Universidade Vigo

ADVANCES IN BIVALVE MOLLUSCS IMMUNE SYSTEM: A GENOMIC APPROACH

PhD THESIS

REBECA MOREIRA SANMARTÍN

VIGO 2015

Universidade Vigo

ADVANCES IN BIVALVE MOLLUSCS IMMUNE SYSTEM: A GENOMIC APPROACH

Thesis submitted by

Rebeca Moreira Sanmartín for the degree of Doctor of the University of Vigo with international mention and as compendium of articles

Doctoral Degree Program in Methodology and Applications in Life Sciences

VIGO 2015

UNIVERSITY OF VIGO

FACULTY OF BIOLOGY

SPANISH NATIONAL RESEARCH COUNCIL

MARINE RESEARCH INSTITUTE

CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Universidade Vigo

AVANCES EN EL ESTUDIO DEL SISTEMA INMUNITARIO DE MOLUSCOS BIVALVOS: UNA APROXIMACIÓN GENÓMICA

Memoria de tesis presentada por

Rebeca Moreira Sanmartín para optar al grado de Doctor por la Universidad de Vigo con mención internacional y por compendio de artículos

Programa de doctorado en Metodología y Aplicaciones en Ciencias de la Vida

VIGO 2015

UNIVERSIDADE DE VIGO

FACULTADE DE BIOLOXÍA

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

INSTITUTO DE INVESTIGACIONES MARINAS

CSIC

Consejo Superior de Investigaciones Científicas

Dña. **BEATRIZ NOVOA GARCÍA**, Doctora en Biología e Investigadora Científica del Consejo Superior de Investigaciones Científicas (CSIC), junto con D. **ANTONIO FIGUERAS HUERTA**, Doctor en Biología y Profesor de Investigación del Consejo Superior de Investigaciones Científicas (CSIC) y D. **PABLO BALSEIRO VIGO**, Doctor en Biología e Investigador en la Universidad de Bergen, INFORMAN: Que la presente memoria adjunta, titulada "ADVANCES IN BIVALVE MOLLUSCS IMMUNE SYSTEM: A GENOMIC APPROACH" / "AVANCES EN EL ESTUDIO DEL SISTEMA INMUNITARIO DE MOLUSCOS BIVALVOS: UNA APROXIMACIÓN GENÓMICA" presentada por Dña. Rebeca Moreira Sanmartín para optar al grado de Doctor por la Universidad de Vigo, con mención internacional y por compendio de artículos, ha sido realizada bajo nuestra dirección y reúne los requisitos necesarios para ser defendida ante el tribunal calificador.

Y para que así conste, se firma la presente en Vigo, a 29 de septiembre de 2015.

Dra. **Beatriz Novoa García** Dr. **Antonio Figueras Huerta** Dr. **Pablo Balseiro Vigo**

Visto bueno del supervisor en la Universidad de Vigo: Dr. **David Posada González**.

With the support of the Spanish Ministerio de Ciencia e Innovación through a FPI research Grant (BES-2009-029765) associated to the project (AGL2008-05111). This work has also partially been funded by the projects: REPROSEED (FP7 245119), BIVALIFE (FP7-KBBE-2010-4/266157) and Genome project of the galician mussel *Mytilus galloprovincialis*(16 10 PXIB 402 096 PR).

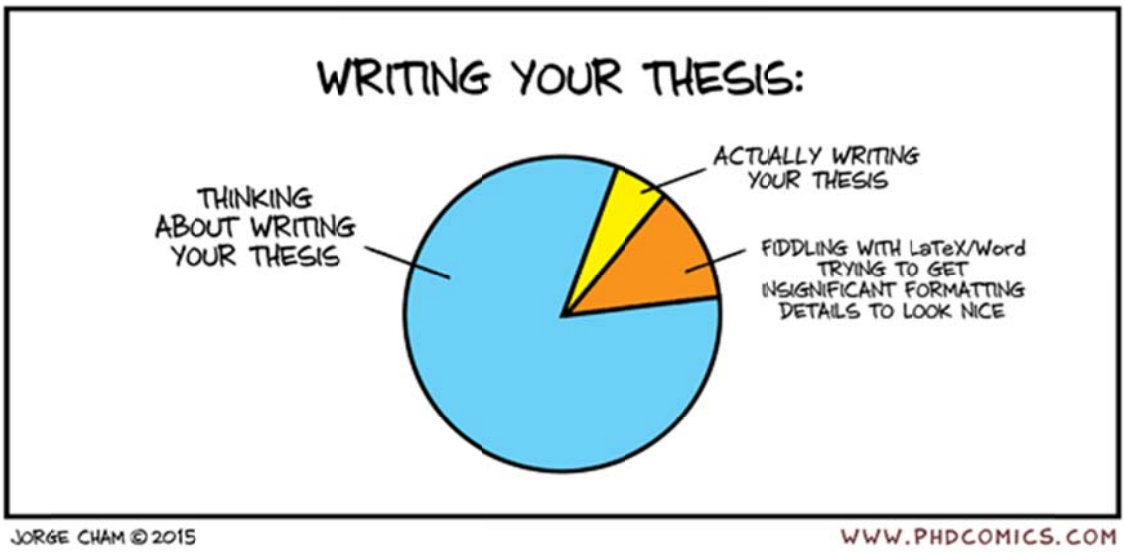
Con el apoyo del Ministerio de Ciencia e Innovación, por la concesión de una beca predoctoral de Formación de Personal Investigador (FPI), asociada al proyecto de investigación AGL2008-05111 "Genes de moluscos bivalvos implicados en la resistencia frente a enfermedades e interacción con agentes patógenos". Además se ha recibido apoyo de los siguientes proyectos: REPROSEED (FP7 245119), BIVALIFE (FP7-KBBE-2010-4/266157) y Proxecto xenoma do mexilón galego *Mytilus galloprovincialis*(16 10 PXIB 402 096 PR).

ACKNOWLEDGEMENTS

Estos años me he sentido muy feliz con este trabajo y aún en estos malos tiempos para la ciencia en los que todos nos preguntamos Cómo Seguir Investigando en Crisis:

When life knocks you down...Calmly get back up, smile, and very politely say:
'You hit like a bitch.'

Después de pasar mucho tiempo "pensando en escribir la tesis" y de aprender, por las malas, lo que significa *procrastination*… Tengo que dar las gracias a mucha gente porque al final me moví al quesito amarillo!



Gracias a Beatriz Novoa y Antonio Figueras, mis directores de tesis, por haber confiado en mí y hacerme un huequito primero en "pato" y después en "I+G". Me habéis enseñado a hacer ciencia y gracias a vosotros he cumplido mi sueño de ser una… ¡investigadora científica!

A Pablo Balseiro, mi codirector de tesis y el que me ha enseñado prácticamente todo lo que sé en técnicas de laboratorio y bioinformática, aunque a veces fuese a base de… "ñññññ, así no!!! Mal, mal, maaaaal!!!". Muchas gracias por todo WikiPablo!

También al resto de los docs del laboratorio: a Laura, Raquel, MMar, PDR, Sonia, Álex, vuestra paciencia y explicaciones sobre BLAST, diseño de primers, citometría, patógenos, infecciones, medios de cultivo, orden y limpieza en el laboratorio… también son parte de esta tesis.

A la gente del despachito, cualquiera que fuese su ubicación, por nuestros "debates" sobre las dudas en protocolos, desarrollo larvario, filogenética, extracción de ARN… pero sobre todo: cómo hacer que las réplicas técnicas salgan clavadas???!!! Primero estuvimos en el despachito del pasillo: gracias a Noe, Monse, Rosi y Sheila, y a Eva, que empezamos juntas el mismo día y eso une! Después de subir y bajar con nuestros bártulos nos instalamos (definitivamente) en el despachito de histología: gracias a Rubén, por todo lo que me ayudaste y lo que me enseñaste en acuarios, no fueron pocas cosas, ni sencillas… a Paolo, grazie a te non mi ho dimenticato del poco italiano che so! Y muy especialmente a Patry, Mónica y Gabriel, que os puedo decir que no nos hayamos dicho ya… Nos hicimos muy buenos amigos, de esos que son para siempre, de esos a los que unen la ciencia, las cañas, el churrasco y el licor café!

Gracias también a mi tutor, David Posada, y a Carlos Canchaya del Departamento de Bioquímica, Genética e Inmunología - Universidad de Vigo. Aunque costó lo suyo tanto los árboles filogenéticos como el RNA-Seq salieron adelante!

Devo ringraziare a tante persone al Dipartimento di Biomedicina Comparata e Alimentazione - Universitá degli Studi di Padova. Prima al Prof. Luca Bargelloni, grazie per la oportunitá di lavorare con tutti voi... due volte! Grazie mille a Massimo, Max e Marianna, con voi ho imparato tuto quelo che so da microarray. Ma devo anche ringraziare a Sere, Lisa, Rafaella, Sara, Francesco, Barbara, Laura, Nadia, Enrico… sicuro lascio a qualcuno... tutti havete fatto del mio soggiorno a Padova un bello ricordo. Grazie per le serate italiane: spritz, birra e pizza. Non mi dimenticaró più dei pranzi sulle scalle d'incendio!

Thanks also to all the authors of the papers I've read all these years. They made me a bit wiser every day. Especially to all of whom I had the luck to meet in conferences or in project meetings, thank you for your questions, recommendations, advice…

A mis "COMPIS" pero muy especialmente a "Las de siempre 👯" Ana, Ire, MFe, MLuz, Vane, porque sois mis mejores amigas, alegrándoos más que yo en los buenos momentos y dándome ánimos sinceros en los malos. Algunas veces riñéndome por no meterle caña a la tesis. Y aunque a veces sois rabudas y me lleváis la contraria en cosas de ciencia… Yo os quiero igual maricas!

A mi familia no tengo palabras, ni espacio suficiente, para agradecerles… TODO!!! A mis padres que siempre han confiado en que voy a ganar el Nobel, o a cambiar el mundo o algo así… (o eso me transmiten cuando me miran!), gracias por enseñarme lo que de verdad importa en la vida (incluido cuidar de nuestra pequeña granja!) y a ser buena persona. Vosotros sí que sois grandes! Salo, ya todo el mundillo científico que me conoce sabe lo que es el pen-vaca! Mira que ha almacenado conocimiento esa vaquita y sus 4G de estómago… Jaco, en breves ya tendrás que llamarme "doc" pero en serio, no de coña, eh? ;D

Amore, a ti te debo lo que soy ahora. Gracias por hacerme rabiar, por hacerme reír, por abrazarme justo cuando lo necesito. Gracias por aguantar mis charlas cuando me apasiona algo, aunque luego me digas "Peque, qué lai eres…". También disfrutaste de tus momentos lais corrigiéndome algunas cosas de la tesis, no lo niegues! Mi vida es mejor contigo, gracias a ti soy feliz!

¿Me dejo a alguien atrás? Seguro… Gracias a cualquiera que haya hecho lo más mínimo por echarme una manita durante estos años, consciente o inconscientemente.

Gracias, gracias, gracias!!!

INDEX

RESUMEN

INTRODUCCIÓN

La acuicultura proporciona prácticamente el 50% de los productos pesqueros mundiales destinados a la alimentación y es, según la Organización de las Naciones Unidas para la Alimentación y la Agricultura (*Food and Agriculture Organization of the United Nations* - FAO), uno de los sectores productivos de crecimiento más rápido en el mundo. La gestión responsable de la acuicultura es un aspecto fundamental a tener en cuenta tanto por criterios de producción como medioambientales y de sostenibilidad. Un aspecto importante en este contexto es la necesidad de controles eficaces de la gestión de la sanidad, porque las enfermedades son el principal cuello de botella para el crecimiento de la acuicultura. Es por eso que la investigación científica adquiere una gran importancia a la hora de aportar soluciones (FAO, 2015).

En Galicia la acuicultura de moluscos bivalvos, especialmente de mejillones y almejas, es muy importante en la economía local. Los bivalvos son un grupo muy interesante no solo en términos de producción (FAO, 2012) sino también ecológicos, ya que filtran el agua y bioacumulan contaminantes, por ello se utilizan como indicadores de contaminación (Wootton *et al*., 2003). Por otra parte, también pueden acumular microorganismos, parásitos y toxinas procedentes de algas tóxicas que afectan a las personas (Morgan *et al*., 2001). Estas son solo unas pocas de las razones por las que es importante el estudio de los bivalvos desde un punto de vista de la biología molecular.

ACUICULTURA DE MOLUSCOS BIVALVOS

La acuicultura existe desde hace miles de años. En Hawai se descubrieron restos de estanques prehistóricos que fueron utilizados para el cultivo de algas, así como para la estabulación y mantenimiento de animales acuáticos (Farber; 1997). También existen representaciones de tilapias (*Oreochromis niloticus*) en estanques en tiempos del antiguo Egipto datadas en el 3500 a. C. (Bardach *et al.*, 1972). En cuanto al cultivo de moluscos, la ostra ya se cultivaba tanto en Japón como en Grecia o Roma hace más de dos mil años (Pillay, 1997). Galicia es un enclave perfecto para el cultivo de moluscos bivalvos, como el mejillón y la almeja, tanto a su orografía como por su situación geográfica, climática y oceanográfica. Estas condiciones favorecen el afloramiento de aguas subsuperficiales ricas en nutrientes, favoreciendo el desarrollo de fitoplancton del que se alimentan los bivalvos.

Almeja

El cultivo de almeja en Galicia se realiza principalmente sembrando semilla en los bancos naturales que necesitan ser recuperados. La extracción se realiza a través del marisqueo en bancos naturales y en parques de cultivo. La extracción de almeja fina (*Ruditapes decussatus*) comenzó en Galicia a finales de los años 20. La ausencia de leyes reguladoras propició la recolección indiscriminada, provocando que los bancos naturales llegasen al borde de la desaparición. Con la actual normativa el volumen de almeja fina que se extrae es bajo comparado con otras especies, sin embargo, supone un importante beneficio debido a su alto precio. La sobreexplotación y los

rendimientos irregulares de la almeja fina, supusieron la introducción del cultivo de almeja japonesa (*Ruditapes philippinarum*) a Europa en 1972, más resistente y de crecimiento más rápido.

Mejillón

El cultivo en suspensión del mejillón mediterráneo (*Mytilus galloprovincialis*), se inició en España en el año 1846 mediante las bateas, balsas donde cuelgan cuerdas en las que se fijan los juveniles y permanecen adheridos hasta que obtienen su talla comercial. Según datos de la FAO se producen anualmente 1.500.000 toneladas de mejillón en todo el mundo. La industria mejillonera gallega, con una producción de entre 150.000 y las 300.000 toneladas anuales, es la segunda potencia a nivel mundial, siendo tan solo superada por China, con una producción media anual de 600.000 toneladas.

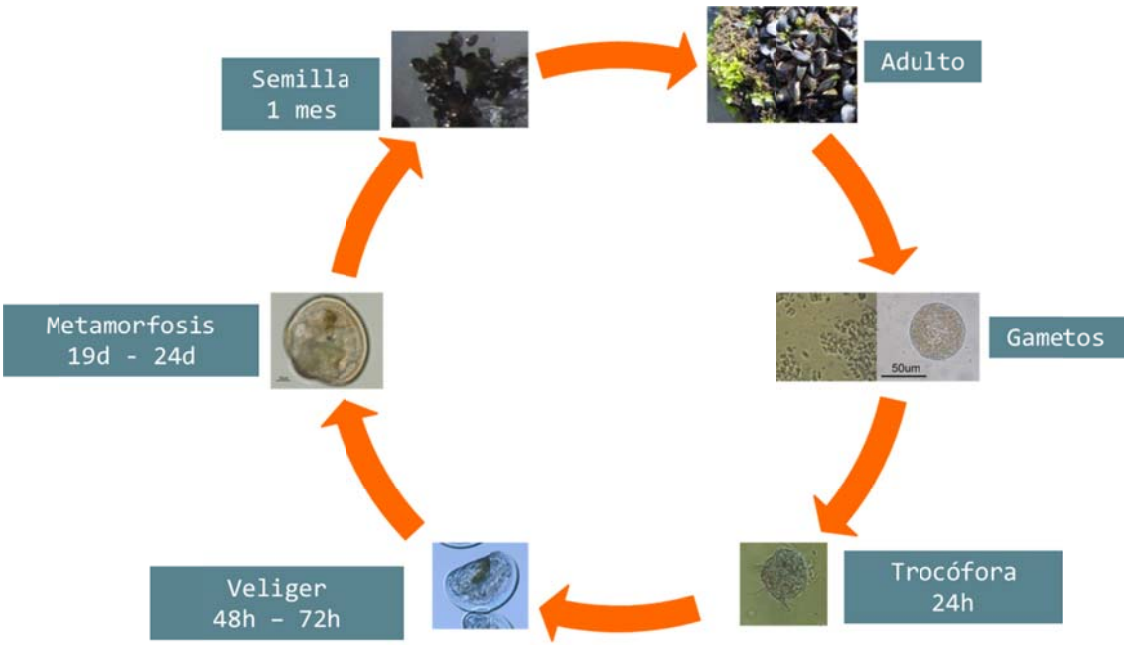


Figura 1. Ciclo vital del mejillón. Fuente: elaboración propia.

SISTEMA INMUNITARIO EN MOLUSCOS BIVALVOS

Según la FAO, el futuro de la producción de moluscos deberá ser planeado como parte de un programa integral que incluya aspectos como la creación de un programa de investigación sobre patologías de moluscos bivalvos. Sin embargo el conocimiento de la respuesta inmune de estas especies sigue siendo limitado y la lucha contra las patologías de moluscos se basa en medidas preventivas y la eliminación de individuos enfermos, ya que se sabe muy poco de las bases moleculares de la respuesta inmune de los bivalvos.

En los vertebrados el sistema inmune consta de respuesta innata o inespecífica y respuesta adquirida, o de memoria, tras un contacto previo con el patógeno. Los bivalvos carecen de respuesta adquirida en sentido estricto (Criscitiello y de

Figueiredo; 2013), sin embargo, están provistos de barreras físicas como la concha y el mucus. Además, poseen un potente sistema inmune innato humoral y celular.

Inmunidad celular

Las células encargadas de la defensa en los moluscos bivalvos son los hemocitos (Bayne CJ *et al*., 1980; Carballal *et al*., 1997; Prado-Alvarez *et al*., 2012). Los hemocitos eliminan las células o partículas extrañas fagocitándolas y destruyéndolas por medio de las enzimas lisosomales y la síntesis de especies reactivas de oxígeno y nitrógeno. Además de la fagocitosis, los hemocitos tienen otras funciones como parte fundamental del sistema inmunitario, como la producción y regulación de los factores humorales, la formación de agregados para atrapar a las partículas no propias (Ratcliffe *et al.*, 1985) o el aislamiento del patógeno mediante encapsulación (Feng, 1988).

Inmunidad humoral

Los componentes de la inmunidad humoral son proteínas o glicoproteínas que se encuentran en el suero, los hemocitos o ambos. Los principales componentes de la defensa humoral son:

- Lisozima y enzimas lisosomales: Liberadas en respuesta a una infección bacteriana, rompen los constituyentes estructurales básicos de la pared de las bacterias, como los peptidoglucanos (Zhao *et al.*, 2010).
- Lectinas: Son receptores de reconocimiento de patrones (*Pattern Recognition Receptors* - PRRs), proteínas de unión a carbohidratos con capacidad de opsonizacion y aglutinación. Favorecen la fagocitosis y la activación del complemento (Humphries and Yoshino, 2003).
- Proteínas de choque térmico: Las proteínas de choque térmico (*Heat Shock Proteins* - HSPs) se sintetizan como respuesta a múltiples factores de estrés y pueden modular la apoptosis y afectar a la transcripción de genes importantes en la respuesta inmunitaria (Parcellier *et al.*, 2003; Guo *et al.*, 2009).
- Sistema profenol-oxidasa: Implicado en la fagocitosis y la encapsulación, así como también producción de factores líticos (Cerenius *et al.*, 2008).
- Proteasas e inhibidores de proteasas: Pueden estar presentes en los patógenos, facilitando la infección (Ordás *et al.*, 2001), por ello muchos organismos, entre ellos los bivalvos, poseen enzimas inhibidoras de proteasas en los lisosomas de los hemocitos, favoreciendo la destrucción de los patógenos tras su fagocitosis (Xue *et al.*, 2006).
- Péptidos antimicrobianos: Los péptidos antimicrobianos (*Antimicrobial Peptides* - AMPs) son proteínas pequeñas ricas en cisteínas con un papel clave en la inmunidad innata. Su mecanismo de acción consiste en alterar la permeabilidad de la membrana de los patógenos, llegando incluso a lisar sus células. Hasta hace poco se conocían cuatro grupos de péptidos antimicrobianos en bivalvos: defensinas, miticinas, mitimicinas y mitilinas (Li *et al.*, 2011). Recientemente se han descrito otros cuatro nuevos grupos: grandes defensinas, mitimacinas (Gerdol *et al*., 2012), miticusinas (*Liao et al*., 2013) y mitiquitinas (Qin *et al*., 2014).

- Complemento: El sistema del complemento se compone de más de 30 proteínas plasmáticas y asociadas a membrana que se activan en cascada después del contacto con distintos patógenos. Estas proteínas favorecen la opsonización, fagocitosis y destrucción de estos patógenos mediante células efectoras (Huber-Lang *et al.*, 2006). Se conoce muy poco sobre el sistema del complemento en moluscos. Los componentes del complemento más estudiados en bivalvos son las proteínas con dominio C1q (Gestal *et al.*, 2010) y el compenente C3 (Prado-Alvarez *et al*., 2009).
- Citoquinas: Son los reguladores más importantes en la inmunidad. Su acción es muy rápida y funcionan a concentraciones es muy bajas, coordinando el desarrollo de la respuesta inmune, amplificando su magnitud y deteniéndola al final del proceso (Benjamini *et al.*, 2000). En moluscos bivalvos se han identificado muy pocas citoquinas, de las que se tiene mayor conocimiento son el homólogo del TNF-α (De Zoysa *et al.*, 2010) y la IL-17 (Roberts *et al.*, 2008).

INVESTIGACIÓN EN EXPRESIÓN GENÉTICA Y SECUENCIACIÓN EN BIVALVOS

El interés en la genómica de los bivalvos es cada vez mayor debido a su gran importancia en la acuicultura y en las ciencias medioambientales marinas. La genómica se centra en el estudio de los genes y sus funciones. Ha cambiado drásticamente desde los años 90, cuando los genes se estudiaban uno por uno, hasta hoy, que se pueden estudiar todos los genes expresados en un momento dado. El siguiente gran paso, la secuenciación de los genomas, supone una herramienta de enorme utilidad, ya que contiene toda la información información codificada que determina la biología de un organismo.

Por tanto la genómica podría servir de ayuda para resolver problemas específicos en la acuicultura de bivalvos, como el control de enfermedades. Mediante la genómica se pueden identificar genes que se expresan ante situaciones de estrés o en presencia de patógenos y así compensar la ausencia de manifestaciones clínicas. Esta es la función de los marcadores genéticos, permitir una identificación correcta y temprana de los animales antes de que las mortalidades aparezcan.

