UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOMEDICINA COMPARATA E ALIMENTAZIONE

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE VETERINARIE

CICLO XXVI

Reproduction and immunology: transcriptomic approaches to improve bivalves farming

31 gennaio 2014

Direttore della Scuola : Ch.mo Prof. Gianfranco Gabai

Supervisore :Ch.mo Prof. Tomaso Patarnello

Dottoranda : Marianna Pauletto

A Guido e Pietro

CONTENTS

<u>SECTION C</u>: "RNA INTERFERENCE TARGETING IKB2 TO INVESTIGATE THE *CRASSOSTREA GIGAS* IMMUNE RESPONSE AGAINST OYSTERS HERPESVIRUS TYPE I"

Abstract

Over the last years, there has been an increasing demand of shellfish for consumption. The ability of traditional capture fisheries to supply bivalves is unlikely to increase significantly because of the drops in natural recruitments of seed, which are mainly due to overexploitation of natural stocks. Production of bivalve seed in hatcheries is a relatively new industry for which empirical approaches were developed, adapting methods across species and measuring the resulting effect in terms of growth and survival. To date, only a few bivalve species of major aquacultural importance in Europe have benefited from newly developed genomic resources (e.g. microarrays and RNA sequencing) and gene silencing approaches, which are both expected to significantly improve the biological knowledge on these commercially important species in the coming years.

The present PhD thesis aimed firstly at increasing, through Next Generation Sequencing, genomic information available for two emerging species, *Venerupis decussata* and *Pecten maximus*.

The second aim was to investigate two main bottlenecks hampering the bivalve production in hatchery: efficiency of reproduction and susceptibility to pathogens. By means of microarray analysis, a gene expression study on *V .decussata* oocytes at different maturation stages was performed and the major biological processes involved in gamete maturation were identified.

A RNA sequencing experiment was conducted on pathogen-challenged hemocytes and unchallenged controls to study the *P. maximus* immune transcriptome. mRNAs encoding proteins with a known immune function were detected and a global analysis of differential expression comparing gene-expression levels in stimulated hemocytes against controls provided evidence on a large set of transcripts involved in *P. maximus* immune response.

Finally, reverse genetic and real-time PCR were implemented to investigate the role of IkB2 in the *Crassostrea gigas* immune response against Ostreid herpesvirus type 1. Following the injection of a dsRNA targeting IkB2, juveniles were infected with OsHV-1 and mRNA levels of four immune genes (IkB1, IkB2, Rel, SOCS) were evaluated in gonads and gills demonstrating their implication in the Pacific oyster antiviral response.

Riassunto

Nel corso degli ultimi anni, si assistito ad una crescita della domanda di molluschi destinati al consumo alimentare, alla quale però la pesca tradizionale non ha saputo far fronte, principalmente a causa dell'impoverimento degli stock naturali e alla conseguente scarsità di seme. La produzione di seme di bivalvi in schiuditoio è un'attività relativamente recente e le attuali metodiche di allevamento non sono altro che protocolli già usati in altre specie, adattati e verificati in termini di crescita e sopravvivenza. Ad oggi, ad aver beneficiato delle recenti risorse genomiche, come *microarry* ed *RNA sequencing*, e delle sofisticate strategie di silenziamento genico sono state solo alcune specie di bivalvi di notevole interesse alimentare a livello europeo. Tuttavia, si pensa che, nel corso dei prossimi anni, questi approcci molecolari possano costituire una risorsa importante per lo studio della biologia di numerose specie di bivalvi non ancora allevate su larga scala ma di crescente interesse commerciale.

Il primo obiettivo della presente tesi di dottorato è stato quello di incrementare, mediante tecniche di *Next Generation Sequencing*, le risorse genomiche attualmente disponibili per due specie emergenti: la vongola verace *Venerupis decussata* e la cappasanta atlantica *Pecten maximus*.

In secondo luogo, si sono considerati due importanti fattori che ostacolano l'allevamento dei bivalvi in schiuditoio: l'efficienza riproduttiva e la suscettibilità ai patogeni. Utilizzando una piattaforma *microarray*, sono stati valutati i livelli di espressione genica di oociti di *V. decussata* a due diversi stadi di maturazione. L'analisi dei dati ha quindi permesso di identificare i principali processi biologici coinvolti nella maturazione dei gameti femminili. Inoltre, al fine di studiare il trascrittoma immunitario di *P. maximus*, è stato effettuato un esperimento di *RNA sequencing* su emociti immuno-stimolati e di controllo. Questo studio ha permesso di identificare trascritti di mRNA che codificano per importanti proteine del sistema immunitario. In aggiunta l'individuazione di geni differenzialmente espressi in emociti stimolati e controlli ha messo in evidenza un set di trascritti potenzialmente implicati nella risposta immunitaria di *P. maximus*.

Infine, allo scopo di valutare il ruolo di IkB2 nella risposta immunitaria di *Crassostrea gigas* all'infezione da herpesvirus di tipo 1, è stato effettuato un esperimento di *RNA interference*. In seguito all'iniezione di un RNA a doppio filamento codificante una porzione del trascritto IkB2, individui giovani di *C. gigas* sono stati infettati con un omogenato contenente il virus HV-1. Infine, per valutare l'importanza di quattro geni

della risposta immunitaria (IkB1, IkB2, Rel, SOCS), i loro livelli di espressione sono stati valutati in due diversi tessuti: le gonadi e le branchie.

1. Introduction

Over the last years, there has been an increasing demand of shellfish for consumption. The ability of traditional capture fisheries to supply bivalves is unlikely to increase significantly because of the drops in natural recruitments of seed, which are mainly due to overexploitation of natural stocks. In fact, about 29.9 percent of natural fish stocks are being harvested to near maximum rates (*Helm et al., 2004*), producing lower yields than their biological and ecological potential. The overexploitation issue has already been discussed during the World Summit on Sustainable Development at Johannesburg (2002), resulting in the need of strict management plans for restoring the overexploited stocks to the level that can produce maximum sustainable yield by 2015, a target that seems unlikely to be met. Moreover, natural spatfall harvesting is vulnerable to adverse impacts of natural, socioeconomic, environmental and technological conditions, which lead inevitably to "bad" and "good" years for seed supply, determining also a tumbling price.

The increasing market demand, the unpredictability of bivalve production and the concurrent impoverishment of the wild sets generated a great interest in hatchery-produced, a solution which could meet spat requirements in the bivalve industry, and could be applicable to the production of high unit value species such as clams, oysters and scallops. Noteworthy, seed supply is already crucial for some species like the hard clam *Mercenaria mercenaria*, for which farmers have to rely entirely on hatcheries because, unlike most bivalves, large quantities of their seed cannot be easily harvested in the wild (*FAO, 2013*). Other species, such as *Crassostrea gigas, Venerupis philippinarum* and *Argopecten irradians* are routinely produced in hatcheries throughout the world and rely only partially on wild spatfall. While hatchery technology is relatively well developed for these species, other bivalves such as *Venerupis decussata*, *Pecten maximus*, *Placopecten magellanicus* and *Ostrea edulis* proved more difficult to rear on a routine basis.

Despite academic knowledge is continuously advancing, the production of bivalve seed in hatcheries and nurseries is a relatively new issue for which empirical approaches were developed, adapting methods across species and measuring the resulting effect in terms of growth and survival. Hatchery production is still affected by several limitations, which are often species-specific and are encountered at different stages in the biological cycle of molluscs, leading to high mortality. The main bottlenecks involve (i) broodstock management and gamete quality, (ii) appropriate methods for larval rearing, (iii) metamorphosis synchronisation and improvement of settlement, (iv) quality of seed in terms of immunity, genetic diversity and sanitary status.

Applied research on physiology provided a better understanding of the effects of biotic (micro-algal diet, bacterial communities, etc.) and abiotic (temperature, salinity, etc.) factors (e.g. *Nicolas et al. 2004; Magnesen et al. 2006*). However, most of these studies only had a limited impact on the bivalve industry. Recent resources that are expected to affect the bivalve aquaculture in the coming years are those based on "omic" studies. In the last decade, it has been demonstrated that "omic" characterization of bivalves revealed many potential applications most notably for pollution biomonitoring (*Suarez-Ulloa et al., 2013*). Yet, the relevence of omics for research on bivalve development is still poorly explored. To date, only few bivalve species of major aquaculture importance in Europe have benefited of omic approaches, like microarrays, high-throughput sequencing technologies and 2-dimensional electrophoresis (2-DE). Recent studies characterized the larval transcriptome in *Crassostrea angulata* (*Qin et al. 2012*) and *Meretrix meretrix* (*Huan et al. 2012*), providing important and useful information on complex and unique early developmental processes. Besides, a 2-DE proteomic approach was exploited in order to identify abundant proteins linked to egg quality in the Pacific oyster *C. gigas*, improving the current knowledge upon the proteomic processes involved in their developmental competence (*Corporeau et al. 2012*). Despite a large quantity of studies exploiting advanced molecular approaches have been already carried out on the aforementioned established species, however there are still several issues to be addressed and further studies are still needed. An example is the phenomenon of oyster mass mortalities, which since 1970s have occurred at a historically unprecedented rate and they are still causing high economic losses in several countries (*Harvell et al. 1999, 2002; Lafferty et al. 2004; Mydlarz et al. 2006*).

Conversely, in the "emerging species", for which hatchery-based production of seed is still limited in Europe, these promising resources have been poorly exploited. Although some studies have been performed, these mainly concern adult tissues rather of development stages. Moreover, such investigations are still limited and not comprehensive enough to formulate integrated theories about bivalves developmental biology.

In the present PhD thesis, by means of Next Generation Sequencing techniques (pyrosequencing and RNA sequencing) a large amount of cDNA sequences were obtained for two emerging species *V. decussata* and *P. maximus*. Moreover gene expression studies were performed to improve the knowledge on the major biological

mechanisms regulating *V. decussata* oocytes maturation and immune response in *P. maximus* and *C. gigas* as well.

References

Corporeau C, Vanderplancke G, Boulais M, Suquet M, Quéré C, Boudry P et al. Proteomic identification of quality factors for oocytes in the Pacific oyster *Crassostrea gigas*. J Proteomics 2012;75:5554-63.

FAO, 2013. Cultured Aquatic Species Information Programme: *Mercenaria mercenaria*. Available at: http://www.fao.org/fishery/culturedspecies/Mercenaria_mercenaria/en (accessed 25.01.14).

Harvell CD, Kim K, Burkholder JM, Colwell RR, Epstein PR, Grimes DJ et al. Emerging marine diseases-climate links and anthropogenic factors. Science 1999;285:1505−1510.

Harvell CD, Mitchell CE, Ward JR, Altizer S, Dobson AP, Ostfeld RS et al. Climate warming and disease risks for terrestrial and marine biota. Science 2002;296:2158−2162.

Helm M, Bourne N and Lovatelli A, editors. 2004. Hatchery culture of bivalves: a practical manual. FAO Fisheries Technical Paper Number 471. Food and Agriculture Organization, Rome, Italy. Available online at: http://www.fao.org/docrep/007/y5720e/y5720e00.htm.

Huan P, Wang H, Liu B. Transcriptomic Analysis of the Clam *Meretrix meretrix* on Different Larval Stages. Mar Biotechnol 2012;14:69-78.

Magnesen T, Bergh Ø, Christophersen G. Yields of great scallop, *Pecten maximus*, larvae in a commercial flow-through rearing system in Norway. Aquacul Int 2006;14: 377-394.

Mydlarz LD, Jones LE, Harvell CD. Innate immunity environmental drivers and disease ecology of marine and freshwater invertebrates. Annu Rev Ecol Evol Syst 2006;37:251-288.

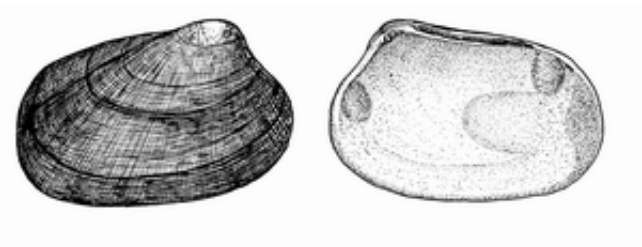
Nicolas J, Corre S, Cochard Jean-Claude. Bacterial population association with phytoplankton cultured in a bivalve hatchery . Microbial Ecol 2004;48: 400-413.

Qin J, Huang Z, Chen J, Zou Q, You W, et al. Sequencing and de novo Analysis of *Crassostrea angulata* (Fujian Oyster) from 8 Different Developing Phases Using 454 GSFlx. PLoS ONE 2012;7: e43653.

Suárez-Ulloa V, Fernández-Tajes J, Manfrin C, Gerdol M, Venier P, Eirín-López JM. Bivalve Omics: State of the Art and Potential Applications for the Biomonitoring of Harmful Marine Compounds. Mar Drugs 2013; 11:4370-4389.

Section A

“Spawned and ovarian oocytes: a gene expression study on *Venerupis decussata*”



1. Background

The clam *Venerupis decussata* is a native European bivalve species and, even if its global aquaculture production is still relatively low in Europe (4.137 tons in 2011; FAO), it has an high economic value. *V. decussata* production is economically important in many Mediterranean countries, mainly Portugal, Italy and Spain but, due to the difficulties in broodstock conditioning and larval rearing (*Hamida et al. 2004*), the culture of this species relies mainly on natural recruitment of seed and it is therefore limited by its availability.

Among the major hurdles reported in the hatchery production of this species, the incapacity in controlling the spawning and the deficient gametes are important issues which highly affect the hatcheries yield. Notably, the spawning success in the European clam is not guaranteed and therefore remains still random. Furthermore the variability of eggs quality, defined as the potential of oocytes to produce a viable progeny (*Kjorsvik et al., 1990*) is still too variable to ensure a technically and economically reliable source of seed.

A further factor which strongly limits the hatchery development of *V. decussata* is the impossibility to obtain fertile gametes by gonadal stripping. This practice is widely used to collect oocytes, before natural spawning, in some bivalve species (i.e. oysters) whose eggs can be fertilized. Indeed, meiotic progression occurring in germ cells is not identically regulated in molluscan species. While full-grown oocytes of all bivalves are blocked in ovaries at prophase I stage, still differences are present. In bivalves such as *Spisula* or *Barnea* spawned oocytes are arrested at prophase I and fertilization occurs at this stage leading to meiosis reinitiation (*Dubé and Guerrier, 1982; Longo, 1983; Colas and Dubé, 1998*). In contrast, bivalves such as *Venerupis* and *Crassostrea* (*Osanai and Kuraishi, 1988*) exit from prophase I and undergo germinal vesicle breakdown (GVBD) after spawning and then are further blocked at the first metaphase (metaphase I). The release from metaphase I is naturally triggered by fertilization or can be artificially induced, however in both cases, it seems that an increase in intracellular [Ca2+] has a pivotal role in the meiosis reinitiation (*Abdelmajid et al., 1993; Guerrier et al. 1993; Moreau et al., 1996*). Although both *Venerupis* and *Crassostrea* oocytes encounter two blockage during meiosis I, their meiotic progression is not completely regulated in the same way. Despite naturally spawned oyster oocytes, like in *Venerupis*, remain at metaphase I and wait for fertilization to reenter meiosis, however, isolated oocytes from ovaries (stripped) remain at prophase and can be fertilized at this stage (*Osanai, 1985*).

Noteworthy, the reasons why *Venerupis* oocytes cannot be fertilized at prophase I remain still unknown.

To date, the mechanisms controlling oocytes maturation in *V. decussata* have been scarcely studied (*Hamida et al., 2004*). Conversely, in other bivalve species meiosis occurring in female gametes was explored and few major factors regulating the oocytes maturation processes were identified. Notably, it was demonstrated that serotonin (5-HT), thought to be the natural inducer of oocyte maturation in bivalves (*Deguchi and Osanai, 1995*), triggers *in vitro* the germinal vesicle breakdown (GVBD) when added to *Spisula, Barnea, Venerupis (philippinarum)* or *Crassostrea* isolated prophase I oocytes (*Osanai, 1985; Osanai and Kuraishi, 1988; Hirai et al., 1988; Brassart et al., 1988; Dubé, 1988; Krantic et al., 1991; Guerrier et al., 1993*). Moreover, it has been suggested that in *Venerupis philippinarum*, the mechanisms by which 5-HT promotes GVBD involve an increase in intracellular [Ca2+], which is thought to be mediated in turn by inositol 1,4,5-trisphosphate receptors (IP3r) and specific 5-HT receptors (*Guerrier et al., 1996; Gobet et al., 1994; Guerrier et al., 1993*). Despite these studies pointed out few interesting factors driving meiosis progression in bivalves, little is known on the complex regulations occurring during gametes maturation in these species. Up to date, only few gene expression and proteomic studies have been carried out in bivalve oocytes (e.g. *Ni et al, 2012; Corporeau et al. 2012*) and a comprehensive picture of the molecular processes characterizing their maturation is still absent.

Remarkably, when talking about gene expression and mRNA in oocytes, an interesting point should be taken into account: at prophase I immature oocytes show a prominent nucleus (the germinal vesicle), which contains de-condensed chromatin (*Tosti, 2006*), thus oocytes at this stage are transcriptionally active until meiosis resumption, when transcription is generally thought to cease (*Heikinheimo and Gibbons, 1998*). However, translation of the stored mRNAs pool continues throughout the final stages of meiosis (*Wassarman, 1996*) to synthetize proteins that are crucial for supporting not only the oocyte maturation (meiotic maturation) but the phase prior to embryonic genome activation and the newly fertilized zygote as well (*Song and Wessel, 2005, Dheilly et al., 2012; Eichenlaub-Ritter and Peschke, 2002*).

Besides gamete maturation, several factors are still limiting the hatchery production of *V. decussata*, research is therefore still needed. The absence of established methods for larval rearing, a scarce knowledge on the best hatching practises to improve metamorphosis synchronization and settlement are only few additional issues of the

European clam seed production which is an activity far from being completely technically controlled. In this context, the improvement of knowledge on broodstock management and gamete maturation's processes seems to be a key step to improve seed production of the emerging bivalve species.

In the present study, combined Next Generation Sequencing (NGS) technologies were employed in order to improve the transcriptomic knowledge on *V. decussata*, thus providing a great number of transcripts for further "omic" studies in this species. A microarray platform (8x60,000) for gene expression analysis was then constructed and a dataset of 15 oocytes samples were analyzed and expression profiles characterizing spawned oocytes with different competence levels and stripped oocytes were evaluated in order to shed light on the molecular processes leading to female gametes maturation and competence.

2. Methods

2.1. V. decussata transcriptome: biological samples, assembly and annotation

Clam tissues, larvae, oocytes and hemocytes were collected in collaboration with the *Instituto Português do Mar e da Atmosfera* (IPMA) and the *Instituto de Investigaciones Marinas* (IIM). All samples were stored in RNAlater at -80°C until RNA purification. RNA was isolated with RNeasy Mini Kit (Qiagen), following the manufacturer instructions and a DNAse treatment (Qiagen) was carried out. Concentration and purity of RNAs were measured using a *NanoDrop ND1000* spectrophotometer (NanoDrop Technologies, Wilmington, USA). The RNA quality was assessed through the Bioanalyzer 2010 instrument (Agilent).

For the RNA sequencing experiments, 13 non normalized libraries (oocytes, larval stages and gonadal stages) were prepared by using Truseq RNA sample prep Kit (Illumina) following the manufacture's instruction and sequencing was carried out with Illumina Hi-Seq 2000 by running two multiplexed lanes 2x100bp paired ends (BGI Tech, Shenzhen, China). All the Illumina reads were analyzed with FastaQC software in order to assess the sequences quality.

Two additional DSN normalized 454 libraries were synthesized, one starting from a pool of oocytes/larval stages/adults tissues/hemocytes and a second one by pooling different hemocytes sampled at different conditions. The two libraries were sequenced with a Roche 454 GS FLX sequencer using the Titanium chemistry (Genomic unit, CCiT-UB, Barcelona, Spain) and a sequences quality report was obtained through the 454 Software Release 2.6.

In order to build a transcriptome scaffold, a mixed strategy was preferred. Different assembler softwares were employed according to the sequences' origin (454 or Illumina). The 454-sequences were assembled through MIRA3 (default parameters), the reads from each of the Illumina libraries were separately assembled by using *CLC Genomic Workbench 5* (default parameters). Finally, to reduce the redundancy, all the obtained contigs plus a set of additional sequences already published in public repositories were merged through CAP3 (default parameters). The contigs not merged in CAP3 were joined to all the meta-contigs, thus obtaining a final transcriptome scaffold.

A functional annotation of the assembled transcriptome was attained through blastx similarity searches conducted against UniProtKB/SwissProt database and several protein databases available on Ensembl Genome Browser (release 68): *Homo sapiens*, *Danio*

rerio, *Drosophila melanogaster*, *Caenorhabtidis elegans*, *Ciona intestinalis*, *Strongylocentrus purpuratus*, *Gasterosteus aculeatus*, *Daphnia pulex* and *Nematostella vectensis*. In addition, to improve the number of annotated contigs, blastx searches against two molluscs databases were attempted: oysterDB (*Crassostrea gigas)* and *Lottia gigantea V1.0.* Alignments with an E-value of at most 1 E-5 were considered significant and only the best hit per each contig was taken into account. Finally, Blast2GO software (*Conesa et al., 2005*) was used to assign Gene Ontology terms (*Ashburner et al., 2000*) to all annotated contigs. Default values in Blast2GO were used to perform the analysis and ontology level 2 was selected to construct the level pie chart. To avoid redundant results, when multiple contigs were annotated with the same UniProtKB/SwissProt accession, only the longest one was selected for the Gene Ontology analysis.

2.2. Microarray platform design

All databases used for the annotation step were consider to reduce the redundancy in annotated contigs. In total 44,333 contigs, found non-redundant in at least one reference database, have been considered for microarray design. Of these, 915 contigs showed ambiguous orientation considering the homologs gene in reference databases and for each of them, two probes with opposite orientations (sense and antisense) were designed. For the remaining 43,418 contigs with known orientation, one probe was designed.

To fill the microarray platform, non-annotated gonads contigs were employed. Basing on the sum of gonads stage reads mapped against the assembled reference transcriptome the 7,376 most highly expressed contigs were considered and for each of them, two probes with opposite orientations (sense and antisense) were designed. Probe design was carried out using the Agilent eArray interface (https://earray.chem.agilent.com/earray/), which applies proprietary prediction algorithms to design 60 mer oligoprobes. A total of 59,951 out of 60,000 probes were successfully obtained, representing 51,709 putative *V. decussata* contigs. The percentage of annotated transcripts represented in the microarray was 85.7%. Probe sequences and other details on the microarray platform can be found in the GEO database (http://www.ncbi.nlm.nih.gov/geo/) under accession number GPL17766.

2.3. Biological samples for the microarray analyses

Clams were sampled in Ria de Aveiro (Western coast of Portugal) and conditioned in common garden from May 2013 to June 2013 (one month) in the experimental bivalve

hatchery of the National Institute of Biological Resources (IPMA) in Tavira, Portugal, to accelerate their gonad development under common rearing facilities. Food regimes consisted of different algal mixtures containing 1/3 *Isochrysis galbana* (clone T-ISO), 1/3 *Skelectonema costatum* (Ria Formosa autochthones clone) and 1/3 *Chaetoceros calcitrans*.

Released oocytes from several females were obtained by a thermal stimulation consisting on the exposition to alternate cycles of 29°C (1 hour) and 5°C (30 minutes) (*Joaquim et al., 2008*). As each female begun to spawn it was removed from the spawning tank and transferred to an individual spawning beaker with filtered seawater at the same temperature (*Joaquim et al., 2008*). Once the spawning was completed, the obtained oocytes were gently washed into a clean glass. The rest of each female spawning was mixed with a sperm suspension (from 7 males) during gentle agitation, aiming to obtain around 10 spermatozoids by oocyte in a microscopic view (*Cesari and Pellizzato, 1990; Joaquim et al., 2008*). Moreover, to evaluate the quality of the collected oocytes, the D-larval rate (ratio between the number of free swimming larvae at 48h post fertilization and the number of starting eggs) of each eggs batch was registered.

In addition, gonads from five females were dissected and oocytes were collected through a practice known as “gametes stripping”. As the name indicates, this procedure involves removal of gamete from gonad tissue. Briefly, fully ripe gonads were slash repeatedly with scalpel and washed with filtered seawater to harvest the gametes. Finally, sex determination and oocytes appearance were achieved through a microscopy examination.

About 20,000 oocytes for each spawning/stripping were collected and filtered in a 40 µm sieve. The oocytes were transferred into an eppendorf tube and, after a short spin, the seawater was removed. To remove the salts, the pellet of oocytes was re-suspended with a solution of ammonium formate 3% which was immediately removed after a short spin. Then the oocytes were included in 1,5 ml of Extract all solution (Eurobio) and preserved in liquid nitrogen until the RNA isolation.

RNA was purified by following the manufacturer instructions and a DNAse treatment was carried out through RTS DNAse Kit (MO-BIO). Samples concentration was measured in a NanoDrop® ND-1000 spectrophotometer and the RNA quality was assessed through the Bioanalyzer 2010 instrument (Agilent).

2.4. Labeling and microarray hybridization

Sample labeling and hybridization were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol with the Low Input Quick Amp Labeling kit. Briefly, for each sample 100 ng of total RNA were linearly amplified and labeled with Cy3-dCTP. A mixture of 10 different viral poly-adenilated RNAs (Agilent Spike-In Mix) was added to each RNA sample before amplification and labeling, to monitor microarray analysis work-flow. Labeled cRNA was purified with the RNAeasy Mini Kit (Qiagen) and samples concentration and specific activity (pmol Cy3/μg cRNA) were measured in a NanoDrop® ND-1000 spectrophotometer. A total of 600 ng of labeled cRNA was prepared for fragmentation adding 5μl 10X Blocking Agent and 1μl of 25X Fragmentation Buffer, heated at 60°C for 30 min, and finally diluted by addition with 25μl 2X GE Hybridization buffer. A volume of 40μl of hybridization solution was then dispensed in the gasket slide and assembled to the microarray slide (each slide containing eight arrays). Slides were incubated for 17 h at 65°C in an Agilent hybridization oven, subsequently removed from the hybridization chamber, quickly submerged in GE Wash Buffer 1 to disassembly the slides and then washed in GE Wash Buffer 1 for approximately 1 minute followed by one additional wash in pre-warmed (37°C) GE Wash Buffer 2.

2.5. Data acquisition, correction and normalization

Hybridized slides were scanned at 2μm resolution using an Agilent G2565BA DNA microarray scanner. Each slide was scanned two times at two different sensitivity levels: XDR Hi 100% and XDR Lo 10%. The two generated images were analyzed together, data were extracted and background subtracted using the standard procedures provided in the Agilent Feature Extraction (FE) Software version 10.7.3.1. To evaluate goodness and reliability of spot intensity estimates the software returns a series of spot quality measures. All control features (positive, negative, etc.), except for Spike-in (Spike-in Viral RNAs), were excluded from subsequent analyses.

The fluorescence values were normalized by performing a quantile normalization in R statistical software. To improve the reliability of the analysis, only the probes for which the reported fluorescence intensity was greater than 10 in at least 5 samples were taken into consideration. A log base 2 transformation was applied to all the expression values and finally, the parametric Combat algorithm was implemented in R in order to adjust for the known between-experiments batch effects (*Johnson et al., 2007*). A total of 31,862 probes were used in all the subsequent analyses.

2.6. Data analysis

The analysis was carried out by dividing the oocytes samples in three experimental groups, each characterized by a different quality in terms of developmental competency. The three groups were the following: unfertile stripped oocytes, spawned oocytes with low hatching rate (LHR, 5%-21% of D-larval rate), spawned oocytes with medium hatching rate (MHR, 40%-47% of D-larval rate). Samples used in the analyses and the relative D-larval rates are reported in Table 1.

A principal component analysis (PCA), using the TMeV 4.5.1 (TIGR MULTIEXPERIMENT VIEWER) (*Saeed et al., 2003*) was applied, to assess the distribution of the studied groups. Hierarchical clustering was performed using TMeV on the whole dataset, to group experimental samples based on similarity of the overall expression profiling. Statistical tests implemented in the program Significance Analysis of Microarray (SAM) were used to identify differentially expressed probes between the stripped oocytes and the two groups of spawned oocytes with different hatching rates.

A one-way ANOVA parametric test was used to identify the probes whose expression changed between the three tested groups, using a p-value cut-off of 0.05 and a minimum fold change (FC) of 1.5.

A more systematic, functional interpretation for significant gene was then obtained using an enrichment analysis from the Database for Annotation, Visualization, and Integrated Discovery (DAVID) software (*Huang et al., 2009*). “KEGG Pathway”, “Biological process”, “Molecular function” and “Cellular component” annotation were carried out by setting the gene count equal to 3 and the ease equal to 0.1. Because DAVID database contains functional annotation data for a limited number of species, it was necessary to link the *V. decussata* transcripts with sequence identifiers that could be recognized in DAVID. This process was accomplished using UniProtKB/SwissProt feature identifiers corresponding to each probe. These identifiers were used to define a “gene list” and a “background” in the bioinformatic tool DAVID, corresponding to differentially transcribed clam genes and to all the transcripts that were represented on the array, respectively.

Stripped oocytes					**Low hatching rate oocytes (LHR)**					**Medium hatching rate oocytes (MHR)**				
P1	P2	P3	P4	P5	X14 (12%)	X30 (21%)	X37 (14%)	X44 (11%)	X59 (5%)	X3 (44%)	X5 (41%)	X9 (40%)	X27 (47%)	X32 (46%)

Table 1. Microarray samples dataset. *The oocytes samples used for the microarray analysis were divided in three experimental groups: stripped oocytes, low hatching rate oocytes and medium*

hatching rate oocytes. The D-larval rates obtained after the fertilization of each spawning were reported in round brackets.

3. Results

3.1. Sequencing results

NGS technologies, RNA sequencing and Roche-454, were applied to sequence several cDNA libraries originating from oocytes, hemocytes, larval stages and adult tissues of *V. decussata*. Sequences across the entire length of the mRNA transcripts expressed in the target tissues were attained and two different assembly algorithms were used to obtain a whole scaffold transcriptome. In order to eliminate the redundant sequences, all the contigs produced by CLC bio (850,812) and MIRA3 (98,180) were submitted to the CAP3 software, obtaining a final scaffold of 503,705 sequences. Sequencing results and the de novo assembly's statistics are summarized in Table 2, while the size distribution of the transcripts after contigs construction is resumed in Figure 1.

A. ILLUMINA Sequencing

Sample	**Lane n°**	**Total reads**
Oocytes (stripped)	1	65 981 907
Oocytes (released)	1	32 722 892
D-larvae (48 hpf)	1	36 465 121
Larvae ready to settle (17 dpf)	1	51 967 552
Larvae settled (21 dpf)	1	47 752 218
1 month old seed (30 dpf)	1	45 800 062
Gonads Stage I	2	41 797 128
Gonads Stage II ♀	2	49 187 996
Gonads Stage II ♂	2	47 737 182
Gonads Stage III ♀	2	44 891 188
Gonads Stage III ♂	2	52 204 184
Gonads Stage IV ♀	2	54 690 184
Gonads Stage IV ♂	2	52 009 646

B. 454 - Sequencing

Sample	**Run**	**Total reads**
Hemocytes naive and in vivo/vitro stimulated	1/2	457 559
Larval stages/Oocytes/Hemocytes/Adult tissues	1/2	471 192

C. Assembly's statistics

CAP3 inputated contigs	990 111
CAP3 metacontigs	69740
CAP3 singletons	433 965
complete scaffold	503 705
total nt	408 246 936
mean lenght	810
max lenght	29 774
min lenght	40

Table 2. Statistics. *Number of reads obtained from the Illumina (A) and 454 (B) sequencing. Total number of contigs inputted to CAP3 software and statistics of the final scaffold (C).*

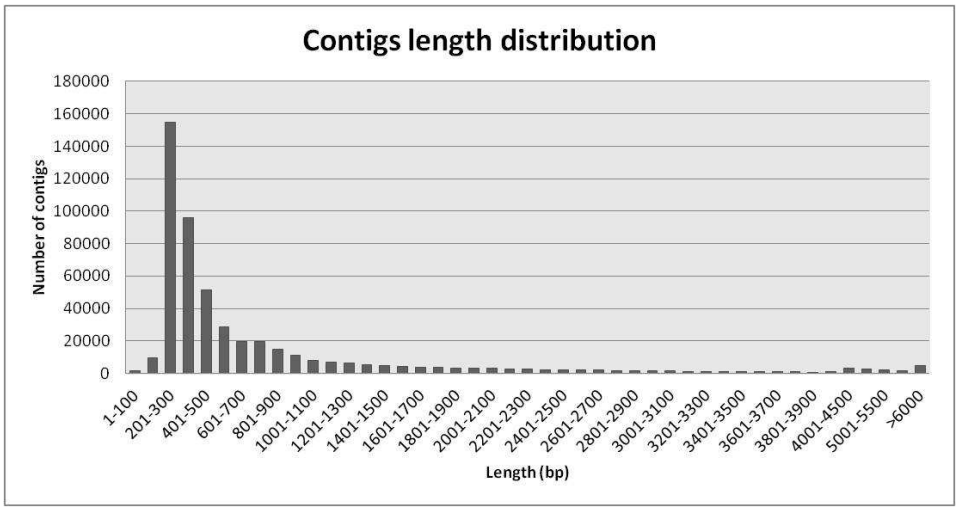


Figure 1. Contigs length distribution. *Histogram reporting the number of contigs (y axis) with a specific length range (x axis)*

3.2. Whole transcriptome analysis

Putative identities of assembled contigs and meta-contigs were obtained by running Blastx similarity searches on several protein databases (see methods). Giving the lack of a reference clam genome as well as the presence of specific sequences features such as high density repetitive regions and increased levels of polymorphism, this approach provided a significant match for 141,728 contigs, leading to a 27% of annotated transcripts.

In order to characterize the annotated transcripts in the *de novo* assembly, Gene ontology (GO) terms were extracted. Cellular Component (GO CC), Molecular Function (GO MF) and Biological Process (GO BP) terms at ontology level 2 are summarized in pie charts (Figure 2). The GO analysis showed that 39% of the GO CC terms were included in the "cell" class and 27% in the "organelle" class. Concerning the GO BP, "cellular process" (23%), "metabolic process" (18%) and "biological regulation" (12%), were the most represented classes. Also the terms "biological regulation", "response to stimulus" and "multicellular organismal process" were reported. Finally, within the GO MF, the two most represented groups were "binding" (49%) and "catalytic activity" (33%).

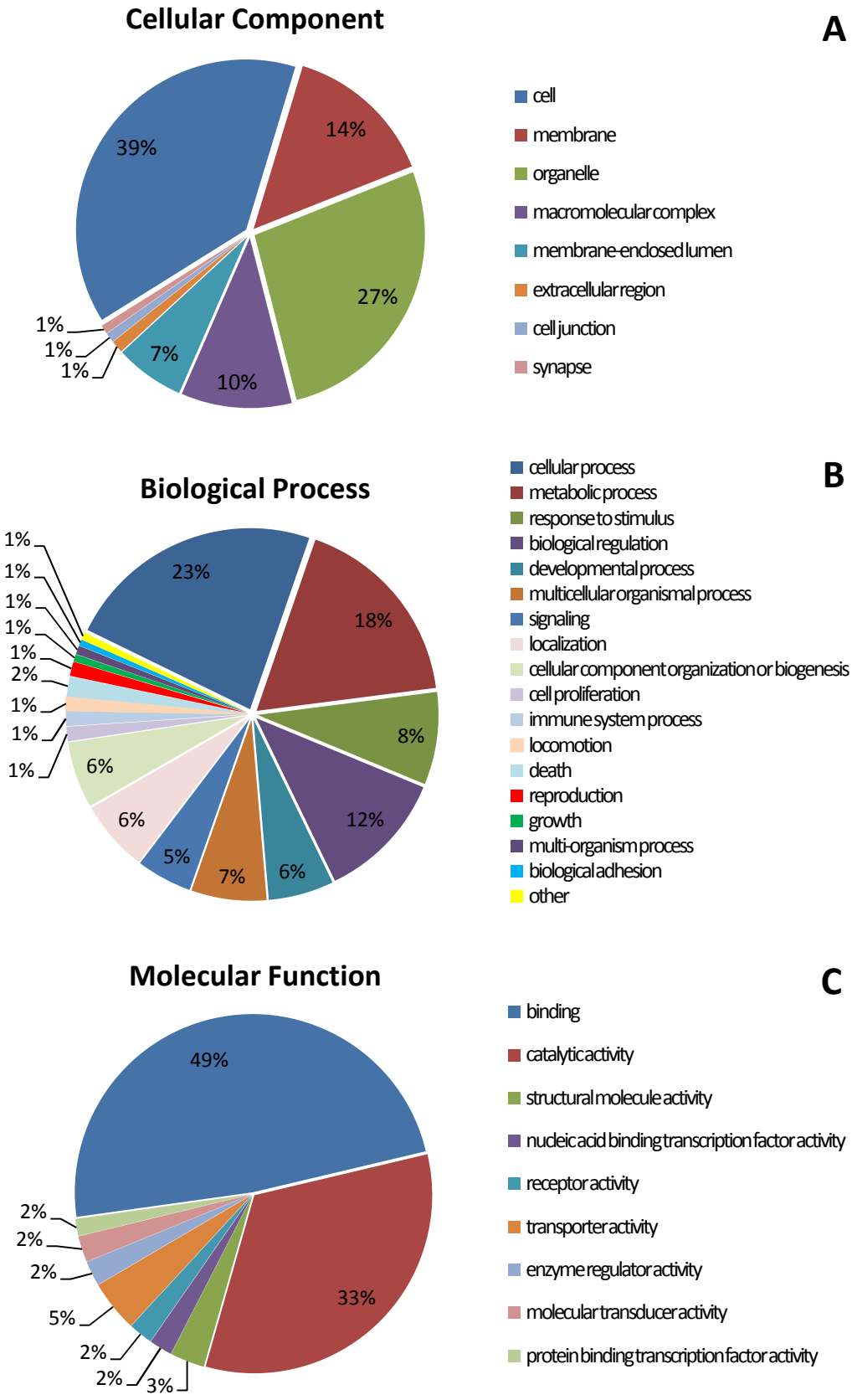


Figure 2. Summary of predicted gene product function and location using level 2 gene ontology terms. *Gene ontology sorted into the three main categories: Cellular Component (A), Molecular Function (B) and Biological Process (C). The GO subcategories are indicated in the legend and the percentage relative to the total number of extracted terms is indicated in the pie chart.*

3.3. Clustering analysis

A Principal Component Analysis (PCA) was applied to the selected gene expression dataset (31,862 probes, see Methods) of the 15 oocytes samples (Figure 3). A clear clustering of the three different groups of samples was observed. Stripped eggs, LHR eggs and MHR eggs were clearly separated along the x axis, which explained 29% of the variation. The expression profiles were also separated along the y axis (11% of the variation), but in this case stripped oocytes and MHR oocytes didn't showed a marked divergence in expression patterns, while the separation of LHF eggs was remarkable. Notably, the expression patterns of the 5 stripped oocytes appeared to be similar, while spawned eggs of the two groups seemed less homogeneous.

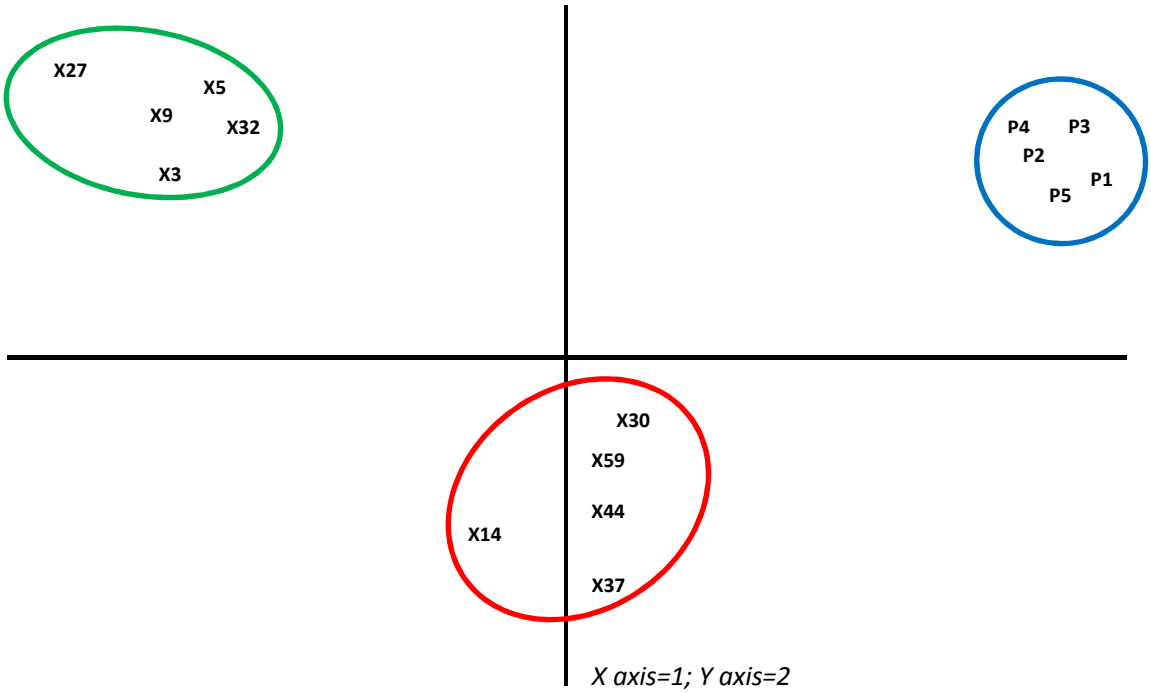


Figure 3. Principal Component Analysis. *Samples belonging to the same experimental group are identified by a coloured circle: green for MHR (medium hatching rate) oocytes, red for LHR (low hatching rate) oocytes and blue for stripped oocytes.*

Hierarchical clustering of the whole set of probes, using Pearson's correlation, is reported in Figure 4. The analysis identified two main clusters: stripped oocytes and spawned oocytes. Within the spawned oocytes cluster, samples were further divided in relation to their mean D-larval rate, leading to a definitive separation of the three experimental groups.

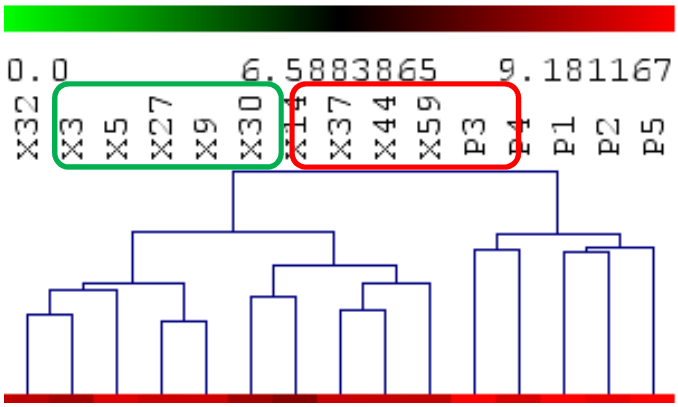


Figure 4. Hierarchical clustering analysis. *Stripped oocytes are identified by the "P" letter, while spawned oocytes are labelled with the "X" letter. Red and green rectangles define LHR (low hatching rate) oocytes and MHR (medium hatching rate) oocytes, respectively.*

3.4. Comparison between stripped and spawned oocytes

In order to compare gene expression profiles between *V. decussata* oocytes, a two-class unpaired Significance Analysis of Microarray (SAM) test was carried out on normalized/log transformed (base 2) data (FDR<1.5%; FC>1.5). The number of significant probes obtained from the two comparisons are summarized in Table 3

	ANALYSIS 1	**ANALYSIS 2**
	Stripped oocytes VS LHR oocytes	*Stripped oocytes VS MHR oocytes*
Up-regulated transcripts	105	4,166
Down-regulated transcripts	372	4,594
Total DEGs (unique transcripts)	477 (217)	8,760 (4,772)

Table 3. Differential expression analysis. *Number of significant probes (FDR<1.5%, FC>1.5) determined through a two-class unpaired SAM (Significance Analysis of Microarray). Unique transcripts refer to the annotation against UniProtKB/SwissProt database. LHR: low hatching rate oocytes; MHR: medium hatching rate oocytes; DEGs: differentially expressed genes.*

A comparison between the significant probes in the two analyses allowed us to identify a set of 439 transcripts which were differentially expressed between stripped and spawned oocytes, independently of the D-larval rate. A putative UniProtKB/SwissProt accession IDs was obtained for 235 probes corresponding to 198 unique genes (reported in Appendix A1) that were differentially expressed between the two oocytes conditions. Among the differentially expressed genes (DEGs), important transcripts encoding regulators of sex steroids synthesis and activity, such as a progestin and adipoQ receptor family member 3 (PAQR3) and steroid 17-alpha-hydroxylase/17,20 lyase (CYP450-C17), were expressed at a higher level in the stripped oocytes. Gene expression of enzymes

involved in the metabolism of ceramide, a signal sphingolipid thought to influence the oocytes maturation and quality, showed significant variations: sphingomyelin phosphodiesterase and neutral ceramidase B were expressed at a higher level in the stripped oocytes, while a putative sphingomyelinase transcript was more abundant in spawned oocytes.

