Accepted Manuscript

Title: IL-17 signaling components in bivalves: comparative sequence analysis and involvement in the immune responses

Author: Umberto Rosani, Laura Varotto, Marco Gerdol, Alberto Pallavicini, Paola Venier

PII: S0145-305X(15)00096-8
DOI: http://dx.doi.org/doi:10.1016/j.dci.2015.05.001
Reference: DCI 2390

To appear in: *Developmental and Comparative Immunology*

Received date: 31-3-2015
Revised date: 30-4-2015
Accepted date: 1-5-2015

Please cite this article as: Umberto Rosani, Laura Varotto, Marco Gerdol, Alberto Pallavicini, Paola Venier, IL-17 signaling components in bivalves: comparative sequence analysis and involvement in the immune responses, *Developmental and Comparative Immunology* (2015), http://dx.doi.org/doi:10.1016/j.dci.2015.05.001.

IL-17 signaling components in bivalves: comparative sequence analysis and involvement in the immune responses

Umberto Rosani[1], Laura Varotto[1], Marco Gerdol[2], Alberto Pallavicini[2], Paola Venier[1]

[1] Dept. of Biology, University of Padua, via U. Bassi 58/b, 35121 Padua, ITALY

[2] Dept. of Life Sciences, University of Trieste, via L. Giorgeri 5, 34127 Trieste, ITALY

Highlights

We traced IL-17 signaling components in 31 bivalves spp and few mollusc genomes

Protein domain searches rescued 59 IL-17, 96 SEFIR and CIKS/CIKSL signatures

Just two homology models sum up the various IL-17 domains in mussels and oysters

IL-17s and proximate signaling components show distinct gene expression trends

Abstract

The recent discovery of soluble immune-regulatory molecules in invertebrates takes advantage of the rapid growth of next generation sequencing datasets. Following protein domain searches in the transcriptomes of 31 bivalve *spp*. and in few available mollusk genomes, we retrieved 59 domains uniquely identifying interleukin 17 (IL-17) and 96 SEFIR domains typical of IL-17 receptors and CIKS/ACT1 proteins acting downstream in the IL-17 signaling pathway. Compared to the *Chordata* IL-17 family members, we confirm a separate clustering of the bivalve domain sequences and a consistent conservation pattern of amino acid residues. Analysis performed at transcript and genome level allowed us to propose an updated view of the components outlining the IL-17 signaling pathway in *Mytilus galloprovincialis* and *Crassostrea gigas* (in both species, homology modeling reduced the variety of IL-17 domains to only two 3D structures). Digital expression analysis indicated more heterogeneous expression levels for the mussel and oyster IL-17 ligands than for IL-17 receptors and CIKS/CIKSL proteins. Besides, new qPCR analyses confirmed such gene expression trends in hemocytes and gills of mussels challenged with heat-killed bacteria. These results uphold the involvement of an ancient IL-17 signaling pathway in the bivalve immune responses and, likewise in humans, suggest the possibility of distinctive modulatory roles of individual IL-17s/IL-17 receptors. Overall, the common evidence of pro-inflammatory cytokines and inter-related intracellular signaling pathways in bivalves definitely add complexity to the invertebrate immunity.

Keywords: interleukin 17; bivalves; Mytilus galloprovincialis; Crassostrea gigas; innate immunity.

Abbreviation list

AA, amino acid

cds, protein coding sequence

hpi, hours post-injection

IL-17, interleukin 17 gene or transcript

IL-17R, receptor gene or transcript for interleukin 17

*CIKS/ACT1/*TRAFIP2, gene or transcript for a TRAF-interacting protein involved in regulating responses to vertebrate cytokines

SEFIR, SEF/IL-17R domain present in IL-17R and CIKS/ACT1/TRAFIP2 proteins

M, million

Mg, *Mytilus galloprovincialis*

qPCR, quantitative Real-time PCR

RPKM, Reads Per Kilobase per Million mapped reads

RQ, relative quantification

Introduction

Innate mechanisms based on non-self recognition, signal transduction and finely regulated gene expression, provide invertebrate organisms with an effective protection against possible pathogens such as bacteria and viruses. Invertebrates are not able to mount long-lasting immune responses but rapidly respond to aggressors with multifaceted hemocytes and a large variety of effector molecules such as antimicrobial peptides, heterogeneous lectins, lysozymes, proteases and protease inhibitors (Buchmann, 2014). Reports on the existence of soluble immune-regulatory molecules in invertebrate animals are relatively recent and the possibility of a cytokine-like signaling network definitely adds complexity to the innate immune systems (Malagoli, 2010; Raftos and Nair, 2004). Spätzle in insects, astakine in crustaceans and invertebrate homologs of MIF, TNF and IL-17 might regulate the behavior of target cells in terms of growth and motility, thus orchestrating hematopoiesis, inflammation and immunity likewise their mammalian counterparts (Saenz et al., 2008; Annunziato et al., 2014).

According to their cellular origin and pleiotropic actions, the vertebrate cytokines are classified as interleukins, chemokines and interferons (Turner et al., 2014). Interleukin-17 (IL-17) was first recognized as cytolitic T-cell factor with pro-inflammatory activity (Rouvier et al., 1993) and it is the unique member of an interleukin class with no homology to any other known cytokine family (Moseley et al., 2003; Xu and Cao, 2010). Soluble IL-17 ligands and related membrane receptors (IL-17Rs) exist in humans as gene families and their effects on the target cells depend on differences in the expression levels of individual family members, receptor usage, and on a highly cross-regulated signaling system (Li P. et al., 2014; Wang et al., 2014; Sabat et al., 2013). The IL-17 signaling pathway starts with the binding of IL-17 homo- or hetero-dimers to specific membrane-bound IL-17R complexes. The signal transduction proceeds in the cytosol through the tumor-necrosis factor receptor-associated factor 6 (TRAF6, a key adaptor also in the TLR- and IL-1R-signaling cascades) up to the activation of canonical transcription factors, such as NF-κB, and the expression of cytokines, chemokines and antimicrobial peptides among other target genes (Hartupee et al., 2007; Steinman, 2007). The crucial element of the IL-17 signaling is the SEF/IL-17R (SEFIR) domain which displays similarity to the Toll/IL1 (TIR) domain and mediates the interaction between the cytoplasmic tail of IL-17Rs and CIKS (alias Act1 or TRAFIP2), a proximate adaptor protein with TRAF-binding motifs (Novatchkova et al., 2003; Chang et al., 2006).

Humans possess six IL-17s (named IL-17A/F), five receptors (named IL-17RA/E) and one CIKS protein (Aggarwal and Gurney, 2002; Gaffen, 2009). IL-17s are produced by activated T lymphocytes and other cell types relevant to the host immunity such as the mucosal epithelial cells (Saenz et al., 2008). On the whole, the IL-17 family members are potent pro-inflammatory cytokines involved in host defense, autoimmunity and cancer (Annunziato et al., 2014; Wang et al., 2014). IL-17 is now recognized as central regulator of inflammatory responses in the human brain (Liu et al., 2014).

IL-17 and the related signaling pathway have been considered as exclusive vertebrate features until 2006, when thirty IL-17 and two IL-17R gene models were reported in the genome of the sea urchin *Strongylocentrotus purpuratus* (Hibino et al., 2006). Sequences similar to IL-17, IL-17R and CIKS have been found since then in several non-vertebrate organisms (Wu et al., 2011; Valenzuela-Muñoz and Gallardo-Escárate, 2014; Vizzini et al., 2015). In mollusks, IL-17 was first identified in *C. gigas* in 2008 (Roberts et al., 2008) and subsequently in *P. fucata* (Wu et al., 2013) and *H. rufescens* (Valenzuela-Muñoz and Gallardo-

82 Escárate, 2014). Five IL-17 genes were reported in the *C. gigas* genome (Li et al., 2014) with more recent
83 analyses indicative of eight IL-17s and five IL-17Rs (Zhang et al., 2015).

84 In this work, we update and describe the transcriptional landscapes of IL-17 and related signaling components
85 of *Bivalvia*, focusing in particular on mussel and oyster *spp*. Aiming to validate the *M. galloprovincialis*
86 sequence findings, we also report expression data of six selected IL-17s, three IL-17Rs and the CIKSL adaptor, as
87 measured in hemocytes and gill of mussels injected with a mixture of heat-killed bacteria.

Materials and Methods

Identification of protein domains for interleukins, IL-17s, IL-17Rs and CISK in *Mytilus* and other mollusks

To obtain a *Mytilus galloprovincialis* (*Mg*) reference transcriptome, we used a collection of 18,788 ESTs obtained by Sanger sequencing from mixed tissues (Venier et al., 2009) and 642.5 million (M) reads obtained by paired-end Illumina sequencing from digestive gland (97 Mreads, SRA ID: PRJNA88481 (Gerdol et al., 2014), gills (189.4 Mreads, unpublished) and from mantle, muscle, gill and immuno-stimulated hemocytes (108 Mreads, SRA ID: SRP033481). A global *de-novo* assembly was performed with Trinity (Grabherr et al., 2011) and CLC Genomic Workbench v.7 (Qiagen, Germany). The draft assembly of the mussel genome was downloaded from NCBI (GenBank ID: APJB000000000.1) (Nguyen et al., 2014).

Available RNA sequencing datasets were *de-novo* assembled separately for 31 bivalve species with the CLC Genomics Workbench using the same parameters as above (Table 1). In the case of *Crassostrea gigas*, we processed together other 704 Mreads (Illumina Hiseq2000, 2x50 bp paired end reads, SRA IDs: SRR334212-20, ENA ID: E-MTAB-2552 or unpublished) to check and expand the already available gene predictions. The genome drafts of *Crassostrea gigas*, *Pinctada fucata*, *Lottia gigantea* and *Aplysia californica* were also downloaded from the public repositories with their gene predictions (details in **Table 1**).

Protein coding sequences (cds) were predicted in each species-specific assembly using Transdecoder (Haas et al., 2013). Selected HMM profiles (Interleukin-17: PF06083, SEFIR: PF08357, other interleukin domains: PF00715, PF00727, PF02025, PF02059, PF03039, PF00340, PF00489, PF00726, PF01415, PF02372, PF03487, PF09238, PF09240, PF10420, PF00048, PF14565, PF15036, PF15095,PF02394, PF05566, PF07400, PF06529, PF15037, PF15177, PF15209, PF15225, PF12233) were preliminary retrieved from the PFAM database release 27 (Finn et al., 2014) and used for HMMER scanning analysis (Eddy, 2011). The resulting positive hits were then screened for the presence of putative signal peptide and trans-membrane domains by SignalP-3.0 (Petersen et al. 2011) and TMHMM v.2.0 (Krogh et al. 2001), respectively.

