Sensing disease and danger: a survey of vertebrate PRRs and  
875 their origins. Dev Comp Immunol. 2011;35(9):886-897.

876 Hilton KB, Lambert LA. Molecular evolution and characterization of hepcidin gene products in  
877 vertebrates. Gene. 2008;415(1-2):40-48.

878 Hodgkinson CF, Simpson EE, Beattie JH, O'Connor JM, Campbell DJ, Strain JJ, Wallace JM.  
879 Preliminary evidence of immune function modulation by thyroid hormones in healthy men and  
880 women aged 55-70 years. J Endocrinol. 2009;202(1):55-63.

881 Huang L, Li G, Mo Z, Xiao P, Li J, Huang J. De Novo assembly of the Japanese flounder  
882 (*Paralichthys olivaceus*) spleen transcriptome to identify putative genes involved in immunity.  
883 PLoS One. 2015;10(2):e0117642.

884 Huong Giang DT, Van Driessche E, Vandenberghe I, Devreese B, Beeckmans S. Isolation and  
885 characterization of SAP and CRP, two pentraxins from *Pangasianodon* (*Pangasius*)  
886 *hypophthalmus*. Fish Shellfish Immunol. 2010;28(5-6):743-753.

887 Hwang SD, Asahi T, Kondo H, Hirono I, Aoki T. Molecular cloning and expression study on  
888 Toll-like receptor 5 paralogs in Japanese flounder, *Paralichthys olivaceus*. Fish Shellfish  
889 Immunol. 2010;29(4):630-638.

890 Jeanes A, Coulthard LG, Mantovani S, Markham K, Woodruff TM. Co-ordinated expression of  
891 innate immune molecules during mouse neurulation. *Mol Immunol.* 2015;68:253-260.

892 C.A. Janeway, P. Travers, M. Walport, M. Shlomchik, *Immunobiology: the immune system in*  
893 *health and disease*, fifth ed., Garland Publishing, New York, 2001.

894 Kannaki TR, Reddy MR, Verma PC, Shanmugam M. Differential Toll-like receptor (TLR)  
895 mRNA expression patterns during chicken embryological development. *Anim Biotechnol.*  
896 2015;26(2):130-135.

897 Kawagoe T, Sato S, Jung A, Yamamoto M, Matsui K, Kato H, Uematsu S, Takeuchi O, Akira S.  
898 Essential role of IRAK-4 protein and its kinase activity in Toll-like receptor-mediated immune  
899 responses but not in TCR signaling. *J Exp Med.* 2007;204(5):1013-1024.

900 Kim YK, Shin JS, Nahm MH. NOD-Like Receptors in Infection, Immunity, and Diseases. *Yonsei*  
901 *Med J.* 2016;57(1):5-14.

902 Lai R, Lomas LO, Jonczyk J, Turner PC, Rees HH. Two novel non-cationic defensin-like  
903 antimicrobial peptides from haemolymph of the female tick, *Amblyomma hebraeum*. *Biochem J.*  
904 2004;379:681e5.

905 Laing KJ, Purcell MK, Winton JR, Hansen JD. A genomic view of the NOD-like receptor family  
906 in teleost fish: identification of a novel NLR subfamily in zebrafish. *BMC Evol Biol.* 2008;8:42.

907 Lam SH, Sin YM, Gong Z, Lam TJ. Effects of thyroid hormone on the development of immune  
908 system in zebrafish. *Gen Comp Endocrinol.* 2005;142(3):325-335.

909 Lange S, Bambir S, Dodds AW, Magnadóttir B. The ontogeny of complement component C3 in  
910 Atlantic cod (*Gadus morhua* L.)--an immunohistochemical study. *Fish Shellfish Immunol.*  
911 2004;16(3):359-367.

912 Løvoll M, Johnsen H, Boshra H, Bøggwald J, Sunyer JO, Dalmo RA. The ontogeny and  
913 extrahepatic expression of complement factor C3 in Atlantic salmon (*Salmo salar*). *Fish Shellfish*  
914 *Immunol.* 2007;23(3):542-552.