Interestingly, also a putative vitellogenin (Vg), the major egg yolk protein precursor, was more abundant in released oocytes.

Moreover two transcripts encoding proteins suggested to be involved in regulating the maternal mRNAs transcription were differentially expressed. Putative translin-associated protein X and a oocyte zinc finger protein XlCOF28 were reported to be more and less abundant in ovarian oocytes, respectively.

Also the regulation of intracellular calcium levels seemed to highly differentiate the ovarian from the spawned oocytes: putative homologs of *C. gigas* calcium-activated chloride channel regulator 4 (two probes), regucalcin and calmodulin were more abundant in stripped oocytes, while sodium/calcium exchanger 3 had an higher expression level in spawned oocytes.

All putative annotated DEGs were used to define a gene list for functional annotation with DAVID. Enrichment analysis showed 7 CC terms, 16 BP terms, 9 MF terms and 1 KEGG to be significantly over-represented (Appendix A2). The only significant KEGG pathway was “ribosome” (dme03010), represented by 3 more abundant transcripts in the stripped oocytes: 60S ribosomal protein L3, 60S acidic ribosomal protein P1 and 40S ribosomal protein S9. “Mitotic cell cycle” (GO:0000278), “translation” (GO:0006412), “WNT receptor signalling pathway” (GO:0016055) and “dephosphorylation” (GO:0016311) were the most represented in the enriched BP terms. Among the DEGs belonging to the biological processes above, 5 M-phase inducer phosphatases (MPIP), a MPIP-like protein, frizzled 8, four-jointed protein and protein WNT-4 were found. With regard to the CC terms, one of the most represented ones was “ribonucleoprotein complex” (GO:0030529), which in oocytes is involved in the storing and compartmentalization of inactive mRNA incompletely polyadenylated, thus rendering them inactive until the resumption of meiosis (*de Moor and Richter, 1997; Mendez et al., 2000*). Significant MF terms concerned activities which are classically involved in the cell cycle regulation such as peptidase (GO:0008233) and phosphatase (GO:0016791) activities, and the action of molecules which contribute to the structural integrity (GO:0005198).

3.5. Comparison between LHR and MHR oocytes

In order to identify the probes whose expression changed in relation to different oocytes competence, a SAM quantitative correlation analysis was carried out on normalized/log-transformed data (FDR<5%). A total of 930 differentially expressed probes were detected, 453 of them were more abundant in oocytes with the highest D-larval rate and 477 in LHR oocytes. A putative annotation in the UniProtKB/SwissProt database was obtained for 587 probes that were differentially regulated in relation to the oocytes competence. The protein names corresponding to the 555 unique UniProtKB/SwissProt accession IDs, out of the 587 matches, were reported in Appendix A3. Genes encoding proteins putatively involved in the oocytes and embryo development were detected. The whole set of annotated transcripts were used to define a gene list for functional annotation with DAVID. Enrichment analysis showed 31 BP terms, 17 CC terms, 18 MF terms and 9 KEGG to be significantly over-represented (Appendix A4). Notably, with 19 genes the BP term "biological adhesion" (GO:0022610) was the most represented. A further BP term significantly enriched was "male gamete generation" (GO:0048232) represented by putative homologs positively correlated with the D-larval rate such as diaphanous and sperm-associated antigen 6 (Spag6). An additional enriched KEGG was "negative regulation of cell cycle" (GO:0045786) represented, among other, by a cyclin-dependent kinase 20 (two probes positively correlated with the oocytes quality).

As concern the CC terms, in addition to the "extracellular region" (GO:0005576), enriched also in the comparison between stripped and spawned oocytes, the cellular compartment showing a consistent enrichment was those sustaining the vesicular trafficking. Genes annotated with CC terms like "Golgi-associated vesicle membrane" (GO:0030660), "early endosome" (GO:0005769) and "vesicle coat" (GO:0030120) were listed among the genes whose expression changed in relation to the oocytes competence on larval development. Interestingly, the 136 genes belonging to the enriched MF term "ion binding" (GO:0043167) reflected the important role of charge regulation in oocytes development, most notably of calcium (GO:0005509). Furthermore, the importance of calcium was confirmed by the oocytes quality-correlated mRNA expression of sodium/calcium exchanger members 2 and 3, calcineurin (higher expression in MHR oocytes) and IP3r, resulting in the enriched KEGG pathway "calcium signalling pathway" (rno04020). IP3r and calcineurin, together with adenylate cyclase led also to the over-representation of the KEGG pathway "oocyte meiosis" (rno04114). Moreover, within the Panther database, the most represented enriched pathways (p-value<0.01) was "WNT

signalling pathway" (P00057), a molecular pathway that also characterized part of the differences between stripped and spawned oocytes. Notably the most interesting probes belonging to this pathway and positively correlated with the oocytes quality were those putatively encoding the WNT proteins receptors frizzled-4 and frizzled-10, and the calcium-dependent cell adhesion molecules cadherins type 4, 7 and 23.

3.6. ANOVA one way analysis

To investigate the set of probes whose expression level increased or decreased in relation to the increment of oocytes competence (stripped oocytes=null; LHR oocytes=low; MHR oocytes=medium), an ANOVA one way analysis was carried out on all samples from the three experimental groups. The analysis allowed to obtain a list of 14,118 probes with a significant change in expression between at least two groups (pval<0.5). A total of 568 probes for which the transcriptional variations were positively correlated to the oocytes competence were detected by selecting probes with a maximum expression value in MHR oocytes, intermediate in LHR oocytes and minimum in stripped oocytes. The same approach was adopted to find a list of 1,505 probes inversely correlated to the competence level (maximum expression value in stripped oocytes, intermediate in MHR oocytes and minimum in LHR oocytes). Finally, to provide a more comprehensive functional interpretation of the genes putatively involved in the biological mechanisms regulating the acquisition of oocytes competence, an enrichment analysis was carried out. Out of 2,073 probes, positively or negatively correlated to the oocytes competence, 1,342 annotated transcripts were used to define a gene list in DAVID. Results reported the over-representation of 3 KEGG, 48 BP terms, 16 CC terms and 31 MF terms (Appendix A5). Among KEGG pathways the higher fold enrichment was shown for "ribosome" (dme03010 and hsa03010), a term found significantly enriched also in the comparison between stripped and spawned oocytes, attesting the important involvement of the ribosomal activities during the oocytes development. Looking at the BP terms, "response to corticosteroid stimulus" (GO:0031960), spindle organization (GO:0007051 and GO:0000022), "steroid metabolic process" (GO:0008202) and sperm-egg recognition (GO:0035036 and GO:0007339) were significantly over-represented. Furthermore, enriched general mechanisms such as "translation" (GO:0006412), adhesion (GO:0007155 and GO:0022610) and oxidation reduction (GO:0055114) were identified.

4. Discussion

The gene expression analysis and the evaluation of DEGs between oocytes at different maturation stage and quality pointed out some interesting results and provided a first overview on the molecular mechanisms responsible for stripped oocytes infertility and oocytes competence. Among the 198 annotated DEGs between stripped and spawned oocytes, several transcripts were implicated in the regulation of the cell cycle. Notably *V. decussata* stripped ovarian oocytes have already replicated their DNA and are blocked at the germinal vesicle stage during late prophase of the first meiotic division. Conversely in spawned oocytes meiosis reinitiates and during metaphase I occurs a second block which is only released upon fertilization or artificial activation (*Lippai et al. 1995*). Several proteins are involved in the regulation of prophase I arrest occurring in ovarian oocytes, the most crucial factor being the maturation promoting factor (MPF) (*Jones, 2004*). MPF, a key G2/M phase regulator in eukaryotic cells consisting of cdc2 kinase (known also as cdk1) and cyclin B (*Nurse, 1990*), is induced during the meiosis resumption and its activity is regulated by the phosphorylation of cdc2 kinase. Cdc2 kinase is activated when M-phase inducer phosphatases (cdc25) dephosphorylate threonin-14 and tyrosine-15 sites (*Millar et al., 1991; Strausfeld et al., 1991; Trunnell et al., 2011*), leading, in turn, to MPF activation and meiosis resumption. In this study, six probes coding for cdc25 or cdc25-like proteins were reported to be more abundant in released oocytes, suggesting a prominent role of these phosphatases in the resumption of the meiotic cell cycle progression in *V. decussata*. The controlling function of cdc25 phosphatases in the meiosis I progression has been proposed in a wide range of species (e.g. *Kim et al., 2010* in *Caenorhabditis elegans*; *Alphey et al., 1992* in *Drosophila*; *Kishimoto, 2011* in startfish; *Oh et al. 2010* in mouse; *Gaffrè et al., 2011* in *Xenopus*). In mouse, both cdc25A and cdc25B were demonstrated to be critical for meiotic maturation and metaphase I spindle formation of oocytes (*Solc et al., 2008*). In mice cdc25B−/− knockout female are sterile because their oocytes cannot exit developmental arrest at meiosis prophase 1 (*Lincoln et al., 2002*), whereas cdc25A−/− mice exhibit early embryonic lethality (*Ray et al., 2007*), indicating that they are required for the control of oocyte meiotic cell cycle and embryonic mitotic cell cycle, respectively. Interestingly, among the six phosphatases differentially expressed between stripped and spawned oocytes, three followed a trend of expression correlated to the D-larval rate and their expression values were lower in stripped oocytes, intermediate in LHR oocytes and higher in MHR oocytes. Unfortunately in the European clam these proteins have not been studied yet and the

number of cdc25 isoforms is still unknown, but these findings allowed to hypothesize that in *V. decussata* more than one cdc25 family members exist and that they may function as distinct but related cell cycle regulators. Moreover, despite a role in meiosis I resumption, it seems that, at least for a few putative members, also the transcript abundance is fundamental for the maturation processes leading to a correct embryo development.

A crucial role in the molecular processes which differentiate stripped and spawned oocytes was predicted also for a few members of the WNT signalling pathway (GO:0016055), a powerful signalling pathways that play crucial roles in the animal life by controlling the genetic programs of embryonic development and adult homeostasis (*Grigoryan et al., 2008*). Recently, in mammalians WNT pathway signalling has been implicated in ovarian development, oogenesis, and early development. Multiple WNT signalling pathway genes are expressed in mouse oocytes and pre-implantation stage embryos, as revealed by microarray analyses (*Wang et al., 2004; Zeng et al., 2004*), and this has led to the hypothesis that WNTs may function in early cell fate determination events (*Wang et al., 2004*). Other studies, however, indicate that WNT signalling pathways are likely not functional in the early embryo (*Kemler et al., 2008*), raising the possibility that expression of WNT signalling genes in oocytes and early embryos are most likely related to functions in oogenesis (e.g., oocyte growth or maturation) (*Zheng et al., 2006*). In the present study the expression levels of three probes putatively encoding a WNT proteins receptor, frizzled-8, were more abundant in released oocytes. This data might suggest that frizzled-8 expression supports the WNT signalling pathway activation by favouring the recognition of WNT proteins. Moreover, our results reported that a putative ortholog of *C. gigas* WNT4 was less abundant in the female gametes extracted from mature gonads, in comparison with the released oocytes, and its mRNA expression tended to increase with the eggs quality. These findings suggest that not only WNT4 transcript could be implied in the *V. decussata* oocytes development but it could also be an important transcript whose abundance affects the oocyte quality, like it has been proposed in mouse. Notably, mice null for WNT4 exhibit sex reversal and a reduced number of oocytes in new-born ovaries (*Jeays-Ward et al., 2004*). In the same species, ovaries of WNT4-mutant females were characterized by a scarce amount of oocytes and these were in the process of degenerating (*Vainio et al., 1999*) and the 80% of WNT4 deficient germs cells failed to enter meiosis (*Naillat et al., 2010*). Finally a putative four jointed protein (Fj) encoding transcript was significantly (FDR<1.5%) more expressed in spawned oocytes than in stripped ones, and its expression, similarly to what reported for

WNT4, has the highest level in MHR oocytes. The role of Fj has been poorly investigated in both vertebrates and invertebrates. However, few studies focusing on this gene have been performed in *D. melanogaster*, where it has been demonstrated that this protein directly interacts with WNT4 and they act synergistically to induce planar polarity during early development (*Lim et al., 2005; Bosveld et al., 2012*). Considering the lack of functional information concerning Fj, it's difficult to establish a specific role of this gene in *V. decussata*, nevertheless it can be suggested that its expression in oocytes maturation is probably linked to the oocytes competence on fertilization and larval development.

Interestingly, the pivotal role suggested for the WNT signalling pathway was confirmed by the expression patterns evaluated in LHR and MHR oocytes, in which the mRNA abundance of a few putative frizzled receptors and cadherins was correlated to the developmental competence level. Notably, in vertebrates a cross-talk between WNT-mediated signals and cadherins have been reported to have a major role in morphogenesis and tissue formation (*Marie and Haÿ, 2013*). Taken together these results allowed us to hypothesize that also in *V. decussata* oocytes, the amount of mRNAs encoding proteins involved in the WNT signalling could be extremely important to permit the complete oocytes maturation and fertilization.

Another interesting finding was the significant change in expression of a putative Vg transcript. Vgs are the major precursor of the egg-yolk proteins, vitellins (Vn), which are stored in developing oocytes and are required for oocyte growth and maturation (*Kanungo et al., 1990; LaFleur et al., 2005*). In addition, yolk storage proteins are traditionally regarded as the energy reserve for nourishment of the developing embryos (*Zhang et al., 2011; Li et al., 1998a, 1998b*). Vgs have been shown to be present in almost all species of oviparous animals ranging from nematodes to vertebrates, and extensive sequence conservation is observed among these groups (*Chen et al., 1997*). To date, full length sequence characterization and gene expression levels of Vg have mainly focused on insects, crustaceans and fish (*LaFleur et al., 1995, Mouchel et al., 1996; Okuno et al., 2002; Tsutsui et al., 2000*). As regards molluscs, only a few studies were reported and Vg transcriptional levels were mainly evaluated in the gonadal tissue (*Osada et al., 2003, 2004; Matsumoto et al., 1997, 2003, 2008*; *Zheng et al. 2012*). In *Patinopecten yessoensis* and *C. gigas*, it has been demonstrated that the synthesis of Vg mRNA occurs in auxiliary cells inside the ovary (hetero-synthetic synthesis pathway), and is controlled by 17beta-estradiol (E2) and vitellogenesis promoting factor (VPF) via estrogen receptor (ER) (*Osada et al., 2003, 2004; Matsumoto et al., 1997, 2003*).

Conversely, studies performed in *Mytilus edulis* and *C. gigas* suggested that yolk proteins are synthesized in developing oocytes through an auto-synthetic pathway (*Pipe, 1987; Suzuki et al, 1992*). However, there has not been direct evidence on Vg synthesis, and it is still unclear whether the synthesis of a major yolk protein occurs through an auto- or hetero-synthetic pathway in bivalve molluscs. The fact that in the present study Vg mRNA was detected in the oocytes, let us thinking that in *V. decussata* Vg is synthesized through an auto-synthetic pathway. Despite that, we cannot exclude that in *V. decussata* a hetero-synthetic synthesis occurs, thus further investigations are needed. In addition, a higher level of Vg transcript was reported in the spawned oocytes. Despite the importance of Vg during oocytes growth/maturation and embryonic development has been demonstrated mainly at a protein level, this result might suggest that also Vg mRNA levels play a pivotal role, probably by providing a reserve of transcripts ready to be translated into functional protein. Thus it can be hypothesized that highest levels of stocked Vg mRNA in released oocytes provide a resource to support maturation and following embryo development. However, since the site of Vg synthesis in molluscs is not yet clarified and major regulations seem to occur at a protein level, gene expression results reported here cannot be sufficient to propose general conclusions regarding the involvement of Vg in *V. decussata* released oocytes.

Nonetheless, the presence of Vg mRNA stocks in bivalves was demonstrated also by *Osada* and colleagues (2004), who observed high level of Vg mRNA in *P. yessoensis* gonads through the spawning stage. A possible explanation they proposed, hypothesized also in *Xenopus* liver cell (*Brock and Shapiro, 1983*), is that Vg mRNA might be stabilized and remained un-translated, resulting in the retained high amount of Vg mRNA. The mechanism supposed to be involved in the retention of Vg mRNA requires the activity of RNA-binding proteins, which are known to function as translational repressors in the cytoplasm of several eukaryotic cells (*Wilkinson and Shyu, 2001; Wickens et al., 2002*). This mechanism is supposed to be crucial in oocytes since, during oogenesis, maternal mRNAs are synthesised and stored in a translationally dormant form and are activated either upon re-entry into the meiotic divisions or after fertilisation (*Song and Wessel, 2005, Dheilly et al., 2012*). In *Xenopus* oocyte maturation and early embryogenesis, translin, a RNA-binding protein, was demonstrated to play a major role to repress maternal mRNA translation (*Castro et al., 2000*). Notably in this study, a translin-associated protein X, supposed to be involved in the nuclear transport of translin in mice (*Aoky et al., 1997; Cho et al., 2004*), was expressed in all the oocytes condition and

particularly abundant in stripped oocytes. The relevance of mRNA translation in oocytes was demonstrated also by the significant variations reported for a putative homolog of the *Xenopus* oocyte zinc finger protein XlCOF28 whose mRNA amount was more abundant in spawned oocytes. The *Xenopus* zinc finger protein XlCOF28 belongs to the family of C_2H_2 zinc finger proteins, which are known to function as RNA-binding molecules (*Hall, 2005*). Noteworthy, it has been recently reported that proteins belonging to this family are required for regulation of maternally supplied mRNAs during oogenesis, oocyte to embryo transition, and early embryogenesis in both vertebrates and invertebrates (*Yamamoto et al., 2013*; *Kaymak and Ryder, 2013*). Consistent with these evidences, it can be hypothesized that C_2H_2 zinc finger proteins take part to the mechanisms which regulate mRNAs silencing in *V. decussata* oocytes.

A further interesting result concerns the differential expression of several transcripts involved in calcium (Ca2+) signalling. There has been a long-standing debate as to whether Ca2+ signals are required for oocyte meiosis and numerous conflicting studies argue that the relationship between Ca2+ and oocyte maturation is complex (*Sun and Machaca, 2004, Tosti et al., 2006*). The role that external calcium, through voltage-gated channels, might play in the induction of GVBD was first reported in molluscs that are both fertilized at the PI stage (*Allen, 1953; Dubè, 1988; Deguchi and Osanai, 1994*), or undergo the second arrest in MI (*Dubé and Guerrier Guerrier, 1982; Cuomo et al., 2005*). It was soon recognized that also the intracellular calcium increase plays a crucial role in almost all species studied independently from their peculiar meiotic arrest (*Deguchi and Osanai, 1994; Juneja et al., 1994; Guerrier et al., 1993*). In particular, the interplay between external and internal calcium currents is evident in *Venerupis*, where a serotonin-induced surge of intracellular calcium was shown to trigger maturation even in the absence of external calcium (*Guerrier et al., 1993*). In the present study, the differences in the amount of mRNA encoding putative calcium-activated chloride channel regulator 4 confirmed that the regulation of intracellular Ca2+ plays an important role in *V. decussata* oocytes maturation. Likewise in the present work, a putative sodium/calcium exchanger 3 was reported to be more abundant in the released oocytes and homologs of *C. gigas* regucalcin and calmodulin showed higher expression levels in ovarian oocytes. Calmodulin and regucalcin are calcium-binding proteins supposed to contribute to the meiosis regulation (*Wasserman and Smith, 1981; van der Voet et al, 2009*) and their mRNA variations suggest an involvement in maintaining the calcium homeostasis in immature oocytes. Interestingly, the regucalcin mRNA was expressed at lowest level in

the MHR oocytes, thus indicating that low levels of this transcript could be related to higher quality eggs. Despite little is known about the molecular regulation of intracellular Ca2+ occurring during oocytes maturation in bivalves, these preliminary results pointed out a few important genes possibly involved in such a complex mechanism.

In the present study, changes in mRNA levels were reported also for putative PAQR3 and CYP450-C17, which were both more abundant in ovarian oocytes. Progestin and adipoQ receptor family is a group of G protein-coupled receptors including membrane progesterone receptors (mPRs) that mediate a variety of rapid cell surface-initiated progesterone actions in the reproductive system involving activation of intracellular signalling pathways. Cytochrome P450-C17 is an enzyme involved in the synthesis of E2 during steroidogenesis. Despite in invertebrates there have been conflicting lines of evidence concerning the existence of enzymes necessary to synthesize vertebrate steroids and related nuclear receptors (*Scott, 2013*), several studies suggested a role of E2 and progesterone in gonadal development, oocytes maturation and spawning in several bivalve species (*Li et al., 1998a; Matsumoto et al., 1997; Osada et al., 2003, 2004; Reis-Henriques and Coimbra, 1990; Varaksina and Varaksin, 1991; Varaksina et al. 1992; Wang and Croll, 2004*). Consistent with these studies, we demonstrated that, during *V. decussata* oocytes maturation, significant variations of transcripts involved in the sex steroids synthesis and activity occurred. In particular the highest PAQR3 and CYP450-C17 mRNAs levels in ovarian oocytes may be associated to an higher E2 synthesis and progesterone activity in comparison with released oocytes. Accordingly, evidences indicate that sex steroids have a pivotal role in the pre-spawning stage since the have stimulatory effects on gamete release in *P. yessoensis* and *Placopecten magellanicus* (*Osada et al., 1992; Wang and Croll, 2003; Wang and Croll, 2007*).

Another metabolic process possibly implied in the oocytes maturation events in *V. decussata* was the regulation of ceramide levels. The enzymes controlling the metabolism of ceramide in oocytes have been poorly studied in molluscs and only recently sequences of genes associated with ceramide metabolism and signalling have been investigated in the Pacific oyster (*Timmins-Schiffman and Roberts, 2012*). Conversely, in vertebrates quite a few studies have been focused on the role of ceramide in oocytes and two main hypotheses have been suggested (*Coll et al., 2007*). First, it has been proposed that the generation of ceramide is a part of the signal transduction pathway activated in response to progesterone and that the increase in ceramide is likely to be functionally important in the resumption of the meiotic cycle (*Strum et al., 1995; Morrill and Kostellow, 1998;*

Buschiazzo et al. 2011). Second, it has been recently demonstrated that ceramide induces the default apoptosis process in oocytes and it has a central role in the age-related decrease of eggs quality (*Perez et al., 2005; Miao et al., 2005*). In the present study, at least three enzymes involved in the ceramide metabolism were found differentially expressed between stripped and spawned oocytes: a ceramide synthase, less abundant in ovarian oocytes, a ceramidase and a spyhgomyelinase, both less abundant in released oocytes. In addition, the expression of the sphingomyelinase and a sphingosine kinase were higher in stripped oocytes and lower in MHR oocytes, suggesting that a low amount of mRNAs encoding these enzyme could be somehow linked to higher eggs quality. Since no data are available on the normal ceramide homeostasis in bivalve oocytes, a comprehensive interpretation of the reported mRNA fluctuations is particularly challenging. Thus we can only legitimately conclude that, also in *V. decussata*, the ceramide metabolism most likely plays an important role in oocytes maturation and competence development.

A second aim of the present study was the identification of molecular mechanisms possibly involved in the development of oocytes competence once fertilization occurred. Since a high variability occurs in the spawned oocytes quality, we hypothesized that differences among the mRNA abundance of some transcripts could affect the oocyte maturation and the acquisition of developmental competence. Notably, mature oocytes of marine invertebrates contain a set of messenger RNAs which are not translated to a significant extent until after fertilization, when they provide most of the metabolic resources and information required for the early development of zygotes (*Watson, 2007*).
The analysis of the expression profiles of LHR and MHR oocytes allowed us to identify a total of 930 probes whose fluorescence intensity was positively or negatively correlated to the eggs quality, measured by the D-larval rate following oocytes fertilization. A functional annotation of those probes was carried out and it pointed out some interesting results.
Important variations occur in the abundance of transcripts encoding proteins regulating the oocyte meiosis, such as adenylate cyclases (ADCY), IP3r-1, IP3r-2 and the catalytic subunit alpha isoform of protein phosphatase 2B. ADCYs are membrane-associated enzymes that catalyse the formation of the secondary messenger cyclic adenosine monophosphate (cAMP) which is thought to have a pivotal role in the meiotic maturation in several species (*Richard, 2007*). In fact, accumulated evidence indicates that the release from prophase I arrest is accompanied by a decrease in oocyte intracellular levels

of cAMP, which is most probably a major prophase-I-arresting factor in amphibian and mammalian oocytes. Despite some exceptional cases seem to exist, such as the oocytes of the bivalve *Spisula solidissima*, in which the onset of meiosis reinitiation is associated with an increase in cAMP levels (*Yi et al., 2002*), it is generally accepted that an almost universal decrease in oocyte cAMP is sufficient to trigger the release of oocytes from prophase I arrest or at least positively affects it. In the present study the mRNA transcripts encoding ADCY type 5 and 6 were more abundant in oocytes with a lower developmental competence. Thus, despite cAMP levels are regulated also by other factors, such as progesterone and mPRs, more likely ADCYs have a key role in cAMP balancing and meiotic progression in *V. decussata*.

Moreover, the variations in the mRNA levels reported for IP3 receptors could provide interesting information concerning the bivalve oocytes maturation. IP3r are membrane glycoprotein complexes acting as Ca2+ channels and are activated by inositol trisphosphate (IP3), which has been strongly implicated in the conversion of external stimuli to intracellular Ca2+ signals (*Yoshida and Imai, 1997*). IP3r have been shown to have a predominant role also in oocytes, in both the formation and propagation of Ca2+ waves at fertilization (*Kume et al., 1993*) and it has been demonstrated that intracellular injections of IP-3 triggered GVBD in *Spisula* (*Bloom et al., 1988*) and *Venerupis* (*Guerrier et al., 1996*). In the present work, putative orthologs of the mammalian IP3r-1 and IP3r-2 were differentially expressed in relation to the oocytes competence. In particular the IP3r-1 was more abundant in the MHR oocytes, while IP3r-2 expression decreased as the oocytes quality increased. The opposite trend of expression reported for the two IP3r types allows us to think that in the European clam oocytes, at least two IP3r are expressed and that not only the presence of this receptor is crucial in determining a complete oocytes maturation and fertilization, but most probably also the isoform identity. While the isoform-specific cellular function of these channel proteins have not been investigated in molluscs, conversely in mammals IP3r isoforms, their subcellular distribution and peculiar functions have been studied (e.g. *Vanlingen et al., 2000; Vermassen et al., 2004*). Thus these preliminary results suggest the importance of IP3r messengers during the oocytes maturation phases occurring after the spawning in *V. decussata*.

A further interesting gene differentially expressed in released oocytes and implied in the meiotic maturation was the one coding for the catalytic subunit alpha isoform of a protein phosphatase 2B, known also as calcineurin. This enzyme is a calcium/calmodulin-

dependent phosphatase and plays a major role in the transduction of calcium signals in a variety of cell types from fungi to vertebrates (*Stewart et al., 1982 and Aramburu et al., 2000*). The significant role of calcineurin in meiotic maturation has been demonstrated in invertebrates (*Takeo et al., 2006; Takeo et al., 2010*) and in lower vertebrates (*Nishiyama et al., 2007; Mochida and Hunt, 2007*), while there is very limited information about calcineurin expression and distribution in mammalian oocytes. In pig oocytes, specific inhibitors of calcineurin affect the *in vitro* maturation of both growing pig oocytes with partial meiotic competence and fully grown pig oocytes with full meiotic competence (*Petr et al., 2013; Tumova et al., 2013*). These calcineurin inhibitors affected the exit of oocytes from the metaphase I of meiotic maturation. However, the role of calcineurin in the regulation of oogenesis and meiosis of gametes of mammalian females remain unclear. In *Drosophila* it has been demonstrated that oocytes lacking family members of positive regulators of calcineurin (RCANs) are arrested in anaphase of meiosis I (*Takeo et al., 2006; Takeo et al., 2010*). Consistent with these evidences, we found that calcineurin mRNA was more abundant in high quality oocytes, paving an important function of this enzyme also in *V. decussata* meiotic maturation.

Beside Ca2+ regulation, also cyclin-dependent protein kinases (cdks) have a crucial role during the oocytes maturation, since they regulate cell cycle transitions (*Morgan, 1995; 1997*). The importance of cdc2 kinase (cdk1) for the meiosis progression has been previously discussed. Interestingly, variations in the mRNA expression of a putative cyclin-dependent kinase 20 (cdk20 or p42) pointed out an additional mark differentiating oocytes with different developmental competence. In fact, the amount of cdk20 RNA messenger showed a positive correlation with the oocytes quality, suggesting an involvement of this kinase in the mechanisms promoting a correct gametes maturation and fertilization. The role of cdk20 has been recently investigated in mammals and it has been demonstrated that it is essential for the phosphorylation and consequent activation of cdk2 (*Liu et al., 2004*). Cdk2 is thought to be essential in the mammalian cell cycle, by driving cells through the G1/S transition (*Heichman and Roberts, 1994*), and in the *Drosophila* early embryogenesis as well (*Knoblich et al, 1994*). However, several lines of evidence indicated that cdk2 was essential in meiosis, but not in mitosis in mice (*Ortega et al., 2003*). Consistent with this hypothesis, cdk2 knockout mice grew normally, but both male and female mice were sterile due to meiotic defects (*Berthet et al., 2003*). Cdk2 activity is thought to be essential also in porcine oocytes, in particular for the first to second meiosis transition (*Sugiura et al., 2005*). Furthermore a recent study in the giant

prawn *Macrobrachium rosenbergii* suggested that cdk2 kinase might have an essential role during the meiotic maturation also in invertebrates (*Chen et al., 2013*). Despite the role of cdk20 is not fully resolved, not even in vertebrates, these studies allow us to propose that its abundance in MHR clam oocytes could be implicated in a major activation of cdk2, thus being one of the several factors which contribute to the their higher developmental competence.

The acquisition of oocytes competence in *V. decussata* seemed to be regulated also by other important factors. One of them was the diaphanous protein, a formin required for cytokinesis in both mitosis and meiosis. It has been demonstrated in *Drosophila* that it has a role in actin cytoskeleton organization and is essential for many, if not all, actin-mediated events involving membrane invagination (*Castrillon and Wasserman, 1994*). Moreover a role as a mediator between signalling molecules and actin organizers at specific phases of the cell cycle (e.g. contractile ring) has been proposed (*Afshar et al., 2000*). The importance of diaphanous during the meiotic maturation has been demonstrated by *Castrillon and Wasserman* (*1994*) that reported that mutated alleles of this gene affected spermatogenesis and oogenesis. Later, *Bione and colleagues* (1998) proposed that the homolog of *Drosophila* diaphanous in human is one of the genes responsible for premature ovarian failure and that it affects the cell divisions that lead to ovarian follicle formation. In molluscs this transcripts has never been reported but, since its mRNA abundance was significantly correlated to *V. decussata* oocytes quality, we can suppose that this protein may have an important role in cytokinesis and other actin-mediated morphogenetic processes that are required in early steps of development.

A further transcript showing the same expression pattern as diaphanous was Spag6, a protein, mostly expressed in testis, thought to be important for structural integrity of the central apparatus in the sperm tail and for flagellar motility in mouse (*Sapiro et al., 2000).* In the same species, this protein has been later implicated in mouse infertility (*Sapiro et al., 2000*) since mice lacking Spag6 produced sperm with marked motility defects, morphological abnormalities, disruption of flagellar structures, including loss of the central pair of microtubules and disorganization of the outer dense fibers. It's intriguing that a similar transcript has been detected in *V. decussata* oocytes with the highest quality. We can suppose that most probably this transcript encodes a protein with functions similar to Spag6 in male gametes and thus is implied in the gamete structural integrity.

Finally, in the present study, a pivotal role of the WNT signalling pathway in the oocyte maturation process was suggested. In fact, the abundance of few transcripts

corresponding to putative frizzled receptors and cadherins was correlated to the developmental competence. Notably, in vertebrates a cross-talk between WNT-mediated signals and cadherins have been reported to have a major role in morphogenesis and tissue formation (*Marie and Haÿ, 2013*). Thus, this results allowed us to hypothesize that also in *V. decussata* released oocytes, the amount of stocked mRNAs encoding proteins involved in the WNT signalling could be extremely important to permit the complete oocytes maturation and fertilization.

5. Conclusion

The gene expression analysis performed in this study allowed to identify a few important mechanism which could have an key role in the process of bivalve oocyte development. Differences in the mRNA expression of some important genes have been detected and the enrichment analysis helped in the interpretation of the DEGs lists, providing a more comprehensive picture of the key processes affecting *V. decussata* oocytes maturation and competence acquisition. Noteworthy, the transcripts which seemed to play a major role in the female gametes maturation and competence acquisition were those encoding proteins involved in the cell cycle progression, calcium regulation and WNT signalling.

Despite the analyses performed within this study provided first interesting findings, however more detailed investigations are still required and alternative molecular approaches should be implemented. For example, to deeply characterize specific events leading to the completion of meiosis and to elucidate the activity of target transcripts, reverse genetic experiments should be exploited.

Notwithstanding the fact that *V. decussata* reproduction is still far from being completely controlled in hatchery, as far as we know, this study represents the most effective attempt to improve the knowledge of oocytes maturation processes in this species.

6. References

Abdelmajid H, Leclerc-David C, Moreau M, Guerrier P, Ryazanov A. Release from metaphase I block in invertebrate oocytes: possible involvement of Ca2+ /calmodulin-dependant kinase III. Int J Dev Biol 1993;37:279-290.

Afshar K, Stuart B, Wasserman SA. Functional analysis of the Drosophila diaphanous FH protein in early embryonic development. Development 2000;127:1887-1897.

Allen RD. Fertilization and artificial activation in the egg of the surf clam, *Spisula solidissima*. Biol Bull 1953;105:213-239.

Alphey L, Jimenez J, White-Cooper H, Dawson I, Nurse P, Glover DM. twine, a cdc25 homolog that functions in the male and female germline of drosophila. Cell 1992;69:977-988.

Aoki K, Ishida R, Kasai M. Isolation and characterization of a cDNA encoding a Translin-like protein, TRAX. FEBS Lett 1997;401:109-112.

Aramburu J, Rao A, Klee CB. Calcineurin: from structure to function. Curr Top Cell Regul 2000;36:237-95.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM et al. Gene ontology: Tool for the unification of biology. Nat Genet 2000;25(1):25-9.

Berthet C, Aleem E, Coppola V, Tessarollo L, Kaldis P. Cdk2 knockout mice are viable. Curr Biol 2003;13:1775-1785.

Bione S, Sala C, Manzini C, Arrigo G, Zuffardi O, Banfi S et al. A human homologue of the Drosophila melanogaster diaphanous gene is disrupted in a patient with premature ovarian failure: evidence for conserved function in oogenesis and implications for human sterility. Am J Hum Genet 1998;62:533-41.

Bloom TL, Szuts EZ, Eckberg WR. Inositol trisphosphate, inositol phospholipid metabolism, and germinal vesicle breakdown in surf clam oocytes. Dev Biol 1998;129:532-540.

Bosveld F, Bonnet I, Guirao B, Tlili S, Wang Z, Petitalot A et al. Mechanical control of morphogenesis by Fat/Dachsous/Four-jointed planar cell polarity pathway. Science 2012;336:724-7.

Brassart M, Duclohier H, Moreau M, Guerrier P. Intracellular pH change does not appear as a prerequisite for triggering activation of *Barnea candida* (Mollusca Pelecypoda) oocytes. Gamete Res 1988;202:59-66.

Brock ML, Shapiro DJ. Estrogen regulates the absolute rate of transcription of the *Xenopus laevis* vitellogenin genes. J Biol Chem 1983;258:5449-55.

Buschiazzo J, Alonso TS, Biscoglio M, Antollini SS, Bonini IC. Nongenomic steroid- and ceramide-induced maturation in amphibian oocytes involves functional caveolae-like microdomains associated with a cytoskeletal environment. Biol Reprod 2011;85:808-22.

Castrillon DH and Wasserman SA. Diaphanous is required for cytokinesis in Drosophila and shares domains of similarity with the products of the limb deformity gene. Development 1994;120:3367-3377.

Castro A, Peter M, Magnaghi-Jaulin L, Vigneron S, Loyaux D, Lorca T et al. Part of *Xenopus* translin is localized in the centrosomes during mitosis. Biochem Biophys Res Commun 2000;276:515-23.

Cesari P, Pellizzato M. 1990. Biology of *Tapes Philippinarum*. In: *Tapes Philippinarum*: Biologia e Sperimentazione. Regione Veneto, Ente di Sviluppo Agricolo, Venice:299 pp. (In Italian and English).

Chen J, Liu P, Li Z, Chen Y, Qiu GF. The cloning of the cdk2 transcript and the localization of its expression during gametogenesis in the freshwater giant prawn, *Macrobrachium rosenbergii*. Mol Biol Rep 2013;40:4781-4790.

Chen JS, Sappington TW, Raikhel AS. Extensive sequence conservation among insect, nematode, and vertebrate vitellogenin reveals ancient common ancestry. J Mol Evol 1997;44:440-451.

Cho YS, Chennathukuzhi VM, Handel MA, Eppig J, Hecht NB. The relative levels of translin-associated factor X (TRAX) and testis brain RNA-binding protein determine their nucleocytoplasmic distribution in male germ cells. J Biol Chem 2004;279:31514-23.

Colas P and Dube F. Meiotic maturation in mollusc oocytes. Semin Cell Dev Bio 1998;9:539-548.

Coll O, Morales A, Fernandez-Checa JC, Garcia-Ruiz C. Neutral sphingomyelinase-induced ceramide triggers germinal vesicle breakdown and oxidant-dependent apoptosis in *Xenopus laevis* oocytes. J Lipid Res 2007;48:1924-1935.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005;21(18):3674-6.

Corporeau C, Vanderplancke G, Boulais M, Suquet M, Quéré C, Boudry P et al. Proteomic identification of quality factors for oocytes in the Pacific oyster *Crassostrea gigas*. J Proteomics 2012;75:5554-63.

Cuomo A, Di Cristo C, Di Cosmo A, Paolucci M, Tosti E. Calcium currents correlate with oocyte maturation during the reproductive cycle in *Octopus vulgaris*. J Exp Zool A 2005;303:193-202.

de Moor CH and Richter JD. The Mos pathway regulates cytoplasmic polyadenylation in *Xenopus* oocytes. Mol Cell Biol 1997;17: 6419-6426.

Deguchi R and Osanai K. Serotonin-induced meiosis reinitiation from the first prophase and from the first metaphase in oocytes of the marine bivalve *Hiatella flacida*: respective changes in intracellular Ca2+ and pH. Dev Biol 1995;171:483-496.

Deguchi R and Osanai K. Meiosis reinitiation from the first prophase is dependent on the levels of intracellular Ca2+ and pH in oocytes of the bivalves *Mactra chinensis* and *Limaria hakodatensis*. Dev Biol 1994;166:587-599.

Dheilly NM, Lelong C, Huvet A, Kellner K, Dubos M-P, et al. Gametogenesis in the Pacific Oyster *Crassostrea gigas*: A Microarrays-Based Analysis Identifies Sex and Stage Specific Genes. PLoS ONE 2012;7: e36353.

Dubé F and Guerrier P: Activation of *Barnea candida* (Mollusca, Pelecypoda) oocytes by sperm or KCl, but not by NH4 Cl, requires a calcium influx. Dev Biol 1982;92:408-417.

Dubé F. The relationships between early ionic events, the pattern of protein synthesis and oocyte activation in the surf clam *Spisula solidissima*. Dev Biol 1988;126:233-241.

Eichenlaub-Ritter U and Peschke M. Expression in in-vivo and in-vitro growing and maturing oocytes: focus on regulation of expression at the translational level. Hum Reprod Update 2002;8:21-41.

FAO, 2013. Cultured Aquatic Species Information Programme: *Ruditapes decussatus*. Available at: http://www.fao.org/fishery/culturedspecies/Ruditapes_decussatus/en (accessed 25.01.14).

Gaffré M, Martoriati A, Belhachemi N, Chambon JP, Houliston E, Jessus C et al. A critical balance between Cyclin B synthesis and Myt1 activity controls meiosis entry in Xenopus oocytes. Development 2011;138:3735-44.

Gobet I, Durocher Y, Leclerc C, Moreau M, Guerrier P. Reception and transduction of the serotonin signal responsible for meiosis reinitiation in oocytes of the japanese clam *Ruditapes philippinarum*. Dev Biol 1994;164:540-549.

Grigoryan T, Wend P, Klaus A, Birchmeier W. Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of beta-catenin in mice. Genes Dev 2008;22:2308-41.

Guerrier P, Durocher Y, Gobet I, Leclerc C, Moreau M. Reception and transduction of the serotonin signal responsible for oocyte meiosis reinitiation in bivalves. Inv Reprod Dev 1996;30:39-45.

Guerrier P, Leclerc-David C, Moreau M. Evidence for the involvement of internal calcium stores during serotonin-induced meiosis reinitiation in oocytes of the bivalve mollusc *Ruditapes philippinarum*. Dev Biol 1993;159: 474-484.

Hall TM. Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol 2005;15:367-73.

Hamida L, Medhioub MN, Cochard JC, Le Pennec M. Evaluation of the effects of serotonin (5-HT) on oocyte competence in *Ruditapes decussatus* (Bivalvia, Veneridae). Aquaculture 2004;239:423-420.

Heichman KA and Roberts JM. Rules to replicate by. Cell 1994;79:557-562.

Heikinheimo O and Gibbons WE. The molecular mechanisms of oocyte maturation and early embryonic development are unveiling new insights into reproductive medicine. Mol Hum Reprod 1998;4:745-56.

Hirai S, Kishimoto T, Kadam AL, Kanatani H, Koide SS. Induction of spawning and oocyte maturation by 5-hydroxytryptamine in the surf clam. J Exp Zool 1988;245:318-321.

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4(1):44-57.

Jeays-Ward K, Dandonneau M, Swain A. Wnt4 is required for proper male as well as female sexual development. DevBiol 2004;276:431-440.

Joaquim S, Matias D, Moreno O. Cultivo de bivalves em maternidade. 2008. Instituto de Investigación Y Formación Agraria Y Pesquera. Consejería de Agricultura Y Pescas. Junta de Andalucia. I.S.B.N.:978-84-8474-252-4:84 pp. (In spanish and portuguese).

Johnson WE, Rabinovic A, Li C. Adjusting batch effects in microarray expression data using Empirical Bayes methods. Biostatistics 2007;8:118-127.

Jones KT. Turning it on and off: M-phase promoting factor during meiotic maturation and fertilization. Mol Hum Reprod 2004;10:1–5.