The protein sequences found positive for one reference domain were subsequently aligned using MUSCLE (Edgar, 2004) and trees were generated with a neighbor joining clustering method and Jukes-Cantor substitution model with 1000 bootstrap replicates using MEGA6 (Tamura et al., 2013). To compute more precise phylogenetic distances, multiple alignments were performed on conserved domain positions previously identified with Gblocks (Castresana, 2000). A position was considered 'conserved' if it was common to 51 % of the locally aligned sequences.

The *Mg* transcript sequences outlining a mussel IL-17 signaling pathway were used as blast query against the genomic mussel contigs in order to recover the corresponding gene structures. Only hits with an E-value lower than 10^{-20} were extracted and *de-novo* assembled to identify overlapping regions. The resulting contigs and singletons were manually combined to obtain the gene sequence; then, correct gene assembly and prediction of alternative splicing events were ascertained by RNA-seq read mapping with appropriate analysis parameters.

The three-dimensional structure of the identified IL-17 domains and proteins was predicted by homology using Phyre2 (Kelley and Sternberg, 2009).

Digital gene expression analysis of *IL-17* signaling components identified in mussels and oysters

Using CLC Genomics Workbench, we analyzed different Illumina RNAseq samples from *Mg* (SRR442031/6 and SRP033481) and from *C. gigas*: different oyster tissues (SRR334212/20), oysters treated with *Vibrio anguillarum*, *V. tubiashii*, *V. aestuarianus*, *V. alginolyticus* and *Micrococcus lysodeikticus* (SRR796582-98), oysters at different developmental stages (SRR334222-59, Zhang et al., 2015) and OsHV-1-positive spat (E-MTAB-2552, Rosani et al., 2014). Details are reported in **Supplementary Table 1**. The *Mg* and *C. gigas* reads were mapped on the gene sequences denoting *IL-17* and proximate signaling components, separately for each species. Sequence length and similarity values were set to 0.75 and 0.95, respectively, whereas mismatch/insertion/deletion penalties were set to 2/3/3. The total number of reads of each dataset were counted and used to calculate digital expression values as RPKM (Reads Per Kilobase per Million mapped reads) (Mortazavi et al., 2008), then datasets were normalized by the total mapped reads and related to the expression levels of Elongation Factor 1α (EF1α).

Mussel immunostimulation and tissue sampling

Native mussels of commercial size (*Mg*, 5.7 ± 0.4 cm shell length) were collected from one outlet of the Venice lagoon (Italy) in May 2014, acclimatized at 23±1 °C in artificial sea water (32 ‰ salinity, 22 °C) and fed with certified food (Plancto® Aqua Medic, Bissendorf, Germany). Mussels were placed in two plastic tanks (50 mussels/tank, 1 liter sea water/mussel) and were injected into the posterior adductor muscle either with 0.1 ml of NaCl-enriched PBS (PBS-NaCl) or with 0.1 ml of a mixture of heat-killed Gram positive and Gram negative bacteria (10^8 CFU/ml). The bacterial cocktail was prepared from equal amounts of *Micrococcus lysodeiticus*, *Vibrio splendidus* and *Vibrio anguillarum,* separately grown overnight in Marine Broth at 22 °C (*Vibrio* spp.) or in Luria-Bertani medium at 30 °C (*M. lysodeiticus*). Bacteria were collected from each culture by centrifugation at 3000 xg for 10 min and their concentration, estimated by optical density at 600 nm, was adjusted by re-suspending them in PBS-NaCl. The mixture was prepared, heated at 65°C for 2 h and the complete inactivation of the bacteria cells was verified by plating on nutrient medium.

Hemolymph was withdrawn from the treated mussels and paired controls at 1, 3, 6, 9, 12, 24 and 48 h post-injection (hpi) with an equal volume of modified Alsever Solution (27 mM sodium citrate, 72 mM NaCl, 113.8 mM D-glucose, 2.6 mM citric acid, pH 7.4). Hemocytes were immediately collected by centrifugation at 800 g, 4 °C for 15 min, re-suspended in 1 ml Trizol reagent (Invitrogen) and stored at -80 °C. Following dissection, gill samples were similarly homogenized in Trizol and stored. At the dissection, the mussels showed intermediate gonad development.

RNA purification and cDNA synthesis

Total RNA was extracted from individual hemolymph and gill samples following the manufacturer's instructions (Trizol, Invitrogen) and additionally purified by precipitation with 8 M LiCl. RNA concentration and quality were ascertained by using the NanoDrop® ND-1000UV spectrophotometer and Agilent 2100 Bioanalyzer (microcapillary electrophoresis on RNA 6000 Nano LabChips, Agilent Technologies). Equal RNA amounts obtained from six individual mussels per time point, both in treated and control groups, were pooled and each RNA pool was treated with DNAse I (Qiagen) to eliminate contaminating genomic DNA. First-strand cDNA synthesis started in the presence of 0.5 µg/µl Oligo(dT)12-18,10 mM dNTPs, 5X First-Strand Buffer and 200 U SuperScriptTM II Reverse Transcriptase (Life Technologies) in 20 µl reaction volume. The reverse transcription

proceeded for 2 h at 42 °C and was stopped by heating at 70°C for 15 min. Each cDNA mix was properly diluted and stored at −80 °C for subsequent SYBR Green quantitative real-time PCR.

Quantitative Real-time PCR (qPCR)

qPCR was carried out with the DyNAmo HS SYBR Green qPCR kit (Thermo Scientific) to amplify 1 µl of purified and diluted first-strand cDNA in a 10 µl of final reaction mixture (1X Master Mix, 1X Rox passive reference dye, 0.5 µM of each specific primer). Primer pairs were designed to amplify selected mussel transcripts using Primer 3 (Rozen and Skaletsky, 2000) with the following guidelines: product size 100-250 bp, melting temperatures Tm 60±1 °C, and G/C ≤55 %. Primer specificity and amplicon sequences were verified by Sanger sequencing. Primer details are reported in **Table 2**. Amplification cycles were performed with an Applied Biosystems 7900HT Fast Real-Time PCR System in a MicroAmp® Fast Optical 96-Well Reaction Plate (Applied Biosystems) under the following conditions: 95 °C for 15 min; 40 cycles of: 95 °C for 30 sec and 60 °C for 1 min. A dissociation curve analysis was performed at the end of the reaction to ensure the specificity of the primers. The qPCRs were carried out in triplicate using the same plate for one selected primer pair (target and endogenous genes) and the complete sample set (immune-stimulated mussels and paired controls, 7 time points). The relative expression ratio of target gene (RQ) was based on the delta–delta Ct method ($2^{-\Delta\Delta Ct}$) (Livak and Schmittgen, 2001). The housekeeping gene Elongation factor 1 alpha (EF1-α) was chosen as endogenous control (Gerdol et al., 2011; Gerdol et al., 2012; Moreira et al., 2014b; Toubiana et al., 2014). A two-tailed paired t test was done to assign statistical significance. Results are given as the mean with standard deviation of three replicates.

Results

Protein domains indicative of IL-17 signaling in molluscan genomes

To investigate the presence of transcripts denoting the IL-17 signaling pathway in marine mollusks we started from the available genomes of two gastropods (*Aplysia californica, Lottia gigantea*) and two bivalves (*Crassostrea gigas*, *Pinctada fucata*), and from the first release of the *Mytilus galloprovincialis* (*Mg*) genome assembly (Nguyen et al., 2014). HMMER scanning analyses led to the identification of IL-17 and SEFIR domains in both gastropod and bivalve *spp*. (only IL-17 was identified in all four genomes, out of 28 domains selected with the term 'Interleukin' from the PFAM database). The number of IL-17-positive hits was quite different in the analyzed genomes: none in *A. californica*, 2 in *L. gigantea*, 8 in *C. gigas*, 26 in *Mg* and 27 in *P. fucata*. Also SEFIR domains were identified in all four genomes, spanning from 2 (*A. californica*) to 8 (*Mg*) positive hits.

Protein domains indicative of IL-17 signaling in bivalve transcriptomes

We retrieved several publicly available SRA runs to produce transcriptome assemblies for 31 bivalve species. On the whole, we processed around 3 Greads, obtaining 2.4 millions of transcriptomic contigs (details in Table 1). D*e-novo* assembling process yielded heterogeneous amounts of contigs, partly related to the number of available reads and ranging from a few thousand contigs in *T. granosa* assembly (1.2 Mreads) to 288 K contigs in the *M. galloprovincialis* (642.5 million of assembled reads). Furthermore, tissue representation was heterogeneous in the analyzed datasets, with 16 assemblies originating from mixed tissues, 3 from muscle and other assemblies from gills, gonad, larval stages, foot or mantle. To avoid biases related to these factors, we decided to report and further analyze only transcripts including complete coding sequences (cds). Domain scanning provided 59 and 96 positive hits for the IL-17 and SEFIR domains, respectively (Table 3). Based on the presence of both DEATH and SEFIR domains or phylogenetic analyses, we subdivided all SEFIR-containing proteins in IL-17Rs and CIKS adaptors (CIKSL if they contained both domains). We were not able to retrieve IL-17 or SEFIR domains in 3 of the analyzed organisms (**Table 3**). Indeed, we could not find a complete *IL-17* transcript in 14 assemblies nor SEFIR containing proteins in 12 assemblies, out of the 31 analyzed bivalve *spp*.

Phylogenetic analysis of the IL-17 and SEFIR domains identified in bivalves

Figure 1 is a graphical visualization resulting from the alignment of all IL-17 domains included in the PFAM database (312 hits) together with the ones described in the present work. We observed that vertebrate IL-17b/c/d/e clustered consistently in separated branches whereas vertebrate IL-17a and IL-17f appeared intermixed. Notably, thirty-five PFAM sequences (11 of *Strongylocentrotus purpuratus*, 6 of *Branchiostoma floridae,* 6 of *Caenorhabditis spp.* and 2 of *Daphnia pulex*) clustered together with the 59 mollusk IL-17s reported in the present work. In more detail, the invertebrate IL-17 branch includes three subdivisions with the mollusk IL-17 domains mainly separated from *Caenorhabditis* IL-17 domains. The clustering of the IL-17s from *Strongylocentrotus purpuratus* and *Branchiostoma floridae* in the protostomes clade probably relates to the fact that these basal deuterostomes hold more primitive IL-17 domains, compared with those of vertebrate IL-17s. Actually, the alignment of protostome IL-17s displays five conserved cysteine residues whereas the IL-17s of vertebrates present four highly conserved cysteines.