915 Lu DQ, Bei JX, Feng LN, Zhang Y, Liu XC, Wang L, Chen JL, Lin HR. Interleukin-1beta gene  
916 in orange-spotted grouper, *Epinephelus coioides*: molecular cloning, expression, biological  
917 activities and signal transduction. *Mol Immunol.* 2008;45(4):857-867.

918 Magnadóttir B. Innate immunity of fish (overview). *Fish Shellfish Immunol.* 2006;20(2):137-151.

Martin-Antonio B, Jiménez-Cantizano RM, Salas-Leiton E, Infante C, Manchado M. Genomic characterization and gene expression analysis of four hepcidin genes in the redbanded seabream (*Pagrus auriga*). Fish Shellfish Immunol. 2009;26(3):483-491.

Mazurais D, Ferrarresso S, Gatta PP, Desbruyères E, Severe A, Corporeau C, Claireaux G, Bargelloni L, Zambonino-Infante JL. Identification of hypoxia-regulated genes in the liver of common sole (*Solea solea*) fed different dietary lipid contents. Mar Biotechnol (NY). 2014;16(3):277-288.

Meada KG, Higgs R, Lloyd AT, Giles S, O'Farrelly C. Differential antimicrobial peptide gene expression patterns during early chicken embryological development. Dev Comp Immunol. 2009;33:516e24.

Meng F, Wang R, Xu T. Identification of 21 novel immune-type receptors in miiuy croaker and expression pattern of three typical inhibitory members. Dev Comp Immunol. 2014;45(2):269-277.

Meyerholz DK, Gallup JM, Grubor BM, Evans RB, Tack BF, McCray PB Jr, Ackermann MR. Developmental expression and distribution of sheep beta-defensin-2. Dev Comp Immunol. 2004;28(2):171-178.

Mulero I, Garcia-Ayala A, Meseguer J, Mulero V. Maternal transfer of immunity and ontogeny of autologous immunocompetence of fish: A minireview. Aquaculture. 2007; 268:244-250.

Nakao M, Tsujikura M, Ichiki S, Vo TK, Somamoto T. The complement system in teleost fish: progress of post-homolog-hunting researches. Dev Comp Immunol. 2011;35(12):1296-1308.

Nam BH, Moon JY, Kim YO, Kong HJ, Kim WJ, Lee SJ, Kim KK. Multiple beta-defensin isoforms identified in early developmental stages of the teleost *Paralichthys olivaceus*. Fish Shellfish Immunol. 2010;28(2):267-274.

Nayak SP, Mohanty BR, Mishra J, Rauta PR, Das A, Eknath AE, Sahoo PK. Ontogeny and tissue-specific expression of innate immune related genes in rohu, *Labeo rohita* (Hamilton). Fish Shellfish Immunol. 2011;30(4-5):1197-1201.

Neves JV, Caldas C, Vieira I, Ramos MF, Rodrigues PN. Multiple Hepcidins in a Teleost Fish, *Dicentrarchus labrax*: Different Hepcidins for Different Roles. J Immunol. 2015;195(6):2696-2709.

Øvergård AC, Fiksdal IU, Nerland AH, Patel S. Expression of T-cell markers during Atlantic halibut (*Hippoglossus hippoglossus* L.) ontogenesis. Dev Comp Immunol. 2011;35(2):203-213.

Palti Y. Toll-like receptors in bony fish: from genomics to function. Dev Comp Immunol. 2011;35(12):1263-72.

Pang JC, Gao FY, Lu MX, Ye X, Zhu HP, Ke XL. Major histocompatibility complex class IIA and IIB genes of Nile tilapia *Oreochromis niloticus*: genomic structure, molecular polymorphism and expression patterns. Fish Shellfish Immunol. 2013;34(2):486-496.

Parma L, Bonaldo A, Massi P, Yufera M, Martinez-Rodriguez G, Gatta PP. Different early weaning protocols in common sole (*Solea solea* L.) larvae: Implications on the performances and molecular ontogeny of digestive enzyme precursors. Aquaculture. 2013; 414:26-35.