Juneja R, Ito E, Koide SS. Effect of serotonin and tricyclic antidepressants on intracellular calcium concentrations in Spisula oocytes. Cell Calcium 1994:15:1-6.

Kanungo J, Petrino TR, Wallace RA. Oogenesis in *Fundulus heteroclitus*. VI. Establishment and verification of conditions for vitellogenin incorporation by oocytes in vitro. J Exp Zool 1990;254:313-321.

Kaymak E, Ryder SP. RNA recognition by the *Caenorhabditis elegans* oocyte maturation determinant OMA-1. J Biol Chem 2013;288:30463-72.

Kemler R, Hierholzer A, Kanzler B, Kuppig S, Hansen K, Taketo MM et al. Stabilization of beta-catenin in the mouse zygote leads to premature epithelial-mesenchymal transition in the epiblast. Development 2008;131:5817–5824.

Kim J, Kawasaki I, Shim YH. cdc-25.2, a *C. elegans* ortholog of cdc25, is required to promote oocyte maturation. J Cell Sci 2010;123:993-1000.

Kishimoto T. A primer on meiotic resumption in starfish oocytes: the proposed signaling pathway triggered by maturation-inducing hormone. Mol Reprod Dev 2011;78:704-7. doi: 10.1002/mrd.21343.

Kjørsvik E, Mangor-Jensen A, Holmefjord I. Egg quality in fishes. Adv Mar Biol 1990;26:71-113.

Knoblich JA, Sauer K, Jones L, Richardson H, Saint R, Lehner CF. Cyclin E controls S phase progression and its down-regulation during Drosophila embryogenesis is required for the arrest of cell proliferation. Cell 1994;77:107-20.

Krantic S, Dubé F, Quirion R, Guerrier P. Pharmacology of the serotonin-induced meiosis reinitiation in *Spisula solidissima* oocytes. Dev Biol 1981;146:491-498.

Kume S, Muto A, Aruga J, Nakagawa T, Michikawa T, Furuichi T et al. The Xenopus IP3 receptor: structure, function, and localization in oocytes and eggs. Cell 1993;73:555-70.

LaFleur GJ Jr, Byrne BM, Kanungo J, Nelson LD, Greenberg RM, Wallace RA. *Fundulus heteroclitus* vitellogenin: The deduced primary structure of a piscine precursor to noncrystalline, liquid-phase yolk protein. J Mol Evol 1995;41:505-21.

LaFleur GJ Jr, Raldúa D, Fabra M, Carnevali O, Denslow N, Wallace RA et al. Derivation of major yolk proteins from parental vitellogenins and alternative processing during oocyte maturation in *Fundulus heteroclitus*. Biol Reprod 2005;73:815-24.

Li Q, Osada M, Suzuki T, Mori K. Changes in vitellin during oogenesis and effect of estradiol-17C on vitellogenesis in the Pacific oyster *Crassostrea gigas*. Invertebr Reprod Dev 1998a;33:87-93.

Li Q, Osada M, Suzuki T, Sato M, Mori K. Degradation of vitellin during embryonic and larval development in the Pacific oyster *Crassotrea gigas*. Invertebr Reprod Dev 1998b;33:1-9.

Lim J, Norga KK, Chen Z, Choi KW. Control of planar cell polarity by interaction of DWnt4 and four-jointed. Genesis 2005;42:150-161.

Lincoln AJ, Wickramasinghe D, Stein P, Schultz RM, Palko ME, De Miguel MP et al. Cdc25b phosphatase is required for resumption of meiosis during oocyte maturation. Nat Genet 2002;30:446-9.

Lippai M, Gobet I, Tomkowiak M, Durocher Y, Leclerc C, Moreau M et al. Thimerosal triggers meiosis reinitiation in oocytes of the Japanese clam *Ruditapes philippinarum* by eliciting an intracellular calcium surge. Int J Dev Biol 1995;39:401-407.

Liu Y, Wu C, Galaktionov K. p42, a novel cyclin-dependent kinase-activating kinase in mammalian cells. J Biol Chem 2004;279:4507-14.

Longo FJ, 1983. Meiotic maturation and fertilization. In: Verdonk, N.H., van den Biggelaar, J.A.M., Tompa, A.S. (Eds.), The Mollusca, 3. Academic Press, New York, pp.48-89.

Marie PJ, Haÿ E. Cadherins and Wnt signalling: a functional link controlling bone formation. Bonekey Rep 2013;2:330.

Matsumoto T, Nakamura AM, Mori K, Kayano T. Molecular characterization of a cDNA encoding putative vitellogenin from the Pacific oyster *Crassostrea gigas*. Zoolog Sci 2003;20:37-42.

Matsumoto T, Osada M, Osawa Y, Mori K. Gonadal estrogen profile and immunohistochemical localization of steroidogenic enzymes in the oyster and scallop during sexual maturation. Comp Biochem Physiol B 1997;118:811-817.

Matsumoto T, Yamano K, Kitamura M, Hara A. Ovarian follicle cells are the site of vitellogenin synthesis in the Pacific abalone Haliotis discus hannai. Comp Biochem Physiol A 2008;149:293-8.

Mendez R, Hake LE, Andresson T, Littlepage LE, Ruderman JV, Richter JD. Phosphorylation of CPE binding factor by Eg2 regulates translation of c-mos mRNA. Nature 2000;404:302-307.

Miao YL, Liu XY, Qiao TW, Miao DQ, Luo MJ, Tan JH. Cumulus cells accelerate aging of mouse oocytes. Biol Reprod 2005;73:1025-1031.

Millar JBA, McGowan CH, Lenaers G, Jones R, Russell P. p80 cdc25 mitotic inducer is the tyrosine phosphatase that activates p34 cdc2 kinase in fission yeast. EMBO J 1991;10:4301-4309.

Mochida A and Hunt T. Calcineurin is required to release Xenopus egg extracts from meiotic M phase. Nature 2007;449:336-340.

Moreau M, Leclerc C, Guerrier P. Meiosis reinitiation in *Ruditapes philippinarum* (Mollusca): Involvement of L-calcium channels in the release of metaphase-I block. Zygote 1996; 4:151-157.

Morgan DO. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol 1997;13:261-91.

Morgan DO. Principles of CDK regulation. Nature 1995;374:131-4.

Morrill GA and Kostellow AB. Progesterone release of lipid second messengers at the amphibian oocyte plasma membrane: role of ceramide in initiating the G2/M transition. Biochem Biophys Res Commun 1998;246:359-363.

Mouchel N, Trichet V, Betz A, Le Pennec JP, Wolff J. Characterization of vitellogenin from rainbow trout (*Oncorhynchus mykiss*). Gene 1996;174:59-64.

Naillat F, Prunskaite-Hyyryläinen R, Pietilä I, Sormunen R, Jokela T, Shan J et al. Wnt4/5a signalling coordinates cell adhesion and entry into meiosis during presumptive ovarian follicle development. Hum Mol Genet 2010;19:1539-50.

Nishiyama T, Yoshizaki N, Kishimoto T, Ohsumi K. Transient activation of calcineurin is essential to initiate embryonic development in *Xenopus laevis*. Nature 2007;449:341-345.

Nurse P. Universal control mechanism regulating onset of M-phase. Nature 1990;344:503-508.

Oh JS, Han SJ, Conti M. Wee1B, Myt1, and Cdc25 function in distinct compartments of the mouse oocyte to control meiotic resumption. J Cell Biol 2010;188:199-207.

Okuno A, Yang WJ, Jayasankar V, Saido-Sakanaka H, Huong DTT, Jasmani S et al. Deduced primary structure of vitellogenin in the giant freshwater prawn, *Macrobrachium rosenbergii*, and yolk processing during ovarian maturation. J Exp Zool 2002;292:417-429.

Ortega S, Prieto I, Odajima J, Martin A, Dubus P, Sotillo R et al. Cyclin-dependent kinase 2 is essential for meiosis but not for mitotic cell division in mice. Nat Genet 2003;35:25-31.

Osada M, Mori K, Nomura T. In vitro effects of estrogen and serotonin on release of eggs from the ovary of the scallop. Nippon Suisan Gakkaishi 1992;58:223-227.

Osada M, Takamura T, Sato H, Mori K. Vitellogenin synthesis in the ovary of scallop, *Patinopecten yessoensis*: control by estradiol-17 β and the central nervous system. J Exp Zool 2003;299:172-179.

Osada M, Tawarayama H, Mori K. Estrogen synthesis in relation to gonadal development of Japanese scallop, *Patinopecten yessoensis*: gonadal profile and immunolocalization of P450 aromatase and estrogen. Comp Biochem Physiol B 2004;139:123-128.

Osanai K and Kuraishi R. Response of oocytes to meiosis-inducing agents in Pelecypods. Bulletin of the Marine Biological Station of Asamushi, 182:45-56.

Osanai K. In vitro induction of germinal vesicle breakdown in oyster oocytes. Bulletin of the Marine Biological Station of Asamushi 1985;18:1-9.

Perez GI, Jurisicova A, Matikainen T, Moriyama T, Kim MR, Takai Y et al. A central role for ceramide in the age-related acceleration of apoptosis in the female germline. FASEB J 2005;19:860-862.

Petr J, Chmelíková E, Zalmanová T, Tůmová L, Kheilová K, Kučerová-Chrpová V et al. Pyrethroids cypermethrin, deltamethrin and fenvalerate have different effects on in vitro maturation of pig oocytes at different stages of growth. Animal 2013;7:134-42.

Pipe RK. Oogenesis in the marine mussel *Mytilus edulis*: An ultrastructural study. Mar Biol 1987;95:405-414.

Ray D, Terao Y, Nimbalkar D, Hirai H, Osmundson EC, Zou X et al. Hemizygous disruption of Cdc25A inhibits cellular transformation and mammary tumorigenesis in mice. Cancer Res 2007;67:6605-6611.

Reis-Henriques MA and Coimbra J. Variations in the levels of progesterone in *Mytilus edulis* uring the annual reproductive cycle. Comp Biochem Physiol A 1990;95:343-348.

Richard FJ. Regulation of meiotic maturation. J Anim Sci 2007;85(13 Suppl):E4-6.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 2003;34:374-378.

Sapiro R, Kostetskii I, Olds-Clarke P, Gerton GL, Radice GL, Strauss III JF. Male infertility, impaired sperm motility, and hydrocephalus in mice deficient in sperm-associated antigen 6. Mol Cell Biol 2002;22:6298-305.

Sapiro R, Tarantino LM, Velazquez F, Kiriakidou M, Hecht NB, Bucan M et al. Sperm antigen 6 is the murine homologue of the *Chlamydomonas reinhardtii* central apparatus protein encoded by the PF16 locus. Biol Reprod 2000;62:511-518.

Scott AP. Do mollusks use vertebrate sex steroids as reproductive hormones? II. Critical review of the evidence that steroids have biological effects. Steroids 2013;78:268-81.

Solc P, Saskova A, Baran V, Kubelka M, Schultz RM, Motlik J. CDC25A phosphatase controls meiosis I progression in mouse oocytes. Dev Biol 2008;317:260-9.

Song JL and Wessel GM. How to make an egg: transcriptional regulation in oocytes. Differentiation 2005;73:1-17.

Stewart AA, Ingebritsen TS, Manalan A, Klee CB, Cohen P. Discovery of a Ca2+- and calmodulin-dependent protein phosphatase: probable identity with calcineurin (CaM-BP80). FEBS Lett 1982;137:80-84.

Strausfield U, Labbe JC, Fesquet D, Caradore JC, Picard A, Sadhu, K et al. Dephosphorylation and activation of a p34 cdc2/cyclin B complex in vitro by human CDC25 protein. Nature 1991;351:242-245.

Strum JC, Swenson KI, Turner JE, Bell RM. Ceramide triggers meiotic cell cycle progression in Xenopus oocytes. A potential mediator of progesterone-induced maturation. J Biol Chem 1995;270:13541-13547.

Sugiura K, Naito K, Tojo H. Cdk2 activity is essential for the first to second meiosis transition in porcine oocytes. J Reprod Dev 2005;51:143-9.

Sun L and Machaca K. Ca2+cyt negatively regulates the initiation of oocyte maturation. J Cell Biol 2004:165:63-75.

Suzuki T, Hara A, Yamaguchi K, Mori K. Purification and immunolocalization of a vitellin-like protein from the Pacific oyster *Crassostrea gigas*. Mar Biol 1992;113:239-245.

Takeo S, Hawley RS, Aigaki T. Calcineurin and its regulation by Sra/RCAN is required for completion of meiosis in Drosophila. Dev Biol; 2010;344:957-967.

Takeo S, Tsuda M, Akahori S, Matsuo T, Aigaki T. The calcineurin regulator Sra plays an essential role in female meiosis in Drosophila. Curr Biol 2006;16:1435-1440.

Timmins-Schiffman E, Roberts S. Characterization of genes involved in ceramide metabolism in the Pacific oyster (*Crassostrea gigas*). BMC Res Notes 2012;13;5:502.

Tosti E. Calcium ion currents mediating oocyte maturation events. Reprod Biol Endocrinol 2006;9:4-26.

Trunnell NB, Poon AC, Kim SY, Ferrell JE Jr. Ultrasensitivity in the Regulation of Cdc25C by Cdk1. Mol Cell 2011;41:263–274.

Tsutsui N, Kawazoe I, Ohira T, Jasmani S, Yang WJ, Wilder MN et al. Molecular characterization of a cDNA encoding vitellogenin and its expression in the hepatopancreas and ovary during vitellogenesis in the kuruma prawn, *Penaeus japonicas*. Zool Sci 200;17:651-660.

Tůmová L, Petr J, Žalmanová T, Chmelíková E, Kott T, Tichovská H et al. Calcineurin expression and localisation during porcine oocyte growth and meiotic maturation. Anim Reprod Sci 2013;141:154-163.

Vainio S, Heikkilä M, Kispert A, Chin N, McMahon AP. Female development in mammals is regulated by Wnt-4 signalling. Nature 1999;397:405-9.

van der Voet M, Berends CW, Perreault A, Nguyen-Ngoc T, Gönczy P, Vidal M et al. NuMA-related LIN-5, ASPM-1, calmodulin and dynein promote meiotic spindle rotation independently of cortical LIN-5/GPR/Galpha. Nat Cell Biol 2009;11:269-77.

Vanlingen S, Sipma H, de Smet P, Callewaert G, Missiaen L, de Smedt H et al. Ca2+ and calmodulin differentially modulate myo-inositol 1,4,5-trisphosphate (IP3)-binding to the recombinant ligand-binding domains of the various IP3 receptor isoforms. Biochem. J 2000;346:275-280.

Varaksina GS and Varaksin AA. Effects of estradiol, progesterone, and testosterone on oogenesis of yezo scallop. Biol Mora 1991;3:61-68.

Varaksina GS, Varaksin AA, Maslennikova LA. The role of gonadal steroid hormones in the spermatogenesis of the scallop *Mizuhopecten yessoensis*. Biol Morya Mar Biol 1992;77–83.

Vermassen E, Parys JB, Mauger JP. Subcellular distribution of the inositol 1,4,5-trisphosphate receptors: functional relevance and molecular determinants. Biol Cell 2004;96:3-17.

Wang C and Croll RP. Effects of sex steroids on in vitro gamete release in the sea scallop. Invertebr Reprod Dev 2003;44:89-100.

Wang C and Croll RP. Estrogen binding sites in the sea scallop: characterization and possible involvement in reproductive regulation. Comp Biochem Physiol B 2007;148:303-13.

Wang QT, Piotrowska K, Ciemerych MA, Milenkovic L, Scott MP, Davis RW et al. A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. Dev Cell 2004;6:133-144.

Wassarman PM, Liu C, Litscher ES. Constructing the mammalian egg zona pellucida: some new pieces of an old puzzle. J Cell Sci 1996;109:2001-4.

Wasserman WJ and Smith LD. Calmodulin triggers the resumption of meiosis in amphibian oocytes. J Cell Biol 1981;89:389-394.

Watson AJ. Oocyte cytoplasmic maturation: A key mediator of oocyte and embryo developmental competence. J Anim Sci 2007;85:E1-E3.

Whitaker M, Patel R:Calcium and cell cycle control. Development 1990;108:525-542.

Wickens M, Bernstein DS, Kimble J, Parker R. A PUF family portrait: 3′UTR regulation as a way of life. Trends Genet 2002;18:150-157.

Wilkinson MF and Shyu AB. Multifunctional regulatory proteins that control gene expression in both the nucleus and the cytoplasm. BioEssays 2001;23:775-787.

Yamamoto TM, Cook JM, Kotter CV, Khat T, Silva KD, Ferreyros M et al. Zar1 represses translation in Xenopus oocytes and binds to the TCS in maternal mRNAs with different characteristics than Zar2 Biochim Biophys Acta 2013;1829:1034-1046.

Yi JH, Lefièvre L, Gagnon C, Anctil M, Dubé F. Increase of cAMP upon release from prophase arrest in surf clam oocytes. J Cell Sci 2002;115:311-20.

Yoshida Y and Imai S. Structure and function of inositol 1,4,5-trisphosphate receptor. Jpn. J Pharmacol 1997;74:125-37.

Zeng F, Baldwin DA, Schultz RM. Transcript profiling during preimplantation mouse development. Dev Biol 2004;272:483-496.

Zhang S, Wang S, Li H, Li L. Vitellogenin, a multivalent sensor and an antimicrobial effector. Int J Biochem Cell Biol 2011;43:303-5.

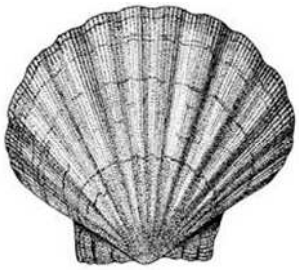
Zheng H, Zhang Q, Liu H, Liu W, Sun Z, Li S at al. Cloning and expression of vitellogenin (Vg) gene and its correlations with total carotenoids content and total antioxidant capacity in noble scallop *Chlamys nobilis* (Bivalve: Pectinidae). Aquaculture 2012;366-367:46-53.

Zheng P, Vassena R, Latham K. Expression and downregulation of WNT signaling pathway genes in rhesus monkey oocytes and embryos. Mol Reprod Dev 2006;73:667-677.

Ni J, Zeng Z, Han G, Huang H, Ke C. Cloning and characterization of the follistatin gene from *Crassostrea angulata a*nd its expression during the reproductive cycle. Comp Biochem Physiol B 2012;163:246-53

Section B

“Deep transcriptome sequencing of *Pecten maximus* hemocytes: a genomic resource for bivalve immunology”



1. Background

The king scallop, *Pecten maximus* (Linnaeus, 1758), also known as great scallop, is a native European species of great commercial importance. Global production is based on both fisheries and aquaculture with 63,702 and 74 tons in the year 2011 respectively (*FAO Fisheries Department, 2013*). These data do not reflect the increasing relevance of scallop aquaculture production. Hatchery-produced seed is more and more used for enhancement programmes of wild scallop populations, notably in France and Ireland. Unfortunately, hatchery production of great scallop is still hampered by rearing conditions that are not optimal yet. One of the major bottlenecks in scallop hatcheries is the high susceptibility of this species to bacterial infections, leading to highly variable survival rates. Especially during the early larval stages, mortalities of up to 100% are often observed. Since the early 90s, to investigate the causes of mass mortality in *P. maximus*, a few studies were conducted. Nicolas et colleagues in 1996 (*Nicolas et al. 1996*) identified four main bacterial strains involved in outbreaks, pointing out that the major causative agents were Vibrios. Few years later, Lambert et *al*. (*Lambert and Nicolas, 1998*) published a study demonstrating the presence of *Vibrio pectenicida* in dying larvae in French hatcheries. More recently, Sandlund et co-workers (*Sandlund et al., 1996*), by means of immunohistochemistry, confirmed that *Vibrio splendidus* and *V. pectenicida* were pathogenic to scallop larvae, and that the *Pseudoalteromonas* strain could act as a secondary opportunistic bacterium (*Sandaa et al., 2008*). Similarly, in farmed northern Chilean scallop *Argopecten purpuratus*, *Vibrio anguillarum* (*Riquelme et al. 1995*) and *Vibrio alginolyticus* (*Riquelme et al. 1996a,1996b)* were reported to cause high mortality. More in general, opportunistic bacteria, mainly *Vibrio spp*, are considered the main cause of mortality in bivalves (*Nicolas et al. 1996*; *Torkildsen et al. 2000, 2002, 2005; Lacoste et al. 2001; Gómez-León et al., 2005; Garnier et al. 2007*) and it is also well documented that, besides bacteria, viruses (e.g. Herpes-like virus, Retroviridae, Picornaviridae), protozoan (e.g. *Marteilia, Bonamia, Perkinsus, Mycrocytos*) and metazoan parasites can cause diseases and severe mortalities in bivalves at different life stages (*Gestal et al., 2008*).

As other invertebrates, they lack a specific immune response and immunological memory, therefore they rely entirely on the innate immune system to fight against pathogens and eventually overcome disease. The immune response of scallop is mediated by both cellular and humoral immune components. The former comprises of various immunocytes, mainly hemocytes, which are circulating cells with phagocytic and

cytotoxic activities, while the latter contains a variety of immune effector such as antimicrobial peptides (AMPs), lysozyme, antioxidant enzymes and heat shock proteins (HSPs). The initial step of immune response is the discrimination of non-self from self, generally known as immune recognition. Bivalves can recognize potential dangerous substances mainly via pattern recognition receptors (PRRs). These PRRs can sense the pathogen-associated molecular patterns (PAMPs), and its binding can trigger intracellular signaling cascades to activate the transcription of immune effectors (*Akira et al., 2006; Lemaitre and Hoffmann, 2007*). In the past decades, various families of conserved PRRs have been identified in molluscs and the important roles in immune recognition have been verified for Toll-like receptors (TLRs), lectins, lipopolysaccharide (LPS) and b-1,3-glucan binding protein (LGBP) and scavenger receptors (SRs). Among them, Toll-like receptors are a family of great interest since they are found throughout the animal kingdom from basal metazoans to vertebrates (*Leulier and Lemaitre, 2008*). TLRs recognize microbial structures via the extracellular leucine-rich-repeat (LRR) domain and signal transduction takes place by the intracellular toll-interleukin-domain (TIR) domain and TIR domain containing adaptor molecules. Investigations of bivalve TLRs identified a single TLR in *Mya arenaria* (*Mateo et al., 2010*) and *Chlamys farreri* (*Qiu et al., 2007*) but recent studies, performed in specie for which deep sequencing techniques have been already exploited, demonstrated that a large TLR repertoire could be present in bivalves. Notably, 4 and 23 TLR members were identified in *C. gigas* (*Zhang et al., 2013*) and *Mytilus galloprovincialis* (*Toubiana et al., 2013*), respectively.

Moreover, to improve the general knowledge on bivalve immunology, several experiments evaluating the immune response against pathogen infections have been reported in different species, mostly in oysters *C. gigas* (e.g. *Genard et al., 2013; Renault et al., 2011*), *C. virginica* (*Dorrington et al., 2011*), and *Ostrea edulis* (*Martin-Gomez et al., 2012*), but also in mussels, clams, and scallops (e.g. *Costa et al., 2009*; *Ji et al., 2013; Araya et al., 2010; Ramses Ramirez-Castillo et al., 2011*). Despite these studies have provoked advances in the knowledge of molecular mechanisms involved in the response against pathogen infections, genomic information regarding bivalve immunity remains still very scarce and fragmentary. In the last years, the employment of genome-scaled sequencing technologies opened the doors to genomics in bivalves species, thus leading to the discovery of an increasing number sequences related to the immune function. To date, high-throughput transcriptomic data on bivalves immunity have been obtained in oysters (*Rosa et al., 2012; Fleury et al., 2009; Wang et al., 2011*), mussels (*Moreira et*

al., 2013; Venier et al., 2009; Philipp et al., 2012), and clams (*Milan et al., 2011; Moreira et al., 2012*), although not all studies were specifically targeted to discovery of immune-relevant genes. As concern scallops, recent experiments conducted in *Patinopecten yessoensis* and *C. farreri* (*Hou et al., 2011; Wang et al., 2013; Chen et al., 2013; Wang et al., 2011*) attest an increasing interest in immunity of Pectinidae. However, very scarce information is available on *P. maximus* immune-related genes and, to date, the limited knowledge of the king scallop immune defense derives mostly from immuno-histochemical and biochemical assays (*Sandlund et al., 1996; Hauton et al., 2001; Hannam et al., 2010*).

Recently, the research activity funded by the European project "ReProSEED" (Research to improve Production of SEED), has been working in order to stabilize and optimize the hatchery production of emerging bivalve species, through a comprehensive strategy linking pure and applied approaches. Critical aspects have been considered, including reproduction, development, rearing conditions, nutrition and susceptibility to pathogens.

In this context, the goal of the present study was to characterize the transcriptome of *P. maximus* hemocytes by means of Illumina RNA sequencing technology. Untreated and stimulated hemocytes with heat-inactivated *V. anguillarum* and several PAMPS were collected and three different cDNA libraries were prepared and sequenced in order to allow the identification of immune-related transcripts and biological processes involved in the great scallop innate response. To our knowledge, this is the first transcriptome analysis of immune-related genes in the great scallop *P. maximus*.

2. Methods

2.1. Challenge and sampling

P. maximus scallops were obtained in the Ría de Vigo (Galicia, Spain). The animals were maintained in open circuit filtered sea water tanks, 20L, at 15°C and fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana*. Prior to the experiments, scallops were acclimatized to aquaria conditions for one month.

For *in vivo* stimulation 3 adult scallops were injected intramuscularly with 100 µl of 10^6 CFU/ml of heat inactivated *Vibrio anguillarum.* Other 3 adult scallops were used as controls and injected with 100 µl of filtered sea water (FSW). Hemolymph (6 mL per scallop) was collected with a 0.8 mm diameter (21G) disposable needle 24 hours after bacterial injection. Hemolymph was then centrifuged at 3000 rpm at 3°C for 5 min, the obtained pellet was re-suspended in 0.5 ml of *RNAlater*® (*Ambion),* maintained one night at 4°C and then stored at -80°C until RNA purification.

For *in vitro* challenge, a sample of 3 mL of hemolymph was withdrawn from the adductor muscle of 10 scallops with a siliconized (sigmacote, SL2, Sigma-Aldrich) syringe and 0.8 mm diameter (21G) disposable needle and pooled in siliconized tubes. Siliconized surfaces significantly reduced the formation of big aggregates, facilitating the subsequent settlement of hemocytes. Hemolymph was 2-fold diluted in FSW and then distributed in 6-well plates, 5 ml per well, in a total of 8 wells, 2 biological replicates for each treatment. Hemocytes were allowed to settle for 60 min at 15°C. Then, hemolymph and not settled hemocytes were discarded and wells were filled with 5 mL of FSW. Hemocytes were incubated overnight at 14°C to allow small aggregates to form a monolayer. The day after, hemocytes were rinsed with 5 mL FSW and stimulated for 3 h at 14°C with Polyinosinic:polycytidylic acid (Poly I:C), LPS, Lipoteichoic acid (LTA) or Zymosan (Zym), one stimulus per well, at a final concentration of 50 µg/mL. All PAMPs were purchased from Sigma. After sampling, cells were pelleted by centrifugation at 3000 rpm at 3°C for 5 min, and stored in 0.5 ml of *RNAlater*® as described before.

2.2. RNA isolation and sequencing

RNA was isolated with *RNeasy Mini Kit* (Qiagen), following the manufacturer instructions and a DNAse treatment (Qiagen) was carried out. Concentration and purity of RNA were measured using a *NanoDrop ND1000* spectrophotometer (NanoDrop Technologies). The RNA quality was assessed through the *Bioanalyzer 2100* instrument

(Agilent Technologies). Finally, RNAs were pooled to obtain three samples each representing a specific condition: *in vivo* stimulation (inactivated *V. anguillarum*), *in vitro* stimulation (mix of PAPMs), and *in vivo* control.

Three non-normalized libraries for RNA sequencing experiments were prepared by using *Truseq RNA sample prep Kit* (Illumina) following the manufacture's instruction and a paired end sequencing 2x100bp was carried out on a Illumina Hi-Seq 2000 (BGI Tech, Shenzhen, China). Raw Illumina sequencing data have been deposited in GenBank (SRA) with the accession numbers SRR1009240, SRR1009241 and SRR1009242. All Illumina reads were analyzed with FastaQC software in order to assess sequence quality.

2.3. Transcriptome assembly and annotation

All reads obtained from each of the Illumina libraries were assembled together using *CLC Genomic Workbench 5*, with default parameters. All assembled contigs were then further clustered using CAP3 with default parameters.

Contigs from the final assembly were annotated through blastx similarity searches conducted against the *UniProtKB/SwissProt* database and selected protein databases available in the Ensembl Genome Browser, respectively for *Homo sapiens*, *Danio rerio,* and *Drosophila melanogaster*. In addition, to further increase the number of annotated contigs, blastx searches were carried out against sequence databases for two mollusc species, *C. gigas* (http://oysterdb.cn/) and *Lottia gigantea V1.0* (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html). Alignments with an e-value of at least 1 e-5 were considered significant and the best hit for each contig was used for annotation. In the case of TLRs, all known bivalve TLR protein sequences were used as queries against the scallop transcriptome in a series of tBlastn searches. Contigs showing a e-value lower than 1 E-10 were manually checked and, when possible, prolonged at 5' and 3' ends by overlapping *P. maximus* sequences which were previously obtained, through a *Roche 454 GS FLX+ System*, starting from a pool of adult tissues, larvae and hemocytes (unpublished data). Open reading frames (ORFs) were inferred and used for further analysis. Putative LRRs, LRR-carboxy terminal (LRR-CT), LRR-amino terminal (LRR-NT), TIRs and trans-membrane domains were obtained using SMART (http://smart.embl-heidelberg.de/) and HMMer software (*Eddy, 2011*), while signal peptide regions were predicted by SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/).

The tools described above have also been employed to acquire structural information concerning differentially expressed transcripts for which the annotation process did not achieve significant hits.

In addition, the Blast2GO software (*Conesa et al., 2005*) was used to assign Gene Ontology terms (*Ashburner et al., 2000*) to annotated contigs. Default values in Blast2GO were used to perform the analysis and ontology level 2 was selected to construct the level pie charts. To avoid redundant results, when multiple contigs were annotated with the same *UniProtKB/SwissProt* accession, only one was used for the Gene Ontology analysis (GO analysis). Finally, a list of keywords and GO immune-related terms were used to find contigs that are putatively involved in immune system function.

2.4. Phylogenetic analyses

Evolutionary relationships were calculated by using a Maximum-Likelihood approach (PhyML 3.1) and trees robustness was assessed with 1000 bootstrap interactions. The phylogenetic analyses were based on multiple sequence alignments obtained using MAFFT 7 alignment program (http://mafft.cbrc.jp/alignment/serve4r/) and further refined to include only reliably aligned positions by using Gblocks (*Castresana, 2000*). A model test (http://darwin.uvigo.es/software/prottest2_server.html) was carried out to select the model of protein evolution that best fits the given set of sequences.

2.5. Mapping and differential expression analysis

Sequence reads obtained from each hemocyte cDNA library were mapped against the whole transcriptome assembly by performing a RNA-seq analysis in *CLC Genomic Workbench 5* (default parameters). Differential expression analysis was carried out by comparing the expression levels in *in vitro* and *in vivo* stimulated hemocytes against controls, using the *NOISeq* package in *R* (*Tarazona et al., 2011*). In order to evaluate the differential expression, the software provides a *q value* corresponding to the probability for contigs of being differential expressed. As suggested by *NOISeq* authors, since no biological replicates were available, *NOIseq-sim* algorithm has been preferred (parameters: *k*= 0.5, *norm*= tmm, *pnr* = 0.2, *nss* = 5, *v* = 0.02, *lc* = 1) and the *q value* threshold has been set to 0.9.

2.6. Enrichment analysis

A functional interpretation of the set of differential expressed genes following the *in vivo* stimulation was obtained by performing an enrichment analysis with the Database for Annotation, Visualization, and Integrated Discovery (DAVID) software (*Huang et al., 2009*). Enriched "Biological process", "Molecular function" and "Cellular component" terms were carried out by setting the gene count equal to 2 and the ease equal to 0.1. Among the accepted databases used for the annotation, the selected one was *UniProtKB/SwissProt* because it provided the higher percentage of annotated transcripts. *UniProtKB/SwissProt* accessions corresponding to differential expressed genes as well as all the transcripts that were represented on the whole hemocytes transcriptome were then used to define a "gene list" and a "background" in the bioinformatics tool DAVID.

3. Results

3.1. De novo transcriptome assembly and functional annotation

Over 200 millions Illumina sequence reads originating from three scallop hemocyte cDNA libraries were assembled into 73,732 contigs. Sequencing results and assembly statistics are summarized in Table 1. Putative annotation of assembled contigs was obtained by sequence similarity. Table 1 reports the number of contigs annotated against individual databases. The highest percentage of matches is obtained against the two molluscan genomes of *C. gigas* and *L. gigantea*. Merging all data, a putative annotation was obtained for 22,858 contigs (31%). Gene Ontology (GO) terms referring to cellular localization (Cell Component, GO_CC), molecular function (GO_MF), and biological process (GO_BP) were associated to annotated contigs using Blast2GO. Figure 1 shows pie charts summarizing the most represented GO_CC, GO_MF and GO_BP.

Illumina library	**Total reads**
Hemocytes in vivo	*67,643,592*
Hemocytes in vivo stimulated (heat-inactivated V. anguillarum)	*70,855,070*
Hemocytes in vitro stimulated (PAMPs)	*77,946,012*
Hemocytes transcriptome assembly	
Total number of contigs	*73,732*
Mean length (bp)	*502.6*
Median length (bp)	*328*
Max length (bp)	*16,205*
Min length (bp)	*200*
Annotation (database)	***n° annotated contigs (%)***
Swiss-prot database	*13,883 (18.8%)*
Ensemble H. sapiens	*15,239 (20.7%)*
Ensemble D. melanogaster	*12,402 (18.2%)*
Ensemble D. rerio	*16,118 (21.9%)*
OysterDB	*20,355 (27.6%)*
L. gigantea v1.0	*18,076 (24.5%)*

Table 1. Statistics. *Summary of the Illumina sequencing, assembly and annotation data.*

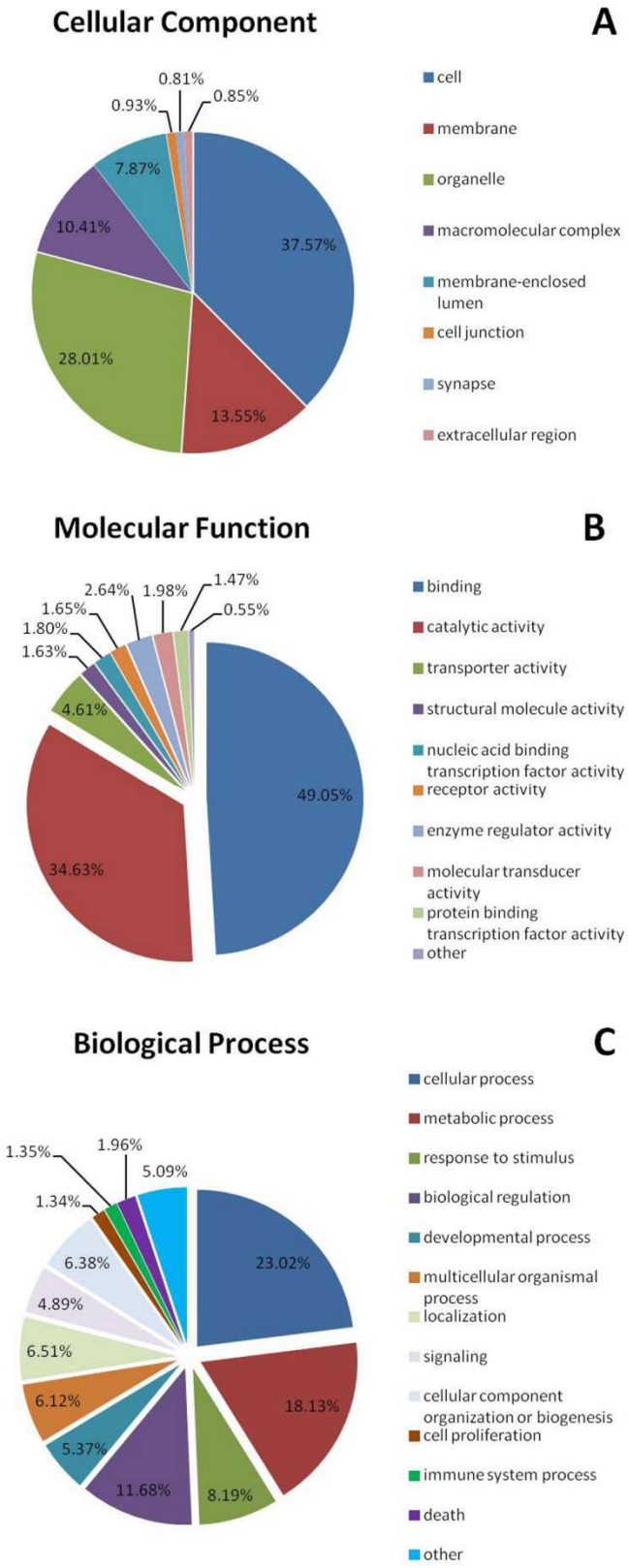


Figure 1. Gene ontology pie charts. *Percentages of occurrences of Cellular component (A), Molecular process (B) and Biological process (C) terms. Besides each pie chart the colours correspondence to GO terms is reported.*

3.2. Mining the scallop transcriptome for immune-relevant genes

A dedicated analysis discovered 934 contigs that encode proteins with a putative role in immune response. These proteins could be grouped into several functional categories, as follows: i) PRRs;
ii) immune effectors; iii) signal transduction; iv) proteins related with apoptosis; v) complement and C1q-like proteins; vi) cytokine-related molecules; vii) cell surface receptors and cell adhesion molecules; viii) other immune-related transcripts (Appendix B1).

PRRs are involved in sensing PAMPs, which are generic molecular signatures for potential pathogens, microbial organisms, and endogenous signals. A total of 109 contigs coding for putative PRRs have been identified, including lectins, TLRs, SRs, glucan-binding proteins and peptidoglycan recognition proteins. Further *in silico* analysis (see section 2.3) revealed at least four complete scallop TLRs. Two multiple cysteine cluster (mcc) TLRs, named here *Pm*TLR1 and *Pm*TLR2, and two single cysteine cluster (scc) TLRs, which were denominated *Pm*TLR3 and *Pm*TLR4 (Figure 2A). The intra-cellular domain of *Pm*TLR2 was characterized by the unusual presence of two TIR domains. The C-terminal one showed a canonical TIR domain signature, which was readily recognized with SMART and the second one, which is located closer to the trans-membrane domain, displayed a lesser degree of conservation compared to canonical TIR domains, yet it could be still identified based on sequence similarity (Figure 2A).

In addition to the four putative full-length scallop TLRs, other eight sequences showed consistent similarities to bivalve TLRs. However, not all predicted domains could be detected, either because they were missing or due to incomplete sequence information. Two partial copies (*Pm*TLR5 and *Pm*TLR6) that contained a TIR domain are shown in Figure 2A. The remaining six partial sequences contained only putative extra-cellular domains with high similarity (e-value lower than 1 E-10) against the corresponding region of bivalve TLRs. These likely represent additional candidates, but require further analysis for reliable classification as scallop TLRs. Scallop TLRs were also classified by considering the architecture of ectodomains as reported in Toubiana et *al.*, 2013, where *M. galloprovincialis* sscTLRs have been divided in three subgroups, based on ectodomain complexity (cluster 1, 2 and 3), and all mcc TLRs have been attributed to a single group (cluster 4) (Figure 2B). It is important to specify that there is no accepted nomenclature for TLRs in molluscs, therefore the number attributed in this study to each TLR is arbitrary.

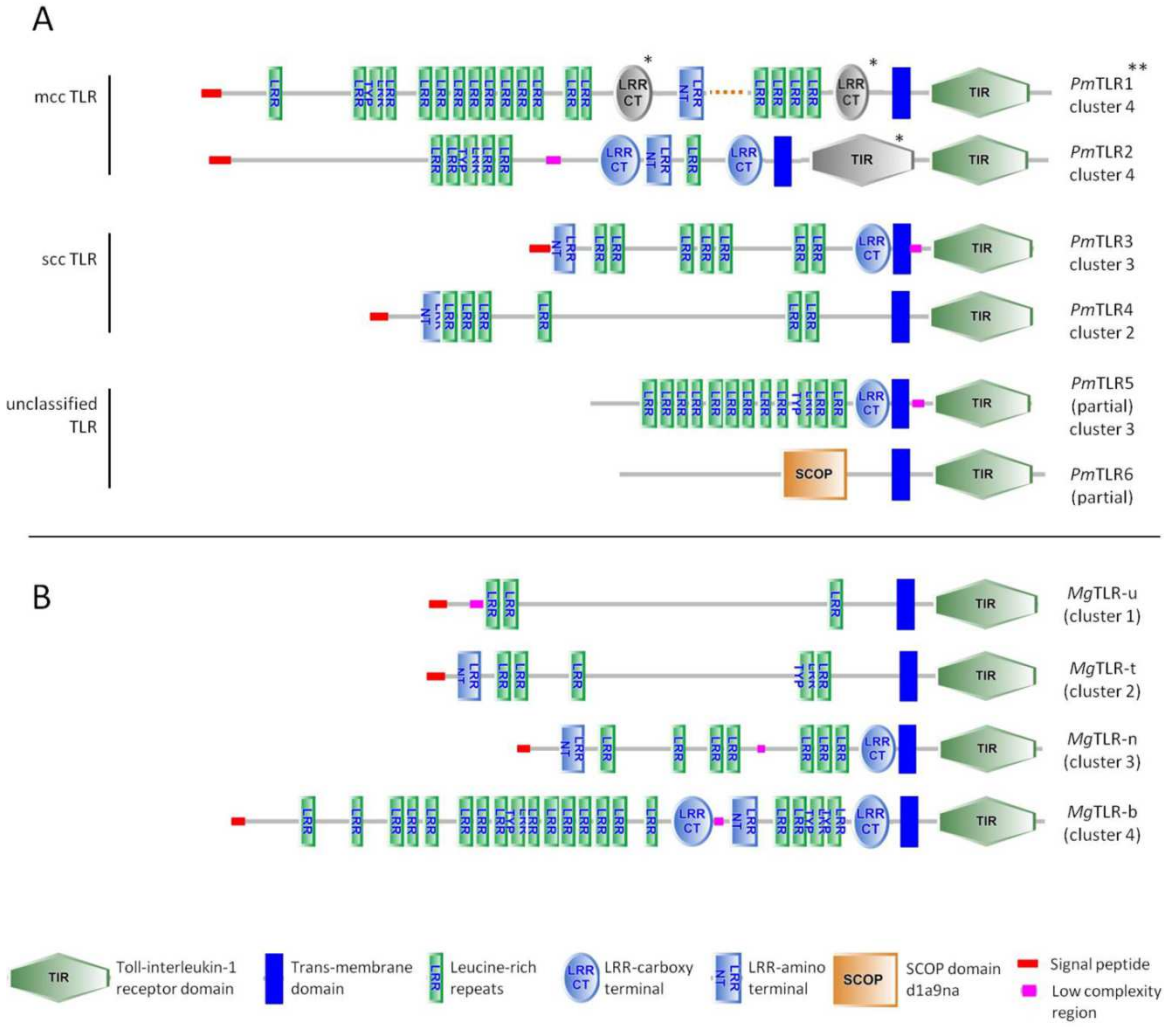


Figure 2. (A) Structure of P.maximus TLRs homologs. *Schematic structure of TLRs identified in P. maximus hemocytes derived using SMART online server. The structure of the partial TLR6 lacks LRR domains but the presence of SCOP domain_d1a9na, a superfamily of less regular structures consisting of variable repeats like LRRs, may suggest that this partial sequence encodes a Toll-like receptor. TLR classification is based on both cysteine cluster multiplicity (scc or mcc TLRs) and ectodomain complexity as reported in Toubiana et al., 2013.* ***(B) Classification based on ectodomain architecture in M. galloprovincialis.*** *Cluster 1 is characterized by the presence of LRR domains only (e.g., TLR-u), cluster 2 by LRR-NT (e.g., TLR-t), cluster 3 by LRR-CT (e.g., TLR-n) and cluster 4 by the presence of mccTLR only (e.g., TLR-b).*

** Domain identity determined by blast similarities searches in SMART. ** Sequence not completely obtained. However the high percentage of aa identity (85%) of the translated contigs number 4021 and 8142 to the N-terminal and C-terminal sections of C. farreri TLR (GenBank: DQ350772) respectively, allowed the hypothesis that they are both partial sequences belonging to the same mRNA transcript. Despite a scan of all the transcriptome resources obtained within this study and in previous sequencing runs, the complete sequence has not been detected.*

In order to analyze the evolutionary relationship of *P. maximus* TLRs in the context of other mollusc TLRs, a phylogenetic tree was constructed (Figure 3). Such analysis showed four large clusters (A-D) and two pairs of sequences located in a basal position (*Pm*TLR6-*Mg*TLRa, *Pm*TLR5-*Cg*TLR3). Notably, only clusters B and C were supported

by a good bootstrap value, for this reason the relationships between the six main branches cannot be completely resolved. With regard to *Pm*TLR5 and *Pm*TLR6, they reasonably represent the orthologs of *Cg*TLR3 and *Mg*TLRa, respectively. Looking at cluster A, *Pm*TLR1 clustered together with *Mg*TLR-b, *Cg*TLR1 and the single TLRs identified (to date) in *C. farreri*, *Haliotis discus discus* and *Hyriopsis cumungii*, suggesting that they are all orthologous copies, *i.e.* they derive from the same TLR gene that was present in the common ancestor of these species. All cluster A TLRs belong to the mcc TLR group (cluster 4). In the same cluster A, *Pm*TLR2 is the ortholog of mussel TLR-i. Cluster B contains a large number (9) of mussel TLRs, while *Pm*TLR4 is the sister group to this set of sequences. The most plausible interpretation is that a single TLR copy homologous to *Pm*TLR4 existed in the ancestral lineage, which underwent several duplication events after the divergence of mussel from scallop. Cluster C is likewise characterized by a multiplicity of mussel TLRs, which likely represent another case of repeated duplications, although it is not clear which is the putative orthologous in other bivalves. Cluster D includes different TLRs from oyster and mussel, and a single scallop gene. Orthology relationships within cluster D appear difficult to establish, yet multiplicity of TLRs is evident again in mussel.