The multiple alignment of the 59 IL-17 domains identified in marine mollusk transcriptomes revealed some highly conserved positions (**Figure 2**). Comparing them with human IL-17-F (AAH70124) we confirmed the high conservation of the amino acid residues in position 5, 6, 18, 20, 26, 28, 31, 37, 55, 62, 72, 73, 81, 92, 108, 110 and 112, including those also involved in the stabilization of the cysteine array. Interestingly, mollusk IL-17 domains lacking CYS-28 are also devoid of CYS-92 (both cysteine residues are involved in a disulfide bond). Moreover, bivalve IL-17 domains display a conserved cysteine in position 62 that putatively binds CYS-5 thus forming a third disulfide bond. The human SER-5 is an invariant cysteine in bivalves and the residue CYS-31, which characterizes the human IL-17 isoforms B, C and E, is highly conserved.

Consistent with the role of SEFIR-SEFIR interactions in the signal transduction activated by IL-17, we identified the SEFIR domain in both IL-17Rs and CIKS proteins (Chang et al., 2006). In fact, the phylogenetic analysis of 96 mollusk SEFIR domains provided a clear-cut discrimination of IL-17 receptors and CIKS adaptor proteins (Figure 3). On one hand, 36 mollusk proteins clustered together with human IL-17Ra and appear in 2 sub-branches, with almost all analyzed species having two IL-17R forms (called IL-17Ra and b). In addition, we observed a third form, named IL-17R-like and discussed below in the text. On the other hand, 37 SEFIR domains clustered together with the human CIKS protein (AAH02823) and appear in 2 sub-branches: the first one included all SEFIR domains (17) associated with the DD or protein sequences probably incomplete at the 5' terminus; the second one included all proteins (20) with only a SEFIR domain whose presence in *Ostreoida, Unionoida, Arcoida and Pectinoida* suggests the existence of both CIKSL and CIKS adaptors in these orders.

Transcripts outlining the IL-17 signaling pathway in mussels

Transcript sequences identifying components of an IL-17 pathway within the genus *Mytilus* were evaluated in more detail. We could retrieve SRA datasets for four mussel species, around one billion reads in total, mainly from *Mg* gills and *M. edulis* larval stages (see Table 1).

The IL-17 domain searches provided 6, 6, 2 and 1 positive hits for *M. galloprovincialis*, *M. edulis*, *M. trossulus* and *M. californianus*, respectively. The SEFIR domain was also identified in the four species, with 4, 5, 4 and 4 hits respectively (Table 3). Overall, we could retrieve the complete cds for 12 IL-17s and 10 SEFIR domain-containing proteins (**Table 4**).

The deduced *Mytilus* IL-17 proteins ranged in length from 165 to 221 amino acids (estimated molecular weight of 19.2 - 25.7 kDa), with a signal peptide of 17-26 amino acid identified in 10 out of 12 transcripts. Moreover, we could notice that one apparently incomplete IL-17 protein (*Mg* IL-17-6) is likely ascribable to a single nucleotide mutation, with the change of the initial methionine in cysteine generating a 25 AA shorter protein.

Transcripts outlining the Il-17 signaling pathway in oyster

Five genes encoding complete IL-17s, with a IL-17 domain and signal peptide, have been reported in the *C. gigas* genome (Li et al., 2014). Scanning the available oyster gene predictions, we could retrieve 8 and 5 positive hits for the IL-17 and SEFIR domains, respectively. To confirm gene annotations, we took advantage of a previous RNA sequencing experiment (unpublished Illumina reads) to produce a whole *C. gigas* transcriptome assembly (see details in M&M). Following domain scanning of the newly built transcriptome, we confirmed, or

updated, all the previously predicted genes and we recovered new positive hits: two *IL-17*s and one protein containing both DD and SEFIR domains (named *Cg*CIKSL). Overall, we can report the presence in *C. gigas* of ten IL-17s, three IL-17Rs and three CIKS adaptors (**Table 5**). Nine out of the ten *IL-17s*, displaying both the typical domain and signal peptide, could be considered as 'functionally' complete IL-17s. Regarding the SEFIR-containing transcripts, we detected one CIKSL, two CIKS adaptors (*CgCIKS_2* and *CgCIKS_3*, presenting only a SEFIR domain), two three IL-17 receptors (*CgIL-17Ra* and *CgIL-17Rb*) and one receptor-like (*CgIL-17R-like*).

To predict IL-17 3D structure we used a homology strategy, retrieving similar models in human IL-17-F (PDB ID: d1jpya) and cysteine-knot motifs present in platelet-derived growth factor-like, transforming growth factor (TGF)-beta, neurotrophin, gonadodropin and noggin. Since we obtained good confidence (>90%) models only for the IL-17 domain regions, we limited our predictions to this region (**Figure 4**). In 5 out of 6 mussel and 9 out of 10 oyster IL-17s we retrieved very similar structures, while for mussel and oyster IL-17-6 we retrieved a peculiar structure characterized with an additional short alpha helix.

Likewise mammalian IL-17Rs, the mussel IL-17Rs possess a transmembrane region and a SEFIR domain located on the cytoplasmic protein tail, necessary for the interaction with CIKS/CIKLS adaptors. Structure prediction based on homology returned a low percentage of covered queries denoting human interleukin receptors a and b and, at lower similarity level, the presence of a TIR domain. Nevertheless a subclass of receptor-like proteins (including *Mg*IL-17R-like and *Cg*IL-17R-like) is devoid of an extracellular region, which is supposedly involved in the recognition and binding of IL-17. These receptor-like proteins, phylogenetically different from the others as reported in Figure 3, present a transmembrane region followed by a SEFIR domain in the first 250 amino acids, but they possess only a very short extracellular tail. Therefore, they are unlikely to be involved in the binding of IL-17 and they might be a regulatory component of the active IL-17 receptor complex, as the SEFIR domain has been frequently implicated in homotypic SEFIR-SEFIR interactions (Novatchkova et al., 2003).

Organization of IL-17 signaling genes in mussel and oyster genomes

We identified the genomic contigs corresponding to IL-17 pathway components in *Mg* and *C. gigas* by blasting the identified transcripts against the mussel and oyster genomes, respectively. To retrieve the gene structures, we subsequently mapped *Mg* and *C. gigas* Illumina reads (see M&M) on the previously selected genomic contigs. Generally, oyster genomic contigs were longer than the mussel ones (N50 19.4 kb in oyster *versus* N50 of 0.8 kb in mussel) and encoded for several genes whereas the genomic mussel contigs included one gene or only a part of it. **Table 6** summarizes our findings.

Actually, we identified the complete gene sequence only for *MgIL-17-4* (by closing completely a 1.5 kb intron) and *MgIL-17-6* (one single exon). Even though the contig size did not allow us to fully cover IL-17 genes, we reconstructed all the mussel IL-17 gene structures, composed by two exons with the first of them encoding just few amino acids (except *MgIL-17-6* characterized by only one exon). The gene structure of mussel and oyster IL-17s appeared similar since two exons, with few amino acid residues (1-11) located on the first one, were also detected in eight out of ten oyster genes rescued in total. Apart the five IL-17s already considered as complete (*CgIL-17-1/5*) (Li et al., 2014), we updated two annotations (*CgIL-17-6* and *8*) and we also recovered the genomic positions of two newly identified IL-17s (*CgIL-17-9*, *CgIL-17-10*). *CgIL-17-1* was also updated according to the modified position of the first exon.

Mussel and oyster IL-17R and CIKS(L) genes displayed a more complex organization with multiple exons. Moreover, the finding of *CgCIKS_2* and *CgCIKS_3* in the same genomic contig would suggest a gene duplication event.

Digital expression levels of IL-17 signaling components in mussel and oyster

The expression levels of IL-17s, IL-17Rs and CIKS(L) transcripts were investigated in *the M. galloprovincialis* and *C. gigas* RNA-seq datasets, including one Illumina RNA-seq from the whole tissues of *C. gigas* spat (ID: E-MTAB-2552) infected by a virulent Ostreid Herpesvirus type 1 (OsHV-1) (Rosani et al., 2014). Data processing was performed with suitable parameters and RPKM values were computed (see Methods Section). In both mussel and oyster datasets, the housekeeping EF1α gene was constantly and highly expressed and, therefore, confirmed as endogenous reference to compare the expression trends of the IL-17 signaling genes.

In the available mussel transcriptomes, the six *IL-17s* were present in different amounts among the five analyzed tissues (mostly below 1/10,000 of the EF1α expression level), with the highest levels detected for *IL-17-3* (6% of EF1α) in immuno-stimulated hemocytes (**Figure 5A**). Only *MgIL-17-3* and *MgIL-17-5* were expressed in hemocytes, the complete set of IL-17 transcripts was evident in digestive gland, gill and mantle and only four different *IL-17* transcripts were expressed in muscle. Compared to IL-17s, the levels of IL-17Rs and CIKSL were on average somewhat higher and more homogeneous among tissues.

In the available oyster tissue transcriptomes, the ten identified *IL-17s* were represented in very different amounts, with the highest levels (~1/100 of EF1α), detected for IL-17-2 in the infected spat and for *IL-17-6* in gills and labial palps of other samples (**Figure 5B**). *IL-17-2*, *IL-17-3*, *IL-17-6* and *IL-17-10* were clearly expressed in the hemocytes whereas *IL-17-1*, *IL-17-8* and *IL-17-10* could not be detected in gills. It is worth noticing the marked expression of the *CgIL-17-1/2/3/4* transcripts in the sample prepared from OsHV-1-infected oysters (black bars in Figure 5B). On average, the oyster IL-17Rs, CIKS and CIKSL were expressed at higher and more homogeneous levels in all analyzed tissue samples, like in mussels (mostly from 1/100 to 1/1,000 of EF1α).

To expand the analysis on the expression levels of oyster IL-17 pathway components, we considered several RNA-seq datasets representing developmental stages (38 samples) and challenges with various bacterial strains(14 samples) (Zhang et al., 2015). The resulting heat map (**Supplementary File 1** and **Table 2**) highlights once more the different expression levels of *IL-17*s compared to receptors and adaptors (*IL-17R*s, *CIKS*s and *CIKSL*). The latter were substantially more expressed than *IL-17*s in the RNA-seq datasets from bacterial challenges and from first stages of the oyster development; *CgIL-17-6* was clearly induced, and, overall, it emerged as the most responsive *IL-17* isoform; *CgIL-17R*-like and *CIKSL* were expressed during the early developmental stages (from egg to 'free swimming' stage) whereas *CgIL-17Ra* and *CIKS* were expressed subsequently (from morula to larval stages). *CgIL-17Rb* was always expressed and *CgCIKS2* was generally detected at low levels.