El mayor avance en el campo de la genómica ha sucedido gracias al desarrollo de las tecnologías de secuenciación de nueva generación (*Next Generation Sequencing Technologies* - NGS). En bivalvos las NGS y los micorarrays se han usado para investigar muchos procesos fisiológicos como la biomineralización de la concha, la reproducción, el desarrollo o la adaptación a cambios climáticos y a la contaminación (revisado en Romero *et al.*, 2012).

Esto significa que cada vez los avances suceden más rápido, en poco tiempo se tendrá mayor cantidad de información relevante sobre la biología molecular de los moluscos bivalvos.

TÉCNICAS DE EXPRESIÓN GENÉTICA Y SECUENCIACIÓN

La biología molecular es un campo que evoluciona constantemente. Han pasado menos de 50 años desde que se descubrió la reacción en cadena de la polimerasa (*Polymerase Chain Reaction* - PCR) hasta que se ha desarrollado la tecnología del RNA-Seq. En esta tesis hemos usado varias de las herramientas para estudiar la expresión genética para aumentar el conocimiento que existe actualmente sobre el sistema inmunológico de almejas y mejillones.

La PCR es la técnica más usada en biología molecular y en la que se basan la gran mayoría de los protocolos que surgieron posteriormente. La PCR fue descubierta y patentada por Kary Mullis en 1983 (Mullis *et al.*, 1987: US 4683195 A). Esta técnica permite obtener una gran cantidad de copias de una secuencia de ADN usando la enzima que se ocupa de su replicación en las células, la ADN polimerasa, que genera dos copias de la secuencia original en cada ciclo de la reacción. Para que la reacción se lleve a cabo se necesitan:

- Una secuencia molde de ADN, que es la que se va a copiar.
- Dos secuencias pequeñas de ADN que coincidan con los extremos del ADN molde, los cebadores.
- Las cuatro bases nucleotídicas de las que se compone el ADN: adenina, timina, citosina y guanina, en forma trifosfato (dATP, dTTP, dCTP, dGTP).
- Una ADN polimerasa estable a altas temperaturas.
- Un tampón que mantenga el pH de la reacción y que aporte magnesio, cofactor imprescindible para la actividad enzimática.

La reacción se lleva a cabo mediante ciclos idénticos en los que la temperatura cambia para que cada fase de la reacción se ejecute de forma óptima, como indica la figura siguiente.

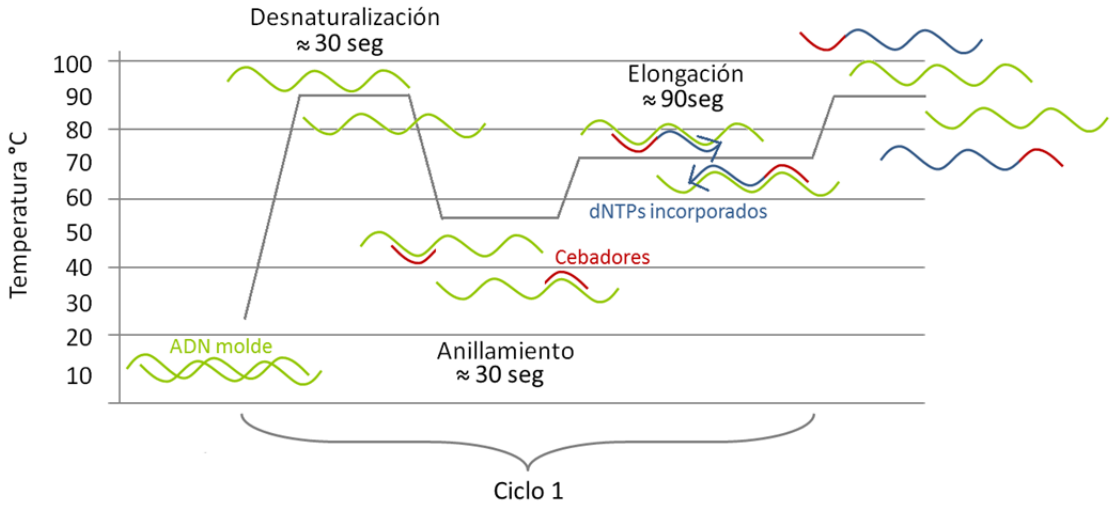


Figura 2. Esquema de la PCR. Fuente: elaboración propia.

Hoy en día, los estudios de expresión genética – el eje central de esta tesis de doctorado – emplean unas de las técnicas más usadas en gran parte de las ramas de la biología, especialmente en fisiología. Normalmente se usan para comparar la activación o inhibición de un gen después de ciertos estímulos como fármacos, infecciones o cambios físicos en el ambiente; aunque se pueden hacer muchísimas más comparaciones: entre distintas especies, tejidos o estadíos del desarrollo, se puede seguir un curso de tiempo o detectar la presencia de células no propias.

En esta tesis se han usado las siguientes técnicas:

- qPCR: La PCR en tiempo real o PCR cuantitativa (*quantitative PCR* – qPCR) sigue el mismo principio que la PCR convencional pero permite al investigador seguir el proceso de amplificación al tiempo que ocurre, por lo que se obtiene una aproximación mucho más precisa del nivel de expresión de un gen. Su mayor ventaja, comparada con la PCR convencional, es que evita la saturación, ya que el resultado de la PCR se obtiene cuando la reacción está terminada y el producto saturado, impidiendo la comparación cuantitativa entre las muestras. El resultado de la qPCR se obtiene cuando se es capaz de detectar una cantidad concreta de ADN, el ciclo en el que esto sucede se llama ciclo umbral (*threshold cycle* - Ct) (Kubista *et al*., 2006).
- Hibridación *in situ*: La hibridación *in situ* (In situ *hybridization* - ISH), fue desarrollada por Joseph G. Gall (Gall and Pardue, 1969). Es una técnica que permite localizar el lugar específico donde se encuentra un gen en células, tejidos o embriones, mediante una sonda de ADN o ARN marcada y complementaria al gen que se estudia. Así es posible determinar el sitio físico en el que se expresan los genes de interés.
- Pirosecuenciación: La pirosecuenciación es una NGS desarrollada por Mostafa Ronaghi y Pål Nyrén (Nyrén P, 2001: US 6 258 568BI). Se basa en la detección de la liberación de los grupos fosfato tras la incorporación secuencial de los nucleótidos durante la replicación.
- Microarrays: Un microarray es un conjunto de fragmentos de ADN unidos a una superficie sólida, normalmente cristal, que sirve para medir el nivel de expresión genética de un gran número de genes simultáneamente (Maskos *et al*., 1992; Schena *et al*., 1995). Cada fragmento de ADN es complementario a una zona de un gen concreto y los microarrays pueden contener fragmentos de decenas de miles de genes. Después de ser hibridado con la muestra que se estudia, previamente marcada con un fluorocromo, cada punto emite fluorescencia. La intensidad de la fluorescencia es proporcional al nivel de expresión del gen que representa.
- RNA-Seq: El RNA-Seq es una tecnología que aúna el potencial de las NGSs y las técnicas de expresión para describir un momento concreto en el transcriptoma de un tejido (Chu y Corey, 2012). Esta técnica permite estudiar un número mucho mayor de genes de lo que permite un microarray y además puede identificar genes nuevos y detectar genes con muy baja expresión, que con otras tecnologías serían imposibles de analizar.

En resumen, el conocimiento de los procesos fisiológicos básicos de los moluscos bivalvos es pobre a nivel molecular. El objetivo para los próximos años es añadir a las bases de datos secuencias específicas de bivalvos. En esta tesis nos centraremos en su sistema inmunitario. Además, también estudiaremos la ontogenia del sistema inmune y los sistemas de defensa en diferentes etapas del desarrollo larvario de los bivalvos. Las recientes tecnologías de secuenciación y expresión serán el punto de arranque para la mejora de los actuales sistemas de acuicultura intensiva de los moluscos bivalvos.

OBJETIVOS

El principal objetivo de esta tesis de doctorado es explorar las posibilidades que nos brindan las herramientas modernas para el estudio de la expresión genética, incluyendo la secuenciación de nueva generación, y así incrementar la base de conocimiento del mecanismo de respuesta inmunitaria de los moluscos bivalvos a nivel molecular.

Los objetivos concretos son los siguientes:

- Estudiar las diferencias en la respuesta inmunitaria de dos especies de almeja que viven en el mismo hábitat (*R. decussatus* y *R. philippinarum*) después de una infección bacteriana con *V. alginolyticus*.
- Descubrir nuevos genes relacionados con el sistema inmunitario usando la pirosecuenciación, para incrementar el número de secuencias nucleotídicas en las bases de datos públicas.
- Diseñar un nuevo microarray para *R. philippinarum* que incluya secuencias de hemocitos relacionadas con el sistema inmunitario y validarlo mediante un experimento de infección con *V. alginolyticus*.
- Entender los primeros pasos para la adquisición de inmunocompetencia en *M. galloprovincialis*, para ello se describirán las funciones inmunológicas y la expresión genética en distintas fases larvarias.
- Usar la tecnología del RNA-Seq para incrementar la información sobre los transcriptomas de *M. galloprovincialis* en las bases de datos y estudiar las diferencias de expresión genética entre hemocitos, músculo, manto y branquias.

PUBLICACIONES

Análisis de la resistencia e inmunidad mediante qPCR en *Ruditapes philippinarum* y *Ruditapes decussatus* después de una infección bacteriana

La almeja fina (*Ruditapes decussatus*) y la almeja japonesa (*Ruditapes philippinarum*), dos especies de bivalvos cultivadas en todo el mundo, se ven afectadas por enfermedades que causan graves pérdidas económicas. El mecanismo molecular de la respuesta inmunitaria de los moluscos bivalvos es escaso y está fragmentado. Las almejas, familia *Veneridae*, son un grupo de interés acuícola especialmente poco estudiado, por lo que hemos examinado todos los marcadores de secuencias expresadas (*Expressed Sequence Tags* - EST) disponibles en las bases de datos públicas para estas dos especies con el objetivo de incrementar el conocimiento sobre los genes con funciones inmunitarias en estos animales.

Después de identificar y clasificar las 3.784 EST no anotadas de *R. decussatus* y las 4.607 de *R. philippinarum* encontradas en GenBank, 424 ESTs de *R. decussatus* y 464 de *R. philippinarum* estaban relacionadas con la respuesta inmunitaria. De todas ellas seleccionamos 13 que eran comunes en las dos especies para estudiar comparativamente la respuesta de *R. decussatus* y *R. philippinarum* después de una infección con *Vibrio alginolyticus*. Para ello se estudiaron los perfiles de expresión de estos 13 genes mediante qPCR. Los resultados muestran que en *R. philippinarum* la respuesta inmunitaria es más rápida que en *R. decussatus*. Además la expresión de los genes activados por NF-κB en *R. decussatus* no parece suficiente como para desencadenar una respuesta inflamatoria ante la infección. Sin embargo, *R. philippinarum* pudo iniciar y regular de manera eficiente la actividad transcripcional de NF-κB, incluso con valores bajos de expresión.

Obtención del transcriptoma de los hemocitos de *Ruditapes philippinarum* usando secuenciación masiva

Como ya se ha mencionado anteriormente, la almeja japonesa es una especie con gran importancia económica en el sector mundial de la acuicultura, pero el desconocimiento de sus mecanismos moleculares de defensa frente a patógenos hace que se produzcan grandes pérdidas económicas al no existir remedios efectivos frente a las enfermedades. El número de secuencias de esta especie en las bases de datos públicas es extremadamente bajo, por lo que hemos decidido secuenciar el mRNA de hemocitos inmunoestimulados de *R. philippinarum* mediante la técnica de pirosecuenciación 454 para identificar los genes involucrados en la defensa contra enfermedades infecciosas.

Esta tecnología de secuenciación masiva nos ha permitido obtener 974.976 lecturas de alta calidad pirosecuenciación 454. Estas lecturas, fragmentos cortos con una media de 250 pares de bases, se ensamblaron formando 51.265 contigs, de los cuales el 44,7% se identificaron con éxito. Entre los 35 contigs más frecuentes había gran cantidad de genes relacionados con la respuesta inmunitaria. Un análisis más detallado demostró la presencia de miembros de varios procesos inmunitarios como la apoptosis, la ruta de señalización de los receptores *toll like* y la cascada del complemento. Encontramos moléculas que no habían sido descritas anteriormente en bivalvos, sobre todo en la cascada del complemento, con casi todos sus componentes presentes.

Este trabajo es el primer análisis del inmunoma de *R. philippinarum* usando la pirosecuenciación 454. Estos resultados suponen una gran fuente de datos para descubrir e identificar genes nuevos que pueden ser la base para la construcción de microarrays y el estudio de la expresión genética así como para la identificación de marcadores genéticos. El descubrimiento de secuencias inmunitarias nuevas fue muy productivo como demuestra la gran variedad de contigs con un posible papel en los mecanismos de defensa de *R. philippinarum*.

Análisis de expresión en hemocitos de *Ruditapes philippinarum* mediante microarrays después de una infección bacteriana

En este trabajo se describe la construcción del primer microarray de *R. philippinarum* que incluye secuencias relacionadas con el sistema inmunitario y su aplicación para el estudio del perfil de expresión de hemocitos de almejas infectadas con *V. alginolyticus* en un curso de tiempo.

Después de ensamblar todas las secuencias disponibles hasta la fecha de R. philippinarum, se usaron un total de 12.156 secuencias anotadas para contruir un microarray de oligos de 8 ×15 k. Los experimentos de infección resultaron en un total de 579 secuencias diferencialmente expresadas. Después de interpretar los resultados de de expresión genética, los téminos de ontología genética asociados a estos genes y los análisis de enriquecimiento, se encontraron distintos mecanismos de respuesta al patógeno. Los genes relacionados con señalización, transcripción y apoptosis, como IL-17D, NF-κB o calmodulina, ya se expresaban 3 horas después de la inoculación; mientras que los genes típicamente relacionados con la respuesta inmunitaria, como PGRPs, FREPs y los péptidos antimicrobianos, aparecen más tarde, a las 8 horas de la inoculación. Tras el inicio de la respuesta inmunitaria, 24 horas después de la inoculación, se encuentran activos un gran número de procesos que son activados para superar la infección. Estos procesos incluyen la expresión de moléculas del citoesqueleto, un indicador del movimiento activo de los hemocitos. Los estudios funcionales realizados prueban un aumento en los niveles de apoptosis, necrosis y migración celular después de la infección. Por último, 72 horas tras la inoculación, todos los procesos vuelven a sus niveles basales, ya que se observa una retroalimentación negativa.

Con esta nueva versión del microarray de oligos de *R. philippinarum* hemos podido establecer la respuesta a lo largo del tiempo en una infección con *Vibrio*. El punto clave para superar la infección parece estar en torno a las 8 horas post infección, ya que se detectan funciones inmunitarias que podrían destruir al patógeno y activar una serie de procesos relacionados con la homeostasis y la defensa. Estos resultados ponen de manifiesto la existencia de una respuesta rápida en los bivalvos, y la efectividad de su sistema inmunitario innato.

Respuesta inmune en larvas de *Mytilus galloprovincialis*: La metamorfosis altera la imnunocompetencia

El mejillón Mediterráneo (*Mytilus galloprovincialis*) es otro bivalvo con gran importancia económica y ecológica. Hemos investigado la ontogenia del sistema inmunitario durante el desarrollo larvario en *M. galloprovincialis*. Lo primero que comprobamos fue la capacidad de fagocitosis en larvas trocóforas y veliger. Los resultados de citometría de flujo y microscopía confocal indican que 24 horas después de la fertilización las larvas pueden fagocitar partículas de zymosan y *Escherichia coli*. Sin embargo la fagocitosis de zymosan no desencadenó la esperada producción de especies reactivas de oxigeno y nitrógeno (ROS y NRS), esto se puede explicar por la alta producción basal de estos compuestos en las larvas. Por otra parte seleccionamos un grupo de genes relacionados con la respuesta inmunitaria (la miticina B, miticina C, mitilina B, el precursor de la Mitimicina 1, el factor inhibidor de la migración de macrófagos, la lisozima, el C1q, la proteína del complejo de ataque de membrana y la proteína relacionada con el fibrinógeno) para analizar su perfil de expresión a lo largo de las diferentes etapas del desarrollo larvario (trocófora, veliger, metamorfosis, fijación y semilla). La expresión de estos genes se incrementa en la transición de larva trocófora a larva pedivelíger. Nuestros resultados también sugieren que la expresión durante los primeros estadíos del desarrollo puede ser de origen materno, ya que el nivel de expresión en oocitos es mayor que en larvas trocóforas. La metamorfosis es la etapa crucial en el desarrollo, ya que es cuando la expresión de los genes de defensa aumenta como respuesta a las señales medioambientales. También realizamos estudios de hibridación *in situ* de larvas completas, cuyos resultados sugieren un importante papel del borde del manto para el desarrollo de la inmunocompetencia en las larvas de moluscos bivalvos.

En conjunto, con nuestros resultados proponemos que los mecanismos de respuesta inmunológica surgen durante el desarrollo temprano de *M. galloprovincialis* para ayudar a las larvas a sobrevivir en ambientes con organismos potencialmente patógenos.

RNA-Seq en *Mytilus galloprovincialis:* Transcriptómica comparativa entre varios tejidos

El *M. galloprovincialis* es un firme candidato para convertirse en el organismo modelo para todos los bivalvos, y posiblemente los moluscos. Por eso es necesario aumentar la comprensión de los mecanismos moleculares involucrados en su fisiología desde un punto de vista tanto cualitativo como cuantitativo. Para empezar a llenar este vacío de conocimientos hemos usado RNA-Seq para estudiar los transcriptomas basales de manto, músculo y branquias y de hemocitos expuestos a diferentes estímulos víricos, bacterianos y fúngicos, ya que nos interesa centrarnos en el sistema inmunitario.

Mediante secuenciación con Illumina, se obtuvieron un total de 393,3 millones de lecturas brutas que fueron ensambladas para obtener 151.320 tránscritos no redundantes. Ya que la ostra es el otro candidato a ser el modelo en bivalvos, los comparamos y confirmamos que las secuencias comunes a los dos representan menos del 10% de sus transcriptomas. En cuanto a mejillón, solo el 55% de estos tránscritos son comunes a todos los tejidos estudiados. Los hemocitos y las branquias comparten el 60% de sus transcriptomas mientras que el manto y el músculo son mucho más parecidos, con el 77% de los tránscritos en común. En los hemocitos aparecen una gran cantidad de secuencias relacionadas con el sistema inmunitario y la defensa. Las branquias expresan moléculas relacionadas con la estructura del tejido y el reconocimiento de lo no propio. En el manto se observan tránscritos relacionados con reproducción y formación de la concha. Por último en el músculo encontramos muchas proteínas relacionadas con el metabolismo del calcio y con funciones contráctiles.

Nuestros análisis muestran que los transcriptomas de estos cuatro tejidos tienen perfiles de expresión propios, no compartidos con otros tejidos, que concuerdan con su estructura y función. Sin embargo, también hemos comprobado que los transcriptomas concretos de cada tejido son complejos y con funciones altamente específicas. Por ejemplo, los hemocitos expresan una cantidad extremadamente alta de péptidos antimicrobianos; las branquias tienen un transcriptoma relacionado con la homeostasis del calcio; en el manto se han encontrado funciones hematopoyéticas, antifúngicas y sensoriales; y, sorprendentemente, el músculo también expresa muchas moléculas relacionadas con la defensa. Además, en el análisis global de todos los tejidos hemos encontrado gran cantidad de secuencias presentes en rutas metabólicas relacionadas con el cáncer. Esta información supone un incremento en la cantidad de información disponible en el campo de la fisiología de los moluscos, especialmente del mejillón.

DISCUSIÓN

Los avances en fisiología cada vez se basan más en la biología molecular y las herramientas genómicas. Desde 1982 hasta hoy en día, el número de bases en el GenBank se ha duplicado cada 18 meses aproximadamente (GenBank Release Notes, 2015). En lo que respecta a los mouscos bivalvos, en 2006 en una búsqueda en GenBank con las palabras clave "Mollusca" + "Bivalvia" se obtenían 32.000 resultados. En ese momento la mayoría de las secuencias se obtenían mediante la investigación gen a gen (Saavedra y Bachere, 2006). A día de hoy (10/09/2015) al buscar secuencias nucleotídicas de bivalvos se obtienen más de 720.000 resultados y el Archivo de Secuencias Cortas (*Short Read Archive* - SRA) tiene 771 resultados de bivalvos. Éste es el resultado de casi 10 años de investigación en la genómica de los bivalvos. La contribución que se ha presentado en esta tesis para almejas y mejillones es la siguiente: 1 colección SRA para *R. philippinarum* (SRX100159), de las 30 que existen actualmente para esta especie, y 7 colecciones SRA para *M. galloprovincialis* (SRP033481), de las 30 existentes de mejillón. Esto supone millones de nuevas lecturas brutas y decenas de miles de secuencias nucleotídicas nuevas que han enriquecido las bases del conocimiento genómico en los bivalvos.

Los resultados generados en esta tesis significan conocimiento básico para iniciar y apoyar otros descubrimientos científicos, no solo en nuestro grupo sino para otros colaboradores e investigadores de otras instituciones en todo el mundo. Hemos publicado otros trabajos interesantes y útiles para la comunidad investigadora en inmunología de bivalvos, por ejemplo, genes de referencia que se pueden usar en qPCR para *R. philippinarum* y *M. galloprovincialis* (Moreira *et al.,* 2014) o la comparación de dos especies de mejillón de la familia *Mytilidae* presentes en hábitats muy diferentes como son *M. galloprovincialis* y *B. azoricus* (Martins *et al.,* 2014). En cuanto al enriquecimiento de las bases de datos públicas con secuencias de bivalvos, en el marco del proyecto ReProSeed se han publicado, y se siguen publicando, transcriptomas como el de la vieira (*Pecten maximus*) (Pauletto *el al.,* 2014). Un resultado más directamente relacionado con esta tesis es la reciente publicación de los resultados de una infección de *R. philippinarum* con *Perkinsus olseni* que se analizó mediante el microarray descrito en el tercer artículo presentado en esta tesis (Romero *et al.,* 2015).

Actualmente, y gracias a la genómica, el conocimiento de las bases moleculares y los mecanismos celulares involucrados en los procesos fisiológicos de interés en acuicultura (la reproducción, el crecimiento y la respuesta inmunitaria) ha aumentado en gran medida. El próximo gran reto será probablemente la búsqueda del papel funcional específico de todos estos nuevos genes y proteínas. También es necesario mejorar el trabajo bioinformático de secuenciación, mapeado, ensamblaje y anotación, ya que el porcentaje de error, aunque bajo, siempre está presente al obtenerse tal cantidad de información (Goldman *et al.,* 2014). El futuro de todas las técnicas de secuenciación y expresión usadas en esta tesis es emocionante y prometedor. Estas herramientas no han dejado de sufrir mejoras técnicas, ya que la ciencia siempre solicita nuevas aplicaciones y actualizaciones para solucionar las infinitas preguntas que surgen cada vez que se hace un nuevo descubrimiento.