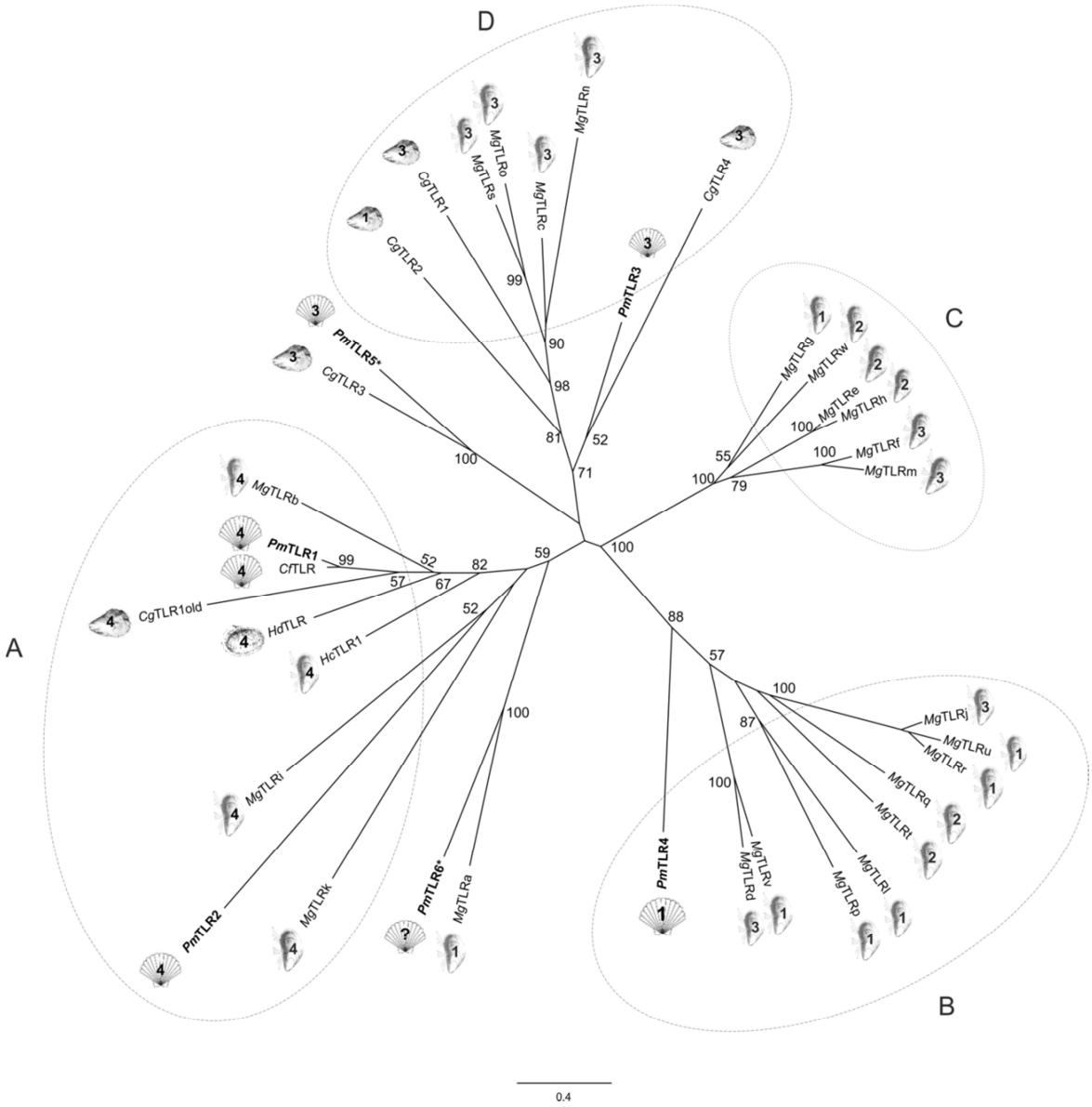


Figure 3. Phylogenetic tree of molluscs TLRs homologs. *Maximum-likelihood tree evaluating the evolutionary relationships of scallop transcripts containing both TIR and trans-membrane domains. The phylogenetic analysis was based on a multiple sequence alignment further refined to include only reliably aligned positions (Gblocks) yielding a total of 149 aligned amino-acids. Only significant bootstrap values for the lineages are shown (>50%). Shells correspond to each species. Numbers contained in the shells (1-4) refer to the cluster classification based on ectodomains organization as reported in M. galloprovincialis (Figure 2B). TLRs homologs identified in this study are in bold and those with a partial sequence were labelled with an asterisk (*). Sequences included in this analysis and GenBank accessions were the following: Chamys farreri (CfTLR: DQ350772), Crassostrea gigas (CgTLR1/2/3/4: KC700617-KC700620; CgTLR1old: ADV16385), Haliotis discus discus (HdTLR: AGJ03555), Hyriopsis cumingii (HcTLR: unpublished sequence) and Mytilus galloprovincialis (MgTLRs: JX173687-90, KC357777-80, KC413022-31).*

As concerns immune effectors, 154 contigs showed a putative functional annotation corresponding to HSP subunits 70 and 90, and antioxidant enzymes. Among the antioxidant enzymes, all glutathione S transferase isoforms (Alpha, Kappa, Theta, Pi, Mu

and Omega), prostaglandin reductase 1 and ceruloplasmin were detected. One contig encoded a putative full-length copy of big defensin (BD). BD is an AMP initially characterized in the horseshoe crab *Tachypleus tridentatus*, and then reported in various molluscan species. The putative *P. maximus* BD (*Pm*BD) is encoded in a polypeptide of 127 amino acids. The deduced protein shows high sequence similarity (70.25%) with *Argopecten irradians* defensin BD (*Ai*BD) and it shares the common features of BD family, including signal peptide (SP), amino-terminal antimicrobial domain (N-AMD) and carboxyl-terminal β-defensin domain (cβDD) (Figure 4).

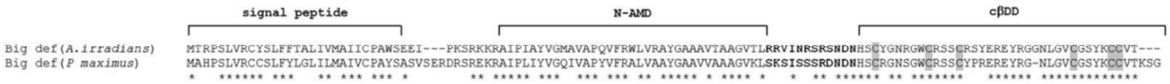


Figure 4. P. maximus big defensin. *Amino acid sequence alignment of big defensins from P. maximus and A. irradians (GenBank: ABC61319). The conserved cysteine residues are evidenced with grey backgrounds. The predicted signal peptide, Hydrophobic N-terminal domain (N-AMD) and C-terminal cysteine-rich domain (CβDD) are indicated. The putative loop between N-AMD and CβDD domains is in bold. Asterisks (*) indicate identical amino acid residues.*

A phylogenetic tree was constructed (Figure 5) based on the multiple alignment of all available mollusc BDs. Among bivalves, a major cluster joined together two groups, one for BDs in Mytilidae and one for Ostreidae BDs. In each of these two sub-clusters, multiple BD copies were observed in *M. galloprovincialis* and *C. gigas*. The closest sequences to the major Ostreidae-Mytilidae cluster were the single BDs found in the Veneridae *V. philippinarum* and in the Arcidae *Scapharca broughtonii*. More distantly related were the two Pectinidae, *A. irradians* and *P. maximus*, whereas the most distant was *H. cumingii*, which belongs to the Paleoheterodonta.

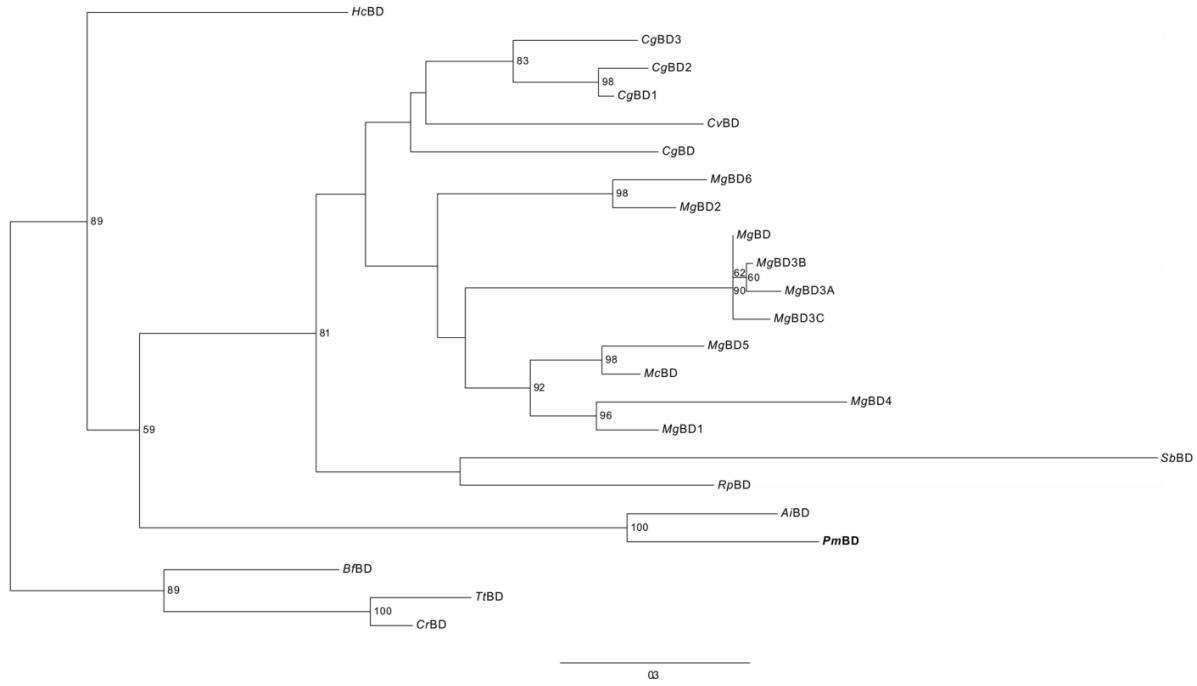


Figure 5. Big defensins phylogenetic tree. *Maximum likelihood tree of several amino acid sequences of big defensins. Only significant bootstrap values are shown (>50%). GenBank accession of the sequences included in this analysis were the following: Crassostrea gigas AEE92768 (CgBD1), AEE92775 (CgBD2), AEE92778 (CgBD3), EKC19784.1 (CgBD) and C. virginica CV133156.1 (CvBD*); Mytilus galloprovincialis AFC37168.1 (MgBD), CCC15007.1 (MgBD1), CCC15008.1 (MgB2), CCC15009.1 (MgBD3a), CCC15010.1 (MgBD3b), CCC15011.1 (MgBD3c), CCC15012.1 (MgBD4), CCC15013.1 (MgBD5), CCC15014.1 (MgBD6) and M. californianus GE759807.1 (McBD*); Hyriopsis cumingii AEP26934.1 (HcBD); Venerupis philippinarum AEK78068.1 (RpBD); Scapharcha broughtonii AFQ02696.1 (SbBD); Argopecten irradians ABC61319 (AiBD); horseshoe crabs Tachypleus tridentatus P80957.2 (TtBD) and Carcinoscorpius rotundicauda CK086629 (CrBD); amphioxus Branchiostoma floridae ADH03419 (BfBD). Asterisks (*) indicate big defensins deduced by translating EST sequences homologous to big defensins in different mollusc species in the GenBank database.*

With regard to signal transduction molecules, 3 contigs showed high similarity with Myeloid differentiation 88 (MyD88), an adaptor molecule involved in the TLR pathway, while 30 contigs encoded TNF receptor associated factors (TRAFs).

As concerns immune proteolytic cascades, 8 contigs were annotated as cathepsins (B, F, L and Z) and additional 49 encoded various proteases and proteases inhibitors.

Transcripts coding for apoptosis-related peptides were 229, including various caspase isoforms (2, 3, 6, 7, 8 and 10), baculoviral IAP repeat-containing proteins, β1-arrestin, autophagy-related proteins, death domain-associated proteins, Bcl-like proteins and several other apoptosis regulators.

While not all complement components could be identified, complement C3 and four proteins similar to C1q-like protein 4 were reported. Cytokine signaling was represented

by 67 contigs, encoding interleukins, interleukin receptors, interferon-related proteins, TNF-receptors, and TNF-ligands. A relevant number (79) of transcripts encoded cell-surface receptors or cell-adhesion molecules, which included putative integrins α and β, TGF-β receptor, and hemicentin-1 (HEM-1), a large extracellular member of the immunoglobulin superfamily.

Other 197 contigs showed significant homology with immune-related proteins such as tetraspanins and Tripartite Motif containing (TRIM) proteins. Tetraspanins, a superfamily of integral membrane proteins, are involved in regulating molecular recognition at the cell surface and in the generation of an efficient immune response, while TRIM proteins participate in the regulation of pathogen-recognition and transcriptional pathways in host defence.

3.3. Unstimulated hemocyte transcriptome analysis

RNA-seq reads originating from unchallenged hemocytes were mapped against the reference transcriptome yielding expression levels for individual transcripts. The list of the 150 genes showing the highest expression value (Appendix B2) comprised constitutive genes such as actins, alpha and beta tubulins, collagen, and key components of protein synthesis such as ribosomal proteins and elongation factors, together with more genes encoding immune-specific proteins. The latter included allograft inflammatory factor 1 (AIF1), HSP70B2, ferritin and ceruloplasmin.

3.4. Differentially expressed genes in stimulated hemocytes

First of all it should be noted that the quantitative data obtained within this study derived from a single replicate, thus the results must be considered as preliminary and need to be confirmed by further analyses. Notwithstanding the fact that the analysis was not statistically supported by the presence of biological replicates, it should be taken in consideration that each sample was a pool of at least three independent samples, thus partly reducing individual variation. Moreover, to minimize the risk of misleading results, the DEGs were calculated through an algorithm (*NOIseq-sim*; see section 2.5) that is appropriate for datasets lacking replicates. This method simulates technical replicates from a multinomial distribution and, even though it relies on an approximation, it permits to identify at least those genes showing the highest fluctuation between conditions.

RNA-seq data from stimulated hemocytes and controls were compared by focusing just on up-regulated transcripts. The analysis revealed respectively 1,576 and 3,347

differentially expressed genes (DEGs) after *in vivo* and *in vitro* stimulation. Annotated DEGs against the Pacific oyster genome database were 373 (23.7%) for the *in vivo* challenge and 611 (18.3%) for *in vitro* one. Since control hemocytes were obtained following an *in vivo* stimulation with FSW, DEGs identified after *in vitro* challenge might show the effects inherent to laboratory culture conditions in addition to those generated by exposure to PAMPs. For this reason, a conservative approach was preferred and only those transcripts which were regulated also after the *in vivo* challenge were considered in the analysis. A common set of 192 transcripts were up-regulated after both stimulations (Fisher's exact test $P < 0.001$). Among these common DEGs there were *Pm*BD, carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1), TRIM45, and regucalcin. Most of common DEGs (167), however, were represented by un-annotated contigs. Detailed analysis of putative open ORFs identified in these 167 contigs revealed that 76 ORFs have a protein region with significant homology to a known protein domain. Of these, noteworthy is contig number 1442, which contained a LCCL domain. The LCCL domain is characterized by a conserved C-terminal motif including a highly conserved hystidine and is thought to be an autonomously-folding domain which has been also proposed to be involved in LPS binding (*Trexler et al., 2000*). Eight additional ORFs encoded by common DEGs carried a significantly recognizable signal peptide (SP). The presence of SP indicates that the protein is targeted to the extracellular space, either to be exposed on the cell surface or to be secreted. A SP is typically present in AMPs. Two of the ORFs carrying a SP also showed a suggestive similarity with the "defensin" domain when analyzed with SMART, although the associated probability was above the default significance threshold.

A specific analysis of scallop TLR expression after the *in vivo* bacterial challenge showed that *Pm*TLR1 was unchanged, while *Pm*TLR6 expression level was significantly reduced. On the opposite, *Pm*TLR2 was more abundant, although with marginal probability values. No significant change was detected for *Pm*TLR3 and *Pm*TLR4., the putative ortholog of *Mg*TLR-a.

A more systematic evaluation of DEGs after the *in vivo* challenge was carried out using enrichment analysis, revealing over-representation of several biological processes that are involved in innate immunity. Enriched GO terms and PANTHER biological processes (BPs) are reported in Appendix B3. 6 GO_CC, 50 GO_BP, and 25 GO_MF terms showed significant enrichment. Enriched GO_BPs included "innate immune response" (GO:0045087), "defense response" (GO:0006952), and "cytokines-mediated signalling

pathway" (GO:0019221). Using PANTHER database, 14 BPs showed a significant fold-enrichment (FE) > 2. Among them, "cell communication" (BP00274), "cell surface receptor mediated signal transduction" (BP00103), and "receptor mediated endocytosis" (BP00132) revealed a FE >3. Additional relevant BPs were "immunity and defense (BP00148)", "apoptosis" (BP00179), "signal transduction" (BP00102), "cell adhesion" (BP00124), and "G-protein mediated signalling" (BP00104).

4. Discussion

Despite the increasing interest in bivalve immunity due to their economic importance, the catalog of immune genes in this taxonomic group is still quite limited. The main reason is the large phylogenetic distance of molluscs from the major model species such as fruitfly, earthworm, mouse, zebrafish, which makes difficult to isolate specific genes based on sequence similarity across taxa. This is especially true for immune-related genes, which are known to be often highly variable. Deep transcriptome sequencing using next generation sequencing (NGS) is emerging as a powerful approach to rapidly increase sequence information in non-model species. NGS-generated transcritptome information is not entirely immune from the problem of limited sequence conservation. In fact, less than one third of scallop contigs (31%) were found to match a known protein coding gene in other species. Nonetheless, a large set of genes could be identified, which belong to all major functional groups of immune-related proteins (Appendix B1), confirming that the main immune pathways are conserved in bivalves, including pathogen recognition and binding, phagocytosis and microbial membrane lysis, apoptosis, cell-cell communication and signal transduction. Moreover the NGS technique employed in this study (RNA sequencing by Illumina) concurrently provided quantitative data and although the analysis was restricted to one library per condition, the adopted pooling strategy and statistical analysis (*Tarazona et al., 2011*) allowed to obtain preliminary information on transcriptional regulations in response to immune stimulation.

With respect to host/pathogen interactions, the large number of putatively identified PRRs (109) highlights the importance of immune recognition mechanisms in scallop. In the past decades, various families of conserved PRRs have been identified in molluscs and their role in immune recognition has been demonstrated (*Wang et al., 2013*). In the present study, considerable substantial set of transcripts showed high sequence similarity to endothelial cells scavenger receptor (SREC) and macrophage mannose receptor 1 (MRC1). This evidence suggests that several isoforms of SREC- and MRC1-like proteins are expressed in scallop hemocytes. In the Pacific oyster genome, the number of predicted gene scaffolds corresponding to MRC1 and SREC was 11 and 21 respectively, supporting the hypothesis that multiple isoforms of these transcripts exist. Noteworthy, in humans, the multiplicity of scavenger receptors has been recently reported (*Greaves and Gordon, 2009*). Scavenger receptors (SRs) mediate non-opsonic phagocytosis by recognizing a wide range of ligands including microbial surface constituents and intact microbes, and are involved in various processes of host defense, apoptosis, autoimmunity, inflammation

and lipid metabolism (*Berwin et al., 2003; Mukhopadhyay and Gordon, 2004*). At present, knowledge on molecular function of SRs in invertebrates is quite scarce and only in *Drosophila* (*Ramet et al., 2001; Kocks et al., 2005*), in the sea urchin *Strongylocentrotus purpuratus* (*Rast et al., 2006*) and in the scallop *C. farreri* (*Liu et al., 2011*) they have been identified and studied. Here 9 expressed sequences with high similarity to the oyster homolog endothelial cell SR and 2 sequences similar to class F scavenger receptors were identified. Another important receptor that was identified in the scallop transcriptome is the mannose receptor (MR). MR is known to recognize mannose, fucose, and N-acetylglucosamine in a Ca^{2+}-dependent manner and it is expressed by most macrophage populations (*Taylor et al., 2005*), which are the vertebrate counterpart of the hemocytes in invertebrates. In addition to endogenous glycoproteins, MR plays an important role in pathogen recognition, antigen internalization and presentation. In bivalves MR has not been studied in detail, but a putative MR homolog was reported in the Mediterranean mussel *M. galloprovincialis* (*Venier et al., 2011*) and in the Pacific oyster *C. gigas* (*Zhang et al., 2012*).

Another immune pathway that comprised a high number of contigs was apoptosis. Programmed cell death has a key immune-modulatory role and is essential for normal functioning of the immune system (*Hildeman et al., 2007; Birge and Ucker, 2008*). Studies using NGS have recently provided evidence of a link between immune defense and programmed cell death in bivalves (*Moreira et al.. 2012; Romero et al., 2011*). The considerable number (229) of apoptosis-related genes identified in the present study suggests that also in *P. maximus* hemocytes programmed cell death is essential in the maintenance of immune system homeostasis.

TLRs were the focus of a more detailed analysis, since these proteins play a pivotal role in the immune system and the activation of their signalling cascade leads to the translocation of the transcription nuclear-factor-kB (NF-kB) inside the nucleus, where it induces a variety of immune-related effector genes involved in building up the front-line against invading pathogens (*Kopp and Medzhitov, 1999; Cornwell and Kirkpatrick, 2001*). However, great structural and functional divergence between mammalian TLRs and invertebrates Tolls has been reported, thus making TLR-mediated innate immune response an on-going area of controversy (*Leulier and Lemaitre, 2008*). Interestingly, the most striking difference between mammal TLRs and *Drosophila* Tolls (the invertebrate species for which more comprehensive data are available) is the ability to direct recognize their ligands. In the present study, accurate phylogenetic analysis was carried out to

identify putative orthologs for scallop TLRs in other bivalve species. A complex evolutionary pattern for this protein family in molluscs was observed as already reported in other taxonomic groups, with ancient duplications as well as recent increase in TLR copy number for specific lineages. Once the evolutionary framework is established, transcriptional response to immune challenge can be compared across species, although quantitative expression data represent a single time point and a unique biological replicate (n=1 pool of at least three independent samples). *Pm*TLR2 showed a two-fold increase in response to the *in vivo* challenge, in agreement with evidence from its putative ortholog in mussel, *Mg*TLR-i (*Toubiana et al., 2013*). Concordant evidence between the two species, although of opposite sign, was found for *Pm*TLR6. *Pm*TLR6 was strongly down-regulated after *in vivo* challenge and its putative mussel ortholog, *Mg*TLR-a, was under-expressed following *in vivo* stimulation with *V. anguillarum* (*Toubiana et al., 2013*). Noteworthy, down-regulation of TLRs after bacterial challenge has also been reported in hemocytes from the clam *Mya arenaria* (*Mateo at al., 2010*) and in the disk abalone subjected to LPS injection (*Elvitigala et al, 2013*). While further studies are needed in *P. maximus* to analyse TLRs expression at different time points after stimulation, preliminary evidence seems to confirm a divergent pattern for individual TLRs, yet conserved for orthologous copies across bivalve species.

As already mentioned, NGS allowed transcriptome characterization in several bivalve species (e.g. *Philipp et al., 2012; Milan et al., 2011; Moreira et al., 2012; Hou et al., 2011*; *Zhang et al., 2012*). However, so far the most frequently used approach has been based on 454 Roche pyrosequencing technology, which provided long sequence reads (up to 500 bp), but was limited in terms of sequencing depth, as typically 1-2 million reads were obtained for each library. Therefore, transcripts abundance could not be reliably assessed. Illumina sequencing technology produces over 300 millions reads per run, which enables quantitative analysis of gene expression by counting sequence reads mapped against a reference transcriptome. Such approach was applied here first to evaluate highly expressed transcripts in un-stimulated scallop hemocytes. Apart from constitutive genes, several highly represented contigs encoded proteins involved in immune function. AIF-1, a calcium-binding cytokine, which has been recently cloned and studied in oysters (*Li et al., 2013*; *Zhang et al., 2013*). These studies demonstrated that oyster AIF-1 is constitutively expressed in various tissues and enriched in hemocytes. In sea urchin *Sterechinus neumayeri* (*Ovando et al., 2012*) and clam *V. philippinarum* (*Zhang et al., 2011*) there is functional evidence that AIF-1 is involved in hemocyte

activation and inflammatory response. Other highly expressed genes were HSP70, ceruloplasmin (CP) and ferritin (FT). HSPs are ubiquitous proteins that help organisms to modulate stress response and protect them from environmentally-induced cellular damage. A great number of HSPs have been identified from molluscs and their involvement in bivalve immunity is widely proven (*Wang et al., 2013; Liu and Chen, 2013*).. More specifically, some studies showed that HSP70 in bivalves might be regulated not only by thermal stress, but also by bacterial challenges (*Song et al., 2006; Cellura et al., 2006*) and that HSP70 gene is abundantly transcribed also at a physiological state (*Zhou et al., 2013*). CP and FT have both a ferroxidase activity. CP in mammals is produced also by macrophages and its ferroxidase activity, by converting Fe^{2+} to Fe^{3+} plays an important role in iron metabolism and transport and also exhibits ferroxidase-dependent bactericidal activity (*Klebanoff, 1992*). In addition, CP inhibits ferrous ion-mediated production of reactive oxygen species (ROS) and thus it possesses a potent antioxidant activity (*Bakhautdin et al., 2013*). Like CP, FT has antioxidant properties, but it is also involved in iron storage and iron-withdrawal with antimicrobial effects, as demonstrated in the Pacific white shrimp *Litopenaeus vannamei* (*Ruan et al., 2010*) and in bay scallop *A. irradians* (*Li et al., 2012*). High expression of CP and FT suggests that scallop hemocytes, even in absence of specific stimuli, possess strong activation of anti-oxidant defense. Even though ROS production in molluscan hemocytes has been mostly studied for its engagement in killing of encapsulated microorganisms (*Donaghy et al., 2009*), ROS may also be involved in cellular homeostasis and thus be produced in unchallenged hemocytes, as demonstrated in oyster by *Lambert* and co-authors (*Lambert et al., 2003*). While quantitative analysis of gene expression in un-stimulated hemocytes already provided substantial information, the most relevant evidence was obtained when analysing immune-challenged hemocytes, with a large set of significant DEGs identified after stimulation. The list of significant DEGs presented here should be regarded as a starting point for future analysis, since a detailed discussion on all DEGs is beyond the reach of a single paper. Nonetheless, functional enrichment analysis might help summarising the main features of transcriptome changes as part of the scallop immune response.

In *in vivo* challenged hemocytes enrichment of BPs such as “cell communication”, “cell surface receptor mediated signal transduction” and “receptor mediated endocytosis” suggests activation of immune signalling pathways, which are triggered by hemocyte-microbe interaction and amplified by hemocyte-hemocyte communication. For instance,

among DEGs belonging to these BPs, CEACAM1 and MR1 were observed. The role of MR1 has been already discussed above and, by homology with vertebrates, it is implicated in hemocyte response against pathogens through binding and phagocytosis of micro-organisms with surface mannose residues and soluble mannose-containing glycoproteins (*Stein et al., 1992*). Conversely, CEACAM1 is a surface adhesion molecule belonging to the CEA-family, a group of glycoproteins that is poorly studied in invertebrates. In mammalian, CEACAM glycoproteins are expressed by certain epithelial, endothelial, lymphoid and myeloid cells (*Gray-Owen and Blumberg 2006*) and are known to play an important role in adhesion, phagocytosis, oxidative burst, degranulation, and regulation of apoptosis. CEACAMs also function as receptors for pathogenic bacteria and viruses (*Gray-Owen and Blumberg 2006; Pańczyszyn and Wieczorek, 2012*). At least in humans, the latter function deserves a special attention since a variety of Gram-negative bacteria have the ability to target CEACAMs possibly allowing them to colonize tissues by mediating adhesion and by exploiting immunosuppressive functions of CEACAM receptors and thus evading host defense (*Klaile et al., 2013*). In particular, CEACAM members have been widely investigated in light of their binding with *Neisseria spp* surface proteins responsible for bacterial adherence, entry into host cells and interactions with the immune system (*Sadarangani et al., 2011*). In this context, it has been reported that CEACAM expression can be up-regulated by neisserial LPS and pro-inflammatory pathogen-induced cytokines (*Gorvel, 2004*). Moreover, *Klaile* et co-workers (*Klaile et al., 2013*) demonstrated that, in primary normal human bronchial epithelial cells, CEACAM1 mRNA transcript and protein were up-regulated by polyI:C and type I and type II interferons. Therefore, evidence that scallop CEACAM1 mRNA is induced in hemocytes 24h post bacterial challenge, as well as after *in vitro* stimulation, should prompt further studies to clarify its function in bivalve immune response as well as its potential role in pathogen immune-evasion.

Another interesting example of DEG is glutathione S-transferase omega-1 (GSTO1), an enzyme involved in the biotransformation of compounds including toxic substances and oxidative stress products, transport of ligands, and regulation of signalling pathways (*Burmeister et al., 2008*). GSTO1 transcript was found in the Pacific oyster gills and digestive gland (*Boutet et al., 2004*) and in disk abalone tissues (*Wan et al., 2009*), and it was over-expressed in the skin mucus of Atlantic cod challenged with *V. anguillarum* (*Rajan et al., 2013*). The role of antioxidant proteins in immune response has been already mentioned. In the case of GSTO1, considering its specific function, it might be to

counteract reactive oxygen metabolites induced by bacterial infection, as shown in a study where the measurement on enzyme activity complemented gene expression data (*Canesi et al., 2010*). Killing of pathogens by hemocytes is usually accompanied by a sudden release of ROS (*Bugge et al., 2007; Segal et al., 2005*). However, ROS generation would disrupt the homeostasis of redox balance in hemocytes and the induction of GSTO1 more likely counteracts the pathogen-induced oxidative stress.

DEGs identified after *in vitro* challenge should be interpreted with some caution, since the differential expression analysis was carried out by using as control condition unchallenged hemocytes collected *in vivo*. Nevertheless, a set of DEGs was in common with those reported after the *in vivo* challenge. Such overlap might be due to the fact that stimulation included LPS, either as a component of *V. anguillarum* in the *in vivo* challenge, or as purified molecule in the PAMP mixture. It might be possible, however, that the common list of 192 DEGs might be represent part of a "core" response to immune stimuli. In either case, shared DEGs are of special interest. In fact, one of these DEGs encoded a putative BD (*PmBD*), a well-known AMP. In molluscs, data on AMPs come mainly from *M. edulis* and *M. galloprovincialis* defensin, mytilin, myticin and mytimycin (*Mitta et al., 2000; Roch et al. 2008; Vera et al., 2011; Gerdol et al., 2012*). However potential AMPs have been also identified from other bivalve including oysters *Crassostrea madrasensis* and *Saccostrea cucullata* (*Sathyan et al., 2012*), *C. gigas* (*Schmitt et al., 2012*), clams *V. decussata* (*Gestal et al., 2007*) and *M. mercenaria* (*Perrigault et al., 2009*), scallops *C. farreri* (*Li et al., 2007*) and *A. purpuratus* (*Tapia et al., 2011*). Notably, even if some studies have investigated such peptides in scallop, no data are available in *P. maximus* and to our knowledge, this is the first AMP transcript reported for this species. The amino acid sequence alignment with the *A. irradians* BD (Figure 4) evidenced a high percentage of identity and a similar protein structure. The two spatially distinct structural domains that can fold and function in isolation could be recognized. N-AMD comprises half of the molecule and is rich in hydrophobic residues (*Kouno et al., 2009*). As demonstrated by *Saito* and colleagues in *T. tridentatus* (*Saito et al., 1995*), while N-AMD has selective activity against Gram-positive bacteria, the C-βDD domain is active on Gram-negative bacteria and it perfectly corresponds to a β-defensin with six identical cysteines to form three disulfide bridges. Structurally, as in *A. irradians* BD, the C-βDD domain is separated from the N-AMD domain by a small loop that represents a major target for insertion/deletion (indel) mutations within this family. Evolutionary analysis confirmed the close relationship between *A. irradians* and *P.*

maximus BDs, and demonstrated that the two Pectinidae clustered separately from mussels and oysters, even though they all belong to the subclass Pteriomorphia, suggesting that BDs in scallops are highly divergent in terms of sequence evolution.

The large group of Pteriomorphia was further split in three branches dividing clams, mussels and oysters. While the existence of additional homologs that they have not been detected yet cannot be ruled out, multiplicity of BDs could be observed especially in *M. galloprovincialis*, with repeated events of gene duplication, resulting in a wide set of BDs.

There is a clear contrast between *C. gigas* and *M. galloprovincialis*, where multiple BDs were found, and *P. maximus*, where a single BD could be identified. Despite targeted Blast searches were carried out using defensins and AMPs from several species as a query, and ORFs were carefully evaluated to find AMP motifs, no other putative AMPs could be detected. While the hypothesis that additional, canonical AMPs exist in *P maximus*, but were not present in the obtained sequences, cannot be entirely ruled out, it seems unlikely considering the broad representation and large sequencing depth of the scallop transcriptome. Alternatively, a single canonical AMP is present in the arsenal of antimicrobial defenses for the great scallop. It has been already observed that multiplicity of genes encoding immune effectors is often found in species known to be robust against microbial infections, as is the case of Mediterranean mussel, while a limited repertoire might lead to increase susceptibility to pathogens. *P. maximus* is notoriously a highly susceptible species to bacterial infections, which might be linked to the evidence of a single AMP. On the other hand, it could be that non-canonical peptides are expressed in scallop, bearing insufficient sequence conservation to be easily identified as AMPs. In the present study, two ORFs coding for short peptides, characterized by a SP and marginal similarity to defensin-like domains, were discovered. The corresponding mRNAs were significantly up-regulated after either *in vitro* or *in vivo* challenge, similar to what observed for the single *bona fide* scallop AMP, *PmBD*. Several previous studies conducted in different species confirm that over-expression of AMPs follows exposure to immune stimuli. *Zhao* and coworkers (*Zhao et al., 2007*) demonstrated that relative expression level of *A. irradians* BD in hemolymph was up-regulated with a drastic increase 32 h after challenge with *V. anguillarum*. In the Pacific oyster, circulating hemocytes showed significant increase in *C. gigas* BD1 and BD2 12 h after stimulation with a mix of heat-killed Gram-positive and Gram-negative bacteria (*Rosa et al., 2011*). A further study performed in *C. gigas* hemocytes reported no differential representation

of BDs following challenge with diverse virulent *Vibrio spp* in comparison with the corresponding inactivated strains, suggesting that BD expression is induced by the presence of the pathogen itself, in a way that does not dependent on bacterial virulence (*de Lorgeril et al., 2011*). The same trend of induction has been observed also for BD *in vivo* challenged clams *V. philippinarum* (*Zhao et al., 2010*) and *S. broughtonii* (*Li et al., 2012*).

5. Conclusions

Deep sequencing of the scallop immune transcriptome demonstrates the great potential for such an approach to rapidly increase our knowledge on bivalve immunity. An extensive number of transcripts was sequenced, allowing the annotation of most major proteins involved in various mechanisms of innate immune response, such as pathogen recognition and binding, phagocytosis and microbial membrane lysis, apoptosis, cell-cell communication and signal transduction. At the same time, preliminary functional characterization of all expressed transcripts could be obtained through the use of ultra high-throughput NGS technology and the comparison of different hemocyte samples, either un-stimulated or challenged. Besides the identification of conserved immune genes, it was shown that in-depth data mining through specific motif searches might lead to isolate novel immune-relevant molecules.

These finding provided information on the great scallop innate immunity and may contribute to develop strategies for management of diseases and for long-term sustainability of *P. maximus* aquaculture.

6. References

Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell 2006;124(4):783-801.

Araya MT, Markham F, Mateo DR, McKenna P, Johnson GR, Berthe FCJ et al. Identification and expression of immune-related genes in hemocytes of soft-shell clams, *Mya arenaria*, challenged with *Vibrio splendidus*. Fish Shell Immunol 2010;29(4):557-64.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM et al. Gene ontology: Tool for the unification of biology. Nat Genet 2000;25(1):25-9.

Bakhautdin B, Febbraio M, Goksoy E, de la Motte CA, Gulen MF, Childers EP et al. Protective role of macrophage-derived ceruloplasmin in inflammatory bowel disease. Gut 2013;62(2):209-19.

Berwin B, Hart JP, Rice S, Gass C, Pizzo SV, Post SR et al. Scavenger receptor-A mediates gp96/GRP94 and calreticulin internalization by antigen-presenting cells. Embo J 2003;22(22):6127-36.

Birge RB, Ucker DS. Innate apoptotic immunity: The calming touch of death. Cell Death Differ 2008;15(7):1096-102.

Boutet I, Tanguy A, Moraga D. Characterisation and expression of four mRNA sequences encoding glutathione S-transferases pi, mu, omega and sigma classes in the pacific oyster *Crassostrea gigas* exposed to hydrocarbons and pesticides. Mar Biol 2004;146(1):53-64.

Bugge DA, Hegaret H, Wikfors GH, Allam B. Oxidative burst in hard clam (*Mercenaria mercenaria*) haemocytes. Fish Shell Immunol 2007;23(1):188-96.

Burmeister C, Lueersen K, Heinick A, Hussein A, Domagalski M, Walter RD et al. Oxidative stress in *Caenorhabditis elegans*: Protective effects of the omega class glutathione transferase (GSTO-1). Faseb J 2008;22(2):343-54.

Canesi L, Barmo C, Fabbri R, Ciacci C, Vergani L, Roch P et al. Effects of *Vibrio* challenge on digestive gland biomarkers and antioxidant gene expression in *Mytilus galloprovincialis*. Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 2010;152(3):399-406.

Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 2000;17(4):540-552.

Cellura C, Toubiana M, Parrinello N, Roch P. HSP70 gene expression in *Mytilus galloprovincialis* hemocytes is triggered by moderate heat shock and *Vibrio anguillarum*, but not by *V. splendidus* or *Micrococcus lysodeikticus*. Dev Comp Immunol 2006;30(11):984-97.

Chen G, Wang C, Zhang C, Wang Y, Xu Z, Wang C. A preliminary study of differentially expressed genes of the scallop *Chlamys farreri* against acute viral necrobiotic virus (AVNV). Fish Shell Immunol 2013;34(6):1619-27.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005;21(18):3674-6.

Cornwell WD, Kirkpatrick RB. Cactus-independent nuclear translocation of Drosophila RELISH. J Cell Biochem 2001;82(1):22-37.

Costa MM, Prado-Alvarez M, Gestal C, Li H, Roch P, Novoa B et al. Functional and molecular immune response of Mediterranean mussel (*Mytilus galloprovincialis*) haemocytes against pathogen-associated molecular patterns and bacteria. Fish Shell Immunol 2009;26(3):515-23.

de Lorgeril J, Zenagui R, Rosa RD, Piquemal D, Bachere E. Whole transcriptome profiling of successful immune response to vibrio infections in the oyster *Crassostrea gigas* by digital gene expression analysis. Plos One 2011;6(8):e23142.

Donaghy L, Lambert C, Choi K, Soudant P. Hemocytes of the carpet shell clam (*Ruditapes decussatus*) and the manila clam (*Ruditapes philippinarum*): Current knowledge and future prospects. Aquaculture 2009;297(1-4):10-24.

Dorrington T, Villamil L, Gomez-chiarri M. Upregulation in response to infection and antibacterial activity of oyster histone H4. Fish Shell Immunol 2011;30(1):94-101.

Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol 2011; 7(10): e1002195.

Elvitigala DAS, Premachandra HKA, Whang I, Nam BH, Lee J. Molecular insights of the first gastropod TLR counterpart from disk abalone (*Haliotis discus discus*), revealing its transcriptional modulation under pathogenic stress. Fish Shell Immunol 2013;35(2):334-42

FAO, 2013. Species Fact Sheets: *Pecten maximus*. Available at: http://www.fao.org/fishery/species/3516/en (accessed 28.10.13)

Fleury E, Huvet A, Lelong C, de Lorgeril J, Boulo V, Gueguen Y et al. Generation and analysis of a 29,745 unique expressed sequence tags from the pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: The GigasDatabase. Bmc Genomics 2009;10:341.

Garnier M, Labreuche Y, Garcia C, Robert A, Nicolas JL. Evidence for the involvement of pathogenic bacteria in summer mortalities of the pacific oyster *Crassostrea gigas*. Microb Ecol 2007;53(2):187-96.

Genard B, Miner P, Nicolas J, Moraga D, Boudry P, Pernet F et al. Integrative study of physiological changes associated with bacterial infection in pacific oyster larvae. Plos One 2013;8(5):e64534.

Gerdol M, De Moro G, Manfrin C, Venier P, Pallavicini A. Big defensins and mytimacins, new AMP families of the mediterranean mussel *Mytilus galloprovincialis*. Dev Comp Immunol 2012;36(2):390-9.

Gestal C, Costa MM, Figueras A, Novoa B. Analysis of differentially expressed genes in response to bacterial stimulation in hemocytes of the carpet-shell clam *Ruditapes decussatus*: Identification of new antimicrobial peptides. Gene 2007;406(1-2):134-43.

Gestal C, Roch P, Renault T, Pallavicini A, Paillard C, Novoa B et al. Study of diseases and the immune system of bivalves using molecular biology and genomics. Rev. Fish. Sci. 2008:16,131–54.

Gomez-Leon J, Villamil L, Lemos ML, Novoa B, Figueras A. Isolation of *Vibrio alginolyticus* and *Vibrio splendidus* from aquacultured carpet shell clam (*Ruditapes decussatus*) larvae associated with mass mortalities. Appl Environ Microbiol 2005;71(1):98-104.