In vivo* expression levels of IL-17 signaling components in *M. galloprovincialis

The expression levels of six IL-17 ligands (1 to 6), three putative receptors (*IL-17Ra/b* and *IL-17R-like*) and *CIKSL* were analyzed by qPCR in the hemocytes and gills of mussels (*M. galloprovincialis*) sampled at 1, 3, 6, 9, 12, 24 and 48 h after injection of heat-killed bacteria (immunostimulated group) or PBS-NaCl (paired controls). Elongation Factor 1 alpha (*EF1α)* was chosen as reference housekeeping gene.

In control mussels, we observed particularly low expression levels for *Mg*IL-17-2 and *Mg*IL-17-5 in hemocytes and for *Mg*IL-17-4 in gills whereas *Mg*IL-17-6 transcripts were not detectable in hemocytes (data not shown). Figure 6 depicts the expression levels of the six *IL-17* ligands, relative to the time-paired controls, in the hemocytes and gills of mussels injected with a mixture of heat-killed Gram+ and Gram- bacteria and sampled from 1 h to 48 h post-injection (hpi). Except for IL-17-6 (not detectable in hemocytes), the *in vivo* immunostimulation caused significant up-regulation of the other five IL-17 ligands, though with clear differences in the expression patterns, and a general down-regulation at 24 and 48 hpi, more evident in hemocytes than gills. In detail, the IL-17-1 transcript levels peaked at 1 hpi and progressively decreased in both hemocytes and gills (IL-17-6 showed a similar trend in gills). IL-17-2 showed less pronounced induction (the highest levels were recorded at 1 hpi in gills and at 3-6 hpi in hemocytes). Despite IL-17-3 reached expression values even higher than IL-17-1, it showed a later inducibility with peak levels at 9 hpi in hemocytes and at 6 hpi in gills. The expression trend of IL-17-4 and IL-17-5 requires further investigation to confirm the significant over-and under-expression trends recorded in the two analysed tissues.

As well, the bacterial cocktail injected in the mussels significantly up-regulated IL-17Ra and b, IL-17R-like and CIKSL transcripts (**Figure 7**). Overall, the expression of all these genes peaked at 9-12 hpi, though at lower levels in gills than in hemocytes. IL-17Ra was the only transcript still substantially expressed at 24 hpi (in hemocytes).

Discussion

The class of bivalve mollusks includes marine species economically relevant in many coastal areas worldwide, such as clams, mussels, oysters, and scallops. Adequate knowledge on the innate defenses and functional responses of reference species could help us to prevent severe mortality episodes associated with virus and bacteria (Jenkins et al., 2013), other infectious diseases and parasitic infestations (Ramilo et al., 2014) and also to manage health risks posed by the accumulation of human pathogens, bio-toxins and chemical pollutants in these filter-feeding organisms (Varotto et al., 2013; Gerdol et al., 2014; Lassudrie et al., 2015). Owing to the rapid development of advanced sequencing technologies, several mollusk species have been deeply analyzed at transcriptome level (Kawashima et al., 2013; Pauletto et al., 2014; Poynton et al., 2014; Shi and He, 2014; Teaniniuraitemoana et al., 2014; Zhang et al., 2014). However, on account of size and complexity only a few sequencing projects of bivalve genomes have been completed or launched.

Based on advanced sequence analysis, the discovery of immune-regulatory invertebrate molecules is steadily progressing. Gene or transcript sequences predicted to be the macrophage migration inhibitory factor (MIF; pfam01187) have been reported in species of different invertebrate phyla including mollusks (Cui et al., 2011; Li et al., 2011; Parisi et al., 2012). Tumor necrosis factor alpha (TNF-α)-like molecules have been identified in flat and cupped oysters and, in both cases they were involved in the host immune responses (Martín-Gómez et al., 2014; Sun et al., 2014; Gerdol and Venier, 2015; Zhang et al., 2015). The induction of IL-17-like transcripts was reported at first in hemocytes of *C. gigas* stimulated with a mixture of heat-killed bacteria (Roberts et al., 2008) and the presence of IL-17, a third pro-inflammatory cytokine, was subsequently confirmed at transcriptome and genome level in cupped and pearl oysters (Wu et al., 2013; Li J. et al., 2014). We now report the distribution and sequence relationships of 59 IL-17 sequences identified in the transcriptome of 31 bivalve *spp*. and in four mollusk genomes (notably, IL-17 was the only interleukin domain that we could identify in the four analyzed genomes). Irrespective to the low similarity to human IL-17s, the mollusk IL-17 domains revealed a consistent pattern of conserved amino acid residues suggestive of a common cysteine knot 3D structure. Moreover, we identified 96 SEFIR-containing proteins distributed in 59 IL-17Rs and 37 CIKS adaptor molecules. Despite the total number of IL-17 and SEFIR positive hits detected in the bivalve assemblies was partly related to the abundance of available Illumina reads, the common presence of IL-17s, IL-17Rs and CIKS adaptors in bivalve and mollusk *spp.* indicates sequence conservation and confirms the functional relevance of such pathway across species.

Most of the 304 IL-17 domains included in the PFAM database were related to Chordata (90 %), with only 29 of the remaining hits distributed among Echinoderma, Arthropoda, Mollusca and Nematoda. The IL-17s reported in the genomes of *Pinctada fucata* (Wu et al., 2013) and *Crassostrea gigas* (Li J. et al., 2014) are still lacking in the Pfam 27.0 database, but the absence of insect IL-17 domain sequences in such database is even more intriguing, as the IL-17 signaling pathway is clearly involved in the immune response to injuries.

The IL-17 domains of bivalves (and 2 gastropods) are closely related to those of the basal deuterostomes *S. purpuratus, B. floridae* and those of other protostomes, which form altogether a branch separated from the branches including chordates and urochordates such as *C. intestinalis* (clusters IL-17a/b/c/d/e/f in Fig. 1). Limiting the phylogenetic analysis to the conserved IL-17 domain rather than the entire sequence, we obtained a clear-cut view, consistent with the most recent reports (Li J. et al., 2014; Zhang et al., 2015). Accordingly, a common IL-17 ancestor should have preceded the radiation of bivalves occurred around 530 Mya (Plazzi and

Passamonti, 2010). The presence of an IL-17R-like molecule in marine mollusk, characterized by the presence of all receptor features except of the extracellular tail, is curious and led us to suppose a regulative role of such proteins within the IL-17 receptor complex.

The hypothesis of a gene duplication event regarding CgCIKS is not surprising if considering the common expansion of gene families involved in the innate immunity (Zhang et al., 2015). Furthermore, we could infer an earlier evolutionary emergence of invertebrate CIKSLs, including a Death Domain in addition to the SEFIR domain. In fact, the SEFIR domain sequences of mollusk CIKSs and CIKSLs clustered apart from those of IL-17 receptors, and the mollusk CIKSLs clustered with the human CIKS (see Fig. 3). The DD of CIKSLs in the IL-17 signaling pathway is reminiscent of the role of MyD88, holding DD and TIR domains, in promoting the Nf-*k*B activation through IRAK in the TLR signaling pathway. The CIKS-mediated signal transduction could also be amplified into the mitogen-activated kinases (MAPK) signaling pathway via SAPK/JNK, namely, MAPK9 or cJun N-terminal kinase. MAP kinase signaling is central to many cell processes and the possible cross-interaction between multiple signaling pathways in the presence of actively interfering and counteracting microorganisms is highly intriguing (Wu et al., 2011). The IL-17R and CIKS gene families do not have a common evolutionary origin and, notably, prokaryotes possess SEFIR homologs possibly acquired by horizontal gene transfer from CIKSs, rather than IL-17Rs (Wu et al., 2011). The hypothesis that bacterial SEFIR domains can capture by homology host SEFIR-containing proteins, and therefore attenuate the IL-17-signaling, adds complexity to the never-ending story of hosts and their microbial associations. Evolutionary events such as gene duplication and gene fission might have contributed to the loss of the DD in vertebrate CIKS molecules. Additionally, sequence diversification seems to have been occurred also in *Palaeoheterodonta* and *Pteriomorphia*, as we observed the presence of both CIKS- and CIKSL-like proteins in *Ostreidae* but not in *Mytiloida*. Given the presence of two similar CIKS loci within the same genomic contig, a further gene duplication event should have occurred in *C. gigas*.

According to the overall results, the scheme of **Figure 8**, illustrates the relationships between ligands, receptors and downstream elements of the IL-17 signaling pathway in bivalves.

Based on sequence analysis, we could report for the first time the presence of six IL-17s in *M. galloprovincialis*, and 10 IL-17s in *C. gigas*. Like in ascidians (Vizzini et al., 2015) and vertebrate animals, the existence of IL-17 as gene family in bivalves rises questions on the expression levels of individual family members and their functional role *in vivo*. At first, we investigated the relative transcript abundance of the different IL-17s, IL-17Rs (or IL-17R-like, three in both species), CIKSL (one in both species) and CIKS (two in oyster) in Illumina datasets representing a variety of bivalve tissues. Overall, the expression patterns evaluated *in silico* suggested a different regulation of the ligands compared to the receptors and proximate adaptors: in both mussels and oysters, the transcription levels of the different *IL-17s* appeared highly heterogeneous among tissues whereas the levels of *IL-17Rs* and *CIKS/CIKSL* transcript sequences were less fluctuating and never negligible. Likewise in humans, the expression levels of bivalve IL-17s might reflect the specific modulatory role of individual IL-17 family members, via different ligand-receptor combinations and within a regulatory network whose complexity would not be so surprising, in the light of current advances on the immune and neuroendocrine mollusk systems (Stewart et al., 2014; Veenstra, 2010). Among the analyzed transcriptome dataset, the sample referring to *C. gigas* spat infected by a virulent OSHV-1 (Rosani et al., 2014) strongly suggests the induction of IL-17 transcripts (in special manner *Cg*IL-17-3) and of the proximate IL-17 signaling components as part of the

antiviral host response. Moreover, we identified *CgIL-17-6* as the most responsive among the oyster interleukin-17 genes, both during bacterial challenges and in the oyster development.