En lo que respecta al futuro próximo en la investigación de bivalvos, una posibilidad interesante y apenas explotada son las investigaciones epigenéticas, el estudio de los cambios químicos en el genoma que permiten la activación o el silenciamiento diferencial de múltiples genes. La epigenética solo se ha usado en el caso de la ostra del Pacífico (*Crassostrea gigas*) para el estudio sus mecanismos de metilación del ADN y su posible relación con la regulación de la expresión genética (Gavery y Roberts, 2014; Olson y Roberts, 2014). También existe la posibilidad de la secuenciación del genoma, que en el caso de bivalvos se limita a estos dos casos: el genoma de la ostra del Pacífico (*C. gigas*) (Zhang et al., 2012) y el borrador del genoma de la ostra perlífera japonesa (*Pinctada fucata*) (Takeuchi *et al.*, 2012).

Después de estos años de investigación, del enriquecimiento de las bases de datos con secuencias de bivalvos y de la publicación de los genomas de dos bivalvos se han añadido cinco categorías específicas de moluscos a la base de datos de ontología genética (*Gene Ontology* – GO), entre ellas "Formación del biso" y " Formación de la concha" (Kawashima *et al.*, 2013). Todos estos datos serán una información muy útil para futuros trabajos genómicos tanto en bivalvos y como en especies afines. Todo esto es solo en principio, ya que en los próximos años se publicarán más genomas de moluscos y nuevas categorías de GO que harán más productivo el trabajo a medida que la investigación avance.

La obtención del "*ribosome profiling*" (Ingolia *et al.,* 2009) es un nuevo método, también relacionado con la expresión genética, que todavía no ha sido utilizado en bivalvos. Esta tecnología ayuda a explicar aspectos clave de la regulación de la expresión, ya que llena el vacío que existe entre la expresión genética y la cuantificación de proteínas y permite estudiar la regulación post-transcripcional y post-traduccional de la expresión genética (Kuersten *et al.,* 2013).

La velocidad con la que todas estas técnicas aparecen, se actualizan y se vuelven obsoletas es todo un reto para los investigadores. Esto subraya la importancia de la investigación de la expresión y regulación genética en las ciencias naturales, así como la necesidad de la comunidad científica de estar siempre al tanto de las nuevas corrientes y de contar con colaboradores de diversos campos, incluyendo la bioinformática, para seguir avanzando a este ritmo.

CONCLUSIONES

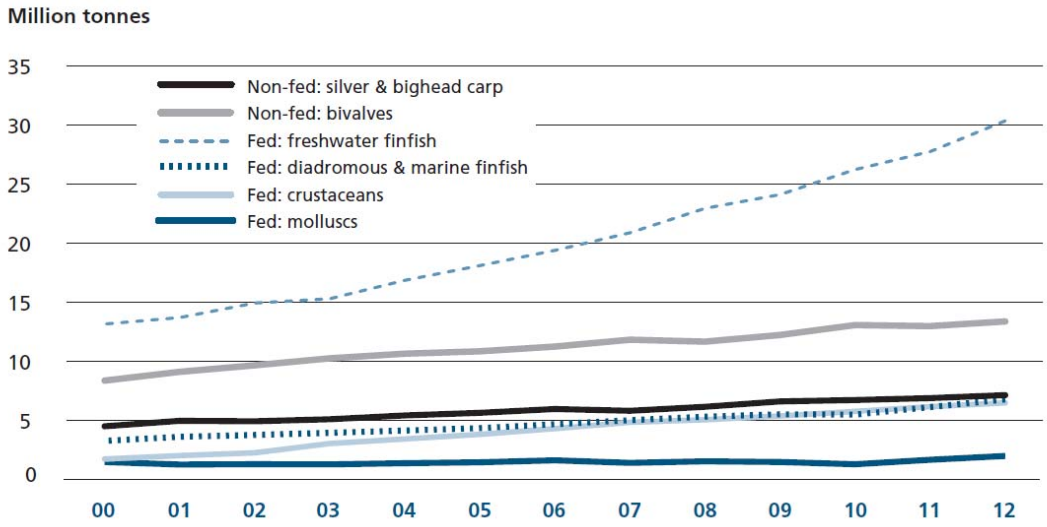
1. Según el análisis de expresión genética por qPCR, la respuesta de *R. philippinarum* ante una infección con *V. alginolyticus* es más rápida y efectiva que la de *R. decussatus.* Además, *R. decussatus* no expresa los factores de transcripción que iniciarían la respuesta inflamatoria. El patrón de expresión de IκB en *R. philippinarum* sugiere una regulación negativa de la respuesta inflamatoria 24 horas después de la infección.
2. Hemos realizado el primer análisis transcriptómico de *R. philippinarum* centrado en su sistema inmunológico. Se han descubierto casi 30.000 nuevas secuencias, de las que más del 10% tienen relación con la inmunidad. Encontramos secuencias relacionadas con el sistema inmune nunca antes descritas en bivalvos como la de genes como C2, C4, C5, C9, AIF, Bax, AKT, TLR6 y TLR13. Hemos caracterizado tres rutas importantes para el sistema inmunológico: la apoptosis, la ruta de señalización de los receptores *toll like* y la cascada del complemento.
3. Hemos diseñado la segunda versión del microarray de oligos de *R. philippinarum*, enriquecido con secuencias del sistema inmunológico. Se ha demostrado su eficacia para estudios de expresión en hemocitos estudiando la secuencia temporal de la respuesta inmunitaria frente a una infección con *V. alginolyticus*:
    - Los genes relacionados con señalización, transcripción y apoptosis se expresan a las pocas horas después de la infección.
    - En torno a las 8 horas post infección se expresan los genes típicos de defensa y respuesta inmunitaria.
    - 24 horas después de la infección se activan una gran cantidad de procesos para superar la infección, incluyendo la quimiotaxis.
    - La regulación negativa de los procesos desencadenados por la infección ocurre 72 horas post infección.
4. La inmunocompetencia de *M. galloprovincialis* surge durante el primer estadío de desarrollo larvario: las larvas trocóforas son capaces de fagocitar. La expresión de genes relacionados con el sistema inmunitario empieza a crecer a partir de la etapa de larva velíger, preparando a las larvas para su fijación. Tanto en la etapa de larva velíger como en la metamorfosis, la expresión de estos genes inmunitarios, característicos de los hemocitos, se localiza en el borde del manto, lo que sugiere un posible papel en la hematopoyesis.
5. Hemos obtenido casi 400 millones de lecturas de Illumina del transcriptoma de distintos tejidos de *M. galloprovincialis*, y por primera vez en bivalvos se han estudiado sus perfiles de expresión comparativos. Hay secuencias relacionadas con el sistema inmunitario en todos los tejidos, lo que podría explicar la gran resistencia de los mejillones al estrés y los patógenos.
    - Los hemocitos tienen un perfil de expresión típicamente defensivo, con una gran cantidad de péptidos antimicrobianos.
    - El manto muestra funciones reproductoras, sensoriales y antifúngicas, y también secuencias relacionadas con la hematopoyesis.

- En el músculo se han encontrado secuencias con funciones contráctiles y también relacionadas con el sistema inmunitario.
- Las branquias tienen muchas secuencias relacionadas con el tejido conectivo y con el reconocimiento de lo no propio, aunque también se han descubierto funciones osmóticas y homeostáticas para este tejido.

INTRODUCTION

BACKGROUND

Aquaculture is the culture of aquatic organisms, in marine or continental waters, that implies human intervention in the breeding process to enhance the production. Nowadays, aquaculture accounts for nearly 50% of the world's food fish production and is, according to the Food and Agriculture Organization of the United Nations (FAO), one of the fastest growing food-producing sectors. It has a food production per hectare very superior to the agriculture or cattle industry. Finned fish represent the main proportion of the aquaculture production followed by molluscs (FAO. The State of World Fisheries and Aquaculture, 2014).



	Inland aquaculture	Mariculture	Quantity subtotal		Value subtotal	
	(Million tonnes)	*(Million tonnes)*	*(Million tonnes)*	*(Percentage by volume)*	*(US$ million)*	*(Percentage by value)*
Finfish	38.599	5.552	44.151	66.3	87 499	63.5
Crustaceans	2.530	3.917	6.447	9.7	30 864	22.4
Molluscs	0.287	14.884	15.171	22.8	15 857	11.5
Other species	0.530	0.335	0.865	1.3	3 512	2.5
Total	**41.946**	**24.687**	**66.633**	**100**	**137 732**	**100**

Figure1. World aquaculture production. Font: FAO. The State of World Fisheries and Aquaculture, 2014.

The responsible management of aquaculture is a key issue following production, environmental and sustainability criterions. All through the intensive aquaculture production, the inefficient use of food and the high production of waste have adversely affected local environments in many regions. Thus, it is imperative to improve these practices to protect water quality, which is essential for optimum health and production in aquaculture. A very important aspect in this context is the need of efficient controls for sanitary management in aquaculture because diseases are one of

the main bottlenecks for the development of aquaculture. That is the reason that the scientific research has a key role to provide solutions (FAO. Fisheries & Aquaculture, 2015).

Mussels and clams are cultured worldwide and they have important worldwide commercial value. Bivalves are an interesting group, not only in terms of aquaculture production but also because they have an important ecological role in the depuration of waters and as environmental sentinels (Wootton *et al*., 2003). Additionally, molluscs include intermediate hosts for serious parasitic human diseases (Morgan *et al.*, 2001). These are only a few reasons to study bivalves from the point of view of the molecular biology. From the polymerase chain reaction (PCR) to the RNAseq technology less than 50 years have passed by but many gene expression tools have been developed. We have used a variety of them to enrich and clarify some aspects of mussel and clam immune system biology.

BIVALVES AQUACULTURE

Aquaculture is the process by which all the stages in the development of the cultured species are controlled, and it is able to provide of all the necessary means for the optimum breeding of cultured animals or plants. Aquaculture has developed as a way to feed the increasing worldwide population. The major purpose of the aquaculture business is to produce human food: more than 50% of the worldwide consumed fish comes from this source. Another uses of aquaculture are the ecosystem repopulation, obtaining pharmaceutical or industrial products and ornamentation (aquaria, nacre, pearls…). Culture of fish, molluscs, crustaceans, echinoderms and algae constitute the main groups of interest in aquaculture.

Aquaculture exists thousands of years ago. In Hawaii, rests of prehistoric ponds were found. They were used for algae culture and for farming and maintenance of aquatic animals (Farber, 1997).

Nile tilapias (*Oreochromis niloticus*) in ponds were frequently represented in ancient Egypt low reliefs (Bardach *et al.,* 1972), especially in the period from 3500 to 3200 b. C. in the predinastic age. Nebamun tomb paintings, dated between 1400 and 1350 b. C., are also very famous.

Figure 2. Egyptian low relief of a Nile tilapia dated in the predinastic age. Font: Relieve - Foro Egipto Viajar e Historia de Egipto.

Figure 3. Nebamun tomb paintings with tilapias and other aquatic species in ponds.
Font: British Museum - Room 61 Tomb-chapel of Nebamun.

Regarding molluscs culture, oysters were yet cultured in Japan, Greece or Rome more than two thousand years ago (Pillay, 1997). Galicia is a perfect place for the culture of bivalve molluscs such as mussel and clams. The reasons are its orography: the *rías*, and its geographic, oceanographic and climatic location: central North Atlantic water upwellings and nutrient dragging in rain periods. All these characteristics favor the development of phytoplankton, essential for bivalves feeding.

Clam

Clam culture in Galicia mainly lies in the seed sowing in natural parks to be restored after the harvesting. It is a particular system as it is difficult to divide clam aquaculture and natural growing. Seed is produced in hatcheries until they gain the optimum size to develop on their own in the intertidal sand. Clam harvesting is made by *marisqueo*: walking the intertidal areas to collect clams with special hand shovels or rakes. Periodically, parks need to be cleaned of predators, algae and mud.

Carpet shell clam (*Ruditapes decussatus*) harvesting started in Galicia in the late 20's. The lack of regulatory laws promoted the indiscriminate collection, almost causing the extinction of the natural populations. The current regulation fixes the amount of carpet shell clam that can be harvested, which is very low compared to other bivalves such as mussels or Manila clam (*Ruditapes philippinarum*). Nevertheless, it means an important profit due to its high market value, especially from the 80's. Carpet shell clam is a scarce and appreciate bivalve. Its production suffers temporary but drastic interruptions by salinity changes, caused by severe rains, and diseases. In Galicia, various pathogens (virus, bacteria such as *Vibrio alginolyticus* or *Vibrio splendidus* and parasites like haplosporidians and *Perkinsus olseni*) have been associated to massive mortalities in parks and natural populations, which had affected local Galician economy (Novoa and Figueras, 2000; Gómez-León *et al.,* 2005).

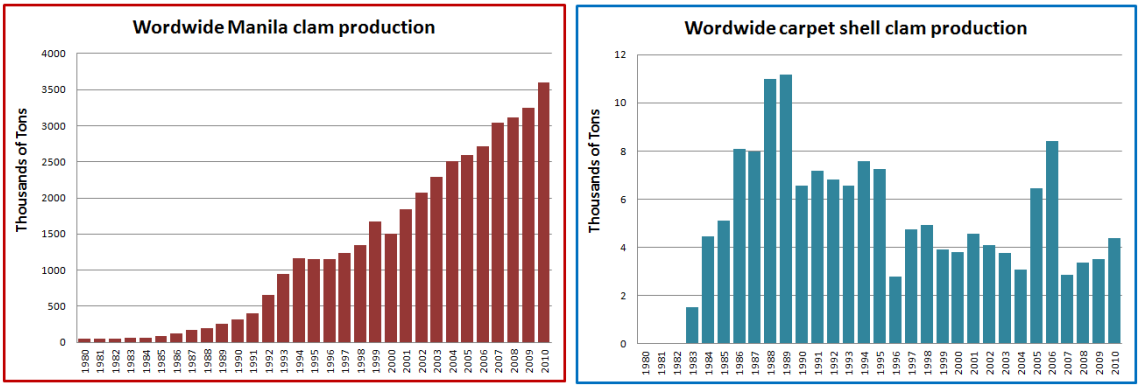


Figure 4. Worldwide production of clam: in red, *R. philippinarum*; in blue, *R. decussatus*.
Font: Own work from FAO data.

Overexploitation and irregular output of local clams, like the carpet shell clam, favored the importation to Europe of Manila clam in 1972. Due to its commercial value, Manila clam has been subsequently introduced and permanently established in many areas. After the production in hatcheries, natural reproduction made possible the species settlement along with the natural populations. Manila clam showed more resistance and faster growth than local species, the carpet shell clam. Consequently, Manila clam represents the main proportion of clam production in Europe and the rest of the world.

Mussel

The first way of myticulture known in Europe was the palisade cultures in France or *bouchots* (word with Celtic origin: *bou* = palisade and *choat* = wood). In 1846 the suspension culture was initiated in Spain (Barnabé, 1991) with *bateas*, raft-like structures with hanging ropes in which juveniles are fixed to grow until they reach commercial size.

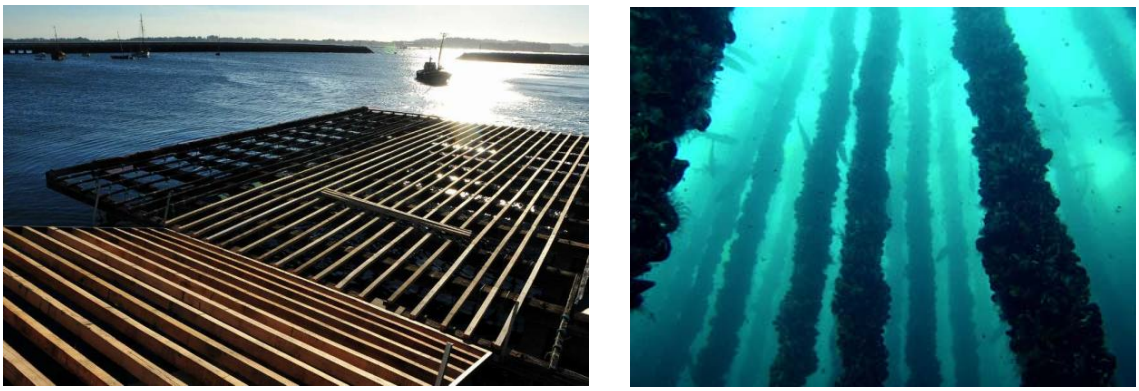


Figure 5. *Bateas* in the *ría* of Vigo, Galicia, NW Spain. Font: Google images.

According to FAO data, each year 1,500,000 tons of mussels are produced worldwide, 48% of which is produced in the European Union. Average production in Galicia is around 150,000 and 300,000 tons, these data place the Galician mussel industry as the second most important (10% of the production in tons and 20% in value), surpassed only by China, with an average year production of 600,000 tons.

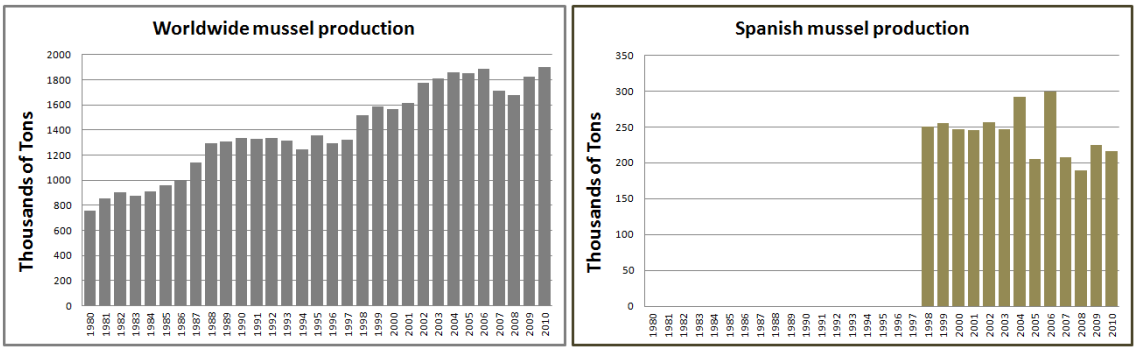


Figure 6. Mussel production: in grey, worldwide production; in brown, Spanish production.
Font: Own work from FAO and Spanish *Ministerio de Agricultura, Alimentación y Medio Ambiente* data.

Mediterranean mussel (*Mytilus galloprovincialis*) is a very resistant species to pathogens and severe environmental conditions (Kurelec *et al.*, 1991; Romero *et al.*, 2014). Myticulture is very stable and has great value in the local economy. *Marteilia refringens* and *Mytilicola intestinalis* are the major pathogens for mussels but none of them have caused severe impact to their production. Currently, the major issue for mussel production are the bio-toxic red tides, which cause enormous economical losses to the mussel farmers. Red tides force the production facilities to close and bann the sale of mussels farmed in the affected areas. Red tides are natural fenomena characterized by microalgal blooms. These blooms have important consequences in marine environments due to the high concentration of toxins released by the phytoplankton. Toxins are safe for seafood, but in humans cause serious digestive disorders and diarrhea (Bricelj and Shumway, 1998).

IMMUNE SYSTEM IN BIVALVE MOLLUSCS

According to the FAO, the future of the molluscan production needs to be planned as an integral program that includes aspects such as the creation of a special research plan on bivalve molluscan pathologies. Nevertheless, the knowledge of the immune response in these species is still limited and the fight against the pathologies is based in preventive strategies and the removal of diseased individuals. Little is known about the molecular basis of the immune response in bivalves, which is the reason that in the last decade many studies have been published about bivalve genomics (Saavedra and Bachère, 2006; Romero *et al.,* 2012).

Vertebrate immune system is characterized by an innate unspecific response and acquired response, with memory after a previous contact with the pathogen. Bivalves lack the acquired response in a narrow sense (Criscitiello and de Figueiredo; 2013), but

in addition of physical barriers such as the shell and the mucus, they have a potent and efficient cellular and humoral innate immune system.

Cellular immunity

Defense cells in bivalve mollusks are the hemocytes. These cells are traditionally classified depending on their morphologic and functional characteristics in granulocytes and hyalinocytes (Bayne CJ *et al*., 1980; Carballal *et al*., 1997; Prado-Alvarez *et al*., 2012). Granulocytes present from scarce to abundant cytoplasmic granules or lysosomes and hyalinocytes are characterized by the absence or very scarce presence of these granules and a high proportion nucleus: cytoplasm.

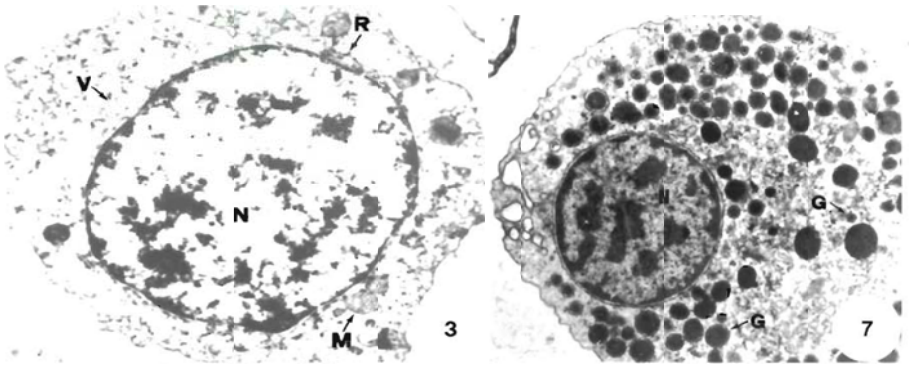


Figure 7. Hyalinocyte and granulocyte of M. *galloprovincialis.* Font: Carballal *et al*., 1997.

Hemocytes trigger their defense mechanisms after pathogen recognition and eliminate these cells or foreign particles by phagocyting and destroying them by these means:

- Lysosomal enzymes
- Reactive oxigen species (ROS) synthesis, known as respiratory burst.
- Nitric oxide production by the nitric oxide synthase after phagocytosis.

In addition to phagocytosis, hemocytes have other functions as central players of the immune system, such as modulate the inflammatory response:

- Production and regulation of humoral factors.
- Formation of aggregates and nodules of hemocyte and non-hemocyte cells when the pathogen invasion is too high. This way a cell-net is constituted and the foreign particles stay trapped (Ratcliffe *et al.,* 1985).
- Pathogen isolation by encapsulation when the pathogen is too big to be phagocyted. Hemocytes and fibroblasts surround and isolate the parasite with concentric layers of cells (Feng, 1988).

Humoral immunity

Humoral immunity is mediated by the proteins and glycoproteins found in the body fluid (the hemolymph), the hemocytes or both. The most important components of the humoral defense are:

- Lysozime and lysosomal enzymes:

There are many lysosomal enzymes present in the hemolymph: lysozyme, β-glucuronidase, alkaline and acid phosphatases, lipase, aminopeptidase and amilase (Chu, 1988).They are released after a bacterial infection. The main responsible in the bacterial defense, with further nutritional functions, is the lysozyme. Lysozyme breaks the bacterial peptidoglycans, the core structural element of the bacterial wall (Zhao *et al.,* 2010).