Patel S, Sørhus E, Fiksdal IU, Espedal PG, Bergh O, Rødseth OM, Morton HC, Nerland AH. Ontogeny of lymphoid organs and development of IgM-bearing cells in Atlantic halibut (*Hippoglossus hippoglossus* L.). Fish Shellfish Immunol. 2009;26(3):385-395.

Pearson W. Finding protein and nucleotide similarities with FASTA. Curr Protoc Bioinformatics. 2004;3:Unit3.9.

Peterson BC, Bosworth BG, Bilodeau AL. Differential gene expression of IGF-I, IGF-II, and toll-like receptors 3 and 5 during embryogenesis in hybrid (channel x blue) and channel catfish. Comp Biochem Physiol A Mol Integr Physiol. 2005;141(1):42-47.

Picchietti S, Abelli L, Guerra L, Randelli E, Proietti Serafini F, Belardinelli MC, Buonocore F, Bernini C, Fausto AM, Scapigliati G. MHC II- $\beta$  chain gene expression studies define the regional organization of the thymus in the developing bony fish *Dicentrarchus labrax* (L.). Fish Shellfish Immunol. 2015;42(2):483-493.

Picchietti S, Guerra L, Buonocore F, Randelli E, Fausto AM, Abelli L. Lymphocyte differentiation in sea bass thymus: CD4 and CD8-alpha gene expression studies. Fish Shellfish Immunol. 2009;27(1):50-56.

Pietretti D, Wiegertjes GF. Ligand specificities of Toll-like receptors in fish: indications from infection studies. Dev Comp Immunol. 2014;43(2):205-222.

Quesada-García A, Valdehita A, Kropf C, Casanova-Nakayama A, Segner H, Navas JM. Thyroid signaling in immune organs and cells of the teleost fish rainbow trout (*Oncorhynchus mykiss*). Fish Shellfish Immunol. 2014;38(1):166-174.

978 Rauta PR, Nayak B, Das S. Immune system and immune responses in fish and their role in  
979 comparative immunity study: a model for higher organisms. *Immunol Lett.* 2012;148(1):23-33.

980 Rise ML, Hall JR, Alcock BP, Hori TS. Dynamic expression profiles of virus-responsive and  
981 putative antimicrobial peptide-encoding transcripts during Atlantic cod (*Gadus morhua*)  
982 embryonic and early larval development. *Gene* 2012;509:232–246.

983 Robertsen B. The interferon system of teleost fish. *Fish Shellfish Immunol.* 2006;20(2):172-191.

984 Rombout JH, Huttenhuis HB, Picchietti S, Scapigliati G. Phylogeny and ontogeny of fish  
985 leucocytes. *Fish Shellfish Immunol.* 2005;19(5):441-455.

986 Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L,  
987 Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient Bayesian phylogenetic inference and model  
988 choice across a large model space. *Syst Biol.* 2012;61(3):539-42.

989 Seppola M, Johnsen H, Mennen S, Myrnes B, Tveiten H. Maternal transfer and transcriptional  
990 onset of immune genes during ontogenesis in Atlantic cod. *Dev Comp Immunol.*  
991 2009;33(11):1205-1211.

992 Shi J, Camus AC. Hepcidins in amphibians and fishes: Antimicrobial peptides or iron-regulatory  
993 hormones? *Dev Comp Immunol.* 2006;30(9):746-755.

994 Sizemore N, Leung S, Stark GR. Activation of phosphatidylinositol 3-kinase in response to  
995 interleukin-1 leads to phosphorylation and activation of the NF-kappaB p65/RelA subunit. *Mol*  
996 *Cell Biol.* 1999;19(7):4798-4805.

997 Smith VJ, Desbois AP, Dyrynda EA. Conventional and unconventional antimicrobials from fish,  
998 marine invertebrates and micro-algae. *Mar Drugs.* 2010;8(4):1213-1262.

999 Sunyer JO, Boshra H, Lorenzo G, Parra D, Freedman B, Bosch N. Evolution of complement as  
1000 an effector system in innate and adaptive immunity. *Immunol Res.* 2003;27(2-3):549-564.

1001 Suzuki N, Suzuki S, Duncan GS, Millar DG, Wada T, Mirtsos C, Takada H, Wakeham A, Itie A,  
1002 Li S, Penninger JM, Wesche H, Ohashi PS, Mak TW, Yeh WC. Severe impairment of interleukin-  
1003 1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature.* 2002;416(6882):750-756.