The analyses performed in the hemocytes and gills of mussels injected with heat-killed bacteria confirmed the involvement of the IL-17 signaling pathway in the immune responses of *M. galloprovincialis*. Overall, the expression levels recorded in the control mussels for IL-17s, IL-17Rs, IL-17R-like and CIKSL seldom exceeded 1/100 of those of the reference EF1α gene. In the immunostimulated mussels, the qPCR analysis revealed markedly different induction levels and time-related trends among the six IL-17s analyzed (depending on isoform and tissue we detected expression peaks already at 1 h or later at 6-9 h post-injection, up to 6-7.5 times related to time-paired controls). For comparison, significant induction of one IL-17 transcript at 3 hpi and subsequent down-regulation at 24-72 hpi were recorded in the hemocytes of Manila clams injected with *Vibrio alginolyticus* (Moreira et al., 2014a) whereas higher induction values were reported for *IL-17* at 6 or 24 hpi in hemocytes of *C. gigas* injected with heat-killed bacteria (Roberts et al., 2008). Given the complex expression trends reported in *C. gigas* hemocytes for five different IL-17s in response to different immunostimulants (Li J. et al 2014), the heterogeneous patterns of IL-17 over-expression observed in this study are not surprising and support diversified roles for the IL-17 gene products.

Compared with the IL-17 ligands, the IL-17 receptors and CIKS/CIKSL proximate adaptors displayed more homogeneous induction trends in the immunostimulated mussels, higher in hemocytes than gills, with maximum expression at 9-12 hpi and subsequent decline (except for IL-17Ra).

The higher induction levels commonly observed in hemocytes related to gills may reflect both the direct effect of the injection in the hemolymph, interstitial in the adductor muscle, and somewhat higher responsiveness of the hemocytes compared to the predominant epithelial-like gill cells. How the expression of these genes is regulated by external and internal stimuli in mussels and which cell types or cell subpopulations are targeted by specific IL-17/IL-17R family members remains to be understood with further study. The identification of mussel cell types targeted by the IL-17 signaling could also provide data about the possible induction of IL-1β e TNFα, and the functional inter-relations between cytokines and intracellular signaling pathways in bivalves.

Acknowledgments

This work is granted by PRIN 2010-11 (20109XZEPR).

456 **References**

457 Aggarwal, S., Gurney, A.L., 2002. IL-17: prototype member of an emerging cytokine family. J. Leukoc. Biol. 71,
458 1–8.
459 Annunziato, F., Romagnani, C., Romagnani, S., 2014. The 3 major types of innate and adaptive cell-mediated
460 effector immunity. J. Allergy Clin. Immunol. doi:10.1016/j.jaci.2014.11.001
461 Buchmann, K., 2014. Evolution of innate immunity: clues from invertebrates via fish to mammals. Mol. Innate
462 Immun. 5, 459. doi:10.3389/fimmu.2014.00459
463 Castresana, J., 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic
464 Analysis. Mol. Biol. Evol. 17, 540–552.
465 Chang, S.H., Park, H., Dong, C., 2006. Act1 Adaptor Protein Is an Immediate and Essential Signaling Component
466 of Interleukin-17 Receptor. J. Biol. Chem. 281, 35603–35607. doi:10.1074/jbc.C600256200
467 Cui, S., Zhang, D., Jiang, S., Pu, H., Hu, Y., Guo, H., Chen, M., Su, T., Zhu, C., 2011. A macrophage migration
468 inhibitory factor like oxidoreductase from pearl oyster Pinctada fucata involved in innate immune
469 responses. Fish Shellfish Immunol. 31, 173–181. doi:10.1016/j.fsi.2011.03.009
470 Eddy, S.R., 2011. Accelerated Profile HMM Searches. PLoS Comput Biol 7, e1002195.
471 doi:10.1371/journal.pcbi.1002195
472 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids
473 Res. 32, 1792–1797. doi:10.1093/nar/gkh340
474 Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm,
475 L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. Nucleic
476 Acids Res. 42, D222–D230. doi:10.1093/nar/gkt1223
477 Gaffen, S.L., 2009. Structure and signalling in the IL-17 receptor family. Nat. Rev. Immunol. 9, 556–567.
478 doi:10.1038/nri2586
479 Gerdol, M., Manfrin, C., De Moro, G., Figueras, A., Novoa, B., Venier, P., Pallavicini, A., 2011. The C1q domain
480 containing proteins of the Mediterranean mussel Mytilus galloprovincialis: A widespread and diverse
481 family of immune-related molecules. Dev. Comp. Immunol. 35, 635–643. doi:10.1016/j.dci.2011.01.018
482 Gerdol, M., De Moro, G., Manfrin, C., Venier, P., Pallavicini, A., 2012. Big defensins and mytimacins, new AMP
483 families of the Mediterranean mussel Mytilus galloprovincialis. Dev. Comp. Immunol. 36, 390–399.
484 doi:10.1016/j.dci.2011.08.003
485 Gerdol, M., De Moro, G., Manfrin, C., Milandri, A., Riccardi, E., Beran, A., Venier, P., Pallavicini, A., 2014. RNA
486 sequencing and de novo assembly of the digestive gland transcriptome in Mytilus galloprovincialis fed
487 with toxinogenic and non-toxic strains of Alexandrium minutum. BMC Res. Notes 7. doi:10.1186/1756-
488 0500-7-722
489 Gerdol, M., Venier, P., 2015. An updated molecular basis for mussel immunity. Fish Shellfish Immunol.
490 doi:10.1016/j.fsi.2015.02.013
491 Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L.,
492 Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F.,
493 Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome
494 assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
495 doi:10.1038/nbt.1883
496 Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B.,
497 Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R.,
498 William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript
499 sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nat. Protoc. 8.
500 doi:10.1038/nprot.2013.084
501 Hartupee, J., Liu, C., Novotny, M., Li, X., Hamilton, T., 2007. IL-17 Enhances Chemokine Gene Expression
502 through mRNA Stabilization. J. Immunol. 179, 4135–4141. doi:10.4049/jimmunol.179.6.4135

503 Hibino, T., Loza-Coll, M., Messier, C., Majeske, A.J., Cohen, A.H., Terwilliger, D.P., Buckley, K.M., Brockton, V.,
504 Nair, S.V., Berney, K., Fugmann, S.D., Anderson, M.K., Pancer, Z., Cameron, R.A., Smith, L.C., Rast, J.P.,
505 2006. The immune gene repertoire encoded in the purple sea urchin genome. Dev. Biol., Sea Urchin
506 Genome: Implications and Insights 300, 349–365. doi:10.1016/j.ydbio.2006.08.065
507 Jenkins, C., Hick, P., Gabor, M., Spiers, Z., Fell, S.A., Gu, X., Read, A., Go, J., Dove, M., OConnor, W., Kirkland,
508 P.D., Frances, J., 2013. Identification and characterisation of an ostreid herpesvirus-1 microvariant
509 (OsHV-1 µ-var) in Crassostrea gigas (Pacific oysters) in Australia. Dis. Aquat. Organ. 105, 109–126.
510 doi:10.3354/dao02623
511 Kawashima, T., Takeuchi, T., Koyanagi, R., Kinoshita, S., Endo, H., Endo, K., 2013. Initiating the Mollusk
512 Genomics Annotation Community: Toward Creating the Complete Curated Gene-Set of the Japanese
513 Pearl Oyster, Pinctada fucata. Zoolog. Sci. 30, 794–796. doi:10.2108/zsj.30.794
514 Kelley, L.A., Sternberg, M.J.E., 2009. Protein structure prediction on the Web: a case study using the Phyre
515 server. Nat. Protoc. 4, 363–371. doi:10.1038/nprot.2009.2
516 Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology
517 with a hidden markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.
518 doi:10.1006/jmbi.2000.4315
519 Lassudrie, M., Wikfors, G.H., Sunila, I., Alix, J.H., Dixon, M.S., Combot, D., Soudant, P., Fabioux, C., Hégaret, H.,
520 2015. Physiological and pathological changes in the eastern oyster Crassostrea virginica infested with
521 the trematode Bucephalus sp. and exposed to the toxic dinoflagellate Alexandrium fundyense. J.
522 Invertebr. Pathol. 126, 51–63. doi:10.1016/j.jip.2015.01.011
523 Li, F., Huang, S., Wang, L., Yang, J., Zhang, H., Qiu, L., Li, L., Song, L., 2011. A macrophage migration inhibitory
524 factor like gene from scallop Chlamys farreri: Involvement in immune response and wound healing.
525 Dev. Comp. Immunol. 35, 62–71. doi:10.1016/j.dci.2010.08.009
526 Li, J., Zhang, Y., Zhang, Y., Xiang, Z., Tong, Y., Qu, F., Yu, Z., 2014. Genomic characterization and expression
527 analysis of five novel IL-17 genes in the Pacific oyster, Crassostrea gigas. Fish Shellfish Immunol. 40,
528 455–465. doi:10.1016/j.fsi.2014.07.026
529 Li, P., Spolski, R., Liao, W., Leonard, W.J., 2014. Complex interactions of transcription factors in mediating
530 cytokine biology in T cells. Immunol. Rev. 261, 141–156. doi:10.1111/imr.12199
531 Liu, Q., Xin, W., He, P., Turner, D., Yin, J., Gan, Y., Shi, F.-D., Wu, J., 2014. Interleukin-17 inhibits Adult
532 Hippocampal Neurogenesis. Sci. Rep. 4, 7554. doi:10.1038/srep07554
533 Livak, K.J., Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR
534 and the 2−ΔΔCT Method. Methods 25, 402–408. doi:10.1006/meth.2001.1262
535 Malagoli D, 2010. Cytokine network in invertebrates: the very next phase of comparative immunology. ISJ 7,
536 146–148.
537 Martín-Gómez, L., Villalba, A., Carballal, M.J., Abollo, E., 2014. Molecular characterisation of TNF, AIF,
538 dermatopontin and VAMP genes of the flat oyster Ostrea edulis and analysis of their modulation by
539 diseases. Gene 533, 208–217. doi:10.1016/j.gene.2013.09.085
540 Moreira, R., Milan, M., Balseiro, P., Romero, A., Babbucci, M., Figueras, A., Bargelloni, L., Novoa, B., 2014a.
541 Gene expression profile analysis of Manila clam (Ruditapes philippinarum) hemocytes after a Vibrio
542 alginolyticus challenge using an immune-enriched oligo-microarray. BMC Genomics 15, 267.
543 doi:10.1186/1471-2164-15-267
544 Moreira, R., Pereiro, P., Costa, M.M., Figueras, A., Novoa, B., 2014b. Evaluation of reference genes of Mytilus
545 galloprovincialis and Ruditapes philippinarum infected with three bacteria strains for gene expression
546 analysis. Aquat. Living Resour. 27, 147–152. doi:10.1051/alr/2014015
547 Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian
548 transcriptomes by RNA-Seq. Nat. Methods 5, 621–628. doi:10.1038/nmeth.1226
549 Moseley, T.A., Haudenschild, D.R., Rose, L., Reddi, A.H., 2003. Interleukin-17 family and IL-17 receptors.
550 Cytokine Growth Factor Rev. 14, 155–174. doi:10.1016/S1359-6101(03)00002-9