- Lectins

Lectins are pattern recognition receptors (PRRs), carbohidrate-binding proteins with opsonizing and agglutinating competence to facilitate phagocytosis and complement activation. They are considered receptors for pathogen-associated molecular patterns (PAMPs). These molecules also mediate chemotaxis (Vasta, 2009). In molluscs lectins are free in the hemolymph or bound to the membrane of hemocytes. Hemocytes possess receptors capable to bind more than one class of lectin (Humphries and Yoshino, 2003).

- Heat shock proteins (HSPs):

HSPs are synthesized as a response to diverse cellular stress factors: raise of temperature, presence of heavy metals, low oxygen levels, oxidative stress, osmotic stress or physical and chemical agents. They can modulate the mitochondrial apoptosis (Parcellier *et al.,* 2003) and regulate the expression of the nuclear factor kappa B (NF-κB), a key transcription factor for the immune response (Guo *et al.,* 2009).

- Prophenoloxidase system (proPO):

proPO is a modified form of the complement response found in some invertebrates. It consists in the transformation of phenolic substances in melanine by a cascade reaction. Melanization is a very important defense mechanism in arthropods and insects. The presence of the phenoloxidase enzyme in bivalves could be involved in phagocytosis, encapsulation and the production of opsonin-like and lytic factors (Cerenius *et al.,* 2008).

- Proteases and protease inhibitors:

Proteases mediate in multiple biologic processes, including the immune response (Bird *et al.,* 2009). Pathogen organisms also possess proteases to escape the host defense mechanisms and to assist the infection (Ordás *et al.,* 2001). For that reason many organisms have developed protease inhibitor enzymes. It has been established the presence of proteases and protease inhibitors in bivalves' hemocytes, to fight the pathogen infection and to favor its destruction after phagocytosis (Xue *et al.,* 2006).

- Antimicrobial peptides (AMPs):

AMPs are small cysteine-rich proteins with a key role in innate immunity. They are highly conserved and present in all the living beings. AMPs principal mechanism of action lies in to alter the permeability of the pathogen plasmatic membrane to lyse cells, but they may also interfere with DNA and protein synthesis, protein folding, and cell wall synthesis. In the last decades AMPs have been extensively studied in mussels and classified into four groups (Li *et al.,* 2011):

  - Defensins: They are especially active against Gram+ and Gram- bacteria but have also shown antifungal activities.
  - Myticins: These molecules show antimicrobial activity against Gram + bacteria but very limited antifungal properties. Myticin C also presents antiviral and immunoregulatory properties.
  - Mytimicins: They have strict antifungal activity.
  - Mytilins: They show broader spectum against distinct pathogens depending on the isoform, including Gram+ and Gram- bacteria, fungi and protozoa.

Recently, another four groups have been described: big defensins, mytimacins (Gerdol *et al*., 2012), myticusins (*Liao et al*., 2013) and mytichitins (Qin *et al*., 2014). Although their spectrum and mode of action against invading pathogens needs further investigation it seems that myticusins and defensins have similar functions, meanwhile mytichitin was active only against Gram + bacteria.

- Complement:

Complement system is comprised of more than 30 plasma and membrane-associated proteins that can be activated by different pathogens in sequential cascade reactions. The key effector molecules, C3a and C5a (anaphylatoxins) are the most potent complement activation products, showing diverse activities on many cell types ranging from chemoattraction to apoptosis. Complement proteins promote opsonisation, phagocytosis and intracellular killing of pathogens by immune effector cells (Huber-Lang *et al.,* 2006). These are the known activation pathways:

  - Classical pathway: It is activated when IgG o IgM immune complexes are formed, they bind to C1q and the activation cascade starts. It is known to occur only in vertebrates.
  - Lectin pathway: It is initiated when PRRs like mannose binding lectins or ficolins recognize and bind PAMPs in the surface of the pathogen, leading to the activation of C3 convertase. It is the most ancestral pathway.
  - Alternative pathway: It is triggered by carbohydrates, lipids and proteins on the pathogen surface. They activate a slow and constant hydrolysis of C3 (C3 tickover) leading to the formation of C3 convertase.
  - Extrinsic protease pathway: It involves direct cleavage of C3 and C5 by a series of proteolytic enzymes released by neutrophils and macrophages in vertebrates.

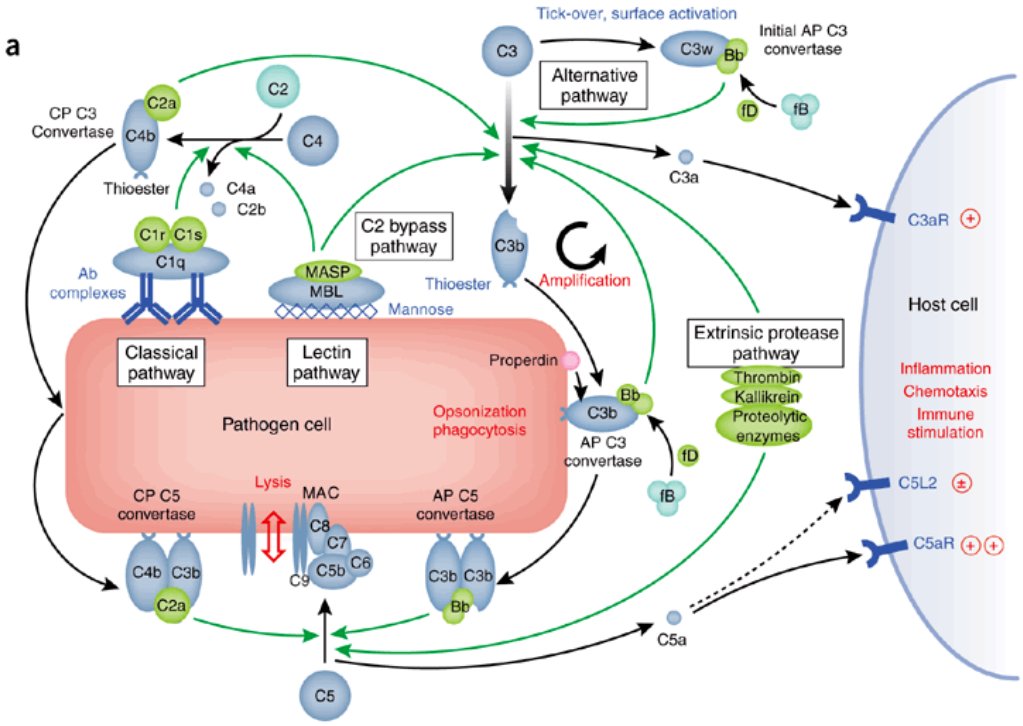


Figure 8. The complement cascade and its activation pathways. Font: Ricklin and Lambris, 2007.

The central component of the complement, C3, has been described in bivalves (Prado-Alvarez *et al.*, 2009), but the C1q domain-containing proteins are the best characterized group in this group of animals. They are produced after an infection and have roles in pathogen recognition, inflammation, apoptosis, autoimmunity, and cellular differentiation. Furthermore, they are activators of the NF-κB pathway (Gestal *et al.*, 2010). Little is known about the complement system in molluscs and, despite the similarities in structure, there is still no evidence that its components are homolog of that in vertebrates.

- Cytokines:

Cytokines are small glycoproteins with regulatory immune functions. They are the most important regulators of innate and acquired immunity. Their mechanism of action is very fast and powerful in the amplification of the immune response despite the extreme low concentration in body fluids. Many cytokines have a pleiotropic effect (Benjamini *et al.*, 2000). Few cytokines have been characterizes in molluscs, among them the TNF-α homologue in the disk abalone (*Haliotis discus discus)* (De Zoysa *et al.*, 2010) or the IL-17 in the Pacific oyster (*Crassostrea gigas*) (Roberts *et al.*, 2008).

GENE EXPRESSION & SEQUENCING RESEARCH IN BIVALVES

The interest in bivalve genomics has emerged from the late 90s due to the importance of bivalves in aquaculture and fisheries and also because of their role in marine environmental science. Genomics focuses on the study of genes and their function. It has changed dramatically from single gene to whole transcriptome approaches. The

next big step, the genome sequence is a very powerful tool as it contains the codified information of the biology of every organism.

Genomic research is basic to solve specific problems in bivalve aquaculture, including disease control. They can identify RNAs or proteins expressed in response to stress or pathogens and compensate the usual absence of clinical manifestations. Additionally, these tools can help to find genetic markers, which show the relation between genes, environmental factors and physiology. All this new information will facilitate the study of bivalves and increase the success of molluscan aquaculture, facilitating the monitoring of production, activation state of pathogens or the immune status of the animals.

In the review by Saavedra and Bachère, 2006, there are some interesting characteristics about the bivalve genome: The DNA content of the haploid bivalve genome could range from 0.65 pg to 5.4 pg, in the middle of the range for the metazoans. Haploid chromosome numbers for bivalves would be between 10 and 23. These chromosomes are very homogeneous in size making difficult cariotyping, which has to be assisted by FISH to obtain accurate results. FISH has also revealed that bivalves are the only invertebrates that carry the vertebrate telomeric sequence. Repetitive long DNA sequences are common in bivalves and some repetitive DNA seems to be related to transposons. Some of these characteristics have been confirmed after the publication of the first bivalve genome, the Pacific oyster (*Crassostrea gigas*) genome (Zhang *et al*., 2012).

Several problems currently affecting bivalves would benefit from the genomics research such as the presence of toxins of phytoplanktonic origin, which causes huge losses to the industry. They produce red tides and cause several diseases in humans after consumption (Bricelj and Shumway, 1998). Genomic tools can be used to study the depuration processes for each toxin. Additionally, the study of the larval stages at cellular and molecular levels, diet design, and growth as well as factors related to disease resistance, stress or pollution are aspects of bivalve physiology that can be studied using genomic techniques.

All this techniques could be supported by other branches of science to obtain substantial results such as: selected strains, well-characterized local races, and mutant phenotypes. Some are already available in the Pacific oyster (reviewed in Hedgecock *et al.,* 2005). On the other hand, cell biology tools, such as cell lines, are not yet available in bivalves slowing down research from a functional and applied point of view.

First genomic studies in bivalves where conducted using homology cloning. But this approach was not successful because of the limited information on bivalve genes in databases. The next step in the genomic approach was to describe the differentially expressed immune-related genes in bivalves using BAC libraries, EST collections from cDNA or subtractive hybridization libraries (SSH). These methodologies have been used to find gene markers related to pathogen resistance, and have led to a great increase in the number of sequences in databases. These collections were also used to design microarrays. Anyway, the major advance in recent genomic research occurred with the development of next-generation sequencing technologies (NGS). In bivalves, the NGS and the microarray technologies have been applied to investigate different physiological processes (reviewed in Romero *et al.,* 2012), such as shell deposition and

repair, biomineralization, environmental monitoring, reproduction, skewed sex ratios of offspring or adaptation to climatic and pollution factors.

Genomic advances happen in an exponential way. In the next incoming years the amount of available information will allow researchers huge progresses regarding molecular biology of bivalve molluscs.

GENE EXPRESSION & SEQUENCING TECHNIQUES

The polymerase chain reaction (PCR) is a laboratory technique of molecular biology that was discovered less than fifty years ago, in 1983, by Kary Mullis (Mullis *et al.*, 1987: US 4683195 A). It is fundamental to many molecular biology projects. Most of the recent achievements in molecular biology were based and inspired by PCR. It is impossible to imagine modern biology without it. It provides a method for amplifying the copy number of DNA molecules using DNA polymerase, which is a first step necessary for multiple different downstream analyses.

In the 1990s accurate quantitative gene expression analysis started to became popular when the first real-time PCR machines were developed. Nowadays, this technology is still the method of choice for performing gene expression analysis in many laboratories across the world. However, the last decade has seen an explosion in new technological innovations that offer faster, cheaper and more accurate analyses.

The PCR is performed on a DNA template, single or double-stranded. Two oligonucleotide primers that flank the DNA sequence to be amplified have to be added to the reaction, as well as the four nucleotide triphosphates (dNTPs: dATP, dTTP, dCTP, dGTP). Finally, a heat-stable polymerase and magnesium ions in the buffer are also imperative. The reaction is performed by temperature cycling: high temperature is applied to separate the strands of the double helical DNA (melting temperature), then temperature is lowered to let primers anneal to the template, and finally the temperature is set around 72 ºC, optimum for the polymerase that extends the primers by incorporating the dNTPs (Mullis *et al.*, 1987: US 4683195 A).

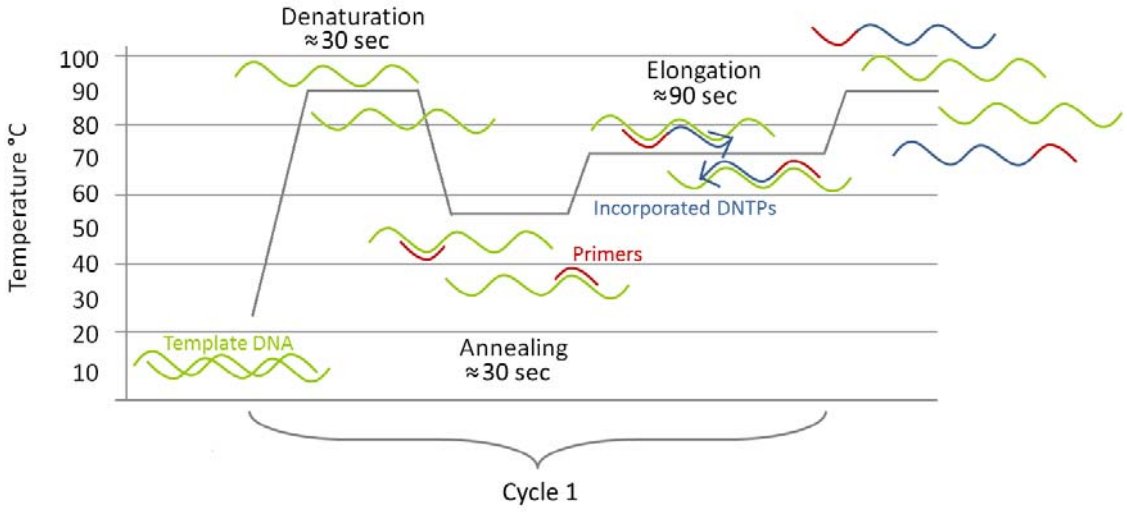


Figure9. Schematic representation of the polymerase chain reaction. Font: Own work.

Initially the PCR was developed for DNA detection and, eventually, quantification, but it was rapidly applied to perform gene expression analysis. Gene expression techniques start with the isolation of the mRNA and its retrotranscription to cDNA. The viral enzyme reverse transcriptase converts single stranded RNA molecules into double stranded DNA. Thus, each individual transcript can be amplified by PCR. The final amount of PCR product is proportional to the initial amount of cDNA in the sample, allowing the estimation and comparison of gene expression levels. Only the genes with identified sequences can be used for expression studies because without the complete or partial expressed sequence tag (EST), PCR cannot be made, as a known DNA sequence it is needed to design primers.

Gene expression studies are nowadays one of the most widespread techniques in almost all biological fields. It was typically used to compare the activation or repression of a gene after certain stimuli such as drugs, infections or physical changes in the environment; but many other comparisons can be made thanks to gene expression tools: among species, tissues or developmental stages, to follow a time course or to detect and quantify the presence of non-self cells.

qPCR

The real-time PCR or quantitative PCR (qPCR) is used to amplify and simultaneously detect and quantify a targeted DNA molecule following the cycling principle of basic PCR. The final amount of product generated by a PCR run is proportional to the amount of starting material in the sample, allowing gene expression levels to be estimated. However, PCR are usually saturated in the last cycles, making comparisons at the end of the reaction meaningless.

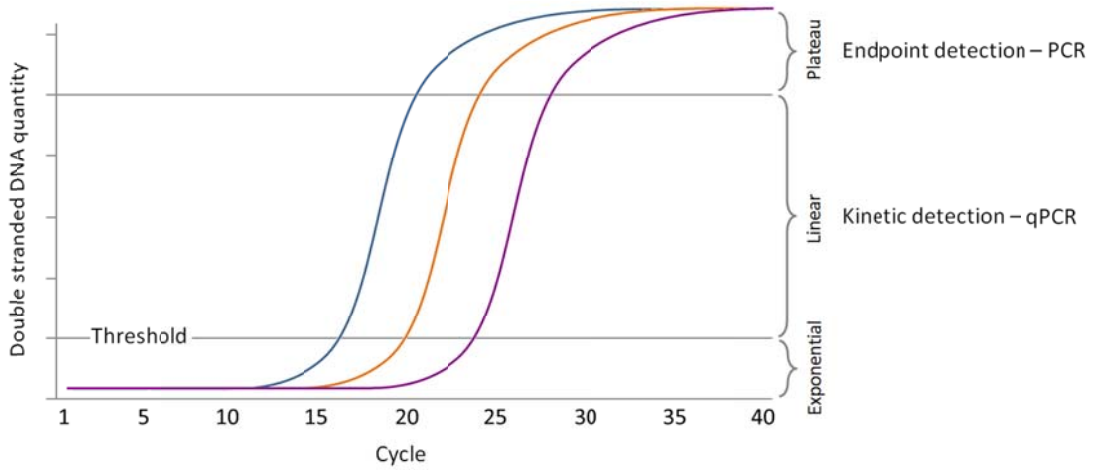


Figure 10. Schematic view of the qPCR reaction, the different color lines represent different samples. Font: Own work.

qPCR allow researchers to follow the PCR process as it progresses, providing an accurate approximation of initial gene expression levels: the reaction takes place in a thermocycler which measures the light emitted by a fluorescent molecule in real time during the exponential phase of the reaction, avoiding the saturation (plateau effect) in the last cycles. The most common methods for the detection of products in qPCR are:

- Non-specific fluorescent dyes, for instance SYBR green, that intercalate with any double-stranded DNA, as the product of PCR.
- Sequence-specific DNA probes that are labelled with a fluorescent reporter which permits detection only after hybridization or hydrolyzation of the probe during the amplification step of the reaction.

The reporter generates a fluorescence signal that is proportional to the amount of product formed in each cycle. During the initial cycles the signal is weak and cannot be distinguished from the background. As the amount of product accumulates the signal increases exponentially. The amplification response curves of the different samples reflect the difference in their initial amounts of cDNA. These curves are parallel in the growth phase of the reaction. The difference between samples is quantified by comparing the number of amplification cycles required for the curves to reach a specific threshold fluorescence level. This is called the Ct (threshold cycle) or Cp (crossing point) (Kubista *et al*., 2006).

However, as the process still relies on DNA amplification, there is the possibility that analyses will be biased by technical variation caused by the efficiency of the PCR. Therefore it is mandatory to stablish the efficiency of each primer pair designed (Pfaffl, 2001). It is also of crucial importance the choice of the most stable reference genes (also known as housekeeping genes) under the conditions that we want to measure to normalize the expression data (Vandesompele *et al*., 2002).

In situ hybridization

In situ hybridization (ISH), first developed by Joseph G. Gall (Gall and Pardue, 1969), uses a labeled DNA or RNA strand (the probe) to localize its complementary specific nucleic acid sequences in morphologically preserved cells, tissues or embryos. It relates gene activity at the DNA or mRNA level with the physical microscopic structure of the studied tissue. In the 1980s ISH methods used radioactively labeled probes to detect expression of genes on histological sections. In the 1990s, the use of chemically labeled probes made the analysis of gene expression patterns quicker and easier. The original conditions have been progressively improved with the goal of obtaining more and more robust and reliable ISH protocols, even at single-cell resolution, including genes expressed at low levels. Probes can now be produced by PCR amplification. Probes are designed including known promoter sequences for digoxygenin labeling purposes. Following hybridization, the transcript is immunohistochemically visualized using an anti-digoxygenin antibody conjugated to alkaline phosphatase, the substrate of which is chromogenic. This way is possible to determine the physical expression patterns of the studied genes.

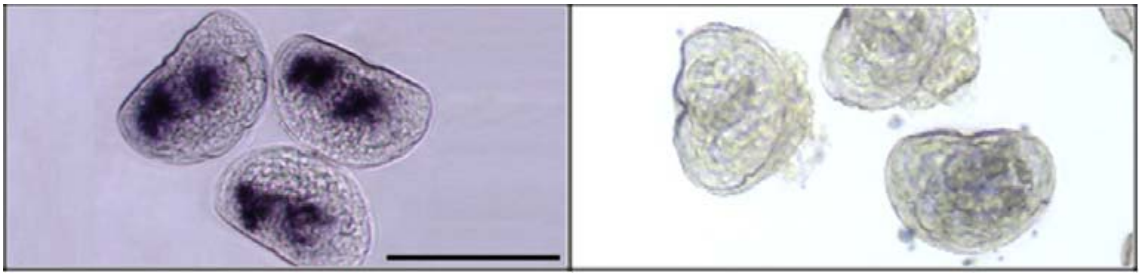


Figure 11. Example of a whole embryo ISH: Stained tissue with actin-labelled pattern and a control sample. Scale bar: 100 µm. Font: Own work.

The ISH technique has several important applications:

- DNA ISH: mapping of chromosomes and direct assignment of mapped regions with heterochromatin or euchromatin. Detection of pathogens in the host.
- FISH (Fluorescence ISH): phylogenetics, karyotyping, chromosomal aberrations and species identification, especially in microbial ecology and taxonomy.
- RNA ISH: Gene expression profiles in cells and tissues.
- WHISH (whole-mount ISH): Establish spatial and temporal gene expression patterns in embryos and early larvae during the time course of embryo differentiation.

Pyrosequencing

New generation sequencing technologies (NGS), also known as deep sequencing, high-throughput sequencing and massively parallel sequencing, have been developed because of the high demand for fast, cheap and accurate genomic and transcriptomic information. The four main sequencing strategies are pyrosequencing (Ronaghi *et al.,* 1998), cyclic reversible termination (Metzker, 2005), sequencing by ligation (Tomkinson *et al.,* 2006) and real-time sequencing (Eid *et al.,* 2009).

Pyrosequencing, the method used in this PhD thesis, is based on the "sequencing by synthesis" principle. Its basis relies on the detection of pyrophosphate release on nucleotide incorporation. The technique was developed by Mostafa Ronaghi and Pål Nyrén at the Royal Institute of Technology in Stockholm in 1996 (Nyrén, 2001: US 6 258 568BI). These are the main steps to perform the pyrosequencing reaction and analysis:

- Generation of a single-stranded template DNA library linked to microbeads.
- Deposition of single microbead-DNA sequences into individual PicoTiterPlate wells and emulsion-based clonal amplification to determine the DNA sequence:
    - DNA polymerase incorporates the complementary nucleotide onto the template. Only one specific nucleotide: A, C, G, or T nucleotides are added at a time. This incorporation releases pyrophosphate (PPi).