1004 Tong C, Zhang C, Zhang R, Zhao K. Transcriptome profiling analysis of naked carp  
1005 (*Gymnocypris przewalskii*) provides insights into the immune-related genes in highland fish. *Fish*  
1006 *Shellfish Immunol.* 2015;46(2):366-377.

1007 van der Weerd K, van Hagen PM, Schrijver B, Heuvelmans SJ, Hofland LJ, Swagemakers SM,  
1008 Bogers AJ, Dik WA, Visser TJ, van Dongen JJ, van der Lelij AJ, Staal FJ. Thyrotropin acts as a  
1009 T-cell developmental factor in mice and humans.

1010 Wang H, Qi P, Guo B, Li J, He J, Wu C, Gul Y. Molecular characterization and expression  
1011 analysis of a complement component C3 in large yellow croaker (*Larimichthys crocea*). Fish  
1012 Shellfish Immunol. 2015;42(2):272-279.

1013 Wang Y, Liu X, Ma L, Yu Y, Yu H, Mohammed S, Chu G, Mu L, Zhang Q. Identification and  
1014 characterization of a hepcidin from half-smooth tongue sole *Cynoglossus semilaevis*. Fish  
1015 Shellfish Immunol. 2012;33(2):213-9.

1016 Welniak LA, Sun R, Murphy WJ. The role of growth hormone in T-cell development and  
1017 reconstitution. J Leukoc Biol. 2002;71(3):381-387.

1018 Wen H, Lan X, Cheng T, He N, Shiomi K, Kajiura Z, Zhou Z, Xia Q, Xiang Z, Nakagaki M.  
1019 Sequence structure and expression pattern of a novel anionic defensin-like gene from silkworm  
1020 (*Bombyx mori*). Mol Biol Rep. 2009;36(4):711-716.

1021 Xia JH, Yue GH. Identification and analysis of immune-related transcriptome in Asian seabass  
1022 *Lates calcarifer*. BMC Genomics. 2010;11:356.

1023 Yada T. Growth hormone and fish immune system. Gen Comp Endocrinol. 2007;152(2-3):353-  
1024 358.

1025 Yoder JA, Turner PM, Wright PD, Wittamer V, Bertrand JY, Traver D, Litman GW.  
1026 Developmental and tissue-specific expression of NITRs. Immunogenetics. 2010;62(2):117-122.

1027 Yu Y, Zhong Q, Li C, Jiang L, Yan F, Wang Z, Zhang Q. Isolation and characterization of Toll-  
1028 like receptor 9 in half-smooth tongue sole *Cynoglossus semilaevis*. Fish Shellfish Immunol.  
1029 2009;26(3):492-499.

1030 Zhu LY, Nie L, Zhu G, Xiang LX, Shao JZ. Advances in research of fish immune-relevant genes:  
1031 a comparative overview of innate and adaptive immunity in teleosts. Dev Comp Immunol.  
1032 2013;39(1-2):39-62.

1033 Zou J, Mercier C, Koussounadis A, Secombes C. Discovery of multiple beta-defensin like  
1034 homologues in teleost fish. Mol Immunol. 2007;44(4):638-647.

1035



1036 **ACKNOWLEDGEMENTS**

1037 This research did not receive any specific grant from funding agencies in the public, commercial, or  
1038 not-for-profit sectors.

1039

1040 **COMPETING INTERESTS**

1041 The authors declare that they have no competing interests.

1042

1043 **SUPPLEMENTARY MATERIAL CAPTIONS**

1044 **Supplementary file 1.** List of immune-related genes identified in *S. solea* larval transcriptome

1045 **Supplementary file 2.** Multiple alignment of teleost CTSL (Figure S2.1) and CRP (Figure S2.2)  
1046 genes.

1047 **Supplementary file 3.** Phylogenetic analysis of teleost MHCII $\alpha$  (Figure S3.1), MHCII $\alpha$  (Figure S3.2)  
1048 and MHCII $\beta$  (Figure S3.3) genes.

1049