Nguyen, T.T.T., Hayes, B.J., Ingram, B.A., 2014. Genetic parameters and response to selection in blue mussel (Mytilus galloprovincialis) using a SNP-based pedigree. Aquaculture 420–421, 295–301. doi:10.1016/j.aquaculture.2013.11.021

Novatchkova, M., Leibbrandt, A., Werzowa, J., Neubüser, A., Eisenhaber, F., 2003. The STIR-domain superfamily in signal transduction, development and immunity. Trends Biochem. Sci. 28, 226–229. doi:10.1016/S0968-0004(03)00067-7

Parisi, M.-G., Toubiana, M., Mangano, V., Parrinello, N., Cammarata, M., Roch, P., 2012. MIF from mussel: Coding sequence, phylogeny, polymorphism, 3D model and regulation of expression. Dev. Comp. Immunol. 36, 688–696. doi:10.1016/j.dci.2011.10.014

Pauletto, M., Milan, M., Moreira, R., Novoa, B., Figueras, A., Babbucci, M., Patarnello, T., Bargelloni, L., 2014. Deep transcriptome sequencing of Pecten maximus hemocytes: A genomic resource for bivalve immunology. Fish Shellfish Immunol. 37, 154–165. doi:10.1016/j.fsi.2014.01.017

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786. doi:10.1038/nmeth.1701

Plazzi, F., Passamonti, M., 2010. Towards a molecular phylogeny of Mollusks: Bivalves' early evolution as revealed by mitochondrial genes. Mol. Phylogenet. Evol. 57, 641–657. doi:10.1016/j.ympev.2010.08.032

Poynton, H.C., Robinson, W.E., Blalock, B.J., Hannigan, R.E., 2014. Correlation of transcriptomic responses and metal bioaccumulation in Mytilus edulis L. reveals early indicators of stress. Aquat. Toxicol. 155, 129–141. doi:10.1016/j.aquatox.2014.06.015

Raftos, D., Nair, S., 2004. Tunicate cytokine-like molecules and their involvement in host defense responses. Prog. Mol. Subcell. Biol. 34, 165–182.

Ramilo, A., González, M., Carballal, M.J., Darriba, S., Abollo, E., Villalba, A., 2014. Oyster parasites Bonamia ostreae and B. exitiosa co-occur in Galicia (NW Spain): spatial distribution and infection dynamics. Dis. Aquat. Organ. 110, 123–133. doi:10.3354/dao02673

Roberts, S., Gueguen, Y., de Lorgeril, J., Goetz, F., 2008. Rapid accumulation of an interleukin 17 homolog transcript in Crassostrea gigas hemocytes following bacterial exposure. Dev. Comp. Immunol. 32, 1099–1104. doi:10.1016/j.dci.2008.02.006

Rosani, U., Varotto, L., Domeneghetti, S., Arcangeli, G., Pallavicini, A., Venier, P., 2014. Dual Analysis of Host and Pathogen Transcriptomes in Ostreid Herpesvirus 1 - Positive Crassostrea gigas. Environ. Microbiol. doi:10.1111/1462-2920.12706

Rouvier, E., Luciani, M.F., Mattéi, M.G., Denizot, F., Golstein, P., 1993. CTLA-8, cloned from an activated T cell, bearing AU-rich messenger RNA instability sequences, and homologous to a herpesvirus saimiri gene. J. Immunol. Baltim. Md 1950 150, 5445–5456.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. Clifton NJ 132, 365–386.

Sabat, R., Witte, E., Witte, K., Wolk, K., 2013. IL-22 and IL-17: An Overview, in: Quesniaux, V., Ryffel, B., Di Padova, F. (Eds.), IL-17, IL-22 and Their Producing Cells: Role in Inflammation and Autoimmunity, Progress in Inflammation Research. Springer Basel, pp. 11–35.

Saenz, S.A., Taylor, B.C., Artis, D., 2008. Welcome to the neighborhood: epithelial cell-derived cytokines license innate and adaptive immune responses at mucosal sites. Immunol. Rev. 226, 172–190. doi:10.1111/j.1600-065X.2008.00713.x

Shi, Y., He, M., 2014. Differential gene expression identified by RNA-Seq and qPCR in two sizes of pearl oyster (Pinctada fucata). Gene 538, 313–322. doi:10.1016/j.gene.2014.01.031

Steinman, L., 2007. A brief history of TH17, the first major revision in the TH1/TH2 hypothesis of T cell–mediated tissue damage. Nat. Med. 13, 139–145. doi:10.1038/nm1551

Stewart, M.J., Favrel, P., Rotgans, B.A., Wang, T., Zhao, M., Sohail, M., O'Connor, W.A., Elizur, A., Henry, J., Cummins, S.F., 2014. Neuropeptides encoded by the genomes of the Akoya pearl oyster Pinctata fucata

and Pacific oyster Crassostrea gigas: a bioinformatic and peptidomic survey. BMC Genomics 15, 840. doi:10.1186/1471-2164-15-840

Sun, Y., Zhou, Z., Wang, L., Yang, C., Jianga, S., Song, L., 2014. The immunomodulation of a novel tumor necrosis factor (*Cg*TNF-1) in oyster Crassostrea gigas. Dev. Comp. Immunol. 45, 291–299. doi:10.1016/j.dci.2014.03.007

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. mst197. doi:10.1093/molbev/mst197

Teaniniuraitemoana, V., Huvet, A., Levy, P., Klopp, C., Lhuillier, E., Gaertner-Mazouni, N., Gueguen, Y., Le Moullac, G., 2014. Gonad transcriptome analysis of pearl oyster Pinctada margaritifera: identification of potential sex differentiation and sex determining genes. BMC Genomics 15. doi:10.1186/1471-2164-15-491

Toubiana, M., Rosani, U., Giambelluca, S., Cammarata, M., Gerdol, M., Pallavicini, A., Venier, P., Roch, P., 2014. Toll signal transduction pathway in bivalves: Complete cds of intermediate elements and related gene transcription levels in hemocytes of immune stimulated Mytilus galloprovincialis. Dev. Comp. Immunol. 45, 300–312. doi:10.1016/j.dci.2014.03.021

Turner, M.D., Nedjai, B., Hurst, T., Pennington, D.J., 2014. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. Biochim. Biophys. Acta 1843, 2563–2582. doi:10.1016/j.bbamcr.2014.05.014

Valenzuela-Muñoz, V., Gallardo-Escárate, C., 2014. Molecular cloning and expression of IRAK-4, IL-17 and I-κB genes in Haliotis rufescens challenged with Vibrio anguillarum. Fish Shellfish Immunol. 36, 503–509. doi:10.1016/j.fsi.2013.12.015

Varotto, L., Domeneghetti, S., Rosani, U., Manfrin, C., Cajaraville, M.P., Raccanelli, S., Pallavicini, A., Venier, P., 2013. DNA Damage and Transcriptional Changes in the Gills of Mytilus galloprovincialis Exposed to Nanomolar Doses of Combined Metal Salts (Cd, Cu, Hg). PLoS ONE 8, e54602. doi:10.1371/journal.pone.0054602

Veenstra, J.A., 2010. Neurohormones and neuropeptides encoded by the genome of Lottia gigantea, with reference to other mollusks and insects. Gen. Comp. Endocrinol. 167, 86–103. doi:10.1016/j.ygcen.2010.02.010

Venier, P., De Pittà, C., Bernante, F., Varotto, L., De Nardi, B., Bovo, G., Roch, P., Novoa, B., Figueras, A., Pallavicini, A., Lanfranchi, G., 2009. MytiBase: a knowledgebase of mussel (M. galloprovincialis) transcribed sequences. BMC Genomics 10, 72. doi:10.1186/1471-2164-10-72

Vizzini, A., Di Falco, F., Parrinello, D., Sanfratello, M.A., Mazzarella, C., Parrinello, N., Cammarata, M., 2015. Ciona intestinalis interleukin 17-like genes expression is upregulated by LPS challenge. Dev. Comp. Immunol. 48, 129–137. doi:10.1016/j.dci.2014.09.014

Wang, K., Kim, M.K., Di Caro, G., Wong, J., Shalapour, S., Wan, J., Zhang, W., Zhong, Z., Sanchez-Lopez, E., Wu, L.-W., Taniguchi, K., Feng, Y., Fearon, E., Grivennikov, S.I., Karin, M., 2014. Interleukin-17 receptor a signaling in transformed enterocytes promotes early colorectal tumorigenesis. Immunity 41, 1052–1063. doi:10.1016/j.immuni.2014.11.009

Wu, B., Jin, M., Gong, J., Du, X., Bai, Z., 2011. Dynamic evolution of CIKS (TRAF3IP2/Act1) in metazoans. Dev. Comp. Immunol. 35, 1186–1192. doi:10.1016/j.dci.2011.03.027

Wu, S.-Z., Huang, X.-D., Li, Q., He, M.-X., 2013. Interleukin-17 in pearl oyster (Pinctada fucata): Molecular cloning and functional characterization. Fish Shellfish Immunol. 34, 1050–1056. doi:10.1016/j.fsi.2013.01.005

Xu, S., Cao, X., 2010. Interleukin-17 and its expanding biological functions. Cell. Mol. Immunol. 7, 164–174. doi:10.1038/cmi.2010.21

Zhang, L., Li, L., Zhu, Y., Zhang, G., Guo, X., 2014. Transcriptome Analysis Reveals a Rich Gene Set Related to Innate Immunity in the Eastern Oyster (Crassostrea virginica). Mar. Biotechnol. 16, 17–33. doi:10.1007/s10126-013-9526-z

Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5. doi:10.1038/srep08693

Legends of Figures, Tables, Supplementary tables and figures

Figure 1. Radial diagram of IL-17 domains available in PFAM (312 sequences, PF06083) or reported in the present work (59 sequences). The latter group is included in the branch highlighted in grey. The tree is built on conserved positions (54 % of the 153 aligned amino acid positions).

Figure 2. Multiple alignment of the IL-17 domains identified in mollusks with human IL-17-F (AAH70124) domain. Sequence IDs recall the organism name as codified in Table 1. Amino acid conservation level is indicated in gray scale (white: 0-33 % of conservation, light gray 34-66 % and dark gray 67-100 %). A black bar plot reports the conservation level of each aminoacidic residue. Putative disulfide bonds are also indicated.