- PPi is converted to ATP, which acts as a substrate for the luciferase-mediated reaction. This reaction generates visible ligth proportional to the number of incorporated dNTPs. The light is detected by a camera and analyzed in a pyrogram.
- Unincorporated nucleotides and ATP are degraded and the reaction can restart with another nucleotide.

- Data analysis using different bioinformatics tools for base calling, quality control, assembly and annotation.

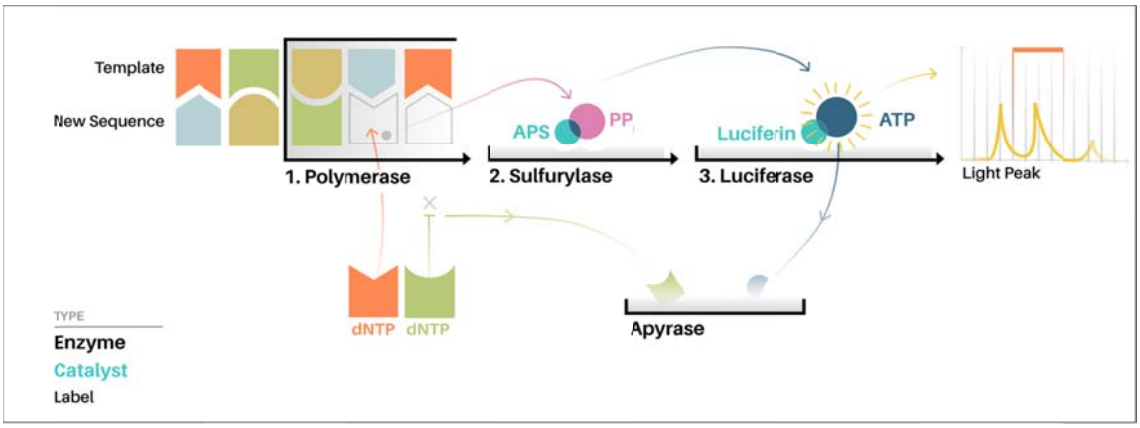


Figure 12. Basis of pyrosequencing reaction. Font: Pompilii "How Pyrosequencing Works".

Pyrosequencing technology was further licensed to 454 Life Sciences. 454 developed an array-based pyrosequencing technology which has emerged as a platform for large-scale DNA sequencing: the GS FLX+ System. This new platform features the unique combination of long reads and high-throughput, making the system well suited for larger genomic projects. The GS FLX System has been at the heart of key breakthrough genomic discoveries and thousands of peer-reviewed publications to date (454 Life Sciences, 2015). High-throughput sequencing is used for a wide range of applications:

- Whole Genome Sequencing: De novo sequencing of complex organisms or multiple bacterial genomes in a single run.
- Transcriptome Sequencing: Full-length de novo sequencing of cDNA libraries of complex organisms.
- Amplicon Sequencing: High-throughput sequencing of 100s to 1,000s samples.
- Metagenomics: Blind genome sequencing of complex environmental samples or pathogen discovery.

It is important to highlight that this was a pioneer technology. Pyrosequencing meant a huge advance in sequencing and it made possible a great amount of genomic projects due to its reduced cost and its high and fast performance.

Microarrays

Microarrays technology was derived from Southern blot, where fragmented DNA is attached to a substrate and then hybridized with a known DNA sequence (Maskos *et al.*, 1992). The first microarrays were manufactured by spotting oligonucleotides onto filter paper. The next step in the technique development was the use of solid surfaces to construct miniaturized microarrays, increasing the resolution of the spots. The gene expression profiling in these newer platforms was first reported in 1995 (Schena *et al.*, 1995). The main use of microarrays is to detect specific mRNA sequences in a sample to find out its expression profiling.

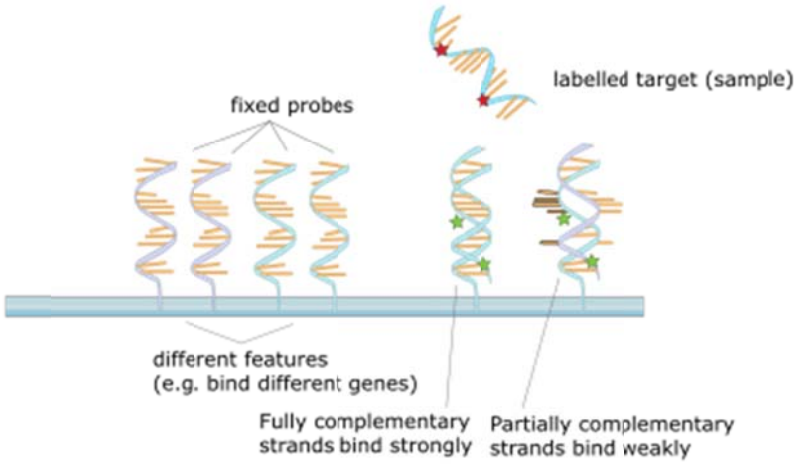


Figure 13. Schematic view of the oligonucleotide probes on the microarray glass slide.
Font: Squidonius "NA hybrid".

An oligo microarray is a set of spots attached to a solid surface, usually glass, to measure the expression level of large numbers of genes simultaneously. Each spot contains small oligonucleotides of a specific section of a gene or mRNA, the probe. An array can contain tens of thousands of probes, so a microarray experiment can study the expression profile of many genes simultaneously in a sample.

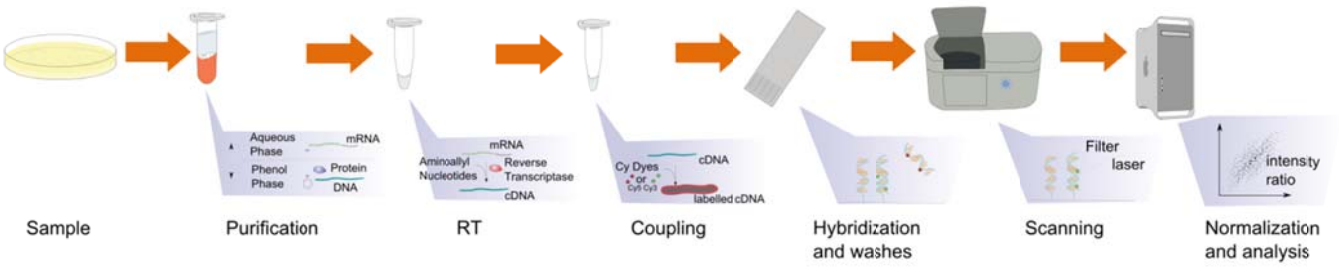


Figure 14. Pipeline of a microarray experiment. Font: Squidonius "Microarray exp horizontal".

To determine the relative abundance of the cDNA or cRNA, the sample has to be labeled and then hybridized to the microarray under high-stringency conditions to increase specifity. Fluorescence intensity for each spot is proportional to the concentration of matching cDNA. The fluorescence is detected by a scanning microscope and translated into a numerical signal that can be further analyzed to find the expression level of each probe in every sample.

RNA-Seq

The last recent years have seen a flourishing of NGS-based methods for genome analysis leading to the discovery terabytes of new genetic information in every single area of science. RNA-Seq (RNA Sequencing), also called Whole Transcriptome Shotgun Sequencing, is a technology that uses the potential of next-generation sequencing to reveal a particular moment in time of a dynamic transcriptome (Chu and Corey, 2012).

The NGS platform chosen for the RNA-Seq analysis in this thesis was Illumina. Illumina dye sequencing was based on inventions of S Balasubramanian and D Klenerman (Bentley *et al.*, 2008; Illumina. History of Illumina Sequencing, 2015). Briefly, this technique consists in:

- Fragmentation of DNA and attachment of two different adaptors.
- Attachment of DNA samples into the flow cell (glass slide with lanes) to the two types of oligos, complementary to the adaptors.

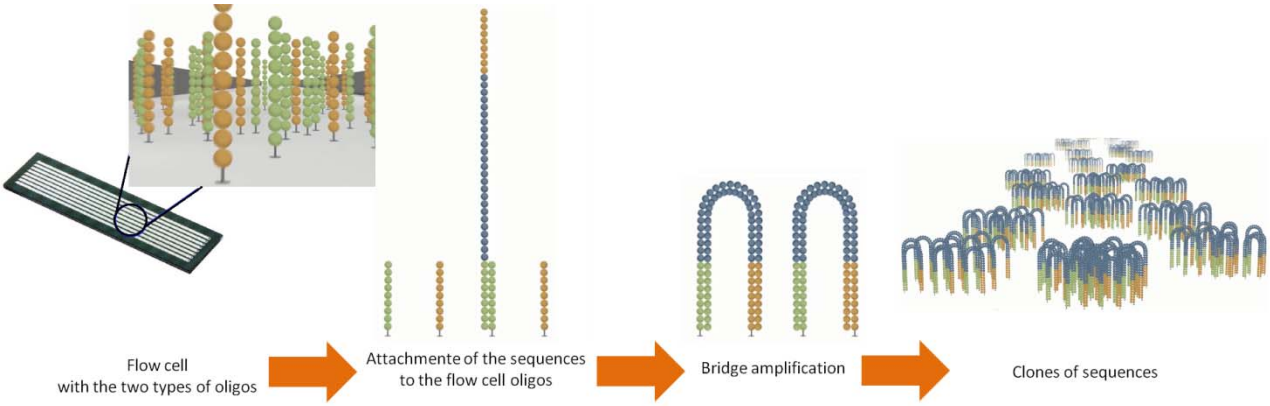


Figure 15. Clonal bridge amplification of the Illumina sequencing.
Font: Own work from Illumina Next-Generation Sequencing (NGS).

- Clonal bridge amplification to perform the sequencing by synthesis of each cluster:
  - The four nucleotides are added, each one labeled with a different fluorochrome and attached with a 3' blocking group, thus, DNA polymerase incorporates one base at a time.
  - A laser excites the dyes of the incorporated nucleotide and a computer detects the color and ascribes the corresponding nucleotide.

- Unincorporated nucleotides are washed away.
- The 3’ terminal blocking group and the dye are removed in a single step.
- The process is repeated until the full DNA molecule is sequenced.

After the bioinformatic data, RNA-Seq compares and analyzes the different samples. This technique facilitates the study of alternative gene spliced transcripts, post-transcriptional modifications, gene fusion, mutations or SNPs and, of couse, gene expression analysis. RNA-Seq can additionally look at different populations of RNA to include total RNA, small RNA, such as miRNA, tRNA, and ribosomal profiling. It can also be used to determine exon/intron boundaries and verify or amend previously annotated 5’ and 3’ gene boundaries.

Compared to microarrays, RNA-Seq technology allows for higher sample throughput than microarrays and can identify novel and low frequency RNAs previously difficult to detect and analyze. Microarrays rely on a good knowledge of an organism's genome or transcriptome and these libraries are not usually available to identify and evaluate rare allele variants. The microarrays are as good as the databases they are designed from, so they have now limited application for some research purposes.

In summary, because the information about basic physiological processes is poor at the molecular level in bivalves, we have pursued the objective of using the presented techniques to improve the knowledge about mussel and clam. Specifically, we expected to enrich the databases with immune-related sequences and shed light into the ontogeny of the immune system in bivalves. These new technologies have a huge potential to improve the modern intensive aquaculture systems for bivalves.

OBJECTIVES

The main objective of this PhD thesis is to increase our knowledge on the immune proccess of bivalves through different gene expression and sequencing tools.

The specific objectives are the following:

- Study the differential immune response of two cohabiting clam species (*R. decussatus* and *R. philippinarum*) after the same bacterial infection (*V. alginolyticus*).
- Discover new immune-related genes using pyrosequencing and increase the number of *R. philippinarum* ESTs in the public databases.
- Improve the existant microarray for *R. philippinarum* developing the version 2.0, enriched with immune-related sequences, and validate it against a *V. alginolyticus* infection.
- Describe the first steps of the immune competence in *M. galloprovincialis* describing the immune functions and gene expression at different larval stages.
- Use the RNA-Seq technology to further enrich databases with *M. galloprovincialis* tissue-specific transcriptomics and research the differential transcription among mussel hemocytes, muscle, mantle and gill.

PUBLICATIONS

GENE EXPRESSION ANALYSIS OF CLAMS *RUDITAPES PHILIPPINARUM* AND *RUDITAPES DECUSSATUS* FOLLOWING BACTERIAL INFECTION YIELDS MOLECULAR INSIGHTS INTO PATHOGEN RESISTANCE AND IMMUNITY

Developmental and Comparative Immunology 36 (2012) 140–149

Contents lists available at SciVerse ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Gene expression analysis of clams *Ruditapes philippinarum* and *Ruditapes decussatus* following bacterial infection yields molecular insights into pathogen resistance and immunity

R. Moreira [a], P. Balseiro [a], A. Romero [a], S. Dios [a], D. Posada [b], B. Novoa [a], A. Figueras [a,*]

[a] Instituto de Investigaciones Marinas (IIM), Consejo Superior de Investigaciones Científicas (CSIC), Eduardo Cabello 6, 36208 Vigo, Spain
[b] Department of Biochemistry, Genetics and Immunology, Universidad de Vigo, Spain

ARTICLE INFO

Article history:
Received 13 May 2011
Revised 24 June 2011
Accepted 25 June 2011
Available online 3 July 2011

Keywords:
Ruditapes decussatus
Ruditapes philippinarum
Vibrio alginolyticus
EST
Immune-related genes
Gene expression
qPCR

ABSTRACT

The carpet shell clam (*Ruditapes decussatus*) and Manila clam (*Ruditapes philippinarum*), which are cultured bivalve species with important commercial value, are affected by diseases that result in large economic losses. Because the molecular mechanism of the immune response of bivalves, especially clams, is scarce and fragmentary, we have examined all Expressed Sequence Tags (EST) resources available in public databases for these two species in order to increase our knowledge on genes related with the immune function in these animals. After automatic annotation and classification of the 3784 not-annotated ESTs of *R. decussatus* and 4607 of *R. philippinarum* found in GenBank, 424 ESTs of *R. decussatus* and 464 of *R. philippinarum* were found to be putatively involved in immune response. These were carefully reviewed and reannotated. As a result, 13 immune-related ESTs were selected and studied to compare the immune response of *R. decussatus* and *R. philippinarum* following a *Vibrio alginolyticus* challenge. Quantitative PCR was performed, and the expression of each EST was determined. The results showed that, in *R. philippinarum,* the immune response seems to be faster than that in *R. decussatus*. Additionally, expression of NF-κB activating genes in *R. decussatus* did not seem to be sufficient to promote an immune response after *Vibrio* infection. *R. philippinarum,* however, was able to trigger and efficiently regulate the transcriptional activity of NF-κB, even when low expression values were reported.

1. Introduction

Due to its increasing commercial value, the culture of the carpet shell clam (*Ruditapes decussatus*) and Manila clam (*Ruditapes philippinarum*) has grown in recent years. However, diseases caused by a wide range of microorganisms are associated with large economic losses (Gómez-León et al., 2005). The majority of bacterial diseases are caused by members of the *Vibrio* genus (reviewed in Gestal et al., 2008; Paillard et al., 2004). In particular, the pathogen used in this study, *Vibrio alginolyticus*, was associated with high mortality rates (up to 73%) of *R. decussatus* larvae and spat in 2001 and 2002 in a commercial hatchery in Spain (Gómez-León et al., 2005). Although there are no reliable data regarding mortalities on seabeds, it is known that *R. philippinarum* is more resistant to physical stress and pathogens than is *R. decussatus* (FAO, 2005; Tanguy et al., 2008). Although bivalves lack a specific immune system, the innate responses seem to be an efficient defense method, involving circulating cells and a large variety of molecular effectors (Canesi et al., 2002; Olafsen, 1995; Ordás et al., 2000a,b; Tafalla et al., 2003).

Information regarding bivalve immune-related genes remains very scarce and fragmentary despite recent advances. Most of the available data for bivalves were obtained on Eastern and Pacific oysters (*Crassostrea virginica* and *Crassostrea gigas*) and mussels (*Mytilus galloprovincialis*) (Fleury et al., 2009; Gueguen et al., 2003; Pallavicini et al., 2008; Venier et al., 2009; Wang et al., 2011). Limited information is available on the *R. decussatus* and *R. philippinarum* immune genes. Gestal et al. (2007) identified several ESTs related to *R. decussatus* immunity by stimulating with a mixture of dead bacterial strains, and Prado-Alvarez et al. (2009a) described immune-related ESTs expressed by *Perkinsus olseni*-infected *R. decussatus*. Only a few transcripts encoded by genes putatively involved in the immune response of *R. philippinarum* against *P. olseni* have been reported (Kang et al., 2006). The European Marine Genomics Network has significantly increased the number of known ESTs from commercial marine mollusk species (Tanguy et al., 2008); however, most of the available resources are not annotated. In the present work, we have analyzed, to the best of our knowledge, all available *R. decussatus* and *R. philippinarum* ESTs.

* Corresponding author. Tel.: +34 986214462; fax: +34 986292762.
E-mail address: antoniofigueras@iim.csic.es (A. Figueras).

0145-305X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dci.2011.06.012

The main goal of this work was to describe and classify immune-related transcripts from *R. decussatus* and *R. philippinarum* and also find differentially expressed genes after bacterial infection in the two species. To our knowledge, this is the first molecular study comparing the immune response of these two clam species to *V. alginolyticus* (strain TA15) infection.

2. Materials and methods

2.1. Selection and identification of ESTs

All non-annotated sequences from *R. decussatus* and *R. philippinarum* were downloaded from the *GenBank* NCBI database, converted into FASTA format and automatically screened with BLASTX (Altschul et al., 1990) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences with an e-value lower than 10e−5 were further annotated manually upon inspection of relevant similarities. After this, Blast2GO (Conesa et al., 2005) was used to apply Gene Ontology terms (Ashburner et al., 2000) to sequences. Default values in Blast2GO were used to perform the analysis, and ontology level 3 was selected to prepare the level pie charts.

Sequences were then analyzed with CAP3 (Huang and Madan, 1999) to find contigs, and nucleotide sequences were translated into protein and analyzed to find conserved domains with ExPASy-PROSITE (Gasteiger et al., 2003) and BLASTP (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Further selection of genes was performed according to the following criteria: presence in the two clam species, importance in the immune system and quality of the sequences (complete proteins/domains and good interspecies alignment).

2.2. Experimental infection with V. alginolyticus

Clams *R. decussatus* and *R. philippinarum* were obtained from a commercial shellfish farm (Vigo, Galicia, Spain) and maintained in open circuit filtered seawater tanks at 15 °C with aeration. They were fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana*. Prior to the experiments, clams were acclimatized to aquaria conditions for 1 week.

Clams ($n = 45$ for each species) were notched in the shell next to the anterior adductor muscle and injected with 100 μl of *V. alginolyticus* (10^8 UFC/ml in filtered seawater) to mimic an intravalvar infection. Controls ($n = 45$ for each species) were injected with 100 μl of filtered seawater. After stimulation, clams were returned to the tanks and maintained at 15 °C until sampling at 3, 6 and 24 h after challenge.

Hemolymph (1 ml) was withdrawn from the adductor muscle of the clams with an insulin syringe. Hemolymph from five individuals was pooled and three biological replicates were taken at each sampling point. Hemolymph was centrifuged at 4 °C at 2500*g* for 15 min. The pellet was resuspended in 1 ml of Trizol (Invitrogen), and RNA was extracted following the manufacturer's protocol. After RNA extraction, samples were treated with *Turbo DNase free* (Ambion) to eliminate DNA. Next, concentration and purity of RNA were measured using a *NanoDrop ND1000* spectrophotometer, and 0.7 μg of each sample was used to obtain cDNA with *SuperScript III Reverse Transcriptase* (Invitrogen). The cDNA obtained was stored at −20 °C until use.

2.3. Real-time quantitative PCR

Specific PCR primers (Table 1) were designed from the selected sequences using the *primer3* program (Rozen and Skaletsky, 2000) according to qPCR restrictions. *Oligo Analyzer 1.0.2* was used to check for dimer and hairpin formation. Efficiency of each primer pair was then analyzed with seven serial fivefold dilutions of cDNA of *R. decussatus* and *R. philippinarum* and calculated from the slope of the regression line of the quantification cycle versus the relative concentration of cDNA (Pfaffl, 2001). A melting curve analysis was also performed to verify that no primer dimers were amplified. If these conditions were not accomplished, new primer pairs were designed.

Table 1
Primer sequences used in quantitative PCR.

Primer name	Species	Secuencia (5′→3′)
Mollusk 18S-F	Both species	GTACAAAGGGCAGGGACGTA
Mollusk 18S-R	Both species	CTCCTTCGTGCTAGGGATTG
F-TRGAL	Both species	CCTAGAAGTGACTGTGCCTTTG
R-TRGAL	Both species	CGTCACAGTGTTGTCCTTTACG
F-HSP22-dec	*R. decussatus*	GCGGTAGCGTGTAGCGAAC
R-HSP22-dec	*R. decussatus*	TCCTTAACCCAGTCCGAAAAC
F-HSP22-phil	*R. philippinarum*	GCAGTCAATTCAAACCTGAAGA
R-HSP22-phil	*R. philippinarum*	TCGATTGCGTTTGGAAGAG
F-APIP-dec	*R. decussatus*	TCATCCCCGCACCATATTG
R-APIP-dec	*R. decussatus*	CTTCACCCTGAACAAGCTCTTTC
F-APIP-phil	*R. philippinarum*	TGTTGTGGTCAGGGACAGAG
R-APIP-phil	*R. philippinarum*	TCCTCCGTCGATTTGTTTTC
F-IAP-dec	*R. decussatus*	CGCTTATCGAGCCATGTAAGA
R-IAP-dec	*R. decussatus*	AAGCCCGAATACGTCAAGG
F-IAP-phil	*R. philippinarum*	TTCATTTGGCGAATGTGAAC
R-IAP-phil	*R. philippinarum*	CGCACGAAAAACACATCACT
F-INFKB	Both species	TTTATGTTTCGCAAGGAAGGA
R-INFKB	Both species	AACCATTTTCAGCCAGCACT
F-FERRITIN-dec	*R. decussatus*	AACGAGGTGGACGTGTTGTT
R-FERRITIN-dec	*R. decussatus*	CGCCTGATTAACGGTTTTCT
F-FERRITIN-phil	*R. philippinarum*	CTTCACAATGTTGCCTGTGG
R-FERRITIN-phil	*R. philippinarum*	CCGTGTCCGCTTCCTAGAC
F-THROM-dec	*R. decussatus*	ATCCGAACCCAAAGGAAACT
R-THROM-dec	*R. decussatus*	CGGCAAGGATAAAATCATCG
F-THROM-phil	*R. philippinarum*	GCCTCATTTGGTCGTGTGATT
R-THROM-phil	*R. philippinarum*	GATCGTTGCCCATTGATTCAA
F-TLR-dec	*R. decussatus*	GTTTTCGCACGAGAAAGCA
R-TLR-dec	*R. decussatus*	CGTTCCAAGAAGGCAACAAT
F-TLR-phil	*R. philippinarum*	AATGTTCTAGCGTTGACGAGAATG
R-TLR-phil	*R. philippinarum*	CGGTATTTATTGTGGCGTTTAGG
F-HSP40	Both species	TGGTGAAGAGGGTTTGAAGAA
R-HSP40	Both species	CCAAAGAAATCCCGGAAAAC
F-C1Q-dec	*R. decussatus*	CATGTGGCTTGACCTCCTTC
R-C1Q-dec	*R. decussatus*	TTACGCAGCCTCAACATCAC
F-C1Q-phil	*R. philippinarum*	TCTTCCCGAGGATACCACTG
R-C1Q-phil	*R. philippinarum*	TTTCCAAGGAGGTCGTATCG
Thioester	Both species	TTCATTCCCAGAACCTGGAC
Thioester	Both species	GCCTCCGAGATCGTTATTTTC
F-LITAF-dec	*R. decussatus*	TGCTGTCTGATTCCGTTCTG
R-LITAF-dec	*R. decussatus*	ACTGGTCCCCACTTCCTCTC
F-LITAF-phil	*R. philippinarum*	TGCTATTGTTGGATGCTGGA
R-LITAF-phil	*R. philippinarum*	CCACACTGCTGGTGACAGTT
F-HMG1-dec	*R. decussatus*	GAAGTCTCCCACCCAAAACA
R-HMG1-dec	*R. decussatus*	CAGCAGTGGACGAGTTGAGA
F-HMG1-phil	*R. philippinarum*	TGGAGGTGGTCGAGGAAAAC
R-HMG1-phil	*R. philippinarum*	GGATGTCCCTGCCTTTGTGT

Real-time quantitative PCR was performed in the *7300 Real Time PCR System* (Applied Biosystems). One microliter of fivefold diluted cDNA template was mixed with 0.5 μl of each primer (10 μM) and 12.5 μl of SYBR green PCR master mix (Applied Biosystems) in a final volume of 25 μl. The standard cycling conditions were 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. All reactions were performed as technical triplicates and an analysis of melting curves was performed in each reaction. The relative expression levels of the genes were normalized using 18S as a reference gene, which was constitutively expressed and not affected by the *Vibrio* challenge, following the Pfaffl method (Pfaffl, 2001).