Gorvel JP. Intracellular Pathogens in Membrane Interactions and Vacuole Biogenesis. New York: Kluwer Academic/Plenum Publishers 2004.

Gray-Owen SD, Blumberg RS. CEACAM1: contact-dependent control of immunity. Nat Rev Immunol 2006;6(6):433-46.

Greaves DR, Gordon S. The macrophage scavenger receptor at 30 years of age: current knowledge and future challenges. J Lipid Res 2009;50:S282-S286.

Hannam ML, Bamber SD, Moody AJ, Galloway TS, Jones MB. Immunotoxicity and oxidative stress in the arctic scallop *Chlamys islandica*: Effects of acute oil exposure. Ecotoxicol Environ Saf 2010;73(6):1440-8.

Hauton C, Hawkins LE, Hutchinson S. Response of haemocyte lysosomes to bacterial inoculation in the oysters *Ostrea edulis* L. and Crassostrea gigas (thunberg) and the scallop *Pecten maximus* (L.). Fish Shell Immunol 2001;11(2):143-53.

Hildeman D, Jorgensen T, Kappler J, Marrack P. Apoptosis and the homeostatic control of immune responses. Curr Opin Immunol 2007;19(5):516-21.

Hou R, Bao Z, Wang S, Su H, Li Y, Du H et al. Transcriptome sequencing and de novo analysis for yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX. Plos One 2011;6(6):e21560.

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4(1):44-57.

Ji C, Wu H, Wei L, Zhao J, Wang Q, Lu H. Responses of *Mytilus galloprovincialis* to bacterial challenges by metabolomics and proteomics. Fish Shell Immunol 2013;35(2):489-98.

Klaile E, Klassert TE, Scheffrahn I, Mueller MM, Heinrich A, Heyl KA et al. Carcinoembryonic antigen (CEA)-related cell adhesion molecules are co-expressed in the human lung and their expression can be modulated in bronchial epithelial cells by non-typable *Haemophilus influenzae*, *Moraxella catarrhalis*, TLR3, and type I and II interferons. Respir Res 2013;14:85.

Klebanoff SJ. Bactericidal effect of Fe(2+), ceruloplasmin, and phosphate. Arch Biochem Biophys 1992;295(2):302-8.

Kocks C, Cho JH, Nehme N, Ulvila J, Pearson AM, Meister M et al. Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in *Drosophila*. Cell 2005;123(2):335-46.

Kopp EB, Medzhitov R. The toll-receptor family and central of innate immunity. Curr Opin Immunol 1999;11(1):13-8.

Kouno T, Mizuguchi M, Aizawa T, Shinoda H, Demura M, Kawabata Set al. A novel beta-defensin structure: Big defensin changes its N-terminal structure to associate with the target membrane. Biochemistry 2009;48(32):7629-35.

Lacoste A, Jalabert F, Malham S, Cueff A, Gelebart F, Cordevant C et al. A *Vibrio splendidus* strain is associated with summer mortality of juvenile oysters *Crassostrea gigas* in the bay of Morlaix (north Brittany, France). Dis Aquat Org 2001;46(2):139-45.

Lambert C, Nicolas JL. Specific inhibition of chemiluminescent activity by pathogenic Vibrios in hemocytes of two marine bivalves: *Pecten maximus* and *Crassostrea gigas*. J Invertebr Pathol 1998;71(1):53-63.

Lambert C, Soudant P, Choquet G, Paillard C. Measurement of *Crassostrea gigas* hemocyte oxidative metabolism by flow cytometry and the inhibiting capacity of pathogenic vibrios. Fish Shell Immunol 2003;15(3):225-40.

Lemaitre B, Hoffmann J. The host defense of *Drosophila melanogaster*. Annu Rev Immunol 2007;25:697-743.

Leulier F, Lemaitre B. Toll-like receptors: taking an evolutionary approach. Nat Rev Genet 2008, 9(3):165-78.

Li C, Song L, Zhao J, Zhu L, Zou H, Zhang H et al. Preliminary study on a potential antibacterial peptide derived from histone H2A in hemocytes of scallop *Chlamys farreri*. Fish Shell Immunol 2007;22(6):663-72.

Li J, Chen J, Zhang Y, Yu Z. Expression of allograft inflammatory factor-1 (AIF-1) in response to bacterial challenge and tissue injury in the pearl oyster, *Pinctada martensii*. Fish Shell Immunol 2013;34(1):365-71.

Li J, Li L, Zhang S, Li J, Zhang G. Three ferritin subunits involved in immune defense from bay scallop *Argopecten irradians*. Fish Shell Immunol 2012;32(2):368-72.

Li M, Zhu L, Zhou CY, Sun S, Fan YJ, Zhuang ZM. Molecular characterization and expression of a novel big defensin (Sb-BDef1) from ark shell, *Scapharca broughtonii*. Fish Shell Immunol 2012;33(5):1167-73.

Liu D, Chen Z. The expression and induction of heat shock proteins in molluscs. Protein Peptide Lett 2013;20(5):601-6.

Liu L, Yang J, Qiu L, Wang L, Zhang H, Wang M et al. A novel scavenger receptor-cysteine-rich (SRCR) domain containing scavenger receptor identified from mollusk mediated PAMP recognition and binding. Dev Comp Immunol 2011;35(2):227-39.

Martin-Gomez L, Villalba A, Abollo E. Identification and expression of immune genes in the flat oyster *Ostrea edulis* in response to bonamiosis. Gene 2012;492(1):81-93.

Mateo DR, Greenwood SJ, Araya MT, Berthe FC, Johnson GR, Siah A. Differential gene expression of gamma-actin, Toll-like receptor 2 (TLR-2) and interleukin-1 receptor-associated kinase 4 (IRAK-4) in *Mya arenaria* haemocytes induced by in vivo infections with two *Vibrio splendidus* strains. Dev Comp Immunol 2010;34(7):710-4.

Milan M, Coppe A, Reinhardt R, Cancela LM, Leite RB, Saavedra C et al. Transcriptome sequencing and microarray development for the Manila clam, *Ruditapes philippinarum*: genomic tools for environmental monitoring. BMC Genomics 2011;12:234.

Mitta G, Vandenbulcke F, Roch P. Original involvement of antimicrobial peptides in mussel innate immunity. FEBS Letters. 2000;486(3):185-90.

Moreira R, Balseiro P, Planas JV, Fuste B, Beltran S, Novoa B, Figueras A. Transcriptomics of in vitro immune-stimulated hemocytes from the manila clam *Ruditapes philippinarum* using high-throughput sequencing. Plos One 2012;7(4):e35009.

Moreira R, Milan M, Balseiro P, Romero A, Babbucci M, Figueras A, Bargelloni L, Novoa B. Gene expression profile analysis of manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibrio alginolyticus* challenge using an immune-enriched oligo-microarray. Fish Shell Immunol 2013;34(6):1666-7.

Mukhopadhyay S, Gordon S. The role of scavenger receptors in pathogen recognition and innate immunity. Immunobiology 2004;209(1-2):39-49.

Nicolas JL, Corre S, Gauthier G, Robert R, Ansquer D. Bacterial problems associated with scallop *Pecten maximus* larval culture. Dis Aquat Org 1996;27(1):67-76.

Ovando F, Gimpel C, Cardenas C, Machado Cunha Da Silva JR, De Lorgeril J, Gonzalez M. Cloning and expression analysis of allograft inflammatory factor type 1 in coelomocytes of antarctic sea urchin (*Sterechinus neumayeri*). J Shellfish Res 2012;31(3):875-83.

Panczyszyn A, Wieczorek M. Role of CEACAM in neutrophil activation. Postepy Higieny i Medycyny Doswiadczalnej 2012;66:574-82.

Perrigault M, Tanguy A, Allam B. Identification and expression of differentially expressed genes in the hard clam, *Mercenaria mercenaria*, in response to quahog parasite unknown (QPX). Bmc Genomics 2009;10:377.

Philipp EER, Kraemer L, Melzner F, Poustka AJ, Thieme S, Findeisen U, Schreiber S, Rosenstiel P. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. Plos One 2012;7(3):e33091.

Qiu L, Song L, Xu W, Ni D, Yu Y. Molecular cloning and expression of a Toll receptor gene homologue from Zhikong Scallop, *Chlamys farreri*. Fish Shell Immunol 2007;22(5):451-66.

Rajan B, Lokesh J, Kiron V, Brinchmann MF. Differentially expressed proteins in the skin mucus of atlantic cod (*Gadus morhua*) upon natural infection with *Vibrio anguillarum*. Bmc Vet Res 2013;9:103.

Ramet M, Pearson A, Manfruelli P, Li XH, Koziel H, Gobel V, Chung E, Krieger M, Ezekowitz RAB. *Drosophila* scavenger receptor cl is a pattern recognition receptor for bacteria. Immunity 2001;15(6):1027-38.

Ramírez-Castillo E, Arellano-Martinez E, Ceballos-Vázquez B, González-Ocampo H, Aguirre-Guzmán G, Luna-González A. Immune response of Lion's paw scallop *Nodipecten subnodosus* (sowerby, 1835) challenged with *Vibrio alginolyticus*. J Vet Med 2011;41(4):425-32.

Rast JP, Smith LC, Loza-Coll M, Hibino T, Litman GW. Genomic Insights into the Immune System of the Sea Urchin. Science 2006;314 (5801):952-6.

Renault T, Faury N, Barbosa-Solomieu V, Moreau K. Suppression substractive hybridisation (SSH) and real time PCR reveal differential gene expression in the pacific cupped oyster, *Crassostrea gigas*, challenged with ostreid *herpesvirus 1*. Dev Comp Immunol 2011;35(7):725-35.

Riquelme C, Hayashida G, Toranzo AE, Vilches J, Chavez P. Pathogenicity studies on a *Vibrio anguillarum*-related (var) strain causing an epizootic in *Argopecten purpuratus* larvae cultured in Chile. Dis Aquat Org 1995;22(2):135-41.

Riquelme C, Hayashida G, Araya R, Uchida A, Satomi M, Ishida Y. Isolation of a native bacterial strain from the scallop *Argopecten purpuratus* with inhibitory effects against pathogenic vibrios. J Shellfish Res 1996a;15(2):369-74.

Riquelme C, Toranzo AE, Barja JL, Vergara N, Araya R. Association of *Aeromonas hydrophila* and *Vibrio alginolyticus* with larval mortalities of scallop (*Argopecten purpuratus*). J Invertebr Pathol 1996b;67(3):213-8.

Roch P, Yang Y, Toubiana M, Aumelas A. NMR structure of mussel mytilin, and antiviral-antibacterial activities of derived synthetic peptides. Dev Comp Immunol 2008;32(3):227-38.

Romero A, Estevez-Calvar N, Dios S, Figueras A, Novoa B. New insights into the apoptotic process in mollusks: Characterization of caspase genes in *Mytilus galloprovincialis*. Plos One 2011;6(2):e17003.

Rosa RD, de Lorgeril J, Tailliez P, Bruno R, Piquemal D, Bachere E. A hemocyte gene expression signature correlated with predictive capacity of oysters to survive vibrio infections. Bmc Genomics 2012;13:252.

Rosa RD, Santini A, Fievet J, Bulet P, Destoumieux-Garzon D, Bachere E. Big defensins, a diverse family of antimicrobial peptides that follows different patterns of expression in hemocytes of the oyster *Crassostrea gigas*. Plos One 2011;6(9):e25594.

Ruan Y, Kuo C, Lo C, Lee M, Lian J, Hsieh S. Ferritin administration effectively enhances immunity, physiological responses, and survival of pacific white shrimp (*Litopenaeus vannamei*) challenged with white spot syndrome virus. Fish Shell Immunol 2010;28(4):542-8.

Sadarangani M, Pollard AJ, Gray-Owen SD. Opa proteins and CEACAMs: Pathways of immune engagement for pathogenic neisseria. FEMS Microbiol Rev 2011;35(3):498-514.

Saito T, Kawabata S, Shigenaga T, Takayenoki Y, Cho JK, Nakajima H et al. A novel big defensin identified in horseshoe-crab hemocytes - isolation, amino-acid-sequence, and antibacterial activity. Journal Biochem 1995;117(5):1131-7.

Sandaa R, Brunvold L, Magnesen T, Bergh O. Monitoring the opportunistic bacteria Pseudoalteromonas sp. LT-13 in a great scallop, *Pecten maximus* hatchery. Aquaculture 2008; 276:14–21.

Sandlund N, Torkildsen L, Magnesen T, Mortensen S, Bergh O. Immunohistochemistry of great scallop *Pecten maximus* larvae experimentally challenged with pathogenic bacteria. Dis Aquat Org 2006;69(2-3):163-73.

Sathyan N, Philip R, Chaithanya ER, Kumar PRA. Identification and molecular characterization of molluskin, a histone-H2A-derived antimicrobial peptide from molluscs. ISRN Molecular Biology, vol. 2012, Article ID 219656, 6 pages.

Schmitt P, de Lorgeril J, Gueguen Y, Destoumieux-Garzon D, Bachere E. Expression, tissue localization and synergy of antimicrobial peptides and proteins in the immune response of the oyster *Crassostrea gigas*. Dev Comp Immunol 2012;37(3-4):363-70.

Segal AW. How neutrophils kill microbes. Annu Rev Immunol 2005;23:197-223.

Song L, Wu L, Ni D, Chang Y, Xu W, Xing K. The cDNA cloning and mRNA expression of heat shock protein 70 gene in the haemocytes of bay scallop (*Argopecten irradians*, lamarck 1819) responding to bacteria challenge and naphthalin stress. Fish Shell Immunol 2006;21(4):335-45.

Stein M, Keshav S, Harris N, Gordon S. Interleukin-4 potently enhances murine macrophage mannose receptor activity - a marker of alternative immunological macrophage activation. J Exp Med 1992;176(1):287-92.

Tapia E, Montes C, Rebufel P, Paradela A, Prieto H, Arenas G. Expression of an optimized *Argopecten purpuratus* antimicrobial peptide in *E. coli* and evaluation of the purified recombinant protein by in vitro challenges against important plant fungi. Peptides 2011;32(9):1909-16.

Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: A matter of depth. Genome Res 2011;21(12):2213-23.

Taylor PR, Martinez-Pomares L, Stacey M, Lin HH, Brown GD, Gordon S. Macrophage receptors and immune recognition. Annu Rev Immunol 2005;23:901-44.

Torkildsen L, Coyne R, Samuelsen OB, Magnesen T, Bergh O. Treatment of the early life stages of scallop (*Pecten maximus*) with antimicrobial agents; searching for an alternative to chloramphenicol. Aquaculture Int 2002;10(5):399-409.

Torkildsen L, Lambert C, Nylund A, Magnesen T, Bergh O. Bacteria associated with early life stages of the great scallop, *Pecten maximus*: impact on larval survival. Aquaculture Int 2005;13(6):575-92.

Torkildsen L, Samuelsen OB, Lunestad BT, Bergh O. Minimum inhibitory concentrations of chloramphenicol, florfenicol, trimethoprim/sulfadiazine and flumequine in seawater of bacteria associated with scallops (*Pecten maximus*) larvae. Aquaculture 2000;185(1-2):1-12.

Toubiana M, Gerdol M, Rosani U, Pallavicini A, Venier P, Roch P. Toll-like receptors and MyD88 adaptors in *Mytilus*: Complete cds and gene expression levels. Dev Comp Immunol 2013;40(2):158-66.

Trexler M, Bányai L, Patthy L. The LCCL module. Eur J Biochem 2000; 267(18):5751-7.

Venier P, De Pitta C, Bernante F, Varotto L, De Nardi B, Bovo G, Roch P, Novoa B, Figueras A, Pallavicini A, Lanfranchi G. MytiBase: A knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. Bmc Genomics 2009;10:72.

Venier P, Varotto L, Rosani U, Millino C, Celegato B, Bernante F, Lanfranchi G, Novoa B, Roch P, Figueras A, Pallavicini A. Insights into the innate immunity of the mediterranean mussel *Mytilus galloprovincialis*. Bmc Genomics 2011;12:69.

Vera M, Martinez P, Poisa-Beiro L, Figueras A, Novoa B. Genomic organization, molecular diversification, and evolution of antimicrobial peptide myticin-C genes in the mussel (*Mytilus galloprovincialis*). Plos One 2011;6(8):e24041.

Wan Q, Whang I, Lee J, Lee J. Novel omega glutathione S-transferases in disk abalone: Characterization and protective roles against environmental stress. Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 2009;150(4):558-68.

Wang A, Wang Y, Gu Z, Li S, Shi Y, Guo X. Development of expressed sequence tags from the pearl oyster, *Pinctada martensii* dunker. Mar Biotechnol 2011;13(2):275-83.

Wang L, Qiu L, Zhou Z, Song L. Research progress on the mollusc immunity in China. Dev Comp Immunol 2013;39(1-2):2-10.

Wang M, Yang J, Zhou Z, Qiu L, Wang L, Zhang H et al. A primitive toll-like receptor signaling pathway in mollusk zhikong scallop *Chlamys farreri*. Dev Comp Immunol 2011;35(4):511-20.

Wang S, Hou R, Bao Z, Du H, He Y, Su H et al. Transcriptome sequencing of zhikong scallop (*Chlamys farreri*) and comparative transcriptomic yesso scallop (*Patinopecten yessoensis*). Plos One 2013;8(5):e63927.

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F et al. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 2012;490(7418):49-54.

Zhang L, Zhao J, Li C, Su X, Chen A, Li T et al. Cloning and characterization of allograft inflammatory factor-1 (AIF-1) from manila clam *Venerupis philippinarum*. Fish Shell Immunol 2011;30(1):148-53.

Zhang Y, He X, Yu F, Xiang Z, Li J, Thorpe KL et al. Characteristic and Functional Analysis of Toll-like Receptors (TLRs) in the lophotrocozoan, *Crassostrea gigas*, Reveals Ancient Origin of TLR-Mediated Innate Immunity. PLoS One 2013;8(10):e76464.

Zhang Y, Li J, Yu F, He X, Yu Z. Allograft inflammatory factor-1 stimulates hemocyte immune activation by enhancing phagocytosis and expression of inflammatory cytokines in *Crassostrea gigas*. Fish Shell Immunol 2013;34(5):1071-7.

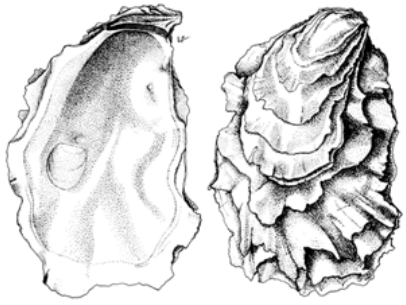
Zhao J, Li C, Chen A, Li L, Su X, Li T. Molecular characterization of a novel big defensin from clam *Venerupis philippinarum*. Plos One 2010;5(10):e13480.

Zhao J, Song L, Li C, Ni D, Wu L, Zhu L et al. Molecular cloning, expression of a big defensin gene from bay scallop *Argopecten irradians* and the antimicrobial activity of its recombinant protein. Mol Immunol 2007;44(4):360-8.

Zhou J, Li CH, Li Y, Su XR, Li TW. cDNA cloning and mRNA expression of heat shock protein 70 gene in blood clam *Tegillarca granosa* against heavy metals challenge. Afr J Biotechnol. 2013(2):2341-2352.

Section C

“RNA interference targeting IkB2 to investigate the *Crassostrea gigas* immune response against oysters herpesvirus type I”



1. Background

Since the mid-1970s, large-scale episodic events such as disease epidemics, mass mortalities, and harmful algal blooms and other population explosions have been occurring in marine environments at a historically unprecedented rate (*Harvell et al. 1999, 2002, Lafferty et al. 2004, Mydlarz et al. 2006*). Ecologically and economically important invertebrates have been affected by large-scale mass mortalities, generally related to infectious diseases (*Mydlarz et al. 2006*). The case of summer mortality syndrome of *Crassostrea gigas* is emblematic. Significant mortality in Pacific oyster has been experienced during the summer months in several countries (e.g. France, United States, Australia and Ireland), and is a major concern for oyster aquaculture. Notably, in France a multidisciplinary approach has been adopted to assess the causes of *C. gigas* summer mortality and to propose solutions that maximize survival, leading to the hypothesis that this phenomenon is multifactorial since it results from a complex interaction between environment, oysters and pathogens (*Samain & McCombie 2008*).

Looking at the environmental factors, hypoxia/anoxia, high concentration of hydrogen sulfur (H_2S), supposedly due to the presence of anaerobic bacteria in sediments, and increase in water temperature have been shown to affect both the growth and the survival of oysters (*Le Moullac et al., 2007b, Le Moullac et al., 2007a*). Most of all, the temperature has been reported to affect the physiology of oysters since it could activate the gametogenesis and thus the reproduction. Noteworthy gametogenesis demands high energy resources and considerably weakens the immune system (*Fleury et al., 2010*), seriously decreasing the oysters survival chances. As evidence, a positive correlation between survival rate and reproductive efforts has been suggested (*Huvet et al., 2010*). Besides physiology, also genetic has been shown to play an important role in determining oysters' mortality or survival (*Degremont et al., 2005, 2007, 2010*). A high heritability was estimated for resistance to summer mortality, which provided an opportunity to develop lines of oysters that were resistant (R) or susceptible (S) to the syndrome. Recently, to investigate the molecular bases underneath the host resistance to summer mortality, a genome-wide expression profiling study of R and S gonads, gills and muscle has been carried out through a microarray platform (*Fleury et al., 2010; Fleury and Huvet, 2012*). Among the 9,058 unigenes analyzed, a great number appeared to be differentially expressed between R and S "lines" during the 3-months period preceding a summer mortality outbreak. Interestingly, the gene ontology analysis of these genes revealed that reproduction allocation, energy metabolism, antioxidant defense and

immune system are constitutive pathways that operate differentially between R and S oysters. In addition, a significant over-representation of immune-related genes at the date preceding the mortality lets the authors thinking that immune defense, mainly through the nuclear factor kB (NF-kB) pathway, is one of the main determinant of resistance in the field.

The third interacting factor is pathogens, above all *Vibrio spp* and herpesvirus. Vibrio genera are commensal bacteria of the digestive tract in oysters (*Colwell and Liston, 1960*) but it has been demonstrated that, in the presence of environmental or physiological stressing conditions, *V. splendidus* and *V. aestuarianus* are two important causative agents of the summer mortality syndrome (*Lacoste et al. 2001; Le Roux et al. 2002; Waechter et al. 2002; Gómez-León et al., 2005; Garnier et al. 2007*). While the relevance of *Vibrio spp* in bivalves mortality outbreaks has been always recognized, the role of herpesvirus appeared crucial since spring 2008. Until this year it was generally accepted that pathogens were mainly opportunistic giving more importance to host and environment in the interaction, high temperature and reproduction stage having the most important causative effects in *C. gigas* summer mortality. Then, in 2008, evidences demonstrated that mortality occurred below the previously established temperature threshold (*Pernet et al., 2012*) and it has been hypothesized that Oyster herpesvirus type I (OsHV1, *Renault et al., 1994*) was most likely the main triggering factor as almost always observed in summer outbreaks. Consistent with this hypothesis recent studies identified a new virus variant, OsHV1µ variant, and recognized it as responsible of several mortality episodes (*Segarra et al., 2010, Cochennec-Laureau et al., 2011*). Nevertheless, the identification of implicated factors still remains an open issue and the complexity of this phenomenon is attested by recent studies suggesting that the principal cause of oyster mass mortality may consist on a co-infection of V. splendidus and OsHV-1, rather than a single pathogen infection (*Dégremont, 2011; Pernet et al, 2012*).In invertebrates, such as oysters, the defense against infectious microorganisms is guided by the innate immune system, which relies on basic mechanisms, conserved throughout much of the animal kingdom (*Hoffmann et al., 1999*), of pathogen recognition and activation of the response. Recent studies have revealed striking similarities in the signaling pathway used by mammals and molluscs to activate their innate immune responses (e.g. *Montagnani et al., 2004, Green and Barnes, 2009; Koutsogiannaki and Kaloyianni, 2010; De Zoysa et al., 2010*). In both cases, infection leads to the induction of the NF-kB pathway, in which the most significant feature is the central role of the

Rel/NF-kB family of transcriptional activator proteins. The Rel/NF-kB family of transcription factors is composed of a set of structurally related and evolutionarily conserved DNA binding proteins (*Baldwin, 1996*) that are involved not only in immune and inflammatory responses but also in the control of a large number of cellular processes such as development, cellular growth and apoptosis (*Hayden et al., 2004*). Since they are evolutionary conserved, several components of the Rel/NF-kB signaling pathway have been cloned and characterized from pearl oysters (*Xiong et al., 2008; Zhang et al., 2009*), abalone (*Jiang et al., 2007*) and squid, which strongly suggest the presence of the Rel/NF-LB pathway in molluscs.

In most cell types, Rel/NF-kB transcription complexes are present as latent, cytoplasmic forms that can be induced to enter the nucleus and activate gene expression (*Ghosh et al., 1998*). The cytoplasmic sequestration of Rel/NF-kB is regulated by a family of IkB inhibitors (*Thompson et al., 1995; Whiteside et al., 1997*). This interaction blocks the ability of NF-kB to bind to DNA and results in the NF-kB complex being primarily in the cytoplasm due to a strong nuclear export signal in IkB. Many of the signals known to activate Rel/NF-kB also lead to the activation of Toll-like receptors, which in turn initiate intra-cellular signaling cascades that culminate in phosphorylation, ubiquitination and degradation of IkB (Figure 1). Serine kinases termed IKKs (IkB kinases) have been shown to phosphorylate IkB in response to signals known to activate the NF-kB cascade (*Régnier et al., 1997; Woronicz et al., 1997; Cohen et al., 1998; Lu et al., 2001*). Once released, NF-kB is located preferentially into the nucleus and activates target genes through binding to 10 bp DNA sites (kB sites) as dimers (*Chen and Ghosh, 1999*). All the Rel/NF-kB proteins are related through a highly conserved 300 amino acid region called the Rel homology domain (RHD) that is involved in DNA binding and dimerization (*Ghosh et al., 1998*). The RHD also contains the nuclear localization site (NLS) that consists of a stretch of four or five basic residues and is responsible for the nuclear translocation of the protein. Further, Rel/NF-kB proteins can be divided into two classes based on sequences C-terminal to the RHD. Members of class I (p100 and p105 in mammals and Relish in *Drosophila*) have long C-terminal domains that contain multiple copies of ankyrin repeats (similar to those present in IkB proteins) which act to inhibit themselves. Thus, members of this class are generally not activators of transcription, except when they form dimers with members of the second class of Rel/NF-kB transcription factors. Members of class II (cRel, RelA, RelB in mammals; Dorsal, Dif in *Drosophila*) contain C-terminal activation domains, which are often not conserved at the

primary sequence level across species, even though they can activate transcription. In its most basic form, therefore, the pathway consists of receptor and receptor proximal signaling adaptor molecules, the IKK complex, IκB proteins and NF-κB dimers.

As concern the Pacific oyster *C. gigas*, several conserved genes from the Rel/NF-kB signaling pathway have been characterized and in Figure 1 similarities of oyster Rel/NF-kB pathway with vertebrates and Drosophila are highlighted. It has been demonstrated that two proteins, an IkB kinase-like (oIKK) and a Rel homolog (Cg-Rel1), shared structural and functional properties with their mammalian and insect counterparts (*Escoubas et al., 1999; Montagnani et al., 2004*). Moreover, the presence of two IkB homologs (Cg-IkB1 and Cg-IkB2) and their ability to inhibit NF-kB/Rel transcriptional activity has been reported (*Zhang et al., 2011*). In addition, Green et co-workers (*Green et al., 2014*) recently evaluated the immune response in circulating hemocytes from poly:IC-injected oysters and, by means of microarray, they demonstrated that oysters rely on a cellular response to minimize viral replication, involving recognition of virus-associated molecular patterns to induce host cells into a Rel/NF-kB-mediated antiviral state. Notably, in this study the mRNA expression of a Toll-like receptor, MyD88, IkB1 and Rel, was significantly induced.

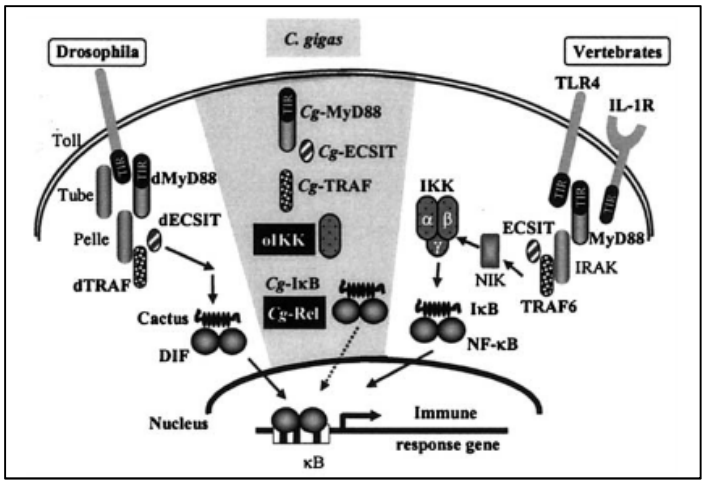


Figure 1. *In vertebrates and in Drosophila the Rel/NF-kB pathways are conserved at the molecular level from transmembrane receptors (Toll, TLR for Toll-Like receptor, IL-1R for interleukin-1 receptor) to Rel proteins (Dif, NF-kB). Ligand binding to TLR leads to the activation of adapter proteins (MyD88, Tube) then to the activation of kinases (Pelle, IRAK, IKK) and to the phosphorylation and degradation of Rel inhibitory proteins (cactus, IkB). The phosphorylated inhibitors are degraded to release Rel transcription factors. The different components characterized in oyster are shown (Cg-MyD88, Cg-TRAF, Cg-ECSIT, oIKK and Cg-IkB). DD, death domain; KD, kinase domain; TIR, Toll/IL-1 receptor homology domain. The figure was reported in Montagnani et al., 2004.*

Despite the Rel/NF-kB family and its regulation is an important field of research, we still have a very little understanding of the complex in vivo dynamics of this pathway in molluscs.

To date, biochemical and molecular approaches have been employed to unravel the Rel/NF-kB pathway functioning and, more in general the mechanisms driving the bivalves immune response. In the last decade, reverse genetic and "omic" tools, most of all transcriptomic, have been widely exploited to investigate bivalve immunology. A promising technique which could help in understanding the specific functions of Rel/NF-kB proteins is the RNA interference (RNAi). RNAi is a highly evolutionally conserved process of post-transcriptional gene silencing by which double-stranded RNA (dsRNA), when introduced into a cell, causes sequence-specific degradation of homologous mRNA sequences (*Almeida et al., 2005*). It was firstly described in plants and *Caenorhabditis elegans* (*Jorgensen et al., 1990*; *Fire et al., 1998*) but later also recognized as process existing in fungi and animals (*Elbashir et al., 2001; Romano et al., 1992*). Following these observations, RNAi-based applications have generated substantial enthusiasm and became a powerful tool for reverse genetic studies. Since classic functional genetic approaches such as mutagenesis are not yet available for bivalve molluscs, RNAi has emerged as a powerful alternative technique for determining the loss-of-function phenotype of a gene (*Elbashir et al., 2001*). Despite that, in molluscs, this technique is still scarcely used and it remains a technical challenge (*Fabioux et al., 2009; Suzuki et al., 2009*). Conversely, RNAi has been widely used *in vitro* and *in vivo* in vertebrate and invertebrate species (*Berns et al., 2004; Robalino et al., 2004; Dash et al., 2008*).

In this study, based on the recent development of RNA interference in *C. gigas* (*Fabioux et al.2009*; *Huvet et al., 2012*), dsRNA targeting the Cg-IkB2 gene was injected *in vivo* into oysters following a OsHV-1 challenge, survival being of the main knock-down phenotype of interest in the present study. The real-time PCR technique was employed to evaluate the expression levels of IkB2 itself and three important immune-related transcripts, Cg-IkB1, Cg-Rel1 and a Cg-SOCS (Suppressor of cytokines signaling) in gonads and gills.

2. Methods

2.1. Biological Material

Healthy naïve oyster spat (i.e. free of mortality, negative for OsHV-1 and Vibrio sp) were produced in May 2012 at the IFREMER (French Research Institute for Exploitation of the Sea) hatchery in Argenton (Finistère, France). The progeny was cultured at the Ifremer station in Bouin (Poitou-Charentes, France) as described in Petton et al. (2013). At the end of November 2012, 900 oysters (with a mean weight of 4 g) were transferred to the IFREMER station in Brest (Finistère, France) and they were placed in 20-µm-filtered seawater at 20°C, until the experiment begun (4 days later).

2.2. dsRNA synthesis

The procedure was similar to that described by Fabioux et al. (2009). A pool of RNA from different oyster tissues was used as a template in a RT-PCR reaction to obtain a cDNA sample. Starting from this cDNA, a PCR reaction allowed to amplify a fragment from position 144 to 456 of Cg-IkB2 cDNA (Genbank: HQ650768). The primer used to obtain the 312 bp fragment were the following: forward: 5'-GGATTTGAACGACCTGGAAG-3', reverse: 5'-GCAGACGACTCGTTTTCATC-3'. The fragment was then sub-cloned into pCR4-TOPO vector (Invitrogen) which was subsequently used to transform chemically competent *Escherichia coli* cells (TOP10 One Shot® competent cells, Invitrogen). Bacteria have been plated in LB solid culture medium containing ampicillin in a final concentration of 75 µg/mL and incubated overnight at 37°C. The following day, colonies were picked, transferred in LB liquid culture medium (amplicillin 75 µg/mL) and incubated overnight at 37°C. The recombinant plasmid was extracted and purified from bacteria by using the Plasmid DNA purification Kit NucleoBond®Xtra Midi (Macherey-Nagel) and the DNA concentration and purity were estimated through the Nanodrop ND -1000 spectrophotometer (Thermo Fisher Scientific). In order to verify the cloning success the plasmid was digested with *EcoRI* (4 hours at 37 °C). The digestion products were analyzed by agarose 1% gel electrophoresis. The plasmid was also sequenced to confirm the effective insertion of the 312 bp fragment of Cg-IkB2 coding sequence. In order to prepare the templates for the fragment transcription, the plasmid was linearized with *Not*I or *Spe*I enzymes (Promega, Madison, WI), then phenol-chloroform extracted and finally ethanol-precipitated and

suspended in RNAse-free water. The linearization products were verified by agarose gel electrophoresis.

Both strands of the purified plasmid were transcribed *in vitro* by using T7 and T3 MEGAscript® Kits (Ambion, Austin, TX, USA) to produce Cg-IkB2 sense and antisense single-stranded RNA (ssRNA). The two strands were phenol-chloroform extracted, ethanol-precipitated and resuspended in RNAse water. In order to obtain a solution of dsRNA 2 µg/µl, equimolar amounts of sense and antisense ssRNA were heated to 100°C for 1 minute and left to cool at room temperature for 10 hours for annealing. Finally the agarose gel electrophoresis allowed ensuring that the synthesized dsRNA corresponded of a single band of 312 bp.

2.3. dsRNA administration and sampling

Oysters were anesthetized in $MgCl_2$ solution (3/5 fresh water 2/5 seawater with 50 g/L $MgCl_2$) overnight (*Suquet et al., 2009*). Anesthetized oyster were injected at T0, either with Cg-IkB2 dsRNA (1 µg for each gram of oyster mean weight) diluted in 100 µl Tris-NaCl saline solution (n=186) or with the same volume of saline solution for the control (n=186). Half of the solution volume (i.e. 50µL) was injected in the gonad and half in the adductor muscle sinus, one of the most important sinuses (*Cheng, 1981*), expecting targeting the circulating haemolymph. After injection, oysters were placed in 30-liters tanks (3 tanks for each condition) containing 1 µm-filtered sea water. Rearing conditions were the one described in *Schikorski et al.* (2011). At 5 and 24 hours after dsRNA injection, dsRNA injected (N=15) and control animals (N=15) were sampled. A solution containing 10^5/µL of herpesvirus OsHV-1 was prepared by dissecting and pooling tissues from naturally infected oysters, as described in *Schikorski et al.* (2011). Twenty-eight h after the dsRNA injection, the shells of the remaining oysters were grinded and animals were intramuscularly injected with 100 µL of a solution containing 10^4/uL particles of herpesvirus OsHV-1. Control oysters were injected with 100 µL of sterile sea water (SSW). Two additional sampling (N=15 in each of the 2 following conditions, dsRNA+OsHV-1 and TrisNaCl+OsHV-1) were performed at 20 and 44 hours post-infection, corresponding to 48 and 72 hours after dsRNA injection, respectively. After each sampling, gonads and gills were immediately dissected, froze and store in liquid nitrogen until subsequent total RNA extraction.

During the dissection at 72 hours, a piece of mantle was also dissected for each collected animal and conserved in ethanol for further DNA extraction and assay, through real-time PCR, of the number of viral copies in each oyster.

Mortality has been visually estimated two times per day (early morning and late afternoon) overall the experiment (10 days). The visual parameter used to determine the mortality was the oysters' ability to close valves once they were removed from the water. Dead/moribund oysters were dissected and a piece of mantle was collected and conserved in ethanol for subsequent DNA extraction and viral load assay.

2.4. RNA isolation and cDNA synthesis

The gonad and gills tissues were individually crushed to a fine powder at 196°C with a Dangoumau mill and about 30 mg of powder was included in 2 ml tubes containing 1,5 ml of Extract-all solution (Eurobio). Total RNA of each sample was isolated according to the manufacturer's instructions. Total RNAs extracted from gonads required an additional precipitation step with sodium acetate in order to purify them from contaminant substances adsorbing at 230 nm. Then, all the samples were treated with RTS DNAse Kit (MO-BIO) according to the manufacturer's instructions. RNAs concentration and quality were measured using a Nanodrop ND -1000 spectrophotometer (Thermo Fisher Scientific) at 260 nm and RNA integrity was assessed through RNA nanochips and Agilent RNA 6000 nanoreagents (Agilent). In order to verify the success of the DNAse treatment (i.e. the absence of DNA carryover), RNA samples were all diluted 1:10 and analyzed in real time PCR by using a couple of primers reported in *Fabioux et al., 2004* and specifically amplifying the elongation factor 1 (Cg-EF1) (see below for real time PCR protocol).

The first-stand synthesis was carried out following the Sigma MMLV (Moloney Murine Leukemia Virus) RT usage recommendations and using the DNAse-treated RNA as template and oligo-dT as primers (as described in *Fabioux et al., 2004*). RNA samples were denatured at 70°C for 10 minutes and retro-transcription was performed at the following thermal conditions: 25°C for 15 min, 42°C for 50 min and 94°C for 5 min.

2.5. Quantification of immune-related gene expression

Real-time RT-PCR was performed with the iQ™ SYBR® Green Supermix (Biorad) on a MyIQ2 Biorad thermal cycler to investigate the expression of immune-related genes, including Cg-IkB1, Cg-IkB2, suppressor of cytokine signaling (Cg-SOCS) and the transcription factor Rel1 (Cg-Rel1) (protocol in *Huvet et al., 2004*). The PCR efficiency

(E) for each primer pair was determined by performing standard curves from serial dilutions to ensure that E ranged from 98% to 102%. The primer sequences used in this assay are shown in Table 1. Relative mRNA levels of the target genes in gonads were calculated by using the comparative Ct method (*Livak and Schmittgen, 2001*) and normalized to Cg-EF1, as no significant differences of Ct values were observed for Cg-EF1 between dsRNA and Tris-NaCl injected oysters over time (One-way Anova: $F=0.59$, $p>0.05$). Whit regard to Cg-EF1 expression values in gills, the variance analysis demonstrated a significant variation between groups (One-way Anova: $F=3.8$, $p<0.05$). Despite that, also in gills the target genes expression was normalized to Cg-EF1, since the post hoc test Tukey HSD showed that differences were restricted to the comparison among oyster collected at 5 and 72 hours post injection, a contrast which is scarcely relevant, considering the goal of the study. All data were given in terms of relative mRNA expression level (Relative Quantification = RQ) as means ± S.D (n=15)

Primer	Sequence	Length	Efficency (%)
Cg-IkB1 Forward	5'-GATATCGCCCTGATCTTGCT-3'	20	99.3
Cg-IkB1 Reverse	5'-AGGTTGGCTCCTGACATCAC-3'	20	
Cg-IkB2 Forward	5'-CAGCATTCACTGACGACGAT -3'	20	98.2
Cg-IkB2 Reverse	5'-TCTGCCTCAGTTTGTCGTTG -3'	20	
Cg-SOCS Forward	5'-ATCAGCCGATTCATCCTCAG -3'	20	100
Cg-SOCS Reverse	5'-TGCTGGAATGTGTAGGCAAC -3'	20	
Cg-Rel1 Forward	5'-GGTTAGACGAGACAAAGACAG -3'	21	98.9
Cg-Rel1 Reverse	5'- GCATCCAGTGAGGAAATGA-3'	19	
Cg-EF1 Forward	5'-GATTGCCACACTGCTCACAT-3'	20	100
Cg-EF1 Reverse	5'-AGCATCTCCGTTCTTGATGC-3'	20	

Table1. *Primers used for the real-time PCR amplification*

2.6. Detection and quantification of OsHV-1

Total DNA was extracted from mantle tissues for real-time PCR assay analyses as described in *Schikorski et al.* (2011). The detection and quantification of OsHV-1 DNA was carried out in triplicate using real-time PCR according to *Pepin et al.* (2008), with specific primers that amplified both reference and µVar types. Briefly, absolute quantification of OsHV-1 DNA copies was carried out by comparing the Ct values obtained for tested samples with a standard curve based on a ten-fold dilution of a stock solution of OsHV-1 genomic DNA (5 × 106 copies/µL) extracted from purified virus particles. The results were expressed as virus DNA copy numbers per ng total DNA. Standardization of OsHV-1 genomic DNA quantity was reached using DNA concentrations measured using an ND-1000 spectrophotometer (Nanodrop Technologies) at 260 nm with the conversion factor of 1 OD = 50 µg/mL DNA (*Schikorski et al., 2011*).

2.7. Statistical analyses

In order to evaluate the survival rate at the different tested conditions, Cox proportional hazards model have been implemented with the statistical software STATGRAPHICS Centurion XVI v. 16.2.03.

Following the logarithmic transformation of the RQ values, a variance analysis (two way ANOVA) was performed in RGui by using a linear mixed model where “time” and “condition” are fixed effects and “tank” is a random effect. Normality and variance of residuals were checked (see Appendix C1 for gonads and Appendix C2 for gills). After ANOVA, a Tukey’s HSD multi comparison test (Confidence interval: 95%; p-value<0.05) was performed to evaluate differences i) between controls and dsRNA-injected at each sampling time and ii) between sampling times within each condition.

As concern the viral load, the estimates of OsHV-1 DNA copies were highly variable, and followed an asymmetric distribution that appeared to be neither normal nor Poisson, even after classical normalization transformations. Values were consequently log-transformed and a non-parametric Kolmogorov-Smirnov test was performed in order to evaluate differences between groups. In addition the viral load was compared to the gene expression data through the assessment of Spearman's rank correlation coefficients. Both statistical analyses were carried out by using the software STATGRAPHICS Centurion XVI v. 16.2.03.

3. Results

3.1. Evaluation of mortality

In infected oysters previously injected with the control solution TrisNaCl, mortality appeared at 44h after the viral infection. At 68h after viral infection the survival probability was 0.71 and at the end of the experiment (260h after viral infection), when the number of dead oysters in each tank was 28, 27 and 24 respectively, the survival probability was less than 2%. Conversely, in the experimental group treated with the dsRNA a total of four oysters died and the survival probability at the end of the experiment was nearly 90%. As concern controls animals (dsRNA and TrisNaCl oysters injected with sterile seawater, SSW), a single mortality event was registered at 128h post infection in the tank containing oysters injected with TrisNaCl/SSW. Figure 2 represents the survival rate calculated through the Cox proportional hazards model.