Figure 3. Phylogenetic analysis of the SEFIR domains identified in the analyzed organisms (96 sequences), together with human IL-17Ra (AAH11624) and human CIKS (AAH02823). The distance tree is calculated on conserved positions (51 % of the 200 aligned positions) with 1000 bootstrap replicates. Bootstrap values are reported and branches with confident values (higher than 700) are in bold. A grey frame highlights the domains related to proteins indicated as IL-17R-like in the text. Black arrows indicate *M. galloprovincialis* proteins, while empty circles indicate human CIKS (AAH02823) and human IL-17Ra (AAH11624).

Figure 4. Model structure of domain of the six mussel and the ten oyster IL-17s (predicted by homology to available protein models or *ab-initio)*. **A.** Superposed models of *Mg*IL-17-1, *Mg*IL-17-2, *Mg*IL-17-3, *Mg*IL-17-4 and *Mg*IL-17-5. **B.** *Mg*IL-17-6. **C.** Superposed models of *Cg*IL-17-1, *Cg*IL-17-2, *Cg*IL-17-3, *Cg*IL-17-5, *Cg*IL-17-7, *Cg*IL-17-8, *Cg*IL-17-9 and *Mg*IL-17-10. **D.** *Cg*IL-17-6.

Figure 5. Gene expression levels of mussel (A) and oyster (B) IL-17s, IL-17Rs and CIKS(L) as measured in different RNA-seq datasets (see Table 1). Gill (GILL), hemolymph (HAE), digestive gland (DG), mantel (iMAN) and muscle (MUS) are represented for *Mg* (A). Hemocytes (HAE), digestive gland (DG), gills (GILL), inner mantle (iMAN), mantle edge (eMAN), male and female gonads (mGON and fGON), labial palp (LAB), muscle (MUS) and the whole tissues or OsHV-1-infected spat (whole SPAT) are represented for oyster *C. gigas* (B). Color symbols are reported once for both the diagrams. Expression data of each dataset were normalized and compared to the Elongation Factor 1α expression levels.

Figure 6. Relative quantification (RQ) of *IL-17* transcripts detectable in the hemocytes (left) and gills (right) of mussels injected with heat-killed bacteria (pooled samples, N=6). Bars indicate transcript levels normalized to the endogenous control EF1α and relative to time-paired (PBS-NaCl injected) controls. Two-tailed paired t test: *=$p < 0.05$; **=$p< 0.01$, mean value ± standard deviation are reported.

Figure 7. Relative quantification (RQ) of *IL-17R* and *CIKSL* transcripts in the hemocytes (left) and gills (right) of mussels injected with heat-killed bacteria. Two-tailed paired t test: *=$p < 0.05$; **=$p< 0.01$, mean value ± standard deviation are reported.

Figure 8. Model of the possible interplay between the IL-17 and TLR signaling pathways in bivalves. IL-17Rs and TLRs are members of the TIR superfamily. Once activated by IL-17 and Pathogen-Associated Molecular

Patterns, respectively, these transmembrane receptors can support the development of an inflammatory response. **Left**: subsequent to IL-17binding, IL-17Rs engage their SEFIR domains to recruit the adaptor/ubiquitin ligase CIKS or CISKL. CISK contains a TRAF-binding motif and can bind TRAF6 or interact with TRAF5 and TRAF2 (the dashed arrows from CIKSL display possible interactions). Peculiarly, CIKSL contains an additional DD-domain which could interact with the DD domain of IRAK to activate TRAF6. CIKS-mediated activation of the MAPK pathway is also expected to promote the inflammation response via the transcription factor AP-1. **Right**: once TLRs are activated, the adaptor protein MyD88 binds to the receptor complex by TIR-TIR interaction to recruit IRAK and TRAF6. Downstream TRAF6, other proteins and the transcription factor NF-*κ*B can lead to the expression of cytokines and antimicrobial peptides, and define a further level for the control of the inflammatory reaction. **Abbreviations**: AP-1, activator protein 1; CIKS, connection to IkB kinase and stress-activated protein kinase; CIKSL, CIKS-like; DD, death-domain; IL-17, interleukin 17; IL-17R, interleukin17 receptor; IRAK, interleukin receptor associated kinase; MAPKs, mitogen-activated protein kinase; MyD88, myeloid differentiation primary response gene; NF-*k*B, nuclear factor-kappa-B; SEFIR, similar expression to fibroblast growth factor genes/interleukin-17 receptor; TIR Toll/Il-1 Receptor domain; TLR, Toll-like receptor; TRAF, tumor necrosis factor receptor-associated factor.

Table 1. Sequence datasets selected for protein domain searches.

Acronym	Project ID	Species	Order	Sample type	Sequencing platform	Read Number [M]	Contig Number	Numbe predicte
aplcal	\	*Aplysia californica*	Gasteropoda	Genome	\	\	\	24,77
lotgig	\	*Lottia gigantea*	Gasteropoda	Genome	\	\	\	23,34
anatra	*PRJNA210944*	*Anadara trapezia*	Arcoida	Mixed tissues	illumina	13.1	42,924	22,52
teggra	*PRJNA159979*	*Tegillarca granosa*	Arcoida	Mixed tissues	Illumina	1.2	22,744	11,53
batazo	*PRJNA79785*	*Bathymodiolus azoricus*	Mytiloida	Gills	454	0.6	22,914	5,78
mytcal	*PRJNA249058*	*Mytilus californianus*	Mytiloida	Mixed tissues	illumina	38.8	84,012	27,27
mytedu	*PRJNA252953PRJNA249058*	*Mytilus edulis*	Mytiloida	Mixed tissues	illumina	390.7	229,951	67,73
mytgal	*PRJNA88481*	*Mytilus galloprovincialis*	Mytiloida	Mixed tissues	illumina	642.5	284,182	81,96
myttro	*PRJNA249058*	*Mytilus trossolus*	Mytiloida	Mixed tissues	illumina	57.5	126,144	38,59
pervir	*PRJNA254094*	*Perna viridis*	Mytiloida	Mixed tissues	Ilumina	281	87,853	30,36
ennten	\	*Ennucula tenuis*	Nuculoida	Not specified	Illumina	8.1	68,090	28,46
cracor	*PRJNA237222*	*Crassostrea corteziensis*	Ostreoida	Not specified	Illumina	13.5	64,284	28,85

cragig	\	*Crassostrea gigas*	Ostreoida	Genome	Illumina	\	388,000	26,86
crahon	*PRJNA223230*	*Crassostrea hongkongensis*	Ostreoida	Mixed tissues	Illumina	24.7	47,357	27,42
cravir	*PRJNA82611*	*Crassostrea virginica*	Ostreoida	Not specified	Illumina	4.8	50,802	29,66
craang	*PRJNA24021*	*Crassostrea angulata*	Ostreoida	Muscle	Illumina	44.8	41,686	25,50
ostchil	*PRJNA249058*	*Ostrea chilensis*	Ostreoida	Mixed tissues	Illumina	8.1	18,770	10,77
ostedu	*PRJNA249058*	*Ostrea edulis*	Ostreoida	Mixed tissues	illumina	46.7	57,360	22,12
ostlur	*PRJNA171849*	*Ostrea lurida*	Ostreoida	Larvae	Illumina	388.5	41,206	16,53
ostste	*PRJNA249058*	*Ostreola stentina*	Ostreoida	Mixed tissues	Illumina	8.2	12,206	5,69
sacglo	*PRJNA253158*	*Saccostrea glomerata*	Ostreoida	Mixed tissues	454	0.8	53,280	21,17
argirr	*PRJNA237853*	*Argopecten irradians*	Pectinoida	Mixed tissues	illumina	76.3	33,994	20,55
mizyes	*PRJNA186890*	*Mizuhopecten yessoensis*	Pectinoida	Mixed tissues	Illumina	94.7	117,295	31,18
pecmax	*PRJNA222492*	*Pecten maximus*	Pectinoida	Hemolymph	illumina	108.2	37,485	18,44
pinfuc	\	*Pinctada fucata*	Pterioida	Genome		\	\	72,58
solvel	\	*Solemya velum*	Solemyoida	Not specified	illumina	33.3	51,268	20,77
neomar	*PRJNA249058*	*Neotrigonia margaritacea*	Trigoniida	Foot	illumina	24.1	55,140	15,68
elicom	*PRJNA194430*	*Elliptio complanata*	Unionoida	Not specified	454	\	89,063	24,52
pyggra	\	*Pyganodon grandis*	Unionoida	Muscle	illumina	100.5	86,957	23,12
unitet	\	*Uniomerus tetralasmus*	Unionoida	Muscle	illumina	113.1	123,694	26,43
villie	*PRJNA75063*	*Villosa lienosa*	Unionoida	Not specified	illumina	81.5	125,347	28,11
corflu	*PRJNA183608*	*Corbicula fluminea*	Veneroida	Mixed tissues	illumina	33.5	27,804	14,50
mermer	*PRJNA257120*	*Meretrix meretrix*	Veneroida	Mantle	illumina	56.6	41,719	21,59
ruddec	*PRJNA218511*	*Ruditapes decussatus*	Veneroida	Gonad	illumina	171.2	88,151	34,21
rudphi	*PRJNA234093*	*Ruditapes philippinarum*	Veneroida	Mixed tissues	Illumina	56.2	68,215	22,85

Table 1 Footer. Acronym, project ID, species and order name, sample type, sequencing technology, millions of available reads, number of assembled contigs and of predicted cds are reported. The four mollusk species with sequenced genome are highlighted in grey. Only for *C. gigas*, both gene predictions and transcriptome assembly were used for the domain searches.

704 **Table 2**. Quantitative RT-PCR analysis of mussel transcripts related to the IL-17 signaling pathway.

ID	GenBank ID	Forward primer (5'-3')	Reverse primer (5'-3')	Amplicon size (bp)
MgIL-17-1	KP988296	ACGACATGCTACAGCCACTG	AGAAGAATCGTGTCCGCATT	247
MgIL-17-2	KP988297	ACGGACATGTGGATCTGTGA	ACGGGATGTGGAAAAACTTG	192
MgIL-17-3	KP988298	ACGGAAGGTTGGACACAGAA	TGTTTCAGCGCTGTATGGAG	177
MgIL-17-4	KP988299	CCTTCGGGATCTTCTGGTTT	AAGCGCCAAGAAATAGCAAA	250
MgIL-17-5	KP988300	TCCATGGTATCCATGTGTGG	TGCACAGAGCCATCTCAGAA	248
MgIL-17-6	KP988301	AGGATGTATAGGTGGGAAGA	GCGACAGGAAGGAGTATTT	128
MgIL-17Ra	KP998600	ACACCACCCATAATGCAGCT	TCTTCGTGTTTGGCTGGAGT	229
MgIL-17Rb	KP998601	CATTCGAGGTTTGTGGTGCC	GCAAAACCAAGGCAGAGACC	129
MgIL-17R-like	KP998602	CGCCAACAAACACATCAAA	CCTCCAAACCACAAAACA	95
MgCIKSL	KP998603	TACCCACGGCAAGCAGTATT	TTGCTTGACCGTGAAACAGC	164
Mg EF1α	AB162021.1	CAAGACCCACAGACAAAGC	GGAGCAAAGGTAACAACCAT	130

705 *Table 2 footer*. Identity codes, primer sequences and expected amplicon size are reported (further details in
706 Table 4).