2.4. Phylogenetic analysis

For each gene, homologous amino acid sequences of vertebrates and invertebrates were gathered and aligned (whole protein,

domains or part of a domain) using the *T-Coffee* server (Notredame et al., 2000) in regular computation mode and using the *t_coffee_msa* multiple alignment method. Ambiguously aligned regions were filtered with *G-Blocks* (Castresana, 2000) choosing all the options for a less stringent selection. ALTER (Glez-Peña et al., 2010) was then used to convert filtered alignments into the Phylip format, and *ProtTest 2.4* (Abascal et al., 2005) was applied to find the best fit model of protein evolution and the substitution model parameters to construct maximum likelihood phylogenetic trees in PhyML (Guindon and Gascuel, 2003); the remaining chosen options were the prespecified but the type of tree improvement, NNI & SPR was chosen for a most robust analysis, 1000 bootstrap replicates to assess phylogenetic confidence. The resulting trees were visualized and edited with *FigTree 1.3.1* (http://tree.bio.ed.ac.uk/software/figtree/) and used to study the phylogenetic position of our sequences compared to other representative species. Branch length estimates were used to study a potential relationship between gene expression and amino acid replacement (Duret and Mouchiroud, 2000; Drummond et al., 2005).

2.5. Statistical analysis

Expression results were represented graphically as the mean ± - the standard deviation of the biological replicates ($n = 3$). Differences between time groups were considered significant at $p \leqslant 0.05$. Statistical analysis was performed with *SPSS Statistics 17.0* through a one-way ANOVA test to know the effect of the time over expression.

The relationship between expression level and branch length on the phylogenetic trees (Table 4) was studied using Kendall's and Spearman's non-parametric correlation analysis. Branch length was defined as the mean branch length from conspecific sequences to their most recent common ancestor.

3. Results

3.1. EST selection

We were able to find 3784 non-annotated ESTs of *R. decussatus* and 4607 of *R. philippinarum* in GenBank (Fig. 1). After removing sequences that belonged to unknown proteins and had e-values higher than 10e−05, the remaining ESTs that were related to the immune system were chosen; 424 ESTs of *R. decussatus* and 464 of *R. philippinarum* (Table 2) were then manually revised.

Gene ontology results showed that ESTs from both species followed the same distribution for the immune EST classification. Regarding cellular components (Fig. 2A and B), a high number of annotated ESTs (about 44%) were included in the group of cell part; membrane-bounded organelle proteins represented 25.42% and 26.47% of ESTs from *R. decussatus* and *R. philippinarum*, respectively; the third most represented group was the organelle part (15.25% and 10.29% of the sequences, respectively). Within molecular function classification (Fig. 2C and D), the most represented group in both species was protein binding activity (over 23.5% of sequences); catalytic activity was the other main group with a high number of ESTs involved in hydrolase processes (14.18% of *R. decussatus* and 22.30% of *R. philippinarum* ESTs). Regarding biological process (Fig. 2E and F), metabolism was the highest represented group (approximately 36% of the sequences in both species); response to stimulus and stress was the second most represented group (9.14% of *R. decussatus* EST and a higher rate of 16.62% for *R. philippinarum*); cell communication and death processes were other important groups with 4–5% of the annotated ESTs each.

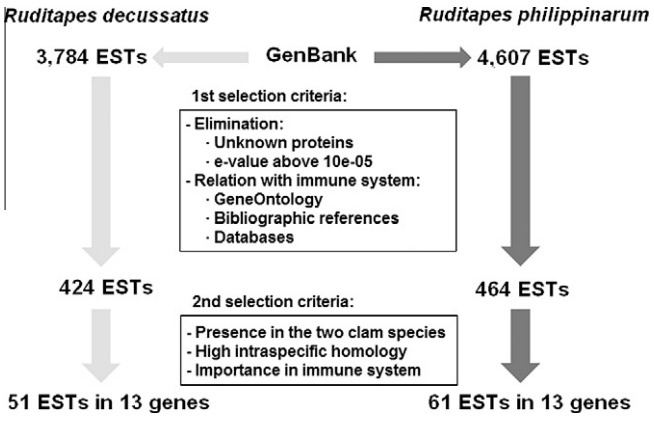


Fig. 1. Scheme of the procedure followed to identify the both clam species ESTs.

From immunity-related ESTs, a final selection was performed according to presence in the two clam species, relevant role in immune response and high intraspecific homology among sequences belonging to the same gene. After this second selection step, the final number of ESTs was 51 for *R. decussatus* and 61 for *R. philippinarum*, corresponding to 13 genes (Table 3) that were grouped in several functional categories, as follows: receptors (host-pathogen interaction), including tandem repeat galectin, Toll-like receptor; heat shock proteins (22 and 40); genes related with complement (C1q domain-containing protein and thioester-containing protein); genes involved in apoptosis (inhibitor of apoptosis protein and APAF1-interacting protein homolog); cytokine-related molecules (LPS-induced TNF-alpha factor and inhibitor of nuclear factor kappa B); and others, such as ferritin, thrombin and high-mobility group 1 protein.

3.2. Expression analysis after V. alginolyticus challenge

Tandem repeat galectins (TRGal) and Toll-like receptors (TLR) function as pattern recognition receptors (PRR) and host-pathogen interacting molecules. The former is soluble and the latter membrane anchored. TRGal expression (Fig. 3A) showed opposite expression patterns after infection of *R. decussatus* and *R. philippinarum*. In *R. decussatus,* the expression tended to diminish over time, whereas in *R. philippinarum,* the TRGal expression increased significantly over time. TLR also showed distinct expression patterns in the two species (Fig. 3B): *R. decussatus* presented a high level of expression 6 h after challenge, whereas in *R. philippinarum*, the expression level increased through the time course, reaching the maximum at 24 h.

The Inhibitor of Apoptosis Protein (IAP) and APAF-1 interacting protein (APIP) are inhibitors of apoptosis. The trends in IAP expression (Fig. 3C) were similar in both species, with the maximum expression observed 3 h after infection and a decrease afterward. However, the magnitude of the decrease was different between both species. Although *R. decussatus* displayed an insignificant and gradually decreasing expression pattern, the diminution was especially evident in *R. philippinarum*, with significant values ($p < 0.05$) with regard to 6 and 24 h after infection. On the other hand, opposite trends between the two species were observed in the case of APIP expression (Fig. 3D).

Another group of genes, including heat shock protein 22 (HSP22), ferritin and thrombin, was considered in this study for their likely role in the immune response and because they commonly function as activators of NF-κB. HSP22 is a small HSP that plays different roles as a molecular chaperone, ferritin captures circulating iron to overcome an infection and thrombin is involved in wound repairing and chemotaxis regulation in the affected area. The TLR previously described also participates in the signal

Table 2
Immune system related ESTs (first selection criteria).

Category and gene identity. BLASTX	Species and number of EST	
	R. decussatus	*R. philippinarum*
LECTINS		
Tandem repeat galectin	6	12
Galectin 4-like protein transcript variant	–	3
Lectin-galC1	–	1
Sialic acid binding lectin	10	15
Putative contactin-associated protein	1	–
Mannose receptor	–	3
C-type lectin 1	–	2
C-type lectin 2	1	–
C-type lectin 4	–	1
C-type lectin 8	1	–
C-type lectin 9	–	1
C-type lectin A	1	4
C-type lectin	–	2
Putative salivary C-type lectin	–	1
Putative perlucin 5	1	–
Putative perlucin 4	1	1
Perlucin-like protein (isoform A/B/C)	–	5
HEAT SHOCK PROTEIN		
Heat shock protein 22 (isoform1/2)	6	7
Heat shock protein 40	2	4
Heat shock protein 60	1	–
Heat shock protein 70	2	7
Heat shock protein 90	–	7
FtsJ homolog 1 (E. coli)	1	–
COMPLEMENT & C1q-LIKE PROTEIN		
C1q domain-containing protein	9	7
Complement component C3	1	2
Complement factor B-like protein	2	1
Mantle gene 4	–	1
Mantle gene 6	3	–
Adiponectin	1	–
Thioester-containing protein	3	6
EP protein precursor. Heavy metal-binding protein HIP	1	4
CATHEPSIN		
Cathepsin B	1	6
Cathepsin C	–	1
Cathepsin D	–	1
Cathepsin H	–	1
Cathepsin I	1	–
Cathepsin L	4	5
Cathepsin S	–	2
Cathepsin Z	2	–
APOPTOSIS		
Brain protein I3	8	–
Autophagy-related protein 3	1	–
Programmed cell death protein	1	1
Death-associated protein	–	2
B-cell translocation gene 1	1	3
c-myc binding protein	1	–
Similar to arrestin domain containing 3	2	–
Securin	1	–
Inhibitor of apoptosis protein-IAP	8	12
BCL2-associatied athanogene (BAG)	1	2
BCL2/adenovirus E1B 19kD-interacting protein 1	1	–
APAF1-interacting protein homolog (APIP)	3	1
Caspase-1	1	–
Caspase-7	1	–
Caspase-8	1	–
Caspase-b	–	1
Separin protein	–	1
Cytochrome c	7	5
PROTEASES AND PROTEASE INHIBITORS		
Serine protease	4	9
Similar to Human Reelin (serine protease)	1	–
Cysteine proteinase preproenzyme	–	1
Serine proteinase inhibitor	–	2
Cysteine protease inhibitor	–	1
Kazal-type serine proteinase inhibitor	3	8
Hemocyte Kazal-type serine proteinase inhibitor	–	1
Metalloproteinase	1	1
Agrin precursor	–	1
Cystatin-A2	–	1
Kunitz-like protease inhibitor	–	1
Signal peptide peptidase	1	–
CYTOKINE SIGNALING		
Interferon-induced protein 2	–	1
Interferon-induced protein 27	–	1
Interferon-induced protein 44	7	3
Interferon-related developmental regulator 1	1	–
LPS-induced TNF-alpha factor	1	4
Tumor necrosis factor-like protein	–	1
Inhibitor of nuclear factor kappaB	2	1
SPRY domain-containing SOCS box protein, putative	1	–
Interleukin enhancer binding factor 2, 45kDa	1	–
IMMUNE DEFENSE AND HOST-PATHOGEN INTERACTION		
Lisozyme	–	7
Ferritin	7	2
Uromodulin	–	1
Peptidoglycan recognition protein (4/S2/S3)	3	–
Peptidoglycan recognition protein precursor	–	1
Beta-glucan recognition protein	–	1
Defensin	–	2
Hemolysin	–	1
Allograft inflammatory factor	–	2
Similar to HLA-B associated transcript 1	2	3
BAT1 homolog	1	–
DEAD (Asp-Glu-Ala-Asp) box polypeptide 49	1	–
Similar to HLA-B associated transcript 3	1	
Saposin	5	–
Splicing variant form of ficolin A	1	–
Myticin C precursor	1	–
Macrophage migration inhibitory factor II	2	–
Coagulation factor IX precursor	1	1
Coagulation factor VII	1	–
Thrombin	2	1
Alpha macroglobulin	1	–
CELL SURFACE RECEPTORS		
G protein-coupled receptor	2	4
Notch 2	–	3
Thrombospondin 1	–	3
Scavenger receptor cysteine-rich protein type 12 precursor	–	2
Stabilin 2	–	1
Toll-like receptor TLR2.1	1	–
Toll-like receptor (AGAP012385-PA)	1	1
SET protein	1	–
LMPX of lamprey	1	–
B-cell receptor associated protein-like protein	1	–
Angiopoietin (2/3/Y1/salivary secreted)	1	7
CELULAR ADHESION		
Fasciclin-like protein	3	–
Hemicentin 1	1	–
OTHER		
High mobility group 1 protein	1	3
Metallothionein	3	–
Hematopoietic stem/progenitor cells protein-like	1	–
c-Jun protein	1	1
Similar to KLHL6 protein	1	–
Similar to lysosomal membrane glycoprotein-2	1	–
Src family associated phosphoprotein 1	1	–
Similar to VNN3 protein	1	–
RuvB-like 2-like protein	–	1
T-cell immunoglobulin and mucin domain containing 2	–	1

(*continued on next page*)

Table 2 (*continued*)

Category and gene identity. BLASTX	Species and number of EST	
	R. decussatus	*R. philippinarum*
Strumpellin	–	1
EF-hand calcium binding domain 10	–	1
Hemagglutinin and aggregation factor	–	1
Catalase	–	1

transduction pathway leading to NF-κB activation. The expression profile of these four genes was similar in both species (Fig. 3B, E, F and G). Whereas in *R. decussatus* the maximum expression occurred at 6 h to drop drastically at 24 h (only significant for HSP22 and TLR), *R. philippinarum* showed an increase in expression through the time course, with the maximum expression detected at 24 h (again significant for HSP22 and TLR).

Fig. 4 shows genes that had no detectable or very low expression in *R. decussatus*. The inhibitor of NF-κB (IκB) retains NF-κB in the cytoplasm through formation of a complex, which prevents translocation of NF-κB to the nucleus and thus blocks its subsequent function as transcription factor. IκB displayed a maximum expression at 24 h in the two species, but the fold change values were very different (1.7 in *R. decussatus* and 24.8 in *R. philippinarum*) (Fig. 4A). Additionally, expression of IκB in *R. decussatus* was much lower than in *R. philippinarum* and remained undetectable until 24 post-infection.

The last group of studied genes (The LPS induced TNF-α factor, LITAF; the high mobility group 1 protein, HMG1; the heat shock protein 40, HSP40; the C1q containing domain protein; C1q and the thioester containing protein; thioester) showed a similar behavior (Fig. 4B–F). The expression of this gene set in *R. decussatus* was undetectable in hemocytes at the sampling points considered. However, in *R. philippinarum*, their expression was higher. LITAF is a transcription factor that mediates TNF-α expression and other cytokines and genes implicated in immune response and apoptosis. HMG1 is a nuclear protein that can act as a cytokine in response to proinflammatory stimuli. HSP40 functions as a molecular chaperone. C1q and thioester are two proteins related to the complement system. The expression pattern in *R. philippinarum* showed a progressive increase over time following infection ultimately reaching the maximum at 24 h, except for thioester (Fig. 4F), which follows an expression pattern defined by a high level of expression at 3 and 24 h and low expression at 6 h.

3.3. Phylogenetic analysis

The phylogenetic position of *R. decussatus* and *R. philippinarum* agreed with that expected, i.e, together with other mollusks, except for some cases in which the lack of information in the sequences resulted in alternative, unsupported relationships with low bootstrap values (thrombin, TLR, IAP and C1q). (Supplementary material, Figs. 1–3).

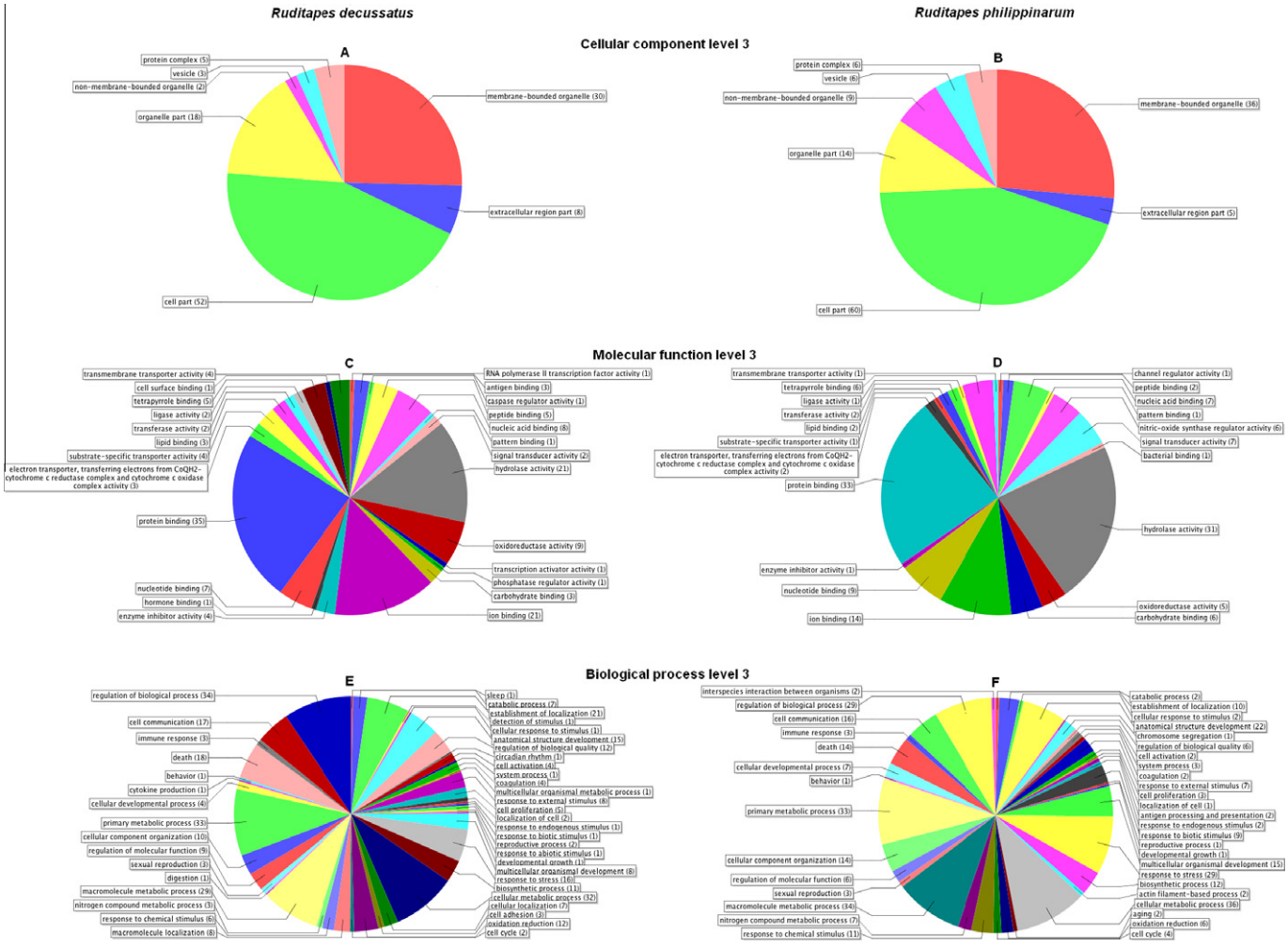


Fig. 2. Classification of *R. decussatus* and *R. philippinarum* ESTs by Gene Ontology Terms. A and B: cellular component level 3. C and D: molecular function level 3. E and F: biological process level 3.

Table 3
Immune system related ESTs selected for expression studies (second selection criteria).

Functional category	BLASTX	ESTs *R. decussatus*	ESTs *R. philippinarum*
Host-pathogen interaction	Tandem repeat galectin	6	12
	Toll-like receptor (AGAP012385-PA)	1	1
Apoptosis	Inhibitor of apoptosis protein	8	12
	APAF1-interacting protein homolog	3	1
Activators of nuclear factor kappa B	Heat shock protein 22 (isoform1/2)	6	7
	Ferritin	7	2
	Thrombin	2	1
Cytokine production	Inhibitor of nuclear factor kappa B	2	1
	LPS-induced TNF-alpha factor	1	4
	High mobility group 1 protein	1	3
Heat shock proteins	Heat shock protein 40	2	4
Complement	C1q domain-containing protein	9	7
	Thioester-containing protein	3	6

Table 4
Mean of branch length until the nearest common ancestor of studied sequences from *R. decussatus* and *R. philippinarum*.

Gene	*R. decussatus*	*R. philippinarum*
TRGal	0.0333	0.1173
TLR	0.6909	1.1502
IAP	0.0442	0.5947
APIP	0.0054	0.0413
HSP22	0.0328	0.1157
Ferritin	0.0095	0.0301
Thrombin	0.8159	1.7741
IkB	0.8057	0.8457
LITAF	0.242	0.5393
HMG1	0.6161	0.24405
HSP40	0.5769	0.5605
C1q	0.8139	1.1739
Thioester	0.0231	0.0832

In general, in these gene trees *R. philippinarum* showed larger branches than *R. decussatus*, except for HMG1 and HSP40. However, we noticed that some loci with identical protein sequences for the two clam species could show different expression patterns (IκB). On the other hand, loci with different protein sequences for both species could display very similar expression patterns (IAP). In fact, we only found a statistically significant relationship between evolutionary change and expression level at the APIP locus.

4. Discussion

Although the generation of ESTs of non-model organisms has been increasing in the last several years, this information frequently remains difficult to access. Identification and annotation of sequences is sometimes difficult, taking into account the important divergence among species and the lack of genetic information for many of them. Considerable efforts, however, have been made in some cases; for example, in the creation of a public EST database for *M. galloprovincialis* (MytiBase) (Venier et al., 2009) and *C. gigas* (Fleury et al., 2009) and the further characterization of the EST collections (Venier et al., 2011).