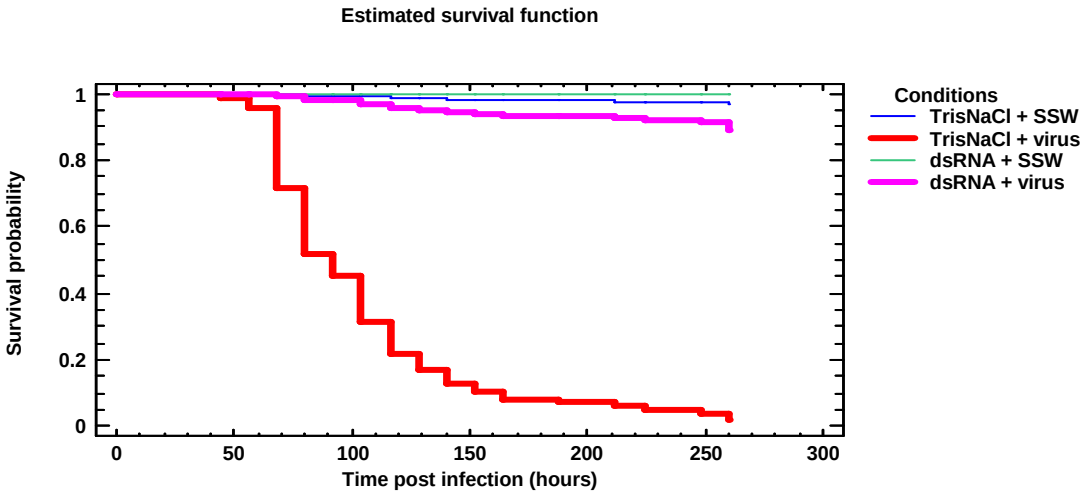


Figure 2*. Cox proportional hazards model of the survival distribution obtained each day in the three tanks used per treatment, for the dsRNA and TrisNaCl groups injected either with the herpesvirus OsHV-1 or Sterile SeaWater (SSW). In the x axis survival time is expressed in hours post Os-HV1 or sterile seawater (SSW) injection.*

3.2. Gene expression

As concern the oysters treated with the IkB2 dsRNA (grey bars in Figure 3B), the expression levels of IkB2 didn't show any significant variation in gonads neither after the dsRNA injection, neither after the herpesvirus injection. The same trend was reported for IkB1, Rel1 and SOCS transcripts, for which no differences appeared over time. Conversely, in the animals treated with the control solution TrisNaCl (black bars in Figure 3), significant changes were reported over the time course of the sampling. As concern IkB1, the mRNA level exhibited a slight decrease, with a Fold Change (FC) of

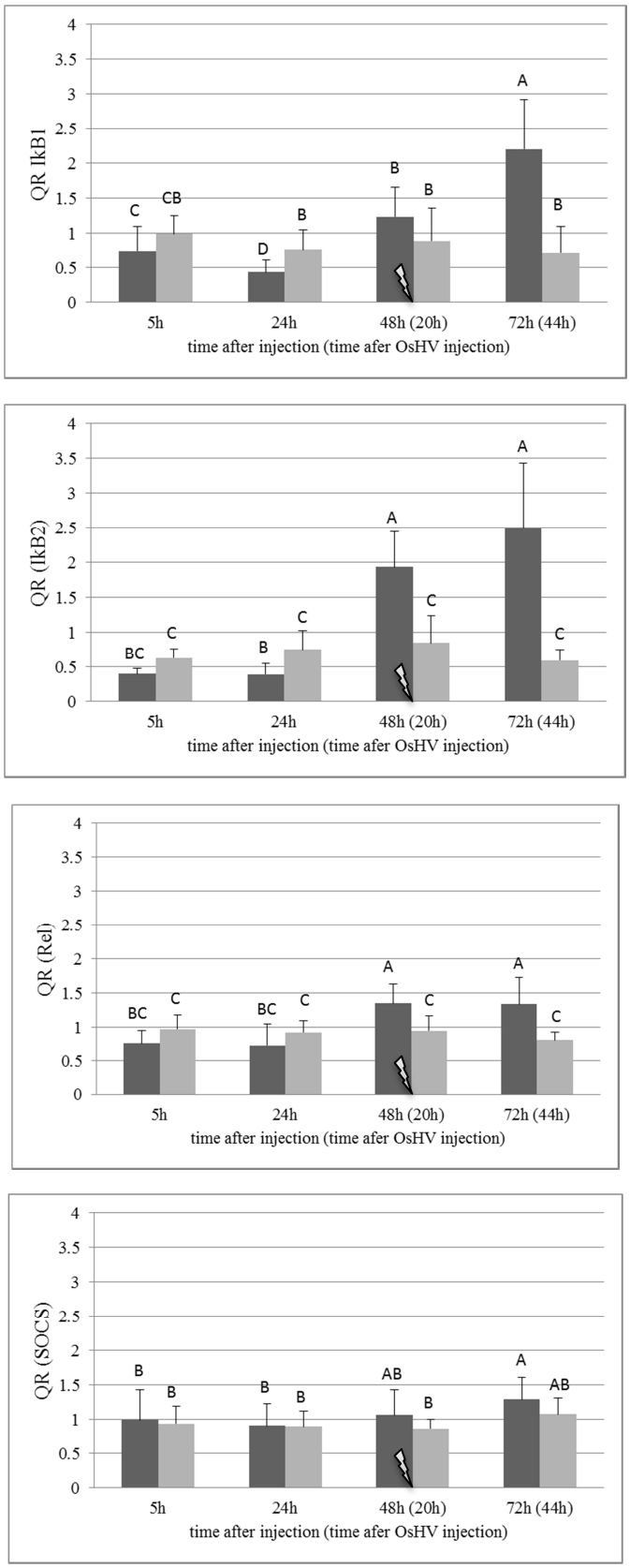


Figure 3. *Relative expression of IkB1 (A), IkB2 (B), Rel1 (C) and SOCS (D) in gonads. Significance evaluated through two way mixed model ANOVA and Tukey's HSD multicomparison test. Black bars represent the oysters injected with TrisNaCl and grey bars represent oysters injected with dsRNA. A common letter for two groups means that post-hoc tests detected no*

significant differences for gene expression levels between them. The lightning symbol indicates the beginning of the viral challenge.

1.72, at 24h after the injection of the control solution TrisNaCl and a substantial increment in the two times following the virus injection. IkB2 showed, at 20h after the virus infection, an

increase in the expression level (FC:4.95) which remained stable at 44h. A similar behavior, but less evident (FC:1.85), was observed for Rel1.

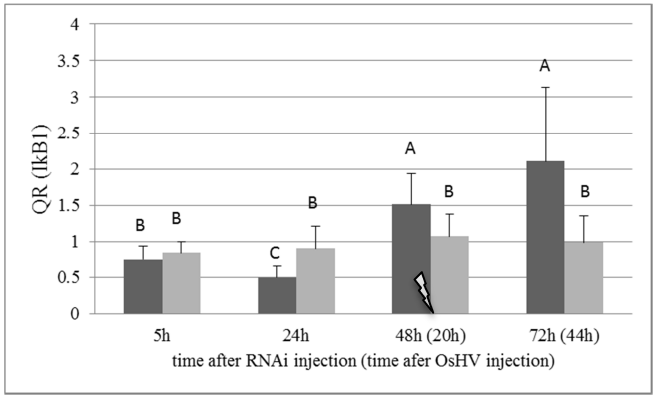
For SOCS gene, the expression appeared up-regulated later, at 44h after the virus injection (statistical significance for the comparison between 24h and 72h after the beginning of the experiment).

Looking at the comparison between the mRNA levels reported in the two conditions at each sampling time, interesting results were evidenced. While at 5h post injection no differences occurred, at 24h post injection the quantity of transcript encoding IkB1 and IkB2 was statistically higher in the gonads of the group of oysters treated with the dsRNA. Notably, once the TrisNaCl-injected oysters were exposed to the herpesvirus, they showed a higher expression of all the target genes, despite a statistical significance was reported only for IkB2 and Rel1 at 20h post infection and for all genes except for SOCS at the last sampling time.

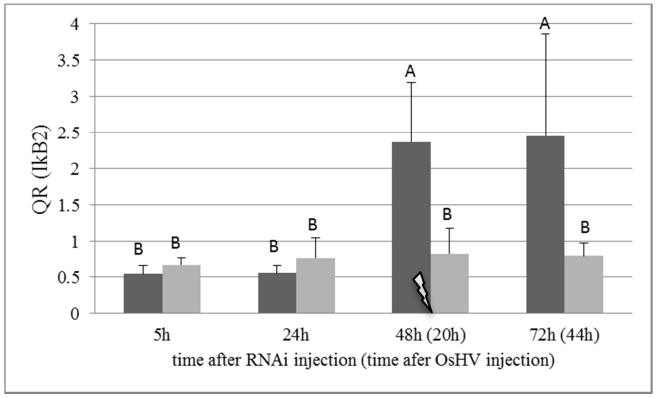
In gills, the 4 target genes showed variation of mRNA levels similar to that seen in gonads. Oysters treated with the IkB2 dsRNA (grey bars in Figure 4), exhibited a slightly significant increase in the expression of Rel1 at 24h after the dsRNA injection. The behavior of the other 3 genes remained stable over time. Instead, the animals treated with the control solution TrisNaCl (black bars in Figure 4), evidenced significant changes in the expression level of all the genes evaluated. IkB1, similarly to what observed in gonads, was slightly down-regulated (FC:1.52) at 24h after the injection of the control solution and considerably up-regulated (FC:3.05) at 20h after the virus injection. Like IkB1 but at a greater extent (FC:4.31), IkB2 was induced after the virus infection and the increment remained stable at 44h post challenge. A similar behavior, but less evident, was observed for SOCS for which a FC of 1.55 was reported. Notably, if compared to what seen in gonads, the Rel1 transcript seemed to have a different regulation: in fact it was 2.76 times more expressed at 20h after the virus injection but its mRNA level immediately decreased at the last sampling time. As concern IkB2, Rel1 and SOCS, any significant variations were recorded before the virus challenge.

With regard to gene expression analyzed at each sampling time, differences between the two experimental groups were reported. While at 5h post injection no significant

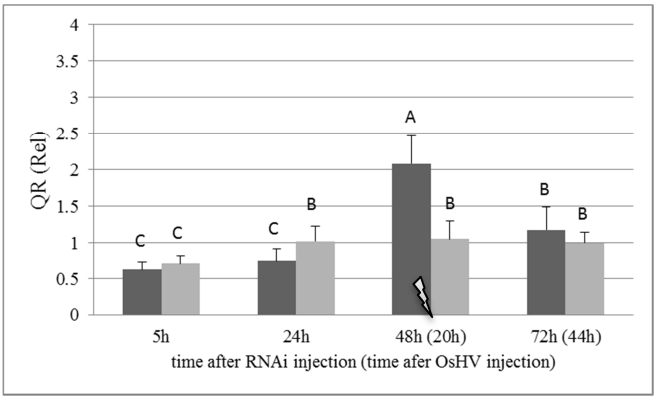
differences were evidenced between the two tested conditions, at 24h the expression of all the target genes was lower in the control group (black bars in Figure 3), but statistically significant only for IkB1 and Rel1. As seen in the gonads, the greater changes occurred after OsHV-1 infection. In fact, in the control animals treated with TrisNaCl, the amount of mRNAs of all the four genes was statistically higher at 20h after the challenge, with a greater difference reported for IkB2 (FC:2.86). At 44h post infection, differences in the expression level of SOCS remained stable, while the variations observed for IkB1 and IkB2 became greater, with a fold change of 2.13 and 3.07, respectively. As concern Rel1 expression at 44h post infection, any statistical variation was highlighted.



A



B



C

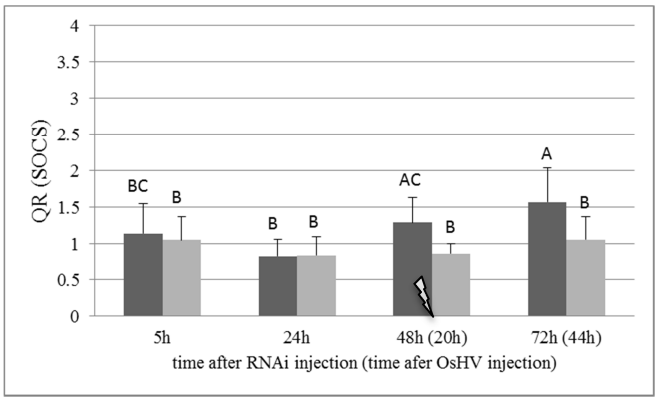


Figure 4. *Relative expression of IkB1 (A), IkB2 (B), Rel1 (C) and SOCS (D) in gills. Significance evaluated through two way mixed model ANOVA and Tukey's HSD multicomparison test. Black bars represent the oysters injected with TrisNaCl and grey bars represent oysters injected with dsRNA. A common letter for two groups means that post-hoc tests detected no significant differences for gene expression levels between them. The lightning symbol indicates the beginning of the viral challenge.*

3.3. Viral load

At 20h post infection the estimated log number of viral genome per ng of total DNA extracted from mantle was 4.71±0.74 and 1.26±1.38 in oysters injected with TrisNaCl (controls) and dsRNA (treated), respectively; a Kolmogorov-Smirnov test demonstrated that this difference was highly significant (K-S test= 0.92; pval < 0.001). At 44h the viral load was 5.90±0.46 and 1.61±0.62 in control and treated oysters, respectively, thus the estimated difference remained high and significant (K-S test= 1; pval < 0.001). Moreover the viral load was evaluated also in the oysters which died during the experiment, demonstrating that the log number of virus particles was significantly higher in the mantle of died control animals in comparison with those reported in died oysters treated with dsRNA (K-S test= 0.88; pval < 0.001).

With regard to the correlation between virus load and gene expression, correlation coefficients are reported in Table 2.

	Gonads				***Gills***			
	IkB1	**IkB2**	**Rel**	**SOCS**	**IkB1**	**IkB2**	**Rel**	**SOCS**
Spearman's rank correlation coefficient	0.82 ***	0.71 ***	0.67 ***	0.54 ***	0.67 ***	0.67 ***	0.33 *	0.58 ***

Table 2. *Spearman test describing the correlation between virus load (measured in mantle) and gene expression (measure in gonads and gills) of the four target genes. *** p-value<0.001; * p-value<0.5.*

4. Discussion

In the present study, it has been reported a RNAi experiment targeting IkB2, an important gene though to play a pivotal role in the Pacific oyster immune defense, and its involvement in the response against OsHV-1 infection has been investigated. Since a strong association between reproductive efforts and resistance to summer mortality was previously suggested in oysters (*Fleury and Huvet, 2012; Fleury et al., 2010; Huvet et al., 2010; Samain and McCombie, 2008*), we originally chose gonad as target tissue to administrate the dsRNA and to evaluate variations in the genes expression levels as well. Indeed, gonad was also demonstrated to be a good tissue to target and administrate dsRNA. Moreover, considering the role of IKB2 in immunity (*Zhang et al. 2011*), we also decided to inject dsRNA into adductor muscle sinus, one of the most important sinuses (*Cheng 1981*), expecting to target hemocytes, cells conveyed by hemolymph and having a key role in the immune defense against pathogens (*Bachère et al., 2004; Le Foll et al., 2010; Venier et al., 2011*). Thus, since gills are bathed by hemocytes (*Gonzalez et al., 2007; Seo et al., 2005*), gills have been sampled in this study as target tissue for the gene expression analysis.

In this study we have shown that the injection of a dsRNA specific to Cg-IkB2 mRNA acts to modify the antiviral response against OsHV-1 infection in *C. gigas*, providing a markedly reduced mortality after OsHV-1 injection (Figure 2). Conversely, oysters treated with the control solution TrisNaCl experienced a high mortality and the viral load, evaluated in the mantle of died animals, was significantly more abundant in the control group than in the dsRNA-injected oysters (at 20h post infection: K-S test= 0.92, pval < 0.001; at 44h post infection: K-S test= 1; pval < 0.001). Furthermore, the significant positive correlation reported, in both gonads and gills, between the viral load and the gene expression levels of all the four genes (Table 2) demonstrated that the mRNA level of these four immune genes increased in response to the spread of OsHV-1 infection. These results allowed us to think that the genes evaluated in the present study are more likely involved in the antiviral response in *C. gigas*. In addition, we can also assert that the death occurred as a consequence of the virus infection, as expected following the challenge employed to reproduce OsHV-1 infection in laboratory conditions (*Schikorski et al., 2011*). Despite the objective of this study was to evaluate the role of IkB2 in the response against OsHV-1 infection, the relative quantification obtained for this transcript was not sufficiently informative to provide substantial information regarding the involvement of this gene in the antiviral response in gonads and gills. While a down-

regulation of IKB2 was expected after the dsRNA injection, nevertheless the expression data didn't show an effective silencing of the IkB2 transcript (Figures 2B and 3B), at least at 24h after the dsRNA injection, which is the only sampling time suitable collected to eventually appreciate a variation induced exclusively by the RNAi. It could be still possible that a IkB2 slightly down-regulation occurred but it was transient and restricted to time-points for which gene expression analyses have not been planned, such as in the first five hours or between 5h and 24h post injection. Nonetheless, the IkB2 silencing could have been immediately counteracted through feedback regulation mechanisms, well established in the Rel/NF-kB pathway (*Tieri et al. 2012*), that rapidly restored the targeted transcript expression. In accordance with this hypothesis, the transient silencing of IkB2 could have temporally released the Rel/NF-kB pathway by liberating the Rel1 transcription factor which was then translocated to the nucleus to activate a cascade of genes involved in the antiviral response. In turn, this activation could have determined the small increment of the Rel1 transcript reported in gills after the dsRNA transfection (grey bars in Figure 4C). Despite the transcription factors belonging to the Rel/NF-kB family are mainly regulated at a post-translational level (e.g. by IkB proteins), few studies showed an up-regulation of their mRNA following the activation of the Rel/NF-kB pathway triggered by experimental infections. In *C. gigas* hemocytes, real-time PCR analyses demonstrated an accumulation of Cg-Rel transcript at 3h and 9h after bacterial challenge (*Montagnani et al., 2004*); in scallop *Chlamys farreri* the temporal expression of Rel mRNA in hemocytes increased at 6h, 24h and 48h after LPS stimulation (*Wang et al., 2011*). Despite this evidence, considering that variations observed for the other target genes were not consistent either before or after the challenge, and that dsRNA-injected oysters were resistant to herpesvirus infection, we can suppose that, if a transient activation of Rel/NF-kB occurred, more likely it could have regulated a set of genes (not evaluated in this study) which were involved in the antiviral response.

A second hypothesis which can be proposed is that the IkB2 silencing occurred later than 24h post injection but we couldn't appreciate it since at 48h post injection we performed the virus challenge, which clearly prevented the evaluation of the IkB2 silencing. Notably, the sampling times of the experiment have been decided basing on evidences in *C. elegans*, in which a first reduction of a targeted gene appeared within 24 hours (*Fire et al., 2008*). Although RNAi generally occurs within 24h of transfection, both onset and duration of RNAi is species-specific and depend on the turnover rate of the protein of interest, as well as the rate of dilution and longevity of the dsRNA (*Mocellin and*

Provenzano, 2004). For example, in a RNAi study performed in *Chlamys farreri*, an efficient knock-down of a Cf-TLR (Toll-like receptor) was appreciated at 48h post dsRNA injection (*Wang et al., 2011*). RNAi effectiveness and timing also depend on the mode of administration and the targeted tissue, this is the reason why in the present study both gonad and muscle sinus were injected. Nonetheless following the dsRNA administration, any relevant differences in the Cg-IkB2 mRNA expression between the two targeted tissues have been reported. Going back to the second hypothesis, if we admit that the IkB2 silencing occurred later than 24h post injection, the interpretation of the sampling times after the viral challenge becomes much more complex since the relative quantification outcomes could be the result of two different kind of regulations, one provoked by the gene silencing and one induced by the OsHV-1 infection.

Finally a third hypothesis could be suggested and it is based on the assumption that the IkB2 knock-down failed, nevertheless the injection of a dsRNA induced in oysters an unspecific immune response which conferred resistance to OsHV-1 infection. It is well based in vertebrates that cells recognize dsRNA (viral or synthetic) via toll-like receptor 3, which triggers a complex signal-transduction pathway resulting in the translocation of transcription factors (NF-kB and interferon regulatory factor) to the cell nucleus and the up-regulation of a number of genes, principally interferon. Interferon exerts an antiviral state in neighbouring cells by inducing the expression of antiviral proteins, such as protein kinase R (PKR), 2',5'-oligoadenylate synthetase (OAS), and Myxovirus resistance protein (Mx) (*Randall and Goodbourn, 2008*). With regard to invertebrates, it was a generally accepted paradigm that they can recognize dsRNA as a virus-associated molecular pattern, resulting in the activation of an innate antiviral response, even in the absence of interferon-signalling pathway (*Robalino et al., 2004*). Recently this paradigm was disproved by the detection of interferon-related genes in genomic and expressed sequence tag (EST) databases (*Huang et al., 2008; Sodergren et al., 2006; Zhang et al., 2012; Philipp et al., 2012*). Moreover, in pearl oyster (*Pinctada fucata*) an homolog of interferon regulatory factor 2 (IRF-2) was characterized (*Huang et al., 2013*) and in *C. gigas* the expression levels of interferon-induced protein 44-like (IFI44) and IRF-2 in response to poly I:C and herpesvirus stimulation were evaluated (*Green and Montagnani, 2013*). Notably, the study conducted by *Green and Montagnani* also demonstrated that the injection of Polyinosinic:polycytidylic acid (poly I:C), an immuno-stimulant structurally similar to a dsRNA, could induce *C. gigas* spats into an antiviral state that provided protection against subsequent OsHV-1 μvar infection. In addition, following the

herpesvirus infection, in the poly I:C injected-oysters the viral load was less abundant than in the control group, as demonstrated here for the dsRNA-injected animals. In parallel, they measured in the whole oyster tissues the expression level of nine putative antiviral genes. In accordance to what we have seen in gills following the dsRNA transfection (grey bars in Figure 4C), the expression level of Cg-Rel was found to be significantly up-regulates at 24h after the poly I:C injection. Furthermore, in channel catfish *Ictalurus punctatus* it has been reported that a poly I:C injection, either directly or through the prior induction of IFN, upregulates gene products that function individually and/or collectively to inhibit virus replication (*Milev-Milovanovic et al., 2009*). Lastly, a very recent study has been performed at Ifremer (Plouzané, France) mimicking the present experiment but with an additional condition represented by oysters treated with a dsRNA targeting the Green Fluorescent Protein (GFP). In the aforementioned study, as observed here, mortality appeared very low for dsRNA-IkB2 injected oysters, compared to the high mortality observed in the control oysters, thus confirming our results. Interestingly, mortality also appeared low and similar in dsRNA-GFP and dsRNA-IkB2 injected oysters reinforcing the third hypothesis. Taken together these results support the idea that the low viral load reported in oysters after the infection was due to the activation of immune-related genes hampering the viral replication and thus conferring resistance to OsHV-1.

Moreover, taken in consideration the expression values reported for IkB2 in gonads and gills (grey bars in Figures 3B and 3B) we can suggest that the induction of an unspecific antiviral response triggered by dsRNA could slightly involve also IkB2, which was more abundant at 24h post dsRNA injection, despite the difference was not statistically significant ($p>0.05$).

Since this study aimed to evaluate the expression of a restricted number of genes though to be regulated by the IkB2 silencing, we didn't obtained data concerning the regulation of several other genes which could have been involved in the unspecific antiviral response more likely provoked by the dsRNA targeting IkB2. However, to gain more information concerning the antiviral response, it could be interesting to evaluate the expression of genes implicated in the different steps occurring during the viral infection such as entry of virus, nuclear translocation, viral genetic mechanisms and particle formation, virus life cycle. Notably, in wild *C. gigas* it has been recently demonstrated that a consistent set of genes was differentially regulated in infected and uninfected

oysters, influencing susceptibility or resistance against the virus infection (*Jouaux et al., 2013*).

With regard to the OsHV1 infected oysters injected with the control solution TrisNaCl, the mild down-regulation of IkB1 occurring in both the target tissues (black bars in Figures 3A and 4A) let us thinking that the injection itself slightly induced an immune response activation at the mRNA levels. Despite we didn't carry out the real-time PCR analyses at T0, we believe that after the injection (5h), the transcription of IkB1 immediately increased, followed by a statistically significant decrease at 24h. A similar trend for both IkB1 and IkB2 was reported by *Zhang* and colleagues (2011) in *C.gigas* hemocytes challenged with *Vibrio alginolyticus*. Moreover, it is well established that the activation of the immune response through Rel/NF-kB pathway leads to a transient induction of IkBs, which are involved in generating auto-regulatory feedback loops in the NF-kB response (*Oeckinghaus and Ghosh, 2009*).

While the injection softly activated the immune system, resulting in a mild up-regulation of the sole IkB1, the OsHV-1 infection definitely had a greater impact on the transcription of both IkB1 and IkB2, either in gonads either in gills (black bars in Figures 3B, 3C, 4B and 4C). As suggested below, IkBs are involved in negative feedback regulation of NF-kB. Despite Cg-IkB1 and Cg-IkB2 have not an identical pattern of expression, it has been already demonstrated that their up-regulation is likely to contribute to appropriate inhibition of NF-kB/Rel activation upon pathogens invasion, which may be a mechanism for keeping balance between protecting the host and killing pathogens in *C. gigas* (*Zhang et al., 2011*). Whereas the overexpression of IkBs following bacterial infections has been widely reported in bivalves (*Zhang et al., 2011; Yang et al., 2011; Kasthuri et al., 2013*), their role during viral infections has been poorly investigated and to our knowledge this is the first study evaluating the response of IkBs following herpesvirus infection in oysters. Noteworthy, an important evidence attesting the involvement of IkBs in the innate immune response was provided by *Fleury and Huvet* (*2012*) and *Zhang et al.* (2011). In this paper the authors showed, by means of microarray, that less abundant IkB1 and IkB2 transcripts are associated to resistance to summer mortality. Moreover, an extra immune-relevant transcript which was less expressed in survived oysters was SOCS, which is expected to limit NF–κB signaling following a classical negative feedback system that regulates cytokine signal transduction. Here we showed that the transcription of IkB1, IkB2 and SOCS have a prominent role in the oyster innate immune response and, in

accordance to what demonstrated by *Fleury and Huvet* (*2012*), the abundance of these transcripts are associated to the survival competence.

Moreover in this study we also evidenced that the transcription of Rel1 is involved in the *C. gigas* antiviral response, as previously demonstrated in this species (*Green and Montagnani, 2013*). In fact, as seen for IkBs and SOCS, Rel1 expression was induced after the viral infection. As discussed above, is well established that this transcription factor is up-regulated after bacterial challenge in *C. gigas* and *C. farreri* (*Montagnani et al., 2004; Wang et al., 2011*) but little is known about its involvement in the response against virus infection. Conversely in vertebrates the role of Rel/NF-kB signaling has been well documented. In humans, NF-kB is activated by multiple families of viruses, including HIV-1, HTLV-1, hepatitis B virus (HBV), hepatitis C virus (HCV), EBV, and influenza virus. This activation may serve several functions: to promote viral replication, prevent virus-induced apoptosis, and mediate the immune response to the invading pathogen (*Hiscott et al., 2001*). Recent findings reported that NF-kB function is limited to a key early phase after virus infection by playing a crucial role in the production of type I interferon (*Wang et al., 2010*), essential molecules for limiting virus replication and promoting clearance.

Take together these results, we suggest that in oysters the role of the Rel/NF-kB pathway in the antiviral response could have striking similarities to what observed in vertebrates.

Finally we reported that the regulation of Rel1 and SOCS is similar but not identical in the two target tissues. In fact, in gills, compared to what seen in gonads, the induction of SOCS transcription appeared earlier (at 20h post infection) and the fold increase of both SOCS and Rel1 was more substantial. This is not surprisingly if we consider that gills were characterized as one of the target organ, second only to the mantle, for the viral replication of the Ostreid herpesvirus 1 implicated in some summer mortality outbreaks (e.g. *Schikorski et al. 2011*).

5. Conclusion

In the present study we discussed the effects of a dsRNA targeting IkB2 in oysters exposed to OsHV-1 infection, an important causative agent of the summer mortality syndrome. Despite it remains still unclear if the IkB2 knock-down was successful limiting our first goal to specific the role of Cg-IKB2, the increased survival rate of dsRNA-injected oysters in comparison with the controls (TrisNaCl-injected) demonstrated that dsRNA molecules more likely induce an antiviral state which prevent the virus replication, as recently observed following poly I:C injection. As concern the relevance of

IkB2 in the Pacific oyster antiviral response, further studies are required to shed light in the mechanisms by which the Rel/NF-kB pathway is regulated. Alternative molecular techniques should be employed to investigate the IkB2 role and avoid a simultaneous non-specific immune response. A promising “gene editing” strategy which seems to be exceptionally precise and efficient, and thus could be implemented in C. gigas immune studies, is the one called “CRISPRs” (Clustered Regularly Interspaced Short Palindromic Repeats). Yet a developing technology, CRISPRs derived from bacteria and has been recently employed in eukaryotes to delete (end eventually substitute) a target DNA region through the action of a specific enzyme called CAS9.

In conclusion, these findings point out some important issues concerning the bivalve innate immunity and confirm the possibility, recently paved, of employing dsRNAs in these species as therapeutical agent against virus which cause severe economic losses in aquaculture. However, despite the enormous potential of this approach, more studies are needed and many obstacles must still be overcome before this technique finds practical application in aquaculture. In addition it has been demonstrated that, following herpesvirus infection, the transcription of IkBs, Rel1 and SOCS is stimulated, providing preliminary data concerning the involvement of the Rel/NF-kB pathway in the Pacific oyster antiviral response.

6. References

Almeida R and Allshire RC. RNA silencing and genome regulation. Trends Cell Biol 2005;15:251.

Bachère E, Gueguen Y, Gonzalez M, de Lorgeril J, Garnier J, Romestand B. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas*. Immunol Rev 2004;198:149-68.

Baldwin AS. The NF-kappa B and I kappa B proteins: New discoveries and insights. Annu Rev Immunol 1996;14:649-683.

Berns K, Hijmans EM, Mullenders J, Brummelkamp TR, Velds A, Heimerikx M et al. A large-scale RNAi screen in human cells identifies new components of the p53 path-way. Nature 2004;428:431–437.

Chen FE and Ghosh G. Regulation of DNA binding by Rel/NF-kappaB transcription factors: structural views. Oncogene 1999;18:6845-6852.

Cheng TC. Bivalves. In: Ratcliffe NA and Rowley AF. Invertebrate Blood Cells I 1981. London, Academic Press, pp. 233-299.

Cochennec-Laureau N, Baud JP, Pepin JF, Benabdelmouna A, Soletchnik P, Lupo C et al. Les surmortalités des naissains d'huîtres creuses, *Crassostrea gigas*: acquis des recherches en 2010. Rapport RST/LER/MPL/11.07 IFREMER, 30p.

Cohen L, Henzel WJ, Baeuerle PA. IKAP is a scaffold protein of the IkappaB kinase complex. Nature 1998;395:292-6.

Colwell RR and Liston J. Microbiology of shellfish. Bacteriological study of the natural flora of Pacific oysters (*Crassostrea gigas*). Appl microbiol 1960;8:104-9.

Dash PK, Tiwari M, Santhosh SR, Parida M, Lakshmana Rao PV. RNA interference mediated inhibition of Chikungunya virus replication in mammalian cells. Biochem Biophys Res Commun 2008;376:718–722.

De Zoysa M, Nikapitiya C, Oh C, Whang I, Lee JS, Jung SJ et al. Molecular evidence for the existence of lipopolysaccharide-induced TNF-α factor (LITAF) and Rel/NF-kB pathways in disk abalone (*Haliotis discus discus*). Fish Shellfish Immunol 2010;28:754–763.

Dégremont L, Bédier E, Boudry P. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). II. Response to selection for survival and its influence on growth and yield. Aquaculture 2010;299:21–29.

Degremont L, Bedier E, Soletchnik P, Ropert M, Huvet A, Moal J et al. Relative importance of family, site, and field placement timing on survival, growth, and yield of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). Aquaculture 2005;249:213-229.

Dégremont L, Ernande B, Bédier E, Boudry P. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). I. Estimation of genetic parameters for survival and growth. Aquaculture 2007;262:41–53.

Degremont L. Evidence of herpesvirus (OsHV-1) resistance in juvenile *Crassostrea gigas* selected for high resistance to the summer mortality phenomenon. Aquaculture 2011;317:94–98.

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 2001;411:494–498.

Escoubas JM, Briant L, Montagnani C, Hez S, Devaux C, Roch P. Oyster IKK-like protein shares structural and functional properties with its mammalian homologues. Febs Lett 1999;453:293-298.

Fabioux C, Corporeau C, Quillien V, Favrel P, Huvet A. In vivo RNA interference in oyster: vasa silencing inhibits germ cell development. FEBS J 2009;276:2566–2573.

Fabioux C, Huvet A, Lelong C, Robert, Pouvreau S, Daniel JY et al. Oyster vasa-like gene as a marker of the germline cell development in *Crassostrea gigas*. Biochem Biophys Res Commun 2004;320:592-8.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 1998;391:806–811.

Fleury E and Huvet A. Microarray analysis highlights immune response of Pacific oysters as a determinant of resistance to summer mortality. Mar biotechnol 2012;14:203-17.

Fleury E, Moal J, Boulo V, Daniel JY, Mazurais D, Henaut A et al. Microarray-Based Identification of Gonad Transcripts Differentially Expressed Between Lines of Pacific Oyster Selected to Be Resistant or Susceptible to Summer Mortality. Mar Biotech 2010;12:326-339.

Garnier M, Labreuche Y, Garcia C, Robert A, Nicolas JL. Evidence for the involvement of pathogenic bacteria in summer mortalities of the pacific oyster *Crassostrea gigas*. Microb Ecol 2007;53:187-96.

Ghosh S, May MJ, Kopp EB. NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol 1998;16:225-60.

Gomez-Leon J, Villamil L, Lemos ML, Novoa B, Figueras A. Isolation of *Vibrio alginolyticus* and *Vibrio splendidus* from aquacultured carpet shell clam (*Ruditapes decussatus*) larvae associated with mass mortalities. Appl Environ Microbiol 2005;71:98-104.

Gonzalez M, Gueguen Y, Desserre G, de Lorgeril J, Romestand B, Bachère E. Molecular characterization of two isoforms of defensin from hemocytes of the oyster *Crassostrea gigas*. Dev Comp Immunol 2007;31:332-9.

Green TJ and Barnes AC. Inhibitor of REL/NF-(K)B is regulated in Sydney rock oysters in response to specific double-stranded RNA and *Vibrio alginolyticus*, but the major immune anti-oxidants EcSOD and Prx6 are non-inducible. Fish & Shellfish Immun 2009;27:260-265.

Green TJ, Montagnani C, Benkendorff K, Robinson N, Speck P. Ontogeny and water temperature influences the antiviral response of the Pacific oyster, *Crassostrea gigas*. Fish Shellfish Immunol 2014;36:151-57.

Green TJ, Montagnani C. Poly I:C induces a protective antiviral immune response in the Pacific oyster (*Crassostrea gigas*) against subsequent challenge with Ostreid herpesvirus (OsHV-1 μvar). Fish Shellfish Immunol 2013;35:382-388.

Harvell CD, Kim K, Burkholder JM, Colwell RR, Epstein PR, Grimes DJ, et al. Emerging marine diseases: climate links and anthropogenic factors. Science 1999 Sept 3;285:1505-10.

Harvell CD, Mitchell CE, Ward JR, Altizer S, Dobson AP, Ostfeld RS et al. Climate warming and disease risks for terrestrial and marine biota. Science 2002;Jun 21;296:2158-62.

Hayden MS and Ghosh S. Signalling to NF-kappaB. Genes Dev 2004;18:2195–2224.

Hiscott J, Kwon H, Genin P. Hostile takeovers: viral appropriation of the NF-kappaB pathway. J Clin Invest 2001;107:143–151.

Huang SF, Yuan SC, Guo L, Yu YH, Li J, Wu T, et al. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome Res 2008;18:1112-1126.

Huang XD, Liu WG, Wang Q, Zhao M, Wu SZ, Guan YY et al. Molecular characterization of interferon regulatory factor 2 (IRF-2) homolog in pearl oyster *Pinctada fucata*. Fish Shellfish Immunol 2013;34:1279-1286.

Huvet A, Fleury E, Corporeau C, Quillien V, Daniel J, Riviere G et al. In Vivo RNA Interference of a Gonad-Specific Transforming Growth Factor-β in the Pacific Oyster *Crassostrea gigas*. Mar Biotechnol 2012;14:402-410.

Huvet A, Herpin A, Dégremont L, Labreuche Y, Samain JF, Cunningham C. The identification of genes from the oyster *Crassostrea gigas* that are differentially expressed in progenies exhibiting opposed susceptibility to summer mortality. Gene 2004;343:211-220.

Huvet A, Normand J, Fleury E, Quillien V, Fabioux C, Boudry P. Reproductive effort of Pacific oysters: a trait associated with susceptibility to summer mortality. Aquaculture 2010;304:95–99.

Jiang Y and Wu X. Characterization of a Rel\NF-kappaB homologue in a gastropod abalone, *Haliotis diversicolor supertexta*. Dev Comp Immunol 2007;31:121–131.

Jouaux A, Lafont M, Blin JL, Houssin M, Mathieu M, Lelong C. Physiological change under OsHV-1 contamination in Pacific oyster *Crassostrea gigas* through massive mortality events on fields. BMC Genomics 2013;14:590.

Kasthuri SR, Whang I, Navaneethaiyer U, Lim BS, Choi CY, Lee J. Molecular characterization and expression analysis of IκB from *Haliotis discus discus*. Fish Shellfish Immunol 2013;34:1596-1604.

Koutsogiannaki S, Kaloyianni M. Signaling molecules involved in immune responses in mussels. Inv Surv J 2010;7:11-21.

Lacoste A, Jalabert F, Malham S, Cueff A, Gelebart F, Cordevant C et al. A Vibrio splendidus strain is associated with summer mortality of juvenile oysters Crassostrea gigas in the bay of Morlaix (north Brittany, France). Dis Aquat Org 2001;46:139-45.

Lafferty KD, Porter JW, Ford SE. Are diseases increasing in the ocean? Annu Rev Ecol Evol Syst 2004;35: 31–54.

Le Foll F, Rioulta D, Boussab S, Pasquiera J, Dagherc Z, Leboulenger F. Characterisation of *Mytilus edulis* hemocyte subpopulations by single cell time-lapse motility imaging. Fish Shellfish Immunol 2010;28:372-86.

Le Moullac G, Fleury PG, Le Coz JR, Moal J, Samain JF. Effect of sediment nearness on the metabolic enzyme activity and energy state of the oyster *Crassostrea gigas*. Aquat Living Resou 2007a;20: 279-286.

Le Moullac G, Queau I, Le Souchu P, Pouvreau S, Moal J, Le Coz JR et al. Metabolic adjustments in the oyster Crassostrea gigas according to oxygen level and temperature. Mar Biol Res 2007b;3:357-366.

Le Roux F, Gay M, Lambert C, Waechter M, Poubalanne S et al. Comparative analysis of *Vibrio splendidus*-related strains isolated during *Crassostrea gigas* mortality events. Aquat Living Resour 2002;15:251–258.

Livak KL and Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the 2-Ct method. Methods 2001;25:402–408.

Lu Y, Wu LP, Anderson KV. (2001). The antibacterial arm of the drosophila innate immune response requires an IkappaB kinase. Genes Dev 2001;15:104–110.

Milev-Milovanovic I, Majji S, Thodima V, Deng Y, Hanson L, Arnizaut A et al. Identification and expression analyses of poly [I:C]-stimulated genes in channel catfish (*Ictalurus punctatus*). Fish Shellfish Immunol 2009;26:811-820.

Mocellin S and Provenzano M. RNA interference: learning gene knock-down from cell physiology. J Transl Med 2004;2:39.

Montagnani C, Kappler C, Reichhart JM, Escoubas JM. Cg-Rel, the first Rel/NF-kappa B homolog characterized in a mollusk, the Pacific oyster *Crassostrea gigas*. Febs Letters 2004;561:75-82.

Mydlarz LD, Jones LE, Harvell CD. Innate immunity environmental drivers and disease ecology of marine and freshwater invertebrates. Annu Rev Ecol Evol Syst 2006;37:251–288.

Napoli C, Lemieux C, Jorgensen R. Introduction of a chimeric chalcone synthase gene into Petunia results in reversible co-suppression of homologous genes in trans. Plant Cell 1990;2:279-289.

Oeckinghaus A and Ghosh S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harb Perspect Biol 2009; 1:a000034.

Pepin JF, Riou A, Renault T. Rapid and sensitive detection of ostreid *herpesvirus 1* in oyster samples by real-time PCR. J Virol Methods 2008;149:269-76.

Pernet F, Barret J, Le Gall P, Corporeau C, Dégremont L, Lagarde F et al. Mass mortalities of Pacific oysters *Crassostrea gigas* reflect infectious diseases and vary with farming practices in the Mediterranean Thau lagoon, France. Aquac Environ Interact 2012;2:215–237.

Petton B, Pernet F, Robert R, Boudry P. Temperature influence on pathogen transmission and subsequent mortalities in juvenile Pacific oysters *Crassostrea gigas*. Aquacult Environ Interact 2013;3:257-273.

Philipp EER, Kraemer L, Melzner F, Poustka AJ, Thieme S, Findeisen U, et al. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. Plos One 2012;7:e33091.

Randall RE and Goodbourn S. Interferons and viruses: an interplay between induction, signalling, antiviral responses and viral countermeasures. J Gen Virol 2008;89:1-47.

Régnier CH, Song HY, Gao X, Goeddel DV, Cao Z, Rothe M. Identification and characterization of an IkappaB kinase. Cell 1997;90:373-83.

Renault T, Ledeuff RM, Cochennec N, Maffart P. Herpesviruses associated with mortalities among Pacific oyster, *Crassostrea gigas*, in France - comparative study. Revue De Medecine Veterinaire 2004;145:735-742.

Robalino J, Browdy CL, Prior S, Metz A, Parnell P, Gross P et al. Induction of antiviral immunity by double-stranded RNA in a marine invertebrate. J Virol 2004;78:10442–10448.

Romano N and Macino G. Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. Mol Microbiol 1992;6:3343-3353.

Samain JF, McCombie H. Summer Mortality of Pacific Oyster *Crassostrea gigas*. Ed. Ifremer/Quae, Versailles 2008.

Schikorski D, Renault T, Saulnier D, Faury N, Moreau P, Pépin JF. Experimental infection of Pacific oyster *Crassostrea gigas* spat by ostreid herpesvirus 1: demonstration of oyster spat susceptibility. Vet Res 2011;7:42-27.

Segarra A, Pepin JF, Arzul I, Morga B, Faury N, Renault T. Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. Virus Res 2010;153: 92-99.

Seo JK, Crawford JM, Stone KL, Noga EJ. Purification of a novel arthropod defensin from the American oyster, *Crassostrea virginica*. Biochem Biophys Res Commun 2007;338:1998-2004.

Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA, Weinstock GM et al. Research article and the genome of the sea urchin *Strongylocentrotus purpuratus*. Science 2006;314:941-952.

Suquet M, De Kermoysan G, Araya RG, Queau I, Lebrun L, Le Souchu P et al. Anesthesia in Pacific oyster, *Crassostrea gigas*. Aquat Living Resour 2009;22:29-34.

Suzuki M, Saruwatari K, Kogure T, Yamamoto Y, Nishimura T, Kato T et al. An acidic matrix protein, Pif, is a key macromolecule for nacre formation. Science 2009;325(5946):1388-90.