707 **Table 3**. Number of identified IL-17, IL-17R and CIKS protein domains.

Species	Family	IL-17 domains	SEFIR domains		
			IL-17R	CIKS(L)	Total number
Aplysia californica	Gasteropoda	**4**	2		**2**
Lottia gigantea	Gasteropoda	**9**	3	1	**4**
Anadara trapezia	Arcoida		1	1	**2**
Tegillarca granosa	Arcoida		1	1	**2**
Bathymodiolus azoricus	Mytiloida		1		**1**
Mytilus californianus	Mytiloida	**1**	2		**2**
Mytilus edulis	Mytiloida	**3**	1	1	**2**
Mytilus galloprovincialis	Mytiloida	**6**	3	1	**4**
Mytilus trossolus	Mytiloida	**2**	2		**2**

Perna viridis	Mytiloida	**3**	4	1	**5**
Ennucula tenuis	Nuculoida		2		**2**
Crassostrea corteziensis	Ostreoida	**1**	2	2	**4**
Crassostrea gigas	Ostreoida	**10**	4	3	**7**
Crassostrea hongkongensis	Ostreoida		2	1	**3**
Crassostrea virginica	Ostreoida		2	4	**6**
Crassostrea angulata	Ostreoida	**2**		2	**2**
Ostrea chilensis	Ostreoida				
Ostrea edulis	Ostreoida		1	2	**3**
Ostrea lurida	Ostreoida			2	**2**
Ostreola stentina	Ostreoida				
Saccostrea glomerata	Ostreoida		1	2	**3**
Argopecten irradians	Pectinoida	**1**	4		**4**
Mizuhopecten yessoensis	Pectinoida	**3**	2	2	**4**
Pecten maximus	Pectinoida	**3**		1	**1**
Pinctada fucata	Pterioida	**4**	2	2	**4**
Solemya velum	Solemyoida	**1**	4	1	**5**
Neotrigonia margaritacea	Trigoniida				
Elliptio complanata	Unionoida	**1**		2	**2**
Pyganodon grandis	Unionoida	**1**	5	1	**6**
Uniomerus tetralasmus	Unionoida	**1**	3	2	**5**
Villosa lienosa	Unionoida		1	1	**2**
Corbicula fluminea	Veneroida	**1**			
Meretrix meretrix	Veneroida		1	1	**2**
Ruditapes decussatus	Veneroida	**1**	3		**3**
Ruditapes philippinarum	Veneroida	**1**			

Table 3 footer. Protein domains were identified in the transcript assemblies of 31 bivalve *spp*. by HMMER analysis or using gene predictions in *A. californica*, *L. gigantea* and *P. fucata*. For *C. gigas*, both gene predictions and transcriptome assembly were used. Organism and family name, number of sequences containing the IL-17 or SEFIR domains are reported (the latter detected in both IL-17 receptors -or receptor like- and CIKS adaptors). Lines reporting the four genomic datasets are shadowed in gray.

Table 4. Transcripts denoting IL-17, IL-17R and CIKSL proteins in *Mytilus spp*.

ID	Description	Transcript length (nt)	Protein length (AA)	Domain (s)	Signal peptide	Trans-membrane region
			M. galloprovincialis			
KP988296	IL-17-1	1131	194	IL-17	yes	no
KP988297	IL-17-2	739	192	IL-17	yes	no
KP988298	IL-17-3	1050	194	IL-17	yes	no
KP988299	IL-17-4	1048	165	IL-17	yes	no
KP988300	IL-17-5	746	221	IL-17	yes	no
KP988301	IL-17-6	885	175	IL-17	no	no
KP998600	IL-17Ra	2665	673	SEFIR	no	yes
KP998601	IL-17Rb	2724	727	SEFIR	yes	yes
KP998602	IL-17R-like	2062	591	SEFIR	no	yes
KP998603	CIKSL	2535	642	DD + SEFIR	no	no
			M. edulis			
HE609101	IL-17-	1498	169	IL-17	yes	no
HE609102	IL-17-3	2246	194	IL-17	yes	no
/	IL-17-6	762	141	IL-17	no	no
/	IL-17R-like	2081	591	SEFIR	no	yes
/	CIKSL	2074	642	DD + SEFIR	no	no
			M. trossulus			
/	IL-17-1	844	185	IL-17	yes	no

/	IL-17-3	719	194	IL-17	yes	no
/	IL-17Rb	1763	541	SEFIR	no	yes
/	IL-17R-like	627	207	SEFIR	no	yes
			M. californianus			
/	IL-17-2	1033	192	IL-17	yes	no
/	IL-17Rb	2818	856	SEFIR	yes	yes
/	IL-17Rb	2188	686	SEFIR	yes	yes

Table 4 footer. GenBank ID, description, length of transcripts and deduced proteins, identified domain(s), predicted signal peptide and transmembrane region are reported.

Table 5. Updated list of the IL-17 signaling components in *C. gigas*.

Contig ID	Description	Transcript length (nt)	Protein length (AA)	Domain (s)	Signal peptide	Trans-membrane region
CGI_10015251*	IL-17-1	600	150	IL17	yes	no
CGI_10004922	IL-17-2	399	132	IL17	yes	no
CGI_10025754	IL-17-3	603	200	IL17	yes	no
CGI_10020734	IL-17-4	573	190	IL17	yes	no
CGI_10027182	IL-17-5	504	167	IL17	yes	no
CGI_10026592*	IL-17-6	1344	185	IL17	yes	no
CGI_10014828	IL-17-7	426	141	IL17	no	no
CGI_10026344*	IL-17-8	500	115	IL17	yes	no
new	IL-17-9	636	181	IL17	yes	no
new	IL-17-10	836	171	IL17	yes	no
CGI_10021486*	IL-17Ra	3216	576	SEFIR	yes	yes
CGI_10015683+4*	IL-17Rb	2969	877	SEFIR	yes	yes
CGI_10002512*	IL-17R-like	2359	542	SEFIR	no	yes
new	*Cg*CIKSL	3130	689	DD+SEFIR	no	no

CGI_10027691	CIKS-2	879	292	SEFIR	no	no
CGI_10027692*	CIKS-3	2758	521	SEFIR	no	no

Table 5 footer. Contig ID and description, transcript and protein length, identified domain(s), predicted signal peptide and transmembrane regions are reported.* cds have been updated from those present in GenBank.

Table 6. Updated list of the IL-17 gene sequences identified in mussel (*Mg*) and oyster (*C. gigas*).

Gene name	Status	Genomic location	Gene structure	Size (kb)	No. of paired reads
		M. galloprovincialis			
IL-17-1	new	APJB012081501	2 exons (5)	3.6	306
IL-17-2	new	APJB011262727	2 exons (2)	>1	164
IL-17-3	new	APJB010402434	2 exons (4)	>1	4,598
IL-17-4	new	*Mg*IL-17-4_gene	2 exons (7)	5.6	1,578
IL-17-5	new	APJB010987114	2 exons (5)	>1	186
IL-17-6	new	*Mg*IL-17-6_gene	1 exon	0.6	82
IL-17Ra	new	*Mg*IL-17Ra_gene	8 exons (3 closed introns)	17.5	28,578
IL-17Rb	new	*Mg*IL-17Rb_gene	11 exons (6 closed introns)	17.5	3,254
IL-17R-like	new	*Mg*IL-17Rc_gene	7 exons (2/6 closed introns)	9.5	57,288
CIKSL	new	*Mg*CIKSL_gene	at least 4 exons	13.3	5,660
		C. gigas			
IL-17-1	updated	JH816757 (18721-20497)	2 exons (9)	1.8	330
IL-17-2	CGI_10004922	JH817760 (61844-62141)	1 exon	0.3	5,506
IL-17-3	CGI_10025754	JH819116 (103341-104268)	2 exons (11)	1	3,238
IL-17-4	updated	JH819141 (199585-201418)	2 exons (9)	1.8	924
IL-17-5	CGI_10027182	JH823231 (1146436-1149275)	3 exons (4)	2.8	102
IL-17-6	updated	JH816764 (246935-248994)	2 exons (5)	2	24,702

IL-17-7	CGI_10014828	JH819178 (222043-222468)	partially covered by RNA-seq	0.5	64
IL-17-8	updated	JH819021 (256768-257369)	2exons (1)	0.6	36
IL-17-9	new	JH816022 (102677-103641)	2 exons (5)	1	684
IL-17-10	new	JH817099 (1632298-1637550)	2 exons (1)	5,3	146
IL-17Ra	updated	Not detected	/	/	/
IL-17Rb	updated	JH818967 (89678-106925)	7 exons	17.2	57,690
IL-17R-like	updated	Multiple contigs	/	/	/
CIKSL	new	JH819194 (566779-587522)	6 exons	20.7	35,336
CIKS_2	CGI_10027691	JH816755 (1235971-1236957)	2 exons	1	10,266
CIKS_3	updated	JH816755 (1245133-1256307)	7 exons	11.2	36,674

Table 6 footer. Gene name, status, genome location, gene structure, dimension in kilobases and number of paired reads supporting the contig sequence are reported. Status indicates if the gene is new, previously recorded (with a contig ID) or if it has been updated. Gene structure indicates the number of exons and for IL-17 genes, in brackets, the number of AA located in the first exon.

Supplementary Figure 1. Heat map summarizing oyster IL-17 pathway components in 52 RNA-seq datasets. Expression values are reported as normalized RPKM in a color scale from low (green) to high (red) expression values. Genes were clustered using a Manhattan complete linkage algorithm (right). The analyzed samples represent OsHV-1-positive oysters (E-MTAB-2552), different bacterial infections (14 samples: SRR796582-98) and developmental stages (38 samples: SRR334222-59). Further details are reported below in Supplementary Table 1. Total gene expression data are available in Supplementary Table 2.

Supplementary Table 1. Details of all analyzed RNA-seq samples (from Zhang et al. 2015 and from Rosani et al.,2014). Sample type, SRA ID, description and number of reads are reported.

Supplementary Table 2. Normalized RPKM values for oyster IL-17 pathway components in the 53 analyzed RNA-seq samples.