To our knowledge, this is the first comparative study of the expression of several immune-related genes between *R. decussatus* and *R. philippinarum*. It is also the first time that genes, such as IAP, APIP, thrombin or HMG1 are described in mollusks. We also report here new TLR expression data that have not previously been well described in this animal group (Qiu et al., 2007).

The gene expression results after a *V. alginolyticus* challenge showed that, for the genes and sampling points considered, *R. decussatus* presented lower expression levels than did *R. philippinarum*. In fact, some genes (IκB, LITAF, HMG1, C1q, thioester and HSP40) displayed very low or undetectable expression in *R. decussatus* hemocytes.

These results are in agreement with previous studies performed in mollusks. It was described that the oyster HSP22 (Zhang et al., 2010) and scallop and abalone LITAF (Zhang et al., 2009; De Zoysa et al., 2010) were constitutively expressed in all studied tissues. However, hemolymph was one of the tissues with the lowest expression of HSP22 and LITAF in oyster, scallop and abalone. Regarding thioester and C1q, it is known that, 3 h after a *Vibrio* challenge, the expression level of C1q in *M. galloprovincialis* hemocytes did not show a significant increase, compared to expression levels in the controls. The expression level of C1q increased rapidly 1 h after a *Vibrio* challenge, but decayed 3 h after infection (Gestal et al., 2010). This could explain why C1q also showed lower expression after 3 and 6 h in both clam species. In the case of thioester, *Chlamys farreri* showed no expression in hemocytes (Zhang et al., 2007). In *R. decussatus* hemocytes, expression of C3 (a thioester containing protein) was 6-fold to 12-fold lower than that in the digestive gland and decayed below the expression in the controls from 1 to 6 h after a *Vibrio* challenge in hemocytes (Prado-Alvarez et al., 2009b).

HMG1 protein, a highly abundant and ubiquitous non-histone nuclear protein, was undetectable by qPCR in *R. decussatus*. However, HMG1 expression was detected in *R. philippinarum* hemocytes, especially 24 h after *Vibrio* challenge. In mammals, HMG1 showed variable expression, depending on the tissue and sampling points (Wang et al., 1999; Sass et al., 2002; Hagiwara et al., 2007). On the other hand, thermal injury, which it is known to increase mRNA levels of proinflammatory cytokines in various tissues, induced HMG1 expression after 24 h in rats (Fang et al., 2002), in agreement with the maximum expression of HMG1 detected at 24 h in *R. philippinarum*.

Expression of IκB was similarly undetected at the initial times post-infection in *R. decussatus*. Nevertheless, a high expression of NF-κB activating genes; TLR (Takeuchi and Akira, 2010), HSP22 (Guo et al., 2009), ferritin (Ruddell et al., 2009) or thrombin (Anrather et al., 1997; Xue et al., 2009) was detected 6 h after challenge in *R. decussatus*. This could be evidence that in *R. decussatus* NF-κB was activated in response to the bacterial challenge between 6 and 24 h post-infection. This also seems to indicate that, although high levels of NF-κB activators were detected, a poor NF-κB response could be achieved, as the low expression of IκB at 24 h seems to suggest. It is well known that IκB is one of the most highly transcribed genes after NF-κB activation (Sun et al., 1993) and could be used as an indicator of NF-κB transcription factor activity. In *R. philippinarum,* we observed a low expression of TLR, HSP22, ferritin and thrombin, but there are indications that an effective immune response is being performed. IκB was highly expressed 24 h post-challenge, probably limiting the reaction generated through NF-κB activity. Moreover, qPCR assays showed that *R. philippinarum* TRGal, TLR, LITAF, HMG1, C1q, thioester and HSPs were highly expressed at 24 h, and all exhibited at least a 2-fold increase with respect to controls. The *R. decussatus* deficiency in IκB expression and the expression profiles of NF-κB related genes might explain its lower resistance to pathogens and stress.

Based on the temporal expression pattern of TRGal, TLR and proteins related to complement system, *R. philippinarum* seems

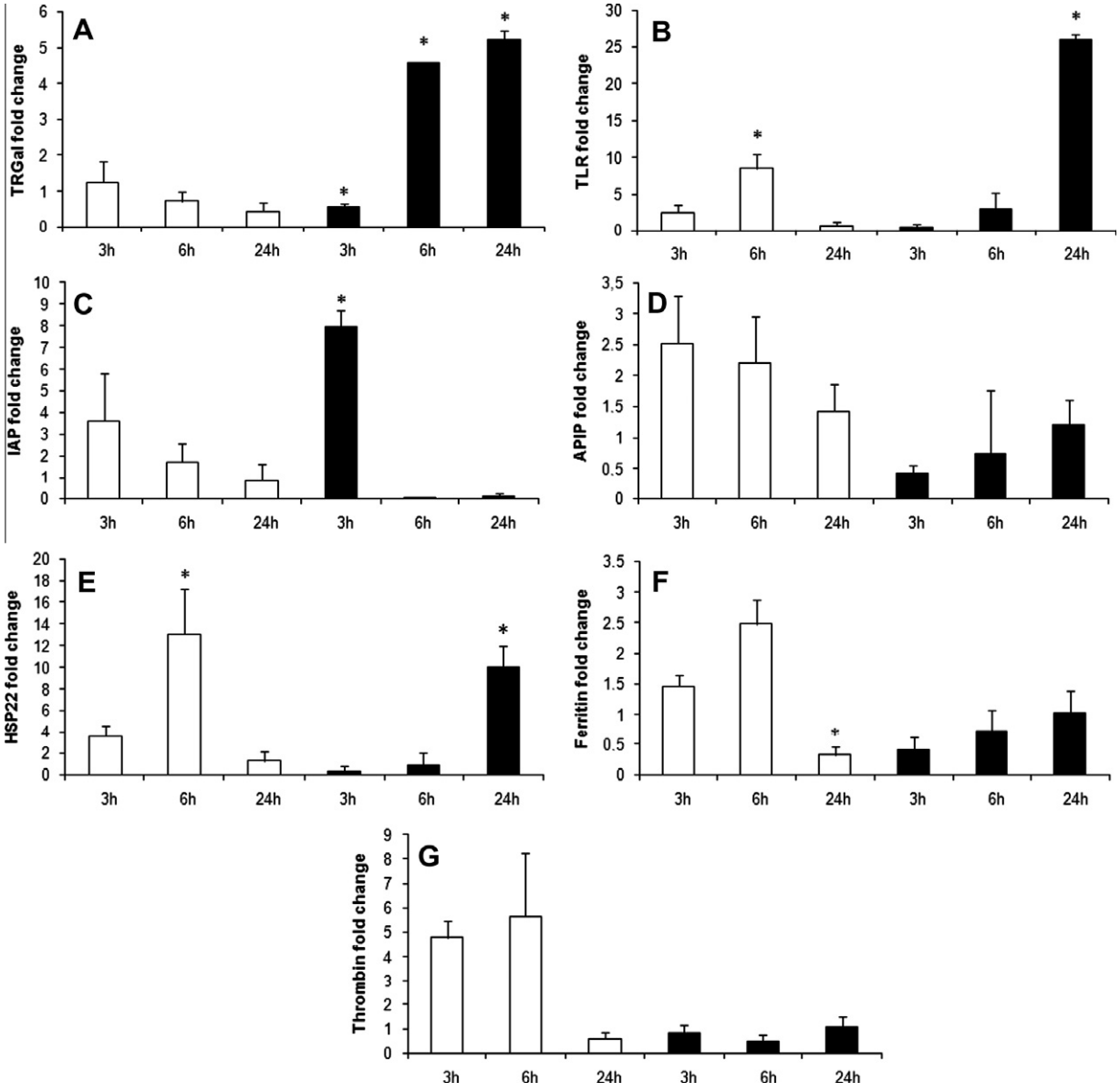


Fig. 3. Expression level of TRGal (A), TLR (B), IAP (C), APIP (D), HSP22 (E), ferritin (F) and thrombin (G) in hemocytes from *R. decussatus* (white bars) and *R. philippinarum* (black bars) at 3, 6 and 24 h after *V. alginolyticus* challenge. All qPCR reactions were performed as technical triplicates and the expression level of analyzed genes was normalized using the 18S rRNA. Fold change units were then calculated by dividing the normalized expression values of hemocytes from infected clams by the normalized expression values of the controls. Each bar represents the mean and standard deviation of three biological replicates. The asterisks indicate statistically significant values ($p < 0.05$) through the time course and for each species independently.

to detect pathogens earlier than did *R. decussatus*. The immune response to *V. alginolyticus* could then be initiated in a few hours, resulting in the negative regulation of the inflammatory response, as reported by the high expression of IκB at 24 h. In fact, it was previously observed in *R. philippinarum* that the highest TRGal mRNA expression was in the mantle and gill, while hemocytes presented the lowest expression. Additionally, TRGal expression increased 24 h after a *Vibrio* challenge (Kim et al., 2008), which is in agreement with our results.

Concerning apoptosis related proteins APIP and IAP, we noticed that whereas *V. alginolyticus* seemed to have no effect on APIP expression (due to the high deviations registered among pools), it elicited a strong effect on early IAP expression, which could lead to the protection of hemocytes from apoptosis. Although IAP and APIP are both related to apoptosis inhibition, their activation against bacterial infection might be quite different, given the differences observed in the expression patterns to the same stimulus for these two genes. IAP was the only studied gene where a common expression pattern in *R. decussatus* and *R. philippinarum* was detected. This could be an indication of the importance of the apoptosis pathway in the immune response in bivalves, as it has recently been reported in *M. galloprovincialis* (Romero et al., 2011). Indeed a survivin–XIAP complex participates directly in NF-κB activation (Altieri, 2010). AMP synthesis is also related to IAP because *Drosophila* iap2 null mutants fail to induce the synthesis of antimicrobial peptides and are highly susceptible to infection by Gram-negative bacteria (Huh et al., 2007).

From these results we can conclude that *R. decussatus* and *R. philippinarum* presented very distinct responses to *V. alginolyticus* both in their gene expression patterns and/or expression levels. This was especially evident in transcription factor genes, such as LITAF and NF-κB (using IκB as indicator of its expression). After a *Vibrio*

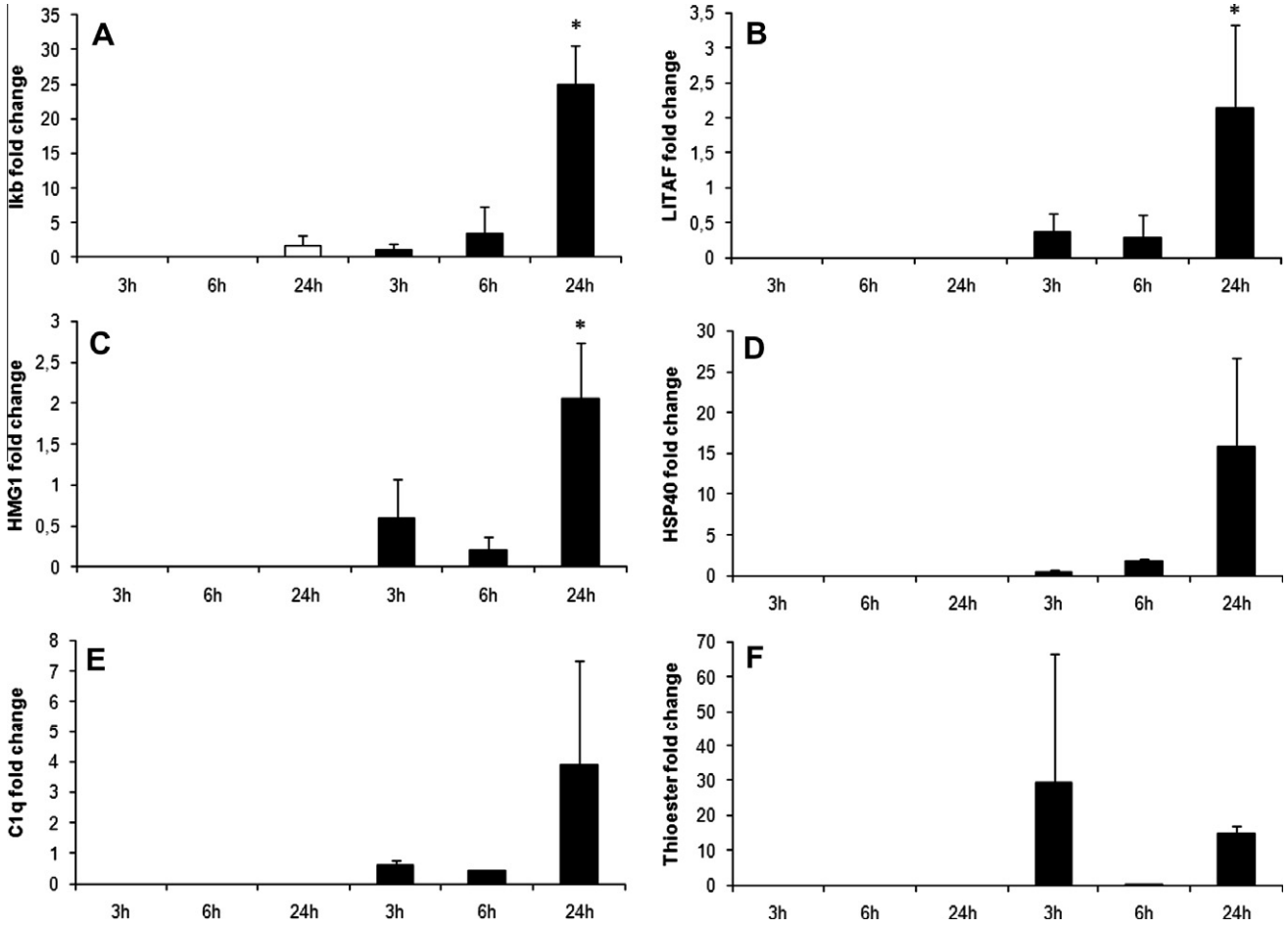


Fig. 4. Expression level of IκB (A), LITAF (B), HMG1 (C), C1q (D), thioester (E) and HSP40 (F) in hemocytes from *R. decussatus* (white bars) and *R. philippinarum* (black bars) at 3, 6 and 24 h after *V. alginolyticus* challenge. All qPCR reactions were performed as technical triplicates, and the expression level of analyzed genes was normalized using the 18S rRNA. Fold change units were then calculated by dividing the normalized expression values of hemocytes from infected clams by the normalized expression values of the controls. Each bar represents the mean and standard deviation of thee biological replicates. The asterisks indicate statistically significant values ($p < 0.05$) through the time course and for each species independently.

infection, *R. decussatus* expression of NF-κB activating genes seemed to be insufficient to promote an immune response, furthermore, *R. decussatus* did not express LITAF. However, even when these activating genes were lowly expressed in *R. philippinarum*, this clam seemed to be able to efficiently trigger NF-κB transcriptional activity and LITAF levels were clearly detected after challenge.

With reference to phylogenetic analysis, it is known that the best predictor of the evolutionary rate of a protein is expression level (Cherry, 2010). Specifically, more highly expressed proteins tend to have lower evolutionary rates. But in this case, and although branch length were larger for *R. philippinarum* in almost all gene trees, we did not find a correlation between expression and amino acid changes in the sequence, but we should consider that some of these sequences are relatively short and may lack information. Only in one case, APIP, a negative correlation was observed: *R. decussatus*, the species with higher expression value had shorter branch length than *R. philippinarum*, consistent with the actual hypothesis.

To our knowledge, the present work constitutes the first approach to understanding the molecular basis of the immune response of two different clam species, one potentially resistant against pathogens and adverse environmental conditions (*R. philippinarum*) and another more susceptible (*R. decussatus*) in wild breeding cultures. Based on the gene expression analysis, *R. philippinarum* response appeared to be more effective and faster than *R. decussatus*. Additionally, *R. decussatus* did not seem to express transcription factors that could initiate an inflammatory response. However, the expression pattern of IκB in *R. philippinarum* suggested a negative regulation of this inflammatory response 24 h after infection.

Further research will improve our understanding of the biological function of all these immune-related genes in clams. Longer sampling times, mortality studies, protein characterization and functional studies, among other assays, will enhance our comprehension of the immune system differences between *R. decussatus* and *R. philippinarum*. Additionally, the knowledge of the molecular mechanisms of the immune response could help to find molecular markers related to pathogen resistance.

Acknowledgements

This work has been funded by the Spanish Ministerio de Ciencia e Innovación (MICINN) (AGL2008-05111). R.M. wishes to acknowledge additional funding from the Spanish MICINN through a FPI Spanish research Grant (BES-2009-029765).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.dci.2011.06.012.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest, selection of bestfit models of protein evolution. Bioinformatics 21, 2104–2105.

Altieri, D.C., 2010. Survivin and IAP proteins in cell-death mechanisms. Biochem. J. 430, 199–205.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Anrather, D., Millan, M.T., Palmetshofer, A., Robson, S.C., Geczy, C., Ritchie, A.J., Bach, F.H., Ewenstein, B.M., 1997. Thrombin activates nuclear factor-kappaB and potentiates endothelial cell activation by TNF. J. Immunol. 159, 5620–5628.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology, tool for the unification of biology. The gene ontology consortium. Nat. Genet. 25, 25–29.

Canesi, L., Gallo, G., Gavioli, M., Pruzzo, C., 2002. Bacteria-hemocyte interactions and phagocytosis in bivalves. Microsc. Res. Technol. 57, 469–476.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Cherry, J.L., 2010. Expression level, evolutionary rate, and the cost of expression. Genome Biol. Evol. 2, 757–769.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO, a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.

De Zoysa, M., Nikapitiya, C., Oh, C., Whang, I., Lee, J.S., Jung, S.J., Choi, C.Y., Lee, J., 2010. Molecular evidence for the existence of lipopolysaccharide-induced TNF-alpha factor (LITAF) and Rel/NF-κB pathways in disk abalone (*Haliotis discus discus*). Fish Shellfish Immunol. 28, 754–763.

Drummond, D.A., Bloom, J.D., Adami, C., Wilke, C.O., Arnold, F.H., 2005. Why highly expressed proteins evolve slowly. Proc. Natl. Acad. Sci. U. S. A 102, 14338–14343.

Duret, L., Mouchiroud, D., 2000. Determinants of substitution rates in mammalian genes, expression pattern affects selection intensity but not mutation rate. Mol. Biol. Evol. 17, 68–74.

Fang, W.H., Yao, Y.M., Shi, Z.G., Yu, Y., Wu, Y., Lu, L.R., Sheng, Z.Y., 2002. The significance of changes in high mobility group-1 protein mRNA expression in rats after thermal injury. Shock 17, 329–333.

FAO, 2005. Cultured aquatic species information programme. *R. philippinarum*. In: Goulletquer, P. (Ed.), Cultured Aquatic Species Information Programme, FAO Fisheries and Aquaculture Department [online], Rome. http://www.fao.org/fishery/culturedspecies/Ruditapes_philippinarum/en. Accessed on: 15 January 2011.

Fleury, E., Huvet, A., Lelong, C., de Lorgeril, J., Boulo, V., Gueguen, Y., Bachère, E., Tanguy, A., Moraga, D., Fabioux, C., Lindeque, P., Shaw, J., Reinhardt, R., Prunet, P., Davey, G., Lapègue, S., Sauvage, C., Corporeau, C., Moal, J., Gavory, F., Wincker, P., Moreews, F., Klopp, C., Mathieu, M., Boudry, P., Favrel, P., 2009. Generation and analysis of a 29, 745 unique expressed sequence tags from the Pacific oyster (*C. gigas*) assembled into a publicly accessible database, the gigas database. BMC Genomics 10, 341.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy, the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784–3788.

Gestal, C., Costa, M., Figueras, A., Novoa, B., 2007. 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 406, 134–143.

Gestal, C., Roch, P., Renault, T., Pallavicini, A., Paillard, C., Novoa, B., Oubella, R., Venier, P., Figueras, A., 2008. Study of diseases and the immune system of bivalves using molecular biology and genomics. Rev. Fish. Sci. 16, 131–154.

Gestal, C., Pallavicini, A., Venier, P., Novoa, B., Figueras, A., 2010. MgC1q, a novel C1q-domain-containing protein involved in the immune response of *Mytilus galloprovincialis*. Dev. Comp. Immunol. 34, 926–934.

Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., Posada, D., 2010. ALTER, program-oriented format conversion of ADN and protein alignments. Nucleic Acids Res. 38, 14–18.

Gómez-León, J., Villamil, L., Lemos, M.L., Novoa, B., Figueras, A., 2005. Isolation of *V. alginolyticus* and *Vibrio splendidus* from aquacultured carpet shell clam (*Ruditapes decussatus*) larvae associated with mass mortalities. Appl. Environ. Micorb. 71, 98–104.

Gueguen, Y., Cadoret, J.P., Flament, D., Barreau-Roumiguière, C., Girardot, A.L., Garnier, J., Hoareau, A., Bachère, E., Escoubas, J.M., 2003. Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *C. gigas*. Gene 303, 139–145.

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

Guo, K., Kang, N.X., Li, Y., Sun, L., Gan, L., Cui, F.J., Gao, M.D., Liu, K.Y., 2009. Regulation of HSP27 on NF-kappaB pathway activation may be involved in metastatic hepatocellular carcinoma cells apoptosis. BMC Cancer 9, 100.

Hagiwara, S., Iwasaka, H., Noguchi, T., 2007. Nafamostat mesilate inhibits the expression of HMGB1 in lipopolysaccharide-induced acute lung injury. J. Anesth. 21, 164–170.

Huang, X., Madan, A., 1999. CAP3, A ADN sequence assembly program. Genome Res. 9, 868–877.

Huh, J.R., Foe, I., Muro, I., Chen, C.H., Seol, J.H., Yoo, S.J., Guo, M., Park, J.M., Hay, B.A., 2007. The Drosophila inhibitor of apoptosis (IAP) DIAP2 is dispensable for cell survival, required for the innate immune response to gram-negative bacterial infection, and can be negatively regulated by the reaper/hid/grim family of IAP-binding apoptosis inducers. J. Biol. Chem. 282, 2056–2068.

Kang, Y.S., Kim, Y.M., Park, K.L., Cho, S.K., Choi, K.S., Cho, M., 2006. Analysis of EST and lectin expression in hemocytes of Manila clams (*Ruditapes phylippinarum*) (Bivalvia, Mollusca) infected with *P. olseni*. Dev. Comp. Immunol. 30, 1119–1131.

Kim, J.Y., Kim, Y.M., Cho, S.K., Choi, K.S., Cho, M., 2008. Noble tandem-repeat galectin of Manila clam *R. philippinarum* is induced upon infection with the protozoan parasite *P. olseni*. Dev. Comp. Immunol. 32, 1131–1141.

Notredame, C., Higgins, D.G., Heringa, J., 2000. T-Coffee, a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.

Olafsen, J.A., 1995. Role of lectins (C-reactive protein) in defense of marine bivalves against bacteria. Adv. Exp. Med. Biol. 371A, 343–348.