Tieri P, Termanini A, Bellavista E, Salvioli S, Capri M, Franceschi C: Charting the NF-κB pathway interactome map. PLoS ONE 2012;7: e32678.

Thompson JE, Phillips RJ, Erdjument-Bromage H, Tempst P, Ghosh S. I kappa B-beta regulates the persistent response in a biphasic activation of NF-kappa B. Cell 1995;80:573-82.

Venier P, Varotto L, Rosani U, Millino C, Celegato B, Bernante F et al. Insights into the innate immunity of the Mediterranean mussel *Mytilus galloprovincialis*. BMC Genomics 2011;12:69.

Waechter M, Le Roux F, Nicolas JL, Marissal E, Berthe F. Characterisation of *Crassostrea gigas* spat pathogenic bacteria. C R Biol 2002;325:231–238.

Wang J, Basagoudanavar SH, Wang X, Hopewell E, Albrecht R, García-Sastre A et al. NF-kappa B RelA subunit is crucial for early IFN-beta expression and resistance to RNA virus replication. J Immunol 2010;185:1720-1729.

Wang M, Yang J, Zhou Z, Qiu L, Wang L, Zhang H et al. A primitive toll-like receptor signaling pathway in mollusk zhikong scallop *Chlamys farreri*. Dev Comp Immunol 2011;35:511-520.

Whiteside ST, Epinat JC, Rice NR, Israël A. I kappa B epsilon, a novel member of the I kappa B family, controls RelA and cRel NF-kappa B activity. EMBO J 1997;16:1413-26.

Woronicz JD, Gao X, Cao Z, Rothe M, Goeddel DV. IkappaB kinase-beta: NF-kappaB activation and complex formation with IkappaB kinase-alpha and NIK. Science 1997;278:866-9.

Xiong X, Feng Q, Chen L, Xie L, Zhang R. Cloning and characterization of an IKK homologue from pearl oyster, *Pinctada fucata*. Dev Comp Immunol 2008;32:15–25.

Yang Q, Yang Z, Li H. Molecular characterization and expression analysis of an inhibitor of NF-kappaB (IkappaB) from Asiatic hard clam *Meretrix meretrix*. Fish Shellfish Immunol 2011;31:485-490.

Zhang D, Jiang S, Qiu L, Su T, Wu K, et al. Molecular characterization and expression analysis of the IkappaB gene from pearl oyster *Pinctada fucata*. Fish Shellfish Immunol 2009;26:84–90.

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F et al. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 2012;490: 49-54.

Zhang Y, He XC, Yu ZN. Two homologues of inhibitor of NF-kappa B (I kappa B) are involved in the immune defense of the Pacific oyster, *Crassostrea gigas*. Fish Shellfish Immunol 2011;30:1354-1361.

General conclusions

In the present thesis NGS expression analyses and RNA interference were employed to improve the understanding on transcriptional events characterizing three bivalve species: *V. decussata*, *P. maximus* and *C. gigas*. The obtained results provided important evidence on molecular mechanisms regulating immune response and gamete maturation in bivalves, thus improving the knowledge on these commercially important species. This work demonstrated that NGS approaches and data mining, besides increasing the genomic knowledge on bivalves, might lead to isolate novel immune-relevant molecules. Moreover gene expression analyses may shed light into complex molecular networks implied in the bivalve immune response, and possibly resistance, against pathogens occurring in both wild and farm conditions. Moreover, this kind of studies could provide important information to improve the efficiency of reproduction in species for which hatchery-based production is more difficult on a routine basis.

In conclusion, this thesis represents an example of how genomic analysis and gene expression studies may contribute to control reproduction and gametes quality, to develop strategies for management of diseases and for long-term sustainability of bivalve aquaculture.

Appendix A1

A. Up-regulated genes
A-kinase anchor protein 9
Apolipophorins
Transforming growth factor-beta-induced protein ig-h3
Complement C1q-like protein 4
Ceramide synthase 6
Contactin
Chordin-like protein 2
DC-STAMP domain-containing protein 1
Protein dispatched homolog 1
Exosome complex component RRP40
Coagulation factor IX
Protein four-jointed
Forkhead box protein E4
Fumarate hydratase, mitochondrial
Frizzled-8
Frizzled-8
Frizzled-8
Indole-3-acetic acid-amido synthetase GH3.4
Zinc finger protein GLIS1
G-protein coupled receptor 64
Serine-enriched protein
Lymphoid-specific helicase
Immunoglobulin superfamily member 10
Kelch-like protein 5
Low-density lipoprotein receptor-related protein 4
Meteorin-like protein
Monocarboxylate transporter 12
Monocarboxylate transporter 14
M-phase inducer phosphatase 1
M-phase inducer phosphatase 1
M-phase inducer phosphatase 3
M-phase inducer phosphatase 3
M-phase inducer phosphatase
Unconventional myosin-X
Sodium/calcium exchanger 3
Poly(rC)-binding protein 3
PH domain leucine-rich repeat-containing protein phosphatase 2
Piwi-like protein 2
Ras-related C3 botulinum toxin substrate 1
Retinal homeobox protein Rx1
Tubulin monoglycylase TTLL3
Cdc25-like protein phosphatase twine
E3 ubiquitin-protein ligase UBR4
Vitellogenin
Protein Wnt-4
Uncharacterized 51.9 kDa protein in rps4-rps11 intergenic region
Protein PF14_0175
UPF0392 protein C33H5.2
Oocyte zinc finger protein XlCOF28
B. Down-regulated genes
Phospholipase D LhSicTox-alphaIA2aviii
Acetylcholinesterase
2-amino-3-carboxymuconate-6-semialdehyde decarboxylase
Aminoacyl tRNA synthase complex-interacting multifunctional protein 1
Allantoicase
Protein arginine N-methyltransferase 7
AP-2 complex subunit sigma
D-arabinose 1-dehydrogenase
Sphingomyelin phosphodiesterase
Asparagine synthetase domain-containing protein 1
Ankyrin repeat, SAM and basic leucine zipper domain-containing protein 1
Ubiquitin-like-conjugating enzyme ATG10
UDP-GalNAc:beta-1,3-N-acetylgalactosaminyltransferase 2
Biliverdin reductase A
Protein BolA
Cadherin-87A
Cathepsin L2
Cathepsin S
Coiled-coil domain-containing protein 150
Coiled-coil domain-containing protein 84
Parafibromin
Cytidine deaminase
Transmembrane protein C5orf28 homolog
Cholinesterase
Cingulin
UPF0696 protein C11orf68 homolog
Calcium-activated chloride channel regulator 1
Calcium-activated chloride channel regulator 4
Ceroid-lipofuscinosis neuronal protein 6 homolog
Putative ATP-dependent Clp protease proteolytic subunit, mitochondrial
Collagen alpha-1(VIII) chain
Collagen alpha-1(XII) chain
Conserved oligomeric Golgi complex subunit 5
COMM domain-containing protein 4
COMM domain-containing protein 4
Steroid 17-alpha-hydroxylase/17,20 lyase
Cytochrome P450 2U1
Crystallin J1A
Crystal protein
Cysteine sulfinic acid decarboxylase
DDB1- and CUL4-associated factor 17
Deoxycytidylate deaminase
DnaJ homolog subfamily B member 5
Probable D-tyrosyl-tRNA(Tyr) deacylase 2
Enoyl-CoA delta isomerase 1, mitochondrial
Epithelial cell-transforming sequence 2 oncogene-like
Elongation of very long chain fatty acids protein 5
Crossover junction endonuclease EME1
Failed axon connections homolog
Low affinity immunoglobulin epsilon Fc receptor
Glutamate decarboxylase-like protein 1
Glutamate decarboxylase-like protein 1
Guanine nucleotide-binding protein subunit beta-2-like 1
GPN-loop GTPase 1
Glutathione peroxidase 2
Glutamate receptor 1
Glutathione S-transferase 1
Glutathione S-transferase 3
Translation factor Guf1, mitochondrial
Hemagglutinin/amebocyte aggregation factor

3-hydroxybutyryl-CoA dehydrogenase
Helicase with zinc finger domain 2
Histone-binding protein N1/N2
5-hydroxyisourate hydrolase
3-hydroxypropionyl-coenzyme A dehydratase
4-hydroxyphenylpyruvate dioxygenase
Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial
Importin subunit alpha-2
Chromosome-associated kinesin KIF4B
LSM domain-containing protein 1-B
Ragulator complex protein LAMTOR3-B
Malate dehydrogenase
Mediator of RNA polymerase II transcription subunit 30
TRAF3-interacting protein 1
MLX-interacting protein
MAGUK p55 subfamily member 6
mRNA turnover protein 4 homolog
Methylthioribulose-1-phosphate dehydratase
Myosin regulatory light polypeptide 9
Myeloid zinc finger 1
Nascent polypeptide-associated complex subunit alpha
Sodium-dependent glucose transporter 1
N-acetylated-alpha-linked acidic dipeptidase-like protein
Neutral ceramidase
Neutral ceramidase B
NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial
Neprilysin
Nardilysin
Nardilysin
U8 snoRNA-decapping enzyme
Nucleoporin Nup37
Protein odr-4 homolog
GDP-fucose protein O-fucosyltransferase 1
Progestin and adipoQ receptor family member 3
PDZ and LIM domain protein 1
Prefoldin subunit 5
Phytanoyl-CoA dioxygenase domain-containing protein 1 homolog
PIH1 domain-containing protein 1
Perlucin-like protein
Protection of telomeres protein 1
Peptidyl-prolyl cis-trans isomerase A
Palmitoyl-protein thioesterase 1
Profilin-4
Prostasin

Periodic tryptophan protein 1 homolog
Regucalcin
60S ribosomal protein L3
60S ribosomal protein L3
60S ribosomal protein L3
60S ribosomal protein L4
60S acidic ribosomal protein P1
60S acidic ribosomal protein P1
40S ribosomal protein S26-B
40S ribosomal protein S26
40S ribosomal protein S3
40S ribosomal protein S9
40S ribosomal protein S9
Solute carrier family 25 member 45
3-oxo-5-alpha-steroid 4-dehydrogenase 1
Selenium-binding protein 1
S-crystallin SL20-1
Vesicle transport protein SEC20
N-lysine methyltransferase setd6
Protein SET
Protein slit
T-complex protein 1 subunit gamma
L-threonine 3-dehydrogenase
L-threonine 3-dehydrogenase
Testis-expressed sequence 10 protein homolog
General transcription factor 3C polypeptide 4
E3 ubiquitin-protein ligase TRIM63
Probable tRNA (uracil-O(2)-)-methyltransferase
Tryptase
Thrombospondin-type laminin G domain and EAR repeat-containing protein
Tetraspanin-9
Translin-associated protein X
Alpha-tocopherol transfer protein-like
Demethylmenaquinone methyltransferase
Ubiquitin carboxyl-terminal hydrolase 20
Ubiquitin carboxyl-terminal hydrolase isozyme L5
Probable uridine nucleosidase 2
U3 small nucleolar RNA-associated protein 18 homolog
Vacuolar protein sorting-associated protein 33A
WD repeat-containing protein 36
Xylose isomerase
Uncharacterized protein YMR196W
YrdC domain-containing protein, mitochondrial
Zinc finger protein 385D
GATA-type zinc finger protein 1

Appendix A2

BP terms	Count	PValue	FE
GO:0016055~Wnt receptor signaling pathway	6	0.00	6.02
GO:0009310~amine catabolic process	5	0.01	5.98
GO:0006470~protein amino acid dephosphorylation	6	0.01	4.10
GO:0016311~dephosphorylation	6	0.03	3.55
GO:0006414~translational elongation	4	0.03	5.78
GO:0006567~threonine catabolic process	2	0.03	62.17
GO:0046395~carboxylic acid catabolic process	5	0.03	4.09
GO:0016054~organic acid catabolic process	5	0.03	4.09
GO:0009063~cellular amino acid catabolic process	4	0.04	5.29
GO:0051329~interphase of mitotic cell cycle	4	0.04	5.29
GO:0051325~interphase	4	0.04	5.18
GO:0006566~threonine metabolic process	2	0.05	41.45
GO:0051187~cofactor catabolic process	3	0.05	8.11
GO:0006412~translation	8	0.05	2.32
GO:0000278~mitotic cell cycle	8	0.08	2.14
GO:0000279~M phase	8	0.09	2.07
GO:0000226~microtubule cytoskeleton organization	5	0.09	2.93
GO:0009066~aspartate family amino acid metabolic process	3	0.10	5.65
CC terms	**Count**	**PValue**	**FE**
GO:0022626~cytosolic ribosome	4	0.00	11.18
GO:0005576~extracellular region	15	0.01	2.13
GO:0005840~ribosome	7	0.01	3.54
GO:0044445~cytosolic part	5	0.02	4.51
GO:0044421~extracellular region part	8	0.04	2.43
GO:0033279~ribosomal subunit	4	0.05	4.89
GO:0030529~ribonucleoprotein complex	10	0.09	1.82
MF terms	**Count**	**PValue**	**FE**
GO:0005198~structural molecule activity	12	0.00	3.06
GO:0003735~structural constituent of ribosome	7	0.00	4.91
GO:0004725~protein tyrosine phosphatase activity	6	0.01	4.93
GO:0004721~phosphoprotein phosphatase activity	7	0.01	3.64
GO:0070011~peptidase activity, acting on L-amino acid peptides	10	0.04	2.15
GO:0016831~carboxy-lyase activity	3	0.04	8.92
GO:0008233~peptidase activity	10	0.05	2.04
GO:0016791~phosphatase activity	7	0.06	2.44
GO:0004175~endopeptidase activity	7	0.07	2.43
KEGG pathways	**Count**	**PValue**	**FE**
dme03010:Ribosome	3	0.00	30.13

Appendix A3

A. Genes negatively correlated with D-larval rate
Protein AATF
Abl interactor 1
Angiotensin-converting enzyme
Activated CDC42 kinase 1
Acetyl-coenzyme A synthetase, cytoplasmic
Adenylate cyclase type 5
Adenylate cyclase type 5
Adenylate cyclase type 6
Androgen-induced gene 1 protein
Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3
Membrane primary amine oxidase
Membrane primary amine oxidase
Allene oxide synthase-lipoxygenase protein
Anucleate primary sterigmata protein B
ADP-ribosylation factor-like protein 1
Arylsulfatase B
Activating signal cointegrator 1 complex subunit 1
Alkali-sensitive linkage protein 1
Autophagy-related protein 9A
L-azetidine-2-carboxylic acid acetyltransferase
Transforming growth factor-beta-induced protein ig-h3
Brefeldin A-inhibited guanine nucleotide-exchange protein 1
Baculoviral IAP repeat-containing protein 7
BUD13 homolog
Complement component C1q receptor
Cholesterol 25-hydroxylase-like protein 1, member 2
Dipeptidyl peptidase 1
Cathepsin D
COBW domain-containing protein 2
Coiled-coil domain-containing protein 103
Coiled-coil domain-containing protein 39
Coiled-coil domain-containing protein 51
Cysteine-rich DPF motif domain-containing protein 1
Centrosomal protein of 192 kDa
Ceramide synthase 6
Cytoskeleton-associated protein 5
Putative uncharacterized protein C12orf63
Chloride channel CLIC-like protein 1
H(+)/Cl(-) exchange transporter 3
CAP-Gly domain-containing linker protein 3
UPF0568 protein C14orf166 homolog
Cytosolic non-specific dipeptidase
Collagen alpha-3(VI) chain
Collagen alpha-1(XII) chain
Coatomer subunit alpha
Coatomer subunit alpha
UPF0585 protein C16orf13 homolog A
Cytochrome P450 2U1
Copine-8
Uncharacterized protein C17orf53 homolog
Uncharacterized protein C17orf53 homolog
Crooked neck-like protein 1
Transcriptional repressor CTCF
Cysteine synthase 1
Disheveled-associated activator of morphogenesis 1
Doublecortin domain-containing protein 5
Malonyl-CoA decarboxylase, mitochondrial
DC-STAMP domain-containing protein 2
DDRGK domain-containing protein 1
DENN domain-containing protein 3
Estradiol 17-beta-dehydrogenase 11
Dehydrogenase/reductase SDR family member 11
Putative ATP-dependent RNA helicase DHX33
Probable dimethyladenosine transferase
Disrupted in renal carcinoma protein 2 homolog
Dynamin-1-like protein
DPH3 homolog
Dolichol-phosphate mannosyltransferase
Dentin sialophosphoprotein
Dynein heavy chain 1, axonemal
Dynein heavy chain 8, axonemal
Dynein beta chain, flagellar outer arm
Transcription factor E2F4
Elongation factor 1-beta
Ankyrin repeat domain-containing protein EMB506, chloroplastic
Enolase
Ectonucleotide pyrophosphatase/phosphodiesterase family member 2
Ectonucleotide pyrophosphatase/phosphodiesterase family member 3
Ectonucleotide pyrophosphatase/phosphodiesterase family member 3
Ectonucleoside triphosphate diphosphohydrolase 1
Ero1-like protein
Electron transfer flavoprotein subunit beta
Protein FAM63A
Fumarylacetoacetase
Focal adhesion kinase 1
Protein FAN
F-box only protein 7
FCH and double SH3 domains protein 2
Focadhesin
Tyrosine-protein kinase Fyn
UDP-glucose 4-epimerase
Galactokinase
GRIP and coiled-coil domain-containing protein 2
Glutamine synthetase 2 cytoplasmic
G-protein coupled receptor 98
G patch domain-containing protein 1
Glutathione S-transferase 1
Gametocyte-specific factor 1
3-hydroxybutyryl-CoA dehydrogenase
Histone deacetylase 8
Heme-binding protein 2
Homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 2 protein
Histone-binding protein N1/N2
Hydroxyacid-oxoacid transhydrogenase, mitochondrial
Hydroxyacid-oxoacid transhydrogenase, mitochondrial
Heat shock protein beta-1
Voltage-gated hydrogen channel 1
Hydantoin utilization protein A
E3 ubiquitin-protein ligase IAP-3

33 kDa inner dynein arm light chain, axonemal
Iduronate 2-sulfatase
Insulin-like growth factor 2 mRNA-binding protein 3
Interferon-induced protein 44-like
Mitochondrial inner membrane protein
Inositol 1,4,5-trisphosphate receptor-like protein A
Isopenicillin N synthase
Inorganic pyrophosphatase
Inositol 1,4,5-trisphosphate receptor type 2
Inositol 1,4,5-trisphosphate receptor type 2
Uncharacterized protein KIAA0895
Arginine kinase
Kielin/chordin-like protein
Kinesin heavy chain
Ribosomal protein S6 kinase alpha-1
Arachidonate 5-lipoxygenase
Lysosomal alpha-mannosidase
NADP-dependent malic enzyme
CDK-activating kinase assembly factor MAT1
DNA replication licensing factor mcm2
Mitotic spindle assembly checkpoint protein MAD2A
Mechanosensory protein 2
Mediator of RNA polymerase II transcription subunit 14
Methyltransferase-like protein 9
Major facilitator superfamily domain-containing protein 7-b
Mitochondrial folate transporter/carrier
E3 ubiquitin-protein ligase MIB2
E3 ubiquitin-protein ligase MIB1
Meiosis-specific nuclear structural protein 1
Monocarboxylate transporter 13
Macrophage mannose receptor 1
Maestro heat-like repeat-containing protein family member 1
Motile sperm domain-containing protein 1
Motile sperm domain-containing protein 2
Methionine-R-sulfoxide reductase B3, mitochondrial
Methylthioribulose-1-phosphate dehydratase
Metaxin-1
Nicotinamide phosphoribosyltransferase
NADPH-dependent diflavin oxidoreductase 1
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 2, mitochondrial
Neuropilin and tolloid-like protein 2
Neurofibromin
Nuclear factor interleukin-3-regulated protein
Nuclear hormone receptor family member nhr-23
Nucleolar protein of 40 kDa
Neurogenic locus notch protein homolog
Neuropeptide FF receptor 2
Nuclear receptor subfamily 1 group I member 2
Peroxisomal NADH pyrophosphatase NUDT12
Nuclear RNA export factor 1
Outer dense fiber protein 3-B
Opioid growth factor receptor-like protein 1
Otoferlin
Calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase 1C
Retinal rod rhodopsin-sensitive cGMP 3',5'-cyclic phosphodiesterase subunit delta
Probable pyridoxine biosynthesis SNZERR

Protein pelota
Prefoldin subunit 3
Post-GPI attachment to proteins factor 2
Geranylgeranyl transferase type-2 subunit beta
GPI transamidase component PIG-S
Placenta-specific gene 8 protein
Perilipin-2
Protein O-linked-mannose beta-1,4-N-acetylglucosaminyltransferase 2
Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase
Polyprenol reductase
Presqualene diphosphate phosphatase
Serine/threonine-protein phosphatase with EF-hands 2
Pre-mRNA-processing factor 6
Pituitary tumor-transforming gene 1 protein-interacting protein
Pregnancy zone protein
Queuine tRNA-ribosyltransferase subunit qtrtd1
Ras-related protein Rab-20
Ras-related protein Rab-7L1
Cell cycle checkpoint protein RAD17
Cell cycle checkpoint protein RAD17
DNA double-strand break repair Rad50 ATPase
RCC1 domain-containing protein 1
Protein Red
60S ribosomal protein L13
60S ribosomal protein L17
60S ribosomal protein L17
39S ribosomal protein L48, mitochondrial
RING finger protein 37
RING finger protein 44
RRP12-like protein
40S ribosomal protein S20
Radial spoke head protein 4 homolog A
Radial spoke head protein 3 homolog
28S ribosomal protein S27, mitochondrial
28S ribosomal protein S29, mitochondrial
Protein RTF2 homolog
Solute carrier family 15 member 4
Sal-like protein 1
Selenium-binding protein 1
Protein transport protein Sec31A
Translocation protein SEC62
Nucleoporin seh1
15 kDa selenoprotein
E3 ubiquitin-protein ligase SHPRH
Sodium/hydrogen exchanger 8
Structure-specific endonuclease subunit SLX4
Protein SMG9
Survival motor neuron protein 1
Smoothened homolog
Solute carrier organic anion transporter family member 4C1
Spermatogenesis-associated protein 5-like protein 1
Pre-mRNA-splicing factor SPF27
Signal recognition particle 54 kDa protein
Serine/threonine-protein kinase 17A
Cytosolic sulfotransferase 1
Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1

Mitochondrial translocator assembly and maintenance protein 41 homolog
TBC1 domain family member 23
Methylcytosine dioxygenase TET2
Tuftelin-interacting protein 11
Transforming growth factor-beta receptor-associated protein 1
TIMELESS-interacting protein
Transmembrane protein 104
Transmembrane protein 41B
Tankyrase-1
E3 ubiquitin-protein ligase TRIM33
Tripartite motif-containing protein 45
Transient receptor potential cation channel subfamily A member 1 homolog
Hamartin
Protein teashirt
Tetratricopeptide repeat protein 21B
Thioredoxin domain-containing protein 16
Uncharacterized protein U88
Ubiquitin-conjugating enzyme E2-17 kDa
Demethylmenaquinone methyltransferase
Ubiquitin carboxyl-terminal hydrolase 15
Ubiquitin carboxyl-terminal hydrolase 47
Ubiquitin carboxyl-terminal hydrolase 4
Ufm1-specific protease 2
DP-glucose:glycoprotein glucosyltransferase 1
E3 ubiquitin-protein ligase UHRF1
Vacuolar protein sorting-associated protein 16 homolog
Vacuolar protein sorting-associated protein 16 homolog
Vacuolar protein sorting-associated protein 4B
Vesicle transport through interaction with t-SNAREs homolog 1A
von Willebrand factor A domain-containing protein 9
pre-mRNA 3' end processing protein WDR33
WD repeat domain-containing protein 83
WD repeat-containing protein 90
Zinc finger protein Xfin
DNA repair protein XRCC1
Putative serine/threonine-protein kinase/receptor R826
Zinc finger CCCH domain-containing protein 14
Zinc finger protein 184
Zinc finger protein 704
Zinc finger protein 740
Zinc finger and SCAN domain-containing protein 29
B. Genes positively correlated with D-larval rate
14-3-3 protein beta/alpha
Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit alpha isoform
SH3 domain-binding protein 5
Putative glucose-6-phosphate 1-epimerase
Protein AAR2 homolog
ATP-binding cassette sub-family B member 7, mitochondrial
Abhydrolase domain-containing protein 16A
Activator of basal transcription 1
Acetylcholine receptor subunit alpha-type acr-16
Neuronal acetylcholine receptor subunit alpha-9-I
Acidic repeat-containing protein

Phosphoacetylglucosamine mutase
Phosphoacetylglucosamine mutase
A-kinase anchor protein 9
Potassium channel AKT3
Aldose reductase
Ankyrin repeat and FYVE domain-containing protein 1
Ankyrin-3
Ankyrin repeat domain-containing protein 7
Ankyrin repeat domain-containing protein 6
Ankyrin repeat domain-containing protein 6
Ankyrin repeat and SAM domain-containing protein 3
Protein arginine N-methyltransferase 2
Annexin A7
Amyloid beta A4 precursor protein-binding family B member 2
AT-rich interactive domain-containing protein 5B
Armadillo repeat-containing protein 8
Aryl hydrocarbon receptor nuclear translocator homolog
Actin-related protein 6
Arylsulfatase J
Abnormal spindle-like microcephaly-associated protein homolog
Abnormal spindle-like microcephaly-associated protein homolog
Abnormal spindle-like microcephaly-associated protein homolog
Beta-1,4-galactosyltransferase 1
BTB and MATH domain-containing protein 38
Tyrosine-protein kinase BAZ1B
B-cell lymphoma/leukemia 11A
Beta,beta-carotene 9',10'-oxygenase
Baculoviral IAP repeat-containing protein 7-A
Baculoviral IAP repeat-containing protein 2
Bloom syndrome protein
BMP-binding endothelial regulator protein
BCL2/adenovirus E1B 19 kDa protein-interacting protein 3
BTB/POZ domain-containing protein 6
Bystin
Cadherin-23
Cadherin-4
Cadherin-7
Uncharacterized protein C2orf16
Cytosolic carboxypeptidase 6
CREB-binding protein
Cyclin-dependent kinase 20
Cyclin-dependent kinase 20
Calcium-dependent protein kinase 4
Ceramide kinase
Ceramide kinase
Uncharacterized protein C11orf63 homolog
Chloride channel protein 2
CAP-Gly domain-containing linker protein 4
Uncharacterized protein C14orf166B
2',3'-cyclic-nucleotide 3'-phosphodiesterase
Centriolin
Phosphopantothenoylcysteine decarboxylase
Cartilage oligomeric matrix protein
Cartilage oligomeric matrix protein
Coatomer subunit delta
Coatomer subunit gamma-2

High affinity copper uptake protein 1
Cytochrome P450 4F6
Cleavage and polyadenylation specificity factor subunit 5
Carnitine O-palmitoyltransferase 1, liver isoform
Cysteine sulfinic acid decarboxylase
Cysteine-rich protein 2-binding protein
Alpha-catulin
Citron Rho-interacting kinase
CUE domain-containing protein 1
Cullin-3
DNA damage-binding protein 2
Probable ATP-dependent RNA helicase DDX43
RNA polymerase II degradation factor 1
17-beta-hydroxysteroid dehydrogenase 14
Deleted in autism protein 1 homolog
Protein diaphanous
DnaJ homolog subfamily C member 10
DnaJ homolog subfamily C member 11
DnaJ homolog subfamily A member 2
DnaJ homolog subfamily C member 8
DNA polymerase alpha catalytic subunit
Dynein light chain, cytoplasmic
Dynein light chain 1, cytoplasmic
Dual specificity tyrosine-phosphorylation-regulated kinase 2
Dystonin
Ecdysone-induced protein 78C
Early endosome antigen 1
EF-hand calcium-binding domain-containing protein 4A
Iron(II)-dependent oxidoreductase EgtB
Transcription elongation factor 1 homolog
Ankyrin repeat domain-containing protein EMB506, chloroplastic
Protein FAM160B1
Protein FAM194A
Protein FAM200A
E3 ubiquitin-protein ligase FANCL
Protein FAN
Fibroblast growth factor 18
Glutamate carboxypeptidase 2
Fucolectin-1
Frizzled-10
Frizzled-4
Frizzled-4
Growth arrest and DNA damage-inducible protein GADD45 alpha
Glutamate decarboxylase-like protein 1
Probable N-acetylgalactosaminyltransferase 9
Gelsolin-like protein 2
D-glucuronyl C5-epimerase
PTB domain-containing engulfment adapter protein 1
Core histone macro-H2A.1
Histone H3.3
Putative nuclease HARBI1
HSPB1-associated protein 1
HSPB1-associated protein 1
Delta-aminolevulinic acid dehydratase
Protein HEXIM2
Hippocampus abundant transcript 1 protein
Hemicentin-1

HRAS-like suppressor 3
Hydroxylysine kinase
Interferon regulatory factor 2-binding protein 2-A
Interferon-induced protein 44
Interferon-induced helicase C domain-containing protein 1
Leukocyte elastase inhibitor
Mitochondrial inner membrane protease subunit 1
Inhibitor of growth protein 4
Integrator complex subunit 1
Integrin alpha-7
Integrin alpha-7
Inositol 1,4,5-trisphosphate receptor type 1
Protease inhibitor bitisilin-3
Uncharacterized protein KIAA0513
Uncharacterized protein KIAA0556
Uncharacterized coiled-coil domain-containing protein KIAA1984
cAMP-dependent protein kinase regulatory subunit
Kv channel-interacting protein 4
KH domain-containing, RNA-binding, signal transduction-associated protein 2
Kinesin-like protein KIF9
Kinesin-like protein KIFC3
Kynurenine 3-monooxygenase
KRR1 small subunit processome component homolog
Ribosomal protein S6 kinase alpha-5
Large neutral amino acids transporter small subunit 4
Protein lin-10
Lipase maturation factor 1
Lon protease homolog 2, peroxisomal
Leucine-rich repeat and death domain-containing protein 1
Leucine-rich repeat flightless-interacting protein 2
Putative helicase Mov10l1
Beta-mannosidase
Beta-mannosidase
Metallo-beta-lactamase domain-containing protein 1
Multidrug resistance protein 3
Mechanosensory protein 2
Multiple epidermal growth factor-like domains protein 10
Malignant fibrous histiocytoma-amplified sequence 1
Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
Mitochondrial genome maintenance exonuclease 1
Alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase B
Microtubule-associated proteins 1A/1B light chain 3C
Carbohydrate-responsive element-binding protein
Mitochondrial inner membrane organizing system protein 1
Monocarboxylate transporter 14
Dual specificity mitogen-activated protein kinase kinase 4
Melatonin receptor type 1B-B
Monocarboxylate transporter 12-B
Myotubularin-related protein 3
Crossover junction endonuclease MUS81
Muscle, skeletal receptor tyrosine protein kinase
Probable exonuclease mut-7 homolog
Myb-like protein X

Myosin-VIIa
Myosin heavy chain, clone 203
NEDD4-binding protein 2-like 2
Sodium-dependent nutrient amino acid transporter 1
Sodium/calcium exchanger 3
Neuron navigator 3
Kinetochore protein NDC80 homolog
Neudesin
NHL repeat-containing protein 2
Nuclear hormone receptor family member nhr-48
Nischarin
Non-specific lipid-transfer protein
Nephrocystin-3
Nitrogen permease regulator 2-like protein
Cytosolic Fe-S cluster assembly factor nubp1
8-oxo-dGDP phosphatase NUDT18
Orcokinin peptides type B
Serine/threonine-protein kinase OSR1
Polyadenylate-binding protein 1
Poly [ADP-ribose] polymerase 14
Platelet-derived growth factor receptor beta
Probable phosphoglucomutase-2
PHD finger protein 3
Phytanoyl-CoA dioxygenase domain-containing protein 1
Phytanoyl-CoA dioxygenase domain-containing protein 1
Pleckstrin homology domain-containing family G member 7
Fibrocystin-L
Serine/threonine-protein kinase PLK2
Serine/threonine-protein kinase PLK4
Plexin-B1
Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform
Intestinal-type alkaline phosphatase
Alkaline phosphatase, tissue-nonspecific isozyme
Alkaline phosphatase, tissue-nonspecific isozyme
PQ-loop repeat-containing protein 3
PR domain zinc finger protein 16
Prominin-1-A
Patched domain-containing protein 3
Receptor-type tyrosine-protein phosphatase eta
Receptor-type tyrosine-protein phosphatase mu
Peroxidasin
R3H domain-containing protein 2
Ras-related protein Rab-1
Double-strand-break repair protein rad21 homolog
Ribonucleoprotein PTB-binding 2
Putative Rab-43-like protein ENSP00000330714
Pre-mRNA-splicing factor RBM22
Retinol dehydrogenase 7
RAD52 motif-containing protein 1
Ras-GEF domain-containing family member 1B
Regulator of G-protein signaling 22
Rho-related GTP-binding protein RhoU
Retinaldehyde-binding protein 1
39S ribosomal protein L4, mitochondrial
39S ribosomal protein L53, mitochondrial
RING finger protein 207
RING finger protein 24
Nuclear receptor ROR-beta

Rap guanine nucleotide exchange factor 4
Ribonuclease P protein subunit p40
Protein RRP5 homolog
Probable RNA-directed DNA polymerase from transposon BS
Electrogenic sodium bicarbonate cotransporter 1
Sodium- and chloride-dependent betaine transporter
Sodium- and chloride-dependent neutral and basic amino acid transporter B(0+)
Sodium- and chloride-dependent glycine transporter 2
Sodium- and chloride-dependent taurine transporter
Sideroflexin-2
Small integral membrane protein 13
Mothers against decapentaplegic homolog 6
Mothers against decapentaplegic homolog 6
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1
Single-strand selective monofunctional uracil DNA glycosylase
Sperm-associated antigen 6
Survival of motor neuron-related-splicing factor 30
Secretion-regulating guanine nucleotide exchange factor
Sporozoite surface protein 2
Sulfotransferase 1A1
Nesprin-1
Transmembrane protein 184B
TGF-beta-activated kinase 1 and MAP3K7-binding protein 3
Very-long-chain enoyl-CoA reductase
Telomerase protein component 1
Testis-expressed sequence 11 protein
Metalloprotease TIKI1
Mitochondrial import inner membrane translocase subunit tim16
Translation machinery-associated protein 16
Transmembrane protein 5
Troponin C, isoform 1
Targeting protein for Xklp2-B
E3 ubiquitin-protein ligase TRIM33
Tripartite motif-containing protein 59
Transient receptor potential cation channel subfamily V member 5
Thrombospondin-4
UPF0184 protein
Ubiquitin carboxyl-terminal hydrolase 15
LIM domain-containing protein unc-97
Intracellular protein transport protein USO1
Williams-Beuren syndrome chromosomal region 27 protein
WD repeat and HMG-box DNA-binding protein 1
WD repeat-containing protein 35
WW domain-containing oxidoreductase
Xanthine dehydrogenase
Xanthine dehydrogenase/oxidase
Xanthine dehydrogenase
Uncharacterized protein STK_23830
Putative ankyrin repeat protein RF_0381
Uncharacterized protein DDB_G0292642
Zinc finger and BTB domain-containing protein 17

Palmitoyltransferase ZDHHC2
Zinc finger protein 37
Zinc finger protein 62

Zinc finger protein 718
Zinc finger protein 43
DNA annealing helicase and endonuclease ZRANB3

Appendix A4

BP terms	Count	PValue	FE
GO:0030514~negative regulation of BMP signaling pathway	4	0.00	17.56
GO:0030510~regulation of BMP signaling pathway	4	0.00	12.55
GO:0050878~regulation of body fluid levels	6	0.01	4.88
GO:0048232~male gamete generation	13	0.02	2.15
GO:0007283~spermatogenesis	13	0.02	2.15
GO:0009166~nucleotide catabolic process	5	0.02	4.77
GO:0006012~galactose metabolic process	3	0.02	13.17
GO:0006997~nucleus organization	5	0.02	4.57
GO:0009187~cyclic nucleotide metabolic process	5	0.02	4.39
GO:0007600~sensory perception	13	0.02	2.04
GO:0046677~response to antibiotic	4	0.03	5.85
GO:0034656~nucleobase, nucleoside and nucleotide catabolic process	5	0.03	4.07
GO:0034655~nucleobase, nucleoside, nucleotide and nucleic acid catabolic process	5	0.03	4.07
GO:0006884~cell volume homeostasis	3	0.04	9.41
GO:0043154~negative regulation of caspase activity	3	0.04	9.41
GO:0007601~visual perception	8	0.04	2.44
GO:0050953~sensory perception of light stimulus	8	0.05	2.41
GO:0010466~negative regulation of peptidase activity	3	0.05	8.23
GO:0044270~nitrogen compound catabolic process	5	0.06	3.43
GO:0009143~nucleoside triphosphate catabolic process	3	0.06	7.32
GO:0016567~protein ubiquitination	7	0.06	2.48
GO:0009791~post-embryonic development	11	0.07	1.89
GO:0007155~cell adhesion	19	0.07	1.54
GO:0007599~hemostasis	4	0.07	4.18
GO:0022610~biological adhesion	19	0.07	1.53
GO:0050954~sensory perception of mechanical stimulus	6	0.07	2.69
GO:0034728~nucleosome organization	5	0.07	3.14
GO:0007517~muscle organ development	9	0.08	2.00
GO:0009968~negative regulation of signal transduction	8	0.08	2.12
GO:0050890~cognition	13	0.08	1.69
GO:0045786~negative regulation of cell cycle	5	0.09	2.97
CC terms	**Count**	**PValue**	**FE**
GO:0030663~COPI coated vesicle membrane	4	0.02	6.16
GO:0030126~COPI vesicle coat	4	0.02	6.16
GO:0000267~cell fraction	28	0.03	1.51
GO:0005625~soluble fraction	10	0.03	2.24
GO:0030137~COPI-coated vesicle	4	0.04	5.39
GO:0044433~cytoplasmic vesicle part	8	0.04	2.53
GO:0000786~nucleosome	4	0.04	5.07
GO:0005626~insoluble fraction	23	0.06	1.49
GO:0030662~coated vesicle membrane	6	0.06	2.81
GO:0030120~vesicle coat	5	0.06	3.26
GO:0032993~protein-DNA complex	5	0.06	3.26
GO:0012506~vesicle membrane	8	0.06	2.24
GO:0005769~early endosome	6	0.06	2.75
GO:0030659~cytoplasmic vesicle membrane	7	0.08	2.32
GO:0005578~proteinaceous extracellular matrix	9	0.09	1.96
GO:0030660~Golgi-associated vesicle membrane	4	0.09	3.75
GO:0005576~extracellular region	26	0.10	1.36
MF terms	**Count**	**PValue**	**FE**
GO:0043167~ion binding	136	0.00	1.24
GO:0043169~cation binding	133	0.00	1.22
GO:0046332~SMAD binding	5	0.00	7.68
GO:0046872~metal ion binding	131	0.00	1.22
GO:0004035~alkaline phosphatase activity	3	0.01	16.12
GO:0005509~calcium ion binding	33	0.02	1.52
GO:0070410~co-SMAD binding	3	0.02	12.90
GO:0070412~R-SMAD binding	3	0.02	12.90
GO:0005328~neurotransmitter:sodium symporter activity	4	0.04	5.06

GO:0005179~hormone activity	3	0.05	8.06
GO:0015293~symporter activity	9	0.05	2.20
GO:0015171~amino acid transmembrane transporter activity	5	0.05	3.47
GO:0005518~collagen binding	3	0.06	7.17
GO:0005326~neurotransmitter transporter activity	4	0.06	4.30
GO:0015298~solute:cation antiporter activity	4	0.06	4.30
GO:0004879~ligand-dependent nuclear receptor activity	5	0.07	3.16
GO:0005343~organic acid:sodium symporter activity	3	0.09	5.86
GO:0015297~antiporter activity	5	0.09	2.90
KEGG pathways	**Count**	**PValue**	**FE**
rno04020:Calcium signaling pathway	5	0.01	6.28
rno04720:Long-term potentiation	4	0.01	8.37
rno04114:Oocyte meiosis	4	0.01	8.37
bta00790:Folate biosynthesis	2	0.08	25.11
rno04270:Vascular smooth muscle contraction	3	0.08	6.28
rno00230:Purine metabolism	3	0.08	6.28
rno04912:GnRH signaling pathway	3	0.09	5.79
rno04540:Gap junction	3	0.09	5.79
rno04020:Calcium signaling pathway	5	0.01	6.28
Panther pathways	**Count**	**PValue**	**FE**
P00057:Wnt signaling pathway	14	0.00	2.66
P00012:Cadherin signaling pathway	7	0.03	2.82

Appendix A5

BP terms	Count	PValue	FE
GO:0007156~homophilic cell adhesion	14	0.00	3.38
GO:0006412~translation	37	0.00	1.88
GO:0016337~cell-cell adhesion	20	0.00	2.44
GO:0006414~translational elongation	12	0.00	3.03
GO:0007155~cell adhesion	40	0.00	1.60
GO:0022610~biological adhesion	40	0.00	1.60
GO:0051258~protein polymerization	9	0.01	3.16
GO:0006576~biogenic amine metabolic process	10	0.01	2.59
GO:0051384~response to glucocorticoid stimulus	7	0.02	3.04
GO:0007005~mitochondrion organization	12	0.03	2.07
GO:0034621~cellular macromolecular complex subunit organization	27	0.03	1.52
GO:0042402~biogenic amine catabolic process	4	0.03	5.43
GO:0030855~epithelial cell differentiation	9	0.03	2.33
GO:0032535~regulation of cellular component size	17	0.04	1.71
GO:0031960~response to corticosteroid stimulus	7	0.04	2.72
GO:0055114~oxidation reduction	53	0.04	1.29
GO:0045022~early endosome to late endosome transport	4	0.04	4.83
GO:0009143~nucleoside triphosphate catabolic process	4	0.04	4.83
GO:0051231~spindle elongation	4	0.04	4.83
GO:0000022~mitotic spindle elongation	4	0.04	4.83
GO:0009310~amine catabolic process	10	0.05	2.09
GO:0008202~steroid metabolic process	15	0.05	1.73
GO:0034622~cellular macromolecular complex assembly	23	0.06	1.47
GO:0022411~cellular component disassembly	6	0.06	2.72
GO:0015904~tetracycline transport	3	0.07	6.52
GO:0042891~antibiotic transport	3	0.07	6.52
GO:0015893~drug transport	4	0.07	3.95
GO:0070585~protein localization in mitochondrion	4	0.07	3.95
GO:0007338~single fertilization	4	0.07	3.95
GO:0042219~cellular amino acid derivative catabolic process	4	0.07	3.95
GO:0006626~protein targeting to mitochondrion	4	0.07	3.95
GO:0007218~neuropeptide signaling pathway	6	0.07	2.61
GO:0007017~microtubule-based process	28	0.07	1.38
GO:0001503~ossification	7	0.08	2.31
GO:0006979~response to oxidative stress	11	0.08	1.78
GO:0007051~spindle organization	8	0.09	2.07
GO:0060348~bone development	7	0.09	2.24
GO:0007272~ensheathment of neurons	5	0.09	2.86
GO:0008366~axon ensheathment	5	0.09	2.86
GO:0006022~aminoglycan metabolic process	5	0.09	2.86
GO:0051341~regulation of oxidoreductase activity	4	0.09	3.62
GO:0007229~integrin-mediated signaling pathway	6	0.10	2.42
GO:0032271~regulation of protein polymerization	7	0.10	2.17
GO:0050999~regulation of nitric-oxide synthase activity	3	0.10	5.43
GO:0035036~sperm-egg recognition	3	0.10	5.43
GO:0032768~regulation of monooxygenase activity	3	0.10	5.43
GO:0007339~binding of sperm to zona pellucida	3	0.10	5.43
GO:0046620~regulation of organ growth	3	0.10	5.43
CC terms	**Count**	**PValue**	**FE**
GO:0033279~ribosomal subunit	14	0.00	2.97
GO:0022626~cytosolic ribosome	9	0.00	4.36
GO:0005840~ribosome	24	0.00	2.11
GO:0022625~cytosolic large ribosomal subunit	6	0.00	6.79
GO:0030529~ribonucleoprotein complex	49	0.00	1.55
GO:0044421~extracellular region part	33	0.00	1.74
GO:0015934~large ribosomal subunit	9	0.00	3.52
GO:0005576~extracellular region	59	0.00	1.45
GO:0005829~cytosol	73	0.01	1.35
GO:0019866~organelle inner membrane	28	0.02	1.57

GO:0005615~extracellular space	16	0.03	1.83
GO:0031012~extracellular matrix	18	0.03	1.73
GO:0005743~mitochondrial inner membrane	26	0.03	1.54
GO:0044445~cytosolic part	12	0.05	1.88
GO:0005578~proteinaceous extracellular matrix	16	0.06	1.65
GO:0031966~mitochondrial membrane	30	0.07	1.37
MF terms	**Count**	**PValue**	**FE**
GO:0003735~structural constituent of ribosome	22	0.00	2.69
GO:0005198~structural molecule activity	43	0.00	1.91
GO:0030246~carbohydrate binding	21	0.00	2.22
GO:0005529~sugar binding	14	0.00	2.50
GO:0015238~drug transporter activity	6	0.01	4.67
GO:0019843~rRNA binding	6	0.01	4.08
GO:0048029~monosaccharide binding	6	0.01	4.08
GO:0016782~transferase activity, transferring sulfur-containing groups	9	0.01	2.72
GO:0008146~sulfotransferase activity	8	0.02	2.90
GO:0005509~calcium ion binding	57	0.02	1.33
GO:0015298~solute:cation antiporter activity	6	0.03	3.27
GO:0004222~metalloendopeptidase activity	11	0.04	2.07
GO:0008237~metallopeptidase activity	17	0.04	1.71
GO:0005518~collagen binding	4	0.04	4.84
GO:0001968~fibronectin binding	3	0.04	8.17
GO:0015297~antiporter activity	8	0.05	2.36
GO:0004175~endopeptidase activity	24	0.06	1.45
GO:0015300~solute:solute antiporter activity	7	0.07	2.38
GO:0015307~drug:hydrogen antiporter activity	3	0.07	6.54
GO:0004062~aryl sulfotransferase activity	3	0.07	6.54
GO:0005220~inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity	3	0.07	6.54
GO:0015520~tetracycline:hydrogen antiporter activity	3	0.07	6.54
GO:0008095~inositol-1,4,5-trisphosphate receptor activity	3	0.07	6.54
GO:0005506~iron ion binding	26	0.07	1.40
GO:0001871~pattern binding	6	0.08	2.51
GO:0030247~polysaccharide binding	6	0.08	2.51
GO:0043167~ion binding	233	0.09	1.07
GO:0004551~nucleotide diphosphatase activity	3	0.10	5.45
GO:0005537~mannose binding	3	0.10	5.45
GO:0042895~antibiotic transporter activity	3	0.10	5.45
GO:0008493~tetracycline transporter activity	3	0.10	5.45
KEGG pathways	**Count**	**PValue**	**FE**
dme03010:Ribosome	4	0.01	8.64
hsa03010:Ribosome	3	0.07	6.48
hsa01040:Biosynthesis of unsaturated fatty acids	3	0.07	6.48
mmu05016:Huntington's disease	5	0.09	2.84

Appendix B1

Category and gene identity (C.gigas proteins)	*Num. contigs*
PATTERN RECOGNITION RECEPTORS (PRRs)	
LECTINS	
Collectin-11	1
Collectin-12	3
Contactin	6
C-type lectin domain family 10 member A	2
C-type lectin domain family 5 member A	1
C-type lectin isoform 1	1
Endoplasmic reticulum lectin 1	1
Lectin	1
Hepatic lectin	1
Fucolectin	1
Fucolectin-5	1
Galectin-4	4
Galectin-9	3
L-rhamnose-binding lectin CSL3	2
C-type mannose receptor 2	1
Macrophage mannose receptor 1	13
Cation-dependent mannose-6-phosphate receptor	5
Perlucin	5
Calreticulin	1
Calnexin	1
Malectin	1
Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	22
Ficolin-2	2
Ficolin-1	1
TOLL-LIKE RECEPTORS	
Toll-like receptor 2	1
Toll-like receptor 2 type-2	2
Toll-like receptor 3	3
Toll-like receptor 4	2
Toll-like receptor 6	1
Toll-like receptor 13	1
Protein toll	6
GLUCAN BINDING PROTEINS	
Beta-1,3-glucan-binding protein	2
PEPTIDOGLYCAN RECOGNITION PROTEINS	
Peptidoglycan-recognition protein SC2	1
Peptidoglycan recognition protein	1
SCAVENGER RECEPTORS	
Endothelial cells scavenger receptor	9
IMMUNE EFFECTORS	
HEAT-SHOCK PROTEINS	
60 kDa heat shock protein, mitochondrial	1
Activator of 90 kDa heat shock protein ATPase-like protein 1	2
Heat shock 70 kDa protein 12A	21
Heat shock 70 kDa protein 12B	17
Heat shock 70 kDa protein 13	1
Heat shock 70 kDa protein 14	1
Heat shock protein 70 B2	11
Heat shock protein 75 kDa, mitochondrial, partial	2
Heat shock protein beta-1	1
Heat shock protein beta-11	1
Heat shock protein HSP 90-alpha 1	2
Hsp70-binding protein 1	1
HSPB1-associated protein 1	1
ftsj methyltransferase domain-containing protein 1	2
S-adenosyl-L-methionine-dependent methyltransferase ftsjd2	1
DNAJ-like protein subfamily A members	2
DNAJ-like protein subfamily B members	16
DNAJ-like protein subfamily C members	36
ANTIMICROBIAL PEPTIDES	
Big defensin	1
ANTIOXIDANT ENZYMES	
Catalase	4
Ceruloplasmin	6
Glutathione peroxidase	1
Glutathione peroxidase 1	1
Glutathione peroxidase 7	1
Glutathione S-transferase A	1
Glutathione S-transferase C-terminal domain-containing protein	2
Glutathione S-transferase kappa 1	1
Glutathione S-transferase omega-1	3
Glutathione S-transferase P 1	1
Glutathione S-transferase theta-1	2
Glutathione S-transferase Y1, partial	1
Microsomal glutathione S-transferase 1	2
Microsomal glutathione S-transferase 2	2
Microsomal glutathione S-transferase 3	2
Prostaglandin reductase 1	4
SIGNAL TRANSDUCTION	
TRAFS AND MyD88	
E3 ubiquitin-protein ligase TRAF7	6
TNF receptor-associated factor 2	2
TNF receptor-associated factor 3	19
TNF receptor-associated factor 6	1
TRAF and TNF receptor-associated protein	2
Myeloid differentiation primary response protein MyD88	3
PROTEASES AND PROTEASE INHIBITORS	
Cathepsin B	2
Cathepsin F	2
Cathepsin L	1
Cathepsin Z	3
A disintegrin and metalloproteinase with thrombospondin motifs 16	5
Disintegrin and metalloproteinase with thrombospondin motifs 2	2
Disintegrin and metalloproteinase with thrombospondin motifs 6	2
Disintegrin and metalloproteinase with thrombospondin motifs 7	2
Disintegrin and metalloproteinase with thrombospondin motifs 9	5
Disintegrin and metalloproteinase domain-containing protein 10	1
Zinc metalloproteinase nas-15	1
Zinc metalloproteinase nas-36	1
Zinc metalloproteinase nas-37	1

Matrix metalloproteinase-17	1
Matrix metalloproteinase-19	6
Cysteine protease ATG4B	1
Putative serine protease F56F10.1	1
Signal peptide peptidase-like 2B	1
Signal peptide peptidase-like 3	1
Thymus-specific serine protease	2
Reelin	6
Agrin	4
Kazal-type serine protease inhibitor domain-containing protein 1	1
Metalloproteinase inhibitor 1	1
Metalloproteinase inhibitor 3	1
Legumain	1
Serpin B8	2
APOPTOSIS	
Apoptosis 1 inhibitor	6
Apoptosis 2 inhibitor	1
Apoptosis inhibitor 5	1
Inhibitor of apoptosis protein	9
Apoptosis inhibitor IAP	5
Baculoviral IAP repeat-containing protein 2	1
Baculoviral IAP repeat-containing protein 3	2
Baculoviral IAP repeat-containing protein 5	1
Baculoviral IAP repeat-containing protein 6	12
Baculoviral IAP repeat-containing protein 7	2
Baculoviral IAP repeat-containing protein 7-A	21
BAG family molecular chaperone regulator 1	2
BAG family molecular chaperone regulator 2	1
BAG family molecular chaperone regulator 4	1
Regucalcin	9
TP53-regulated inhibitor of apoptosis 1	1
Arrestin domain-containing protein 2	3
Arrestin domain-containing protein 3	9
Beta-arrestin-1	2
Apoptosis-inducing factor 1, mitochondrial	1
Apoptosis-inducing factor 3	2
Apoptosis-stimulating of p53 protein 1	4
Apoptosis regulator BAX	3
Programmed cell death protein 2	2
Programmed cell death protein 2-like protein	1
Programmed cell death protein 4	1
Programmed cell death protein 5	2
Programmed cell death protein 6	2
Programmed cell death 6-interacting protein	2
Programmed cell death protein 7	1
Programmed cell death protein 10	1
Autophagy-related protein 2-like protein A	9
Autophagy-related protein 2-like protein B	5
Autophagy-related protein 3	1
Autophagy-related protein 7	2
Autophagy-related protein 9A	2
Autophagy-related protein 10	1
Autophagy-related protein 13	1
Autophagy-related protein 16-1	1
Beclin-1	1
Bcl-2-like protein 1	2
Bcl-2-like protein 13	3
B-cell lymphoma 3-encoded protein	2
B-cell lymphoma 6-like protein	2
B-cell CLL/lymphoma 7 protein family member A	2
B-cell lymphoma/leukemia 11A	1
B-cell lymphoma/leukemia 11B	2
BCL-6 corepressor	2
BCL9-like protein	3
Putative Bcl-2-like protein antagonist/killer 2	1
Caspase	5
Caspase-2	6
Caspase-3	1
Caspase-6	5
Caspase-7	13
Caspase-8	3
Caspase-10	1
CASP8 and FADD-like apoptosis regulator	1
CASP8-associated protein 2	1
Cell death regulator Aven	1
Cell division cycle and apoptosis regulator protein 1	1
APAF1-interacting protein	1
c-myc promoter-binding protein	2
c-myc-binding protein	1
Death domain-associated protein 6	1
Death domain-containing protein CRADD	2
Death effector domain-containing protein	1
Death-associated protein 1	1
Death-associated protein kinase 1	1
Macrophage erythroblast attacher	1
MAP kinase-activating death domain protein	5
Ankyrin repeat and death domain-containing protein 1A	2
Apoptosis regulator R1	2
Ski oncogene	2
Growth arrest and DNA-damage-inducible proteins GADD45 gamma	5
Apoptosis regulatory protein Siva	1
NAD-dependent deacetylase sirtuin-1	3
Tumor proteins	6
COMPLEMENT AND C1q-LIKE PROTEINS	
complement C1q tumor necrosis factor-related protein 6	1
complement C1q-like protein 4	4
complement C3	1
complement component 1 Q subcomponent-binding protein, mitochondrial	1
Heavy metal-binding protein HIP	2
CELL SURFACE RECEPTOR AND CELL ADHESION	
CELL SURFACE RECEPTORS	

Angiopoietin-2	3
Angiopoietin-4	2
B-cell receptor CD22	1
B-cell receptor-associated protein 31	1
Macrophage colony-stimulating factor 1 receptor 2	2
Protein SET	2
Stabilin-1	1
Stabilin-2	2
TGF-beta receptor type-1	1
Thrombospondin-1	1
Thrombospondin type-1 domain-containing protein 4	2
Thrombospondin-3b	1
CELL ADHESION	
Carcinoembryonic antigen-related cell adhesion molecules	6
Hemicentin-1	36
Scavenger receptor class F member 2	2
Integrins (alpha and beta)	16
CYTOKINE-RELATED MOLECULES	
Allograft inflammatory factor 1	2
Alpha-2-macroglobulin receptor-associated protein	1
BTB/POZ domain-containing protein TNFAIP1	1
C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8	2
Gamma-interferon-inducible lysosomal thiol reductase	1
Inhibitor of nuclear factor kappa-B kinase subunit alpha	2
Inhibitor of nuclear factor kappa-B kinase subunit epsilon	3
Interferon regulatory factor 2	4
Interferon regulatory factor 8	2
interferon-induced helicase C domain-containing protein 1	2
interferon-induced protein 44	1
interferon-induced protein 44-like protein	16
interleukin enhancer-binding factor 2-like protein	1
interleukin-1 receptor-associated kinase 1	2
interleukin-1 receptor-associated kinase 4	2
interleukin-17 receptor D	4
interleukin-17D	2
Lipopolysaccharide-induced tumor necrosis factor-alpha factor-like protein	2
Nuclear factor interleukin-3-regulated protein	2
SPRY domain-containing protein 3	1
SPRY domain-containing SOCS box protein 3	2
SPRY domain-containing SOCS box protein 4	1
TNFAIP3-interacting protein 2	2
tumor necrosis factor ligand superfamily member 10	2
tumor necrosis factor ligand superfamily member 11	1
tumor necrosis factor receptor superfamily member 16	1
tumor necrosis factor receptor superfamily member 1B	1
tumor necrosis factor, alpha-induced protein 3	2
tumor necrosis factor, alpha-induced protein 8	2
OTHER	
EF-hand calcium-binding domain-containing protein 2	1
EF-hand calcium-binding domain-containing protein 6	1
EF-hand calcium-binding domain-containing protein 7	2
HemAgglutinin/amebocyte aggregation factor	6
LysM and putative peptidoglycan-binding domain-containing protein 2	1
LysM and putative peptidoglycan-binding domain-containing protein 3	1
Lysosome-associated membrane glycoprotein 1	1
Macrophage migration inhibitory factor	1
Metallothionein	2
Monocyte to macrophage differentiation protein	2
Multiple coagulation factor deficiency protein 2-like protein	1
N-terminal EF-hand calcium-binding protein 1	1
Prothrombin, partial	1
Soma ferritin	3
Stress response protein nhaX	3
T-cell immunomodulatory protein	2
Transcription factor AP-1	1
Proto-oncogene tyrosine-protein kinase FER	1
Proto-oncogene tyrosine-protein kinase FYN	1
B-Raf proto-oncogene serine/threonine-protein kinase	1
Caveolin-1	1
Clathrin heavy chain 1	3
Dual oxidase	7
Dual oxidase 2	3
Glutaredoxin-C6	2
Glutaredoxin	1
Glutaredoxin-3	3
Nucleoredoxin	4
Hypoxia-inducible factor 1-alpha inhibitor	1
Hypoxia-inducible factor 1 alpha	3
Plexin-A4	6
Tripartite motif containing (TRIM) proteins	99
Thioredoxin	1
Tetraspanins	29
Syntenin-1	1

Appendix B2

Annotation		in vivo control library	
***C. gigas* ID**	***C. gigas* gene description**	**Total gene reads**	**RPKM**
CGI_10022730	Actin	2513347	83345.7
CGI_10022730	Actin	386694	27604.5
na	na	814491	25171.5
CGI_10024999	Tubulin alpha-1C chain	310666	15610.5
CGI_10012474	Elongation factor 1-alpha	380027	11676.5
CGI_10019752	Tubulin beta chain	95595	8318.5
na	na	103014	7699.5
CGI_10025348	Matrix metalloproteinase-19	161719	7458.1
CGI_10027591	Soma ferritin	82011	6754.3
CGI_10023099	60S ribosomal protein L30	46723	6731.2
CGI_10001493	40S ribosomal protein S24	61021	6284.7
CGI_10013084	60S ribosomal protein L27a	65033	6021.7
CGI_10019752	Tubulin beta chain	122635	5836.3
CGI_10027605	Allograft inflammatory factor 1	104954	5623.4
CGI_10009952	Blastula protease 10	99002	5153.9
CGI_10010647	Heat shock protein 70 B2	136690	5148.6
CGI_10003625	Poly [ADP-ribose] polymerase 1	41480	4665.1
CGI_10004554	60S ribosomal protein L24	64067	4539.5
CGI_10025548	hypothetical protein CGI_10025548	85677	4450.5
CGI_10007104	40S ribosomal protein S8	66992	4394.2
CGI_10010647	Heat shock protein 70 B2	125247	4306.1
CGI_10022234	60S ribosomal protein L13	25884	3869.3
CGI_10000147	60S ribosomal protein L21	42731	3671.5
CGI_10024998	Tubulin alpha-1C chain	21410	3659.3
CGI_10025348	Matrix metalloproteinase-19	70352	3638.6
CGI_10015910	Transforming growth factor-beta-induced protein ig-h3	172319	3378.8
CGI_10018366	na	46733	3243.8
CGI_10003072	60S ribosomal protein L37a	26909	3192.0
na	na	31565	3142.4
na	na	42030	3069.3
CGI_10014767	60S ribosomal protein L23a	51599	3064.2
CGI_10021165	60S ribosomal protein L18, partial	44958	3062.7
CGI_10018105	Hemicentin-1	44788	3051.1
CGI_10008425	hypothetical protein CGI_10008425	40816	3022.3
na	na	47354	2959.5
CGI_10025934	Guanine nucleotide-binding protein subunit beta-2-like 1	66679	2715.3
CGI_10012002	40S ribosomal protein S10	28018	2506.7
CGI_10014112	Y-box factor-like protein, partial	89885	2447.9
CGI_10013164	Tropomyosin	64731	2447.9
CGI_10026412	60S ribosomal protein L18a	30345	2383.9
CGI_10001571	Elongation factor 1-gamma	39565	2340.8
CGI_10022820	na	36594	2317.4
CGI_10015715	na	21761	2280.6
CGI_10006637	Cdc42-like protein	46946	2256.7
CGI_10020449	40S ribosomal protein S7	29244	2245.6
CGI_10021672	na	40459	2238.2
CGI_10021852	40S ribosomal protein SA	53323	2232.4
CGI_10003531	T-cell acute lymphocytic leukemia protein 1	61572	2157.3
CGI_10022083	Matrix metalloproteinase-19	56286	2155.9
na	na	31179	2148.6
na	na	37246	2129.6
CGI_10028140	40S ribosomal protein S23	22030	2063.9
CGI_10006923	Ceruloplasmin	95728	2060.9
CGI_10005445	X-box-binding protein 1	50156	2046.0
CGI_10022058	60S ribosomal protein L7a	52347	2006.6
na	na	27797	1990.3
CGI_10028514	na	15062	1967.8

CGI_10010474	40S ribosomal protein S28	11010	1966.4
CGI_10024065	60S ribosomal protein L23	20046	1939.0
CGI_10010163	Tubulin beta chain	20427	1936.5
CGI_10020592	Proline-rich transmembrane protein 1	33316	1930.4
CGI_10012475	transport protein Sec61 subunit gamma	23667	1919.4
CGI_10020922	Heterogeneous nuclear ribonucleoprotein A/B	47222	1918.1
CGI_10020628	Metalloproteinase inhibitor 1	95242	1890.0
na	na	25101	1867.3
na	na	14282	1855.7
CGI_10012808	60S ribosomal protein L14	22856	1786.7
na	na	16580	1780.7
CGI_10000595	60S ribosomal protein L12	24832	1759.5
CGI_10009326	60S acidic ribosomal protein P1	15371	1753.5
na	na	15889	1745.8
CGI_10025208	Putative ferric-chelate reductase 1-like protein	42587	1734.2
CGI_10028812	60S ribosomal protein L26	18264	1691.1
CGI_10000077	60S ribosomal protein L17	22654	1690.6
CGI_10027192	60S ribosomal protein L5	34372	1681.1
CGI_10019706	40S ribosomal protein S3a	40992	1667.9
CGI_10022234	60S ribosomal protein L13	13401	1651.3
CGI_10026160	Actophorin	57092	1629.3
CGI_10020975	na	37246	1617.6
CGI_10014579	Putative 60S ribosomal protein L37-A	14033	1604.7
CGI_10023721	Murinoglobulin-2	22432	1603.7
na	na	12852	1596.0
CGI_10025410	Fascin	60506	1580.1
CGI_10015747	B2 bradykinin receptor	28014	1566.1
CGI_10012822	Blastula protease 10	27070	1564.7
CGI_10015917	Calreticulin	59432	1532.8
CGI_10011577	60S acidic ribosomal protein P2	14837	1524.8
CGI_10008101	na	19059	1494.8
CGI_10006874	Pregnancy zone protein	89218	1494.4
CGI_10013164	Tropomyosin	33980	1449.3
CGI_10017178	hypothetical protein CGI_10017178	57699	1440.5
CGI_10017178	hypothetical protein CGI_10017178	26275	1438.5
CGI_10025378	Calponin-2	43122	1358.2
CGI_10025348	Matrix metalloproteinase-19	42305	1351.2
CGI_10026239	Ornithine decarboxylase antizyme 1	32339	1348.0
CGI_10022808	Cyclic AMP-dependent transcription factor ATF-5	40106	1342.1
CGI_10019268	Transcription factor BTF3-like protein 4	25734	1333.9
CGI_10010682	60S acidic ribosomal protein P0	16254	1322.7
na	na	9113	1308.9
CGI_10016741	40S ribosomal protein S4, X isoform	37905	1301.3
na	na	8426	1300.3
CGI_10025401	ADP,ATP carrier protein	30139	1297.1
CGI_10026636	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-7	17567	1296.7
CGI_10006923	Ceruloplasmin	27560	1294.8
na	na	9567	1292.4
na	na	67539	1290.8
CGI_10027081	hypothetical protein CGI_10027081	31977	1267.5
CGI_10018335	na	50521	1265.3
CGI_10003923	60S ribosomal protein L38, partial	10808	1257.0
CGI_10010679	hypothetical protein CGI_10010679	16670	1240.1
CGI_10017187	na	8394	1216.6
na	na	7124	1208.9
CGI_10008615	Embryonic polyadenylate-binding protein B	37581	1190.0
CGI_10020851	Putative ribosome biogenesis protein RLP24	20326	1169.2
CGI_10009651	40S ribosomal protein S5	22285	1167.8
CGI_10003396	na	13204	1151.1
CGI_10010682	60S acidic ribosomal protein P0	13816	1137.9
na	na	18119	1126.5

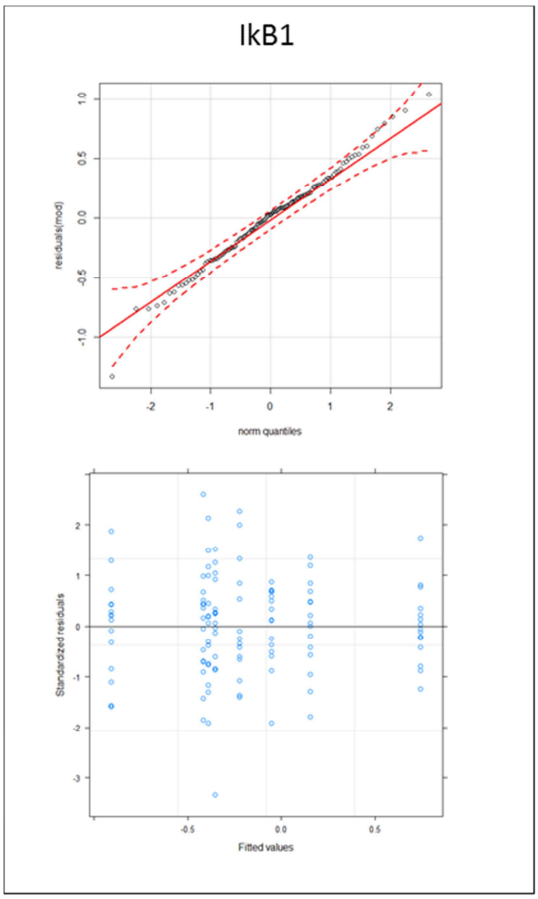
CGI_10027951	Ras-related protein RAP-1b	30361	1124.9
CGI_10010805	Prostaglandin reductase 1	23767	1103.6
CGI_10015086	hypothetical protein CGI_10015086	45173	1103.5
CGI_10020662	Adenosylhomocysteinase A	36215	1098.6
CGI_10020961	40S ribosomal protein S16	13539	1072.5
CGI_10009700	hypothetical protein CGI_10009700	10144	1070.2
na	na	6524	1069.1
na	na	15164	1063.4
na	na	12292	1050.5
CGI_10025609	Actin-related protein 2/3 complex subunit 1A	18679	1039.4
CGI_10007702	40S ribosomal protein S19	8651	1036.5
na	na	12177	1008.1
CGI_10018930	Tubulin alpha-3 chain	32163	999.2

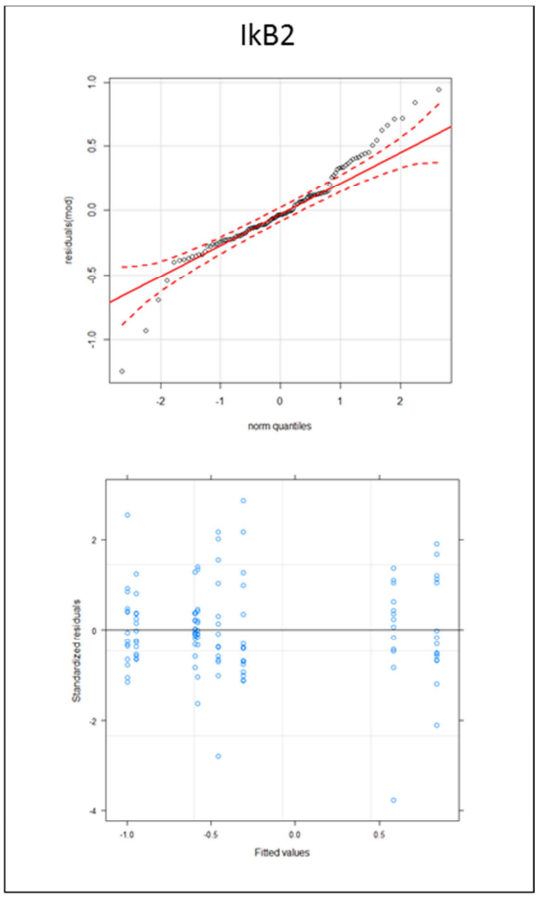
Appendix B3

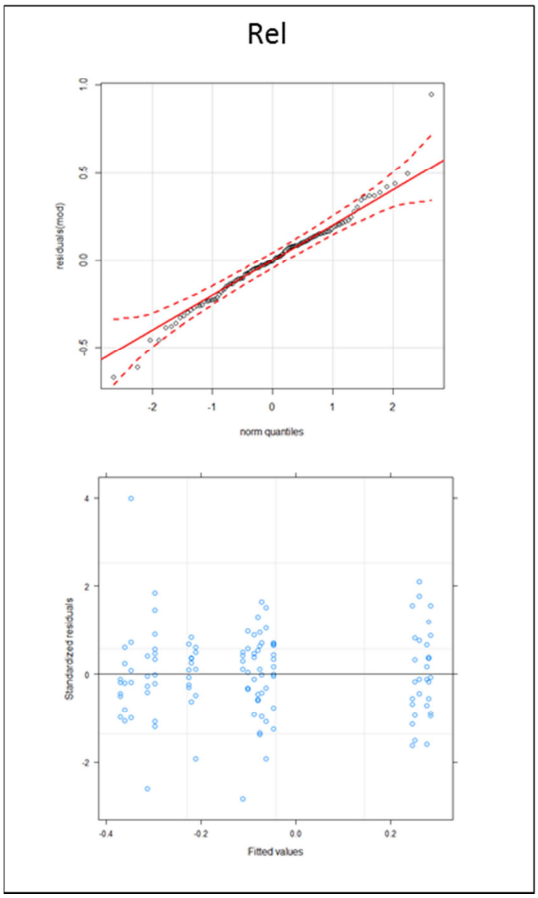
BP terms	Count	PValue	FE
GO:0007155~cell adhesion	17	0.00	4.03
GO:0022610~biological adhesion	17	0.00	4.03
GO:0007160~cell-matrix adhesion	5	0.00	8.72
GO:0031589~cell-substrate adhesion	5	0.00	8.45
GO:0046271~phenylpropanoid catabolic process	3	0.00	27.04
GO:0046274~lignin catabolic process	3	0.00	27.04
GO:0009808~lignin metabolic process	3	0.01	23.18
GO:0009698~phenylpropanoid metabolic process	3	0.01	18.03
GO:0019748~secondary metabolic process	5	0.01	5.41
GO:0007259~JAK-STAT cascade	3	0.02	14.75
GO:0007166~cell surface receptor linked signal transduction	14	0.02	2.01
GO:0042335~cuticle development	3	0.02	12.48
GO:0019221~cytokine-mediated signaling pathway	3	0.02	12.48
GO:0007600~sensory perception	6	0.02	3.65
GO:0007186~G-protein coupled receptor protein signaling pathway	7	0.03	3.08
GO:0034614~cellular response to reactive oxygen species	3	0.03	11.59
GO:0050877~neurological system process	10	0.03	2.3
GO:0055114~oxidation reduction	17	0.03	1.73
GO:0050890~cognition	7	0.03	2.87
GO:0030818~negative regulation of cAMP biosynthetic process	2	0.04	54.08
GO:0030803~negative regulation of cyclic nucleotide biosynthetic process	2	0.04	54.08
GO:0030800~negative regulation of cyclic nucleotide metabolic process	2	0.04	54.08
GO:0030815~negative regulation of cAMP metabolic process	2	0.04	54.08
GO:0030809~negative regulation of nucleotide biosynthetic process	2	0.04	54.08
GO:0016337~cell-cell adhesion	5	0.04	3.81
GO:0019439~aromatic compound catabolic process	3	0.05	8.54
GO:0042219~cellular amino acid derivative catabolic process	3	0.05	8.54
GO:0030336~negative regulation of cell migration	3	0.05	8.54
GO:0006892~post-Golgi vesicle-mediated transport	3	0.05	8.11
GO:0040013~negative regulation of locomotion	3	0.05	8.11
GO:0030814~regulation of cAMP metabolic process	3	0.05	8.11
GO:0045087~innate immune response	4	0.05	4.7
GO:0042423~catecholamine biosynthetic process	2	0.05	36.05
GO:0045980~negative regulation of nucleotide metabolic process	2	0.05	36.05
GO:0006952~defense response	6	0.05	2.9
GO:0034599~cellular response to oxidative stress	3	0.06	7.73
GO:0051271~negative regulation of cell motion	3	0.06	7.37
GO:0007156~homophilic cell adhesion	3	0.07	6.76
GO:0019853~L-ascorbic acid biosynthetic process	2	0.07	27.04
GO:0002696~positive regulation of leukocyte activation	3	0.08	6.49
GO:0030799~regulation of cyclic nucleotide metabolic process	3	0.08	6.49
GO:0050867~positive regulation of cell activation	3	0.08	6.49
GO:0016044~membrane organization	8	0.08	2.14
GO:0007268~synaptic transmission	5	0.08	3.04
GO:0006140~regulation of nucleotide metabolic process	3	0.08	6.24
GO:0006575~cellular amino acid derivative metabolic process	5	0.09	2.97
GO:0007565~female pregnancy	3	0.09	5.79
GO:0000302~response to reactive oxygen species	3	0.09	5.79
GO:0006470~protein amino acid dephosphorylation	5	0.10	2.85
GO:0016358~dendrite development	3	0.10	5.59
MF terms	**Count**	**PValue**	**FE**
GO:0031177~phosphopantetheine binding	4	0.00	16.21
GO:0016597~amino acid binding	5	0.00	8.5
GO:0043176~amine binding	5	0.00	7.53
GO:0048037~cofactor binding	12	0.00	2.74
GO:0019842~vitamin binding	8	0.00	3.8
GO:0016682~oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor	3	0.01	26.35
GO:0008471~laccase activity	3	0.01	26.35

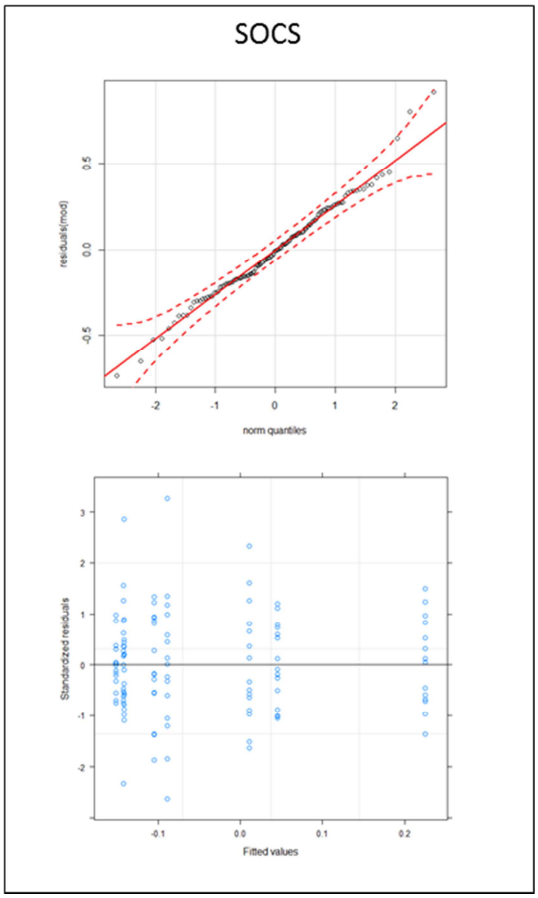
GO:0005509~calcium ion binding	16	0.01	2.11
GO:0043169~cation binding	58	0.01	1.31
GO:0043167~ion binding	58	0.01	1.3
GO:0016679~oxidoreductase activity, acting on diphenols and related substances as donors	3	0.01	17.56
GO:0008066~glutamate receptor activity	3	0.01	15.81
GO:0046872~metal ion binding	56	0.02	1.28
GO:0000036~acyl carrier activity	3	0.03	11.29
GO:0001642~group III metabotropic glutamate receptor activity	2	0.04	52.69
GO:0005507~copper ion binding	4	0.04	5.27
GO:0016651~oxidoreductase activity, acting on NADH or NADPH	4	0.04	5.14
GO:0016655~oxidoreductase activity, acting on NADH or NADPH, quinone or similar compound as acceptor	3	0.05	8.32
GO:0004725~protein tyrosine phosphatase activity	5	0.05	3.56
GO:0031406~carboxylic acid binding	5	0.05	3.56
GO:0001640~adenylate cyclase inhibiting metabotropic glutamate receptor activity	2	0.06	35.13
GO:0046920~alpha(1,3)-fucosyltransferase activity	2	0.09	21.08
GO:0004348~glucosylceramidase activity	2	0.09	21.08
GO:0016831~carboxy-lyase activity	3	0.09	5.85
GO:0003950~NAD+ ADP-ribosyltransferase activity	3	0.09	5.85
Panther BP terms	**Count**	**PValue**	**FE**
BP00124:Cell adhesion	12	0.00	5.11
BP00103:Cell surface receptor mediated signal transduction	14	0.00	3.55
BP00120:Cell adhesion-mediated signaling	8	0.00	6.01
BP00102:Signal transduction	24	0.00	2.05
BP00199:Neurogenesis	9	0.00	4.67
BP00246:Ectoderm development	9	0.00	4.39
BP00274:Cell communication	11	0.00	3.17
BP00193:Developmental processes	16	0.00	2.2
BP00104:G-protein mediated signaling	6	0.02	3.61
BP00132:Receptor mediated endocytosis	4	0.03	5.86
BP00148:Immunity and defense	9	0.03	2.37
BP00129:Endocytosis	6	0.03	3.29
BP00287:Cell motility	5	0.05	3.53
BP00179:Apoptosis	6	0.10	2.41

Appendix C1

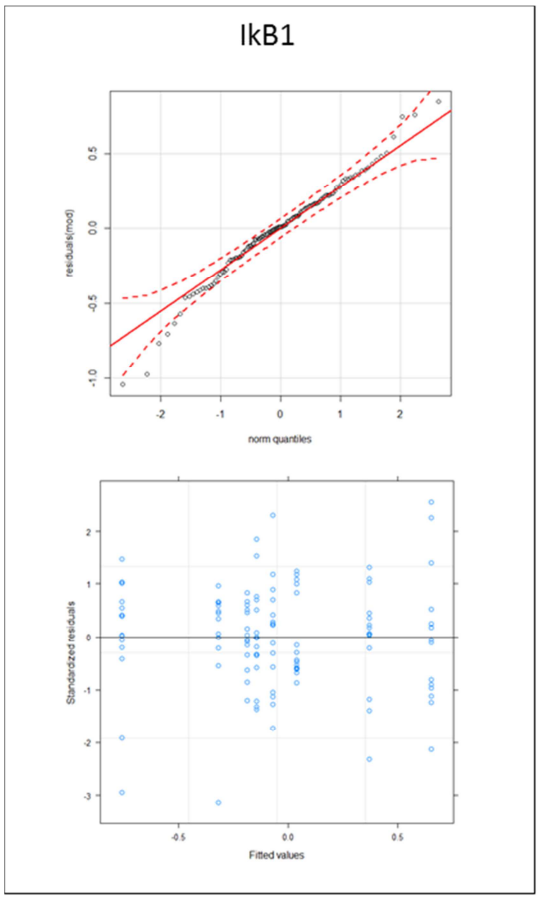

IkB1
residuals(mod)
norm quantiles
Standardized residuals
Fitted values

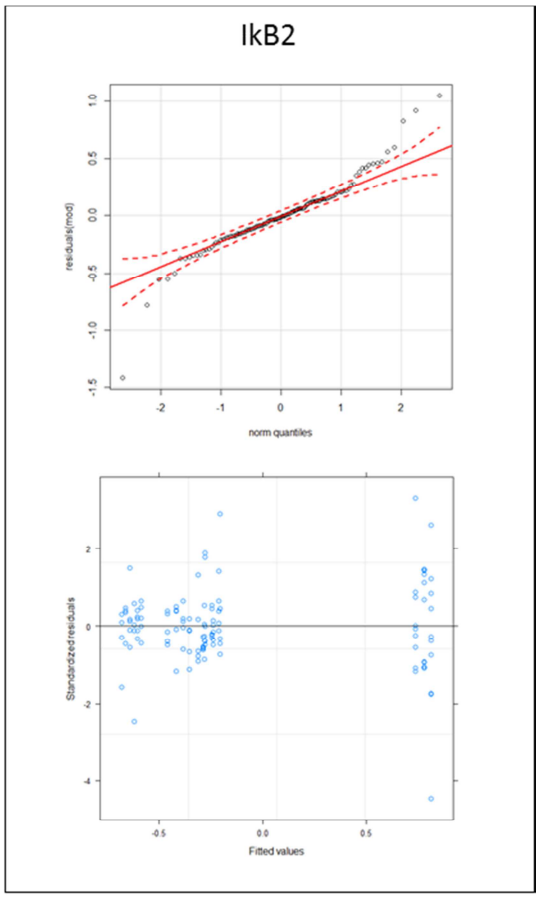

IkB2
residuals(mod)
norm quantiles
Standardized residuals
Fitted values

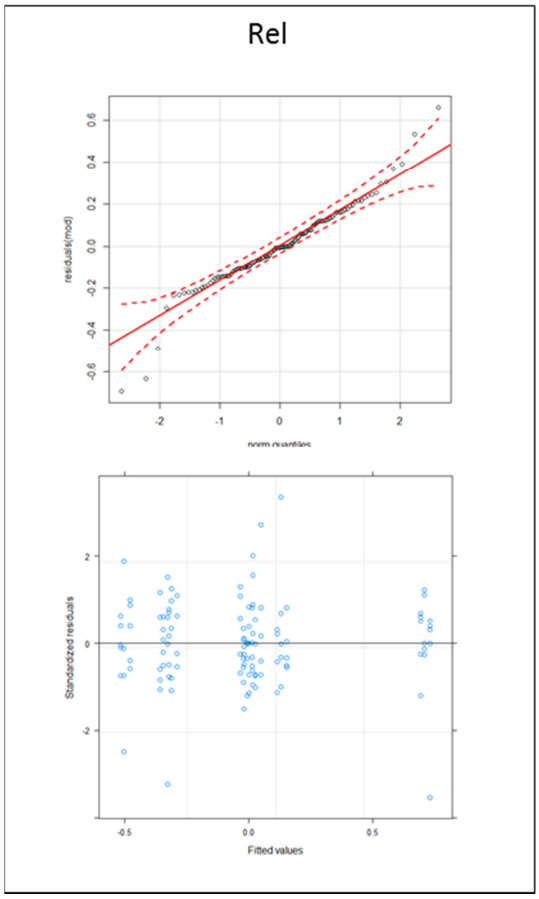

Rel
residuals(mod)
norm quantiles
Standardized residuals
Fitted values

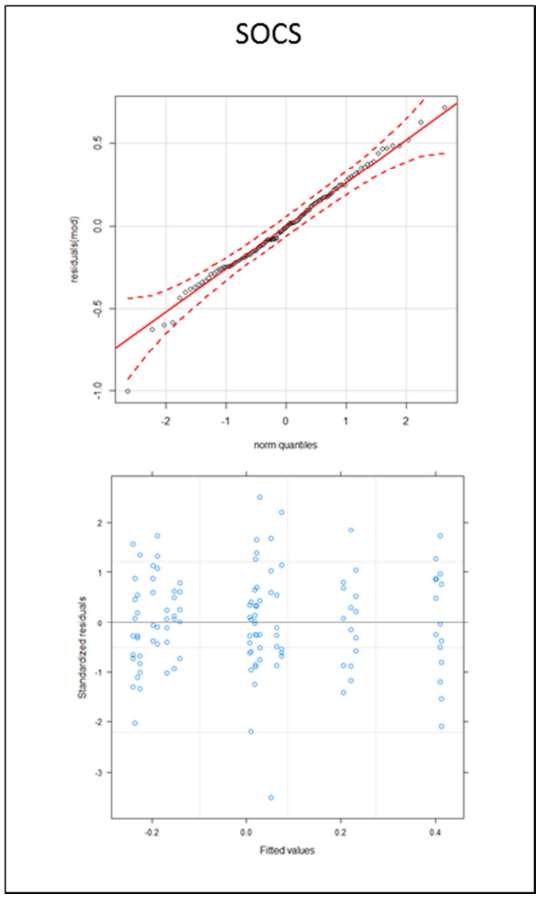

SOCS
residuals(mod)
norm quantiles
Standardized residuals
Fitted values

Appendix C2


IkB1
residuals(mod)
norm quantiles
Standardized residuals
Fitted values


IkB2
residuals(mod)
norm quantiles
Standardized residuals
Fitted values


Rel
residuals(mod)
Standardized residuals
Fitted values


SOCS
residuals(mod)
norm quantiles
Standardized residuals
Fitted values

Al Prof. Tomaso Patarnello che ha reso possibile questa esperienza.

A Luca, per la fiducia che ha sempre riposto in me e per le numerose opportunità offerte in questi tre anni.

Al mio "capetto" Massimo che mi aperto le porte della trascrittomica e mi ha iniziata al magico mondo dei microarray, "La scienza esatta e intramontabile" (!!!).

A Lisa e Serena, perché confrontarmi con loro è sempre stata occasione di crescita professionale e umana. Insostituibili!

A Rafaella, Sara, Francesco e Massimiliano perché ognuno di loro, ha saputo arricchire questo percorso.

Ad Arnaud per avermi dato la preziosa opportunità di lavorare nel suo gruppo.

A Charlotte, per le numerose chiacchierate scientifiche (…e non!) che mi hanno permesso di imparare il francese.

A Jean-Yves, Philippe e Virgile per le importanti competenze che hanno saputo trasmettermi.

A Mauro, perché sotto la sua guida ho fatto i primi passi nel mondo della ricerca e perché oggi è fiero di vedermi camminare.

A tutti i colleghi che ho avuto il piacere di incontrare in questi anni e che in qualche modo hanno contribuito alla mia formazione.