Ordás, M.C., Novoa, B., Figueras, A., 2000a. Modulation of the chemiluminescence response of Mediterranean mussel (*Mytilus gallorpovincialis*) haemocytes. Fish Shellfish Immunol. 10, 611–622.

Ordás, M.C., Ordás, A., Beloso, C., Figueras, A., 2000b. Immune parameters in carpet shell clams naturally infected with *Perkinsus atlanticus*. Fish Shellfish Immunol. 10, 597–609.

Paillard, C., Leroux, F., Borrego, J.J., 2004. Bacterial disease in marine bivalves, review of recent studies. Trends and evolution. Aquat. Living Resour. 17, 477–498.

Pallavicini, A., Costa, M.M., Gestal, C., Dreos, R., Figueras, A., Venier, P., Novoa, B., 2008. Sequence variability of myticins identified in haemocytes from mussels suggests ancient host–pathogen interactions. Dev. Comp. Immunol. 32, 213–226.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, 2002–2007.

Prado-Alvarez, M., Gestal, C., Novoa, B., Figueras, A., 2009a. Differentially expressed genes of the carpet shell clam *Ruditapes decussatus* against *P. olseni*. Fish Shellfish Immunol. 26, 72–83.

Prado-Alvarez, M., Rotllant, J., Gestal, C., Novoa, B., Figueras, A., 2009b. Characterization of a C3 and a factor B-like in the carpet-shell clam, *Ruditapes decussatus*. Fish Shellfish Immunol. 26, 305–315.

Qiu, L., Song, L., Xu, W., Ni, D., Yu, Y., 2007. Molecular cloning and expression of a toll receptor gene homologue from Zhikong Scallop, *C. farreri*. Fish Shellfish Immunol. 22, 451–466.

Romero, A., Estévez-Calvar, N., Dios, S., Figueras, A., Novoa, B., 2011. New Insights into the apoptotic process in Mollusks: characterization of caspase genes in *Mytilus galloprovincialis*. PLoS One 6, e17003.

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

Ruddell, R.G., Hoang-Le, D., Barwood, J.M., Rutherford, P.S., Piva, T.J., Watters, D.J., Santambrogio, P., Arosio, P., Ramm, G.A., 2009. Ferritin functions as a proinflammatory cytokine via iron-independent protein kinase C zeta/nuclear factor kappaB-regulated signaling in rat hepatic stellate cells. Hepatology 49, 887–900.

Sass, G., Heinlein, S., Agli, A., Bang, R., Schumann, J., Tiegs, G., 2002. Cytokine expression in three mouse models of experimental hepatitis. Cytokine 19, 115–120.

Sun, S.C., Ganchi, P.A., Ballard, D.W., Greene, W.C., 1993. NF-kappa B controls expression of inhibitor I kappa B alpha: evidence for an inducible autoregulatory pathway. Science 259, 1912–1915.

Tafalla, C., Gómez-León, J., Novoa, B., Figueras, A., 2003. Nitric oxide production by carpet shell clam (*Ruditapes decussatus*) hemocytes. Dev. Comp. Immunol. 27, 197–205.

Takeuchi, O., Akira, S., 2010. Pattern recognition receptors and inflammation. Cell 140, 805–820.

Tanguy, A., Bierne, N., Saavedra, C., Pina, B., Bachère, E., Kube, M., Bazin, E., Bonhomme, F., Boudry, P., Boulo, V., Boutet, I., Cancela, L., Dossat, C., Favrel, P., Huvet, A., Jarque, S., Jollivet, D., Klages, S., Lapègue, S., Leite, R., Moal, J., Moraga, D., Reinhardt, R., Samain, J.F., Zouros, E., Canario, A., 2008. Increasing genomic information in bivalves through new EST collections in four species, development of new genetic markers for environmental studies and genome evolution. Gene 408, 27–36.

Venier, P., De Pittà, C., Bernante, F., Varotto, L., De Nardi, B., Bovo, G., Roch, P., Novoa, B., Figueras, A., Pallavicini, A., Lanfranchi, G., 2009. Mytibase, a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics 10, 72.

Venier, P., Varotto, L., Rosani, U., Millino, C., Celegato, B., Bernante, F., Lanfranchi, G., Novoa, B., Roch, P., Figueras, A., Pallavicini, A., 2011. Insights into the innate immunity of the Mediterranean mussel *M. galloprovincialis*: from Mytibase to RNA transcript profiling. BMC Genomics 12, 69.

Wang, H., Bloom, O., Zhang, M., Vishnubhakat, J.M., Ombrellino, M., Che, J., Frazier, A., Yang, H., Ivanova, S., Borovikova, L., Manogue, K.R., Faist, E., Abraham, E., Andersson, J., Andersson, U., Molina, P.E., Abumrad, N.N., Sama, A., Tracey, K.J., 1999. HMG-1 as a late mediator of endotoxin lethality in mice. Science 285, 248–251.

Wang, A., Wang, Y., Gu, Z., Li, S., Shi, Y., Guo, X., 2011. Development of Expressed Sequence Tags from the Pearl Oyster, *Pinctada martensii* Dunker. Mar. Biotechnol. 13, 275–283.

Xue, J., Thippegowda, P.B., Hu, G., Bachmaier, K., Christman, J.W., Malik, A.B., Tiruppathi, C., 2009. NF-kappaB regulates thrombin-induced ICAM-1 gene expression in cooperation with NFAT by binding to the intronic NF-kappaB site in the ICAM-1 gene. Physiol. Genomics 38, 42–53.

Zhang, H., Song, L., Li, C., Zhao, J., Wang, H., Gao, Q., Xu, W., 2007. Molecular cloning and characterization of a thioester-containing protein from Zhikong scallop *C. farreri*. Mol. Immunol. 44, 3492–3500.

Zhang, D., Jiang, J., Jiang, S., Ma, J., Su, T., Qiu, L., Zhu, C., Xu, X., 2009. Molecular characterization and expression analysis of a putative LPS-induced TNF-alpha factor (LITAF) from pearl oyster *Pinctada fucata*. Fish Shellfish Immunol. 27, 391–396.

Zhang, L., Wang, L., Song, L., Zhao, J., Qiu, L., Dong, C., Li, F., Zhang, H., Yang, G., 2010. The involvement of HSP22 from bay scallop *Argopecten irradians* in response to heavy metal stress. Mol. Biol. Rep. 37, 1763–1771.

Supplementary data

Supplementary data associated with this article can be found at:
http://www.sciencedirect.com/science/article/pii/S0145305X11001789

TRANSCRIPTOMICS OF *IN VITRO* IMMUNE-STIMULATED HEMOCYTES FROM THE MANILA CLAM *RUDITAPES PHILIPPINARUM* USING HIGH-THROUGHPUT SEQUENCING

OPEN ACCESS Freely available online

PLoS one

Transcriptomics of *In Vitro* Immune-Stimulated Hemocytes from the Manila Clam *Ruditapes philippinarum* Using High-Throughput Sequencing

Rebeca Moreira[1], Pablo Balseiro[1], Josep V. Planas[2], Berta Fuste[3], Sergi Beltran[3], Beatriz Novoa[1], Antonio Figueras[1]*

1 Instituto de Investigaciones Marinas, Consejo Superior de Investigaciones Científicas, Vigo, Spain, **2** Departament de Fisiologia i Immunologia, Facultat de Biologia, Universitat de Barcelona i Institut de Biomedicina de la Universitat de Barcelona, Barcelona, Spain, **3** Centros Científicos y Tecnológicos de la UB, Universitat de Barcelona, Barcelona, Spain

Abstract

Background: The Manila clam (*Ruditapes philippinarum*) is a worldwide cultured bivalve species with important commercial value. Diseases affecting this species can result in large economic losses. Because knowledge of the molecular mechanisms of the immune response in bivalves, especially clams, is scarce and fragmentary, we sequenced RNA from immune-stimulated *R. philippinarum* hemocytes by 454-pyrosequencing to identify genes involved in their immune defense against infectious diseases.

Methodology and Principal Findings: High-throughput deep sequencing of *R. philippinarum* using 454 pyrosequencing technology yielded 974,976 high-quality reads with an average read length of 250 bp. The reads were assembled into 51,265 contigs and the 44.7% of the translated nucleotide sequences into protein were annotated successfully. The 35 most frequently found contigs included a large number of immune-related genes, and a more detailed analysis showed the presence of putative members of several immune pathways and processes like the apoptosis, the toll like signaling pathway and the complement cascade. We have found sequences from molecules never described in bivalves before, especially in the complement pathway where almost all the components are present.

Conclusions: This study represents the first transcriptome analysis using 454-pyrosequencing conducted on *R. philippinarum* focused on its immune system. Our results will provide a rich source of data to discover and identify new genes, which will serve as a basis for microarray construction and the study of gene expression as well as for the identification of genetic markers. The discovery of new immune sequences was very productive and resulted in a large variety of contigs that may play a role in the defense mechanisms of *Ruditapes philippinarum*.

Citation: Moreira R, Balseiro P, Planas JV, Fuste B, Beltran S, et al. (2012) Transcriptomics of *In Vitro* Immune-Stimulated Hemocytes from the Manila Clam *Ruditapes philippinarum* Using High-Throughput Sequencing. PLoS ONE 7(4): e35009. doi:10.1371/journal.pone.0035009

Editor: Jason Mulvenna, James Cook University, Australia

Received October 21, 2011; **Accepted** March 8, 2012; **Published** April 19, 2012

Funding: This work has been funded by the Spanish Ministerio de Ciencia e Innovación (MICINN) (AGL2008-05111) and by the EU Project REPROSEED (245119). R.M. wishes to acknowledge additional funding from the Spanish MICINN through a Formación de Personal Investigador Spanish research grant (BES-2009-029765). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: antoniofigueras@iim.csic.es

Introduction

The Manila clam (*Ruditapes philippinarum*) is a cultured bivalve species with important commercial value in Europe and Asia, and its culture has expanded in recent years. Nevertheless, diseases produced by a wide range of microorganisms, from viruses to metazoan parasites, can result in large economical losses. Among clam diseases, the majority of pathologies are associated with the *Vibrio* and *Perkinsus* genera [1–3]. Although molluscs lack a specific immune system, the innate response involving circulating hemocytes and a large variety of molecular effectors seems to be an efficient defense method to respond to external aggressions by detecting the molecular signatures of infection [4–8]; however, not many immune pathways have been identified in these animals.

Although knowledge of bivalve immune-related genes has increased in the last few years, the available information is still scarce and fragmentary. Most of the data concern mussels and Eastern and Pacific oysters [9–14], and very limited information is available on the expressed immune genes of *R. philippinarum*. Recently, the expression of 13 immune-related genes of *Ruditapes philippinarum* and *Ruditapes decussatus* were characterized in response to a *Vibrio alginolyticus* challenge [15]. Also, a recent 454 pyrosequencing study was carried out by Milan *et al.* [16], who sequenced two normalized cDNA libraries representing a mixture of adult tissues and larvae from *R. philippinarum*. Even more recently Ghiselli et al. [17], have de novo assembled the *R. philippinarum* gonad transcriptome with the Illumina technology. Moreover, a few transcripts encoded by genes putatively involved in the clam immune response against *Perkinsus olseni* have been reported by cDNA library sequencing [18]. Currently (19/12/

2011), there are 5,662 ESTs belonging to *R. philippinarum* in the GenBank database.

The European Marine Genomics Network has increased the number of ESTs for marine mollusc species particularly for ecologically and commercially important groups that are less studied, such as mussels and clams [19]. Unfortunately, most of the available resources are not annotated or well described, limiting the identification of important genes and genetic markers for future aquaculture applications. The use of 454-pyrosequencing is a fast and efficient approach for gene discovery and enrichment of transcriptomes in non-model organisms [20]. This relatively low-cost technology facilitates the rapid production of a large volume of data, which is its main advantage over conventional sequencing methods [21].

In the present work, we undertook an important effort to significantly increase the number of *R. philippinarum* ESTs in the public databases. Specially, the aim of this work was to discover new immune-related genes using pyrosequencing on the 454 GS FLX (Roche-454 Life Sciences) platform with the Titanium reagents. To achieve this goal, we sequenced the transcriptome of *R. philippinarum* hemocytes previously stimulated with different pathogen-associated molecular patterns (PAMPs) to obtain the greatest number of immune-related transcripts as possible. The raw data are accessible in the NCBI Short Read Archive (Accession number: SRA046855.1).

Results and Discussion

Sequence analysis and functional annotation

The *R. philippinarum* normalized cDNA library was sequenced with 454 GS FLX technology as shown in Figure 1. Sequencing and assembly statistics are summarized in Table 1. Briefly, a total of 975,190 raw nucleotide reads averaging 284.1 bp in length were obtained. Of these, 974,976 exceeded our minimum quality standards and were used in the MIRA assembly. A total of 842,917 quality reads were assembled into 51,265 contigs, corresponding to 29.9 megabases (Mb). The length of the contigs varied from 40 to 5565 bp, with an average length of 582.4 bp and an average coverage of 5.7 reads. Singletons were discarded, resulting in 37,093 contigs formed by at least 2 ESTs, and 26,675 of these contigs were longer than 500 bp. Clustering the contigs resulted in 1,689 clusters with more than one contig. The distribution of contig length and the number of ESTs per contig, as well as the contig distribution by cluster are all shown in Figure 2.

Even though the knowledge of expressed genes in bivalves has increased in the last few years, it is still limited. Indeed, only 41,598 nucleotide sequences, 362,149 ESTs, 24,139 proteins and 704 genes from the class *Bivalvia* have been deposited in the GenBank public database (19/12/11), and the top entries are for the *Mytilus* and *Crassostrea* genera. For *Ruditapes philippinarum*, these numbers are reduced to 5,662 ESTs, 612 proteins and 12 genes. This evidences the lack of information which prompted the recent efforts to increase the number of annotated sequences of bivalves in the databases. For non-model species, functional and comparative genomics is possible after obtaining good EST databases. These studies seem to be the best resource for deciphering the putative function of novel genes, which would otherwise remain "unknown".

NCBI Swissprot, NCBI Metazoan Refseq, the NCBI non-redundant and the UniprotKB/Trembl protein databases were chosen to annotate the contigs that were at least 100 bp long (49,847). The percentage of contigs annotated with a cut off e-value of 10e-3 was 44.7%. Contig sequences and annotations are included in Table S1. Of these contigs, 3.26% matched sequences from bivalve species and the remaining matched to non-Bivalvia mollusc classes (4.13%), other animals (81.38%), plants (2.58%), fungi (1.78%), protozoa (1.50%), bacteria (4.95%), archaea (0.20%), viruses (0.21%) and undefined sequences (0.01%). As shown in Figure 3A, the species with the most sequence matches was *Homo sapiens* with 3,106 occurrences. The first mollusc in the top 35 list was *Lymnaea stagnalis* at position 11. The first bivalve, *Meretrix lusoria*, appeared at position 17. *R. philippinarum* was at position 25 with 124 occurrences. Notably, a high percentage of the sequences had homology with chordates, arthropods and gastropods (Figure 3B and C), and only 343 contigs matched with sequences from the *Veneroida* order (Figure 3D). These values can be explained by the higher representation of those groups in the databases as compared to bivalves and the quality of the annotation in the databases, which has been reported in another bivalve transcriptomic study [22]. The data shown highlight, once again, the necessity of enriching the databases with bivalve sequences.

A detailed classification of predicted protein function is shown for the top 35 BLASTx hits (Figure 4A). The list is headed by actin with 903 occurrences, followed by ferritin, an angiopoietin-like protein and lysozyme. An abundance of proteins directly involved in the immune response was predicted for this 454 run; ferritin, lysozyme, C1q domain containing protein, galectin-3 and hemagglutinin/amebocyte aggregation factor precursor are immune-related proteins present on the top 35 list.

Ferritin has an important role in the immune response. It captures circulating iron to overcome an infection and also functions as a proinflammatory cytokine via the iron-independent nuclear factor kappa B (NF-κB) pathway [23]. Lysozyme is a key

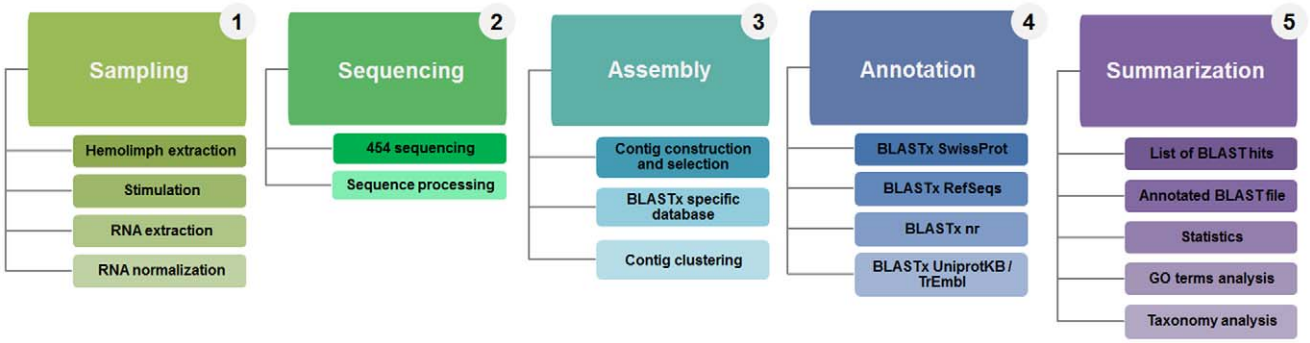


Figure 1. Flow chart summarizing work tasks and the data processing pipeline.
doi:10.1371/journal.pone.0035009.g001

protein in the innate immune responses of invertebrates against Gram-negative bacterial infections and could also have antifungal properties. In addition, it provides nutrition through its digestive properties as it is a hydrolytic protein that can break the glycosidic union of the peptidoglycans of the bacteria cell wall [24]. The C1q domain containing proteins are a family of proteins that form part of the complement system. The C1q superfamily members have been found to be involved in pathogen recognition, inflammation, apoptosis, autoimmunity and cell differentiation. In fact, C1q can be produced in response to infection and it can promote cell survival through the NF-κB pathway [25]. Galectin-3 is a central regulator of acute and chronic inflammatory responses through its effects on cell activation, cell migration, and the regulation of apoptosis in immune cells [26]. The hemagglutinin/amebocyte aggregation factor is a single chain polypeptide involved in blood coagulation and adhesion processes such as self-nonself recognition, agglutination and aggregation processes. The hemagglutinin/amebocyte aggregation factor and lectins play important roles in defense, specifically in the recognition and destruction of invading microorganisms [27].

Other proteins that are not specifically related to the immune response but could play a role in defense mechanisms include the following: angiopoietin-like proteins, apolipoprotein D and the integral membrane protein 2B. In other animals, angiopoietin-like proteins (ANGPTL) potently regulate angiogenesis, but a subset also function in energy metabolism. Specifically, ANGPTL2, the most represented ANGPTL, promotes vascular inflammation rather than angiogenesis in skin and adipose tissues. Inflammation occurs via the a5b1 integrin/Rac1/NF-κB pathway, which is evidenced by an increase in leukocyte infiltration, blood vessel permeability and the expression of inflammatory cytokines (tumor necrosis factor-α, interleukin-6 and interleukin-1b) [28]. Apolipoprotein D (apoD) has been associated with inflammation. Pathological and stressful situations involving inflammation or growth arrest have the capacity to increase its expression. This effect seems to be triggered by LPS, interleukin-1, interleukin-6 and glucocorticoids and is likely mediated by the NF-κB pathway, as there are several conserved NF-κB binding sites in the apoD promoter (APRE-3 and AP-1 binding sites are also present). The highest affinity ligand for apoD is arachidonic acid, which apoD traps when it is released from the cellular membrane after inflammatory stimuli and, thus, prevents its subsequent conversion in pro-inflammatory eicosanoids. Within the cell, apoD could modulate signal transduction pathways and nuclear processes such as transcription activation, cell cycling and apoptosis. In summary, apoD induction is specific to ongoing cellular stress and could be part of the protective components of mild inflammation [29–31]. Finally, the short form of the integral membrane protein 2B (ITM2Bs) can induce apoptosis via a caspase-dependent mitochondrial pathway [32].

To avoid redundancy, the longest contig of each cluster was used for Gene Ontology terms assignment. A total of 23.05% of the representative clusters matched with at least one GO term. Concerning cellular components (Figure 4B), the highest percentage of GO terms were in the groups of cell and cell part with 25.9% in each; organelle and organelle part represented 19.67% and 11.38%, respectively. Within the molecular function classification (Figure 4C), the most represented group was binding with 49.25% of the terms, which was followed by catalytic activity (29.12%) and structural molecular activity (4.60%). With regard to biological process (Figure 4D), cellular and metabolic processes were the highest represented groups with 16.78% and 12.43% of the terms, respectively, which was followed by biological regulation (10.18%).

Comparative analysis

Similarities between the *R. philippinarum* transcriptome and another four bivalve species sequences were analyzed by comparative genomics (*Crassostrea gigas* of the family Ostreidae, *Bathymodiolus azoricus* and *Mytilus galloprovincialis* of the family Mytilidae and *Laternula elliptica* of the family Laternulidae). This analysis could identify specific transcripts that are conserved in these five species. A Venn diagram was constructed using unique sequences from these databases according to the gene identifier (gi id number) of each sequence in its respective database: 207,764 from *C. gigas*, 76,055 from *B. azoricus*, 121,318 from *M. galloprovincialis* and 1,034,379 from *L. elliptica*. *C. gigas* was chosen because is the most represented bivalve species in the public databases. The other three species are bivalves that have been studied in transcriptomic assays.

Figure 5 shows that of the total 29,679 clusters, 72% were found exclusively in the *R. philippinarum* group, while only 7.59% shared significant similarity with all five species. The number of coincidences among other groups was very low (4.14% to 0.31% of sequences), suggesting that 21,454 new sequences were discovered within the bivalve group. The percentage of new sequences is very high compared to previous transcriptomic studies

Table 1. Summary of assembly and EST data.

Sequences before filtering	
Number of reads	975,190
Total Megabases	277.05
Average read length (bp)	284.1
N50 read length (bp)	356
Sequences after filtering	
Number of reads	974,976
Total Megabases	250.36
Average read length	256.78
N50 read length	338
Assembly statistics	
Number of reads assembled	842,957
Number of contigs	51,265
Total consensus Megabases	29.9
Average contig coverage	5.7
Average contig length	582.4
N50 contig length	677
Range contig length	40–5,565
Number of contigs >99 pb	49,847
Number of contigs >500 pb	26,675
Number of contigs with 2 reads	14,172
Number of contigs with >2 reads	37,093
Number of clusters	29,679
Number of clusters with 1 contig	27,990
Number of clusters with >1 contig	1,689
Percentage of contigs annotated	44.7
Percentage of annotated contigs by SwissProt	81.3
Percentage of annotated contigs by nr	16.2
Percentage of annotated contigs by RefSeq	2.5
Percentage of annotated contigs by UniprotKB/Trembl	0

doi:10.1371/journal.pone.0035009.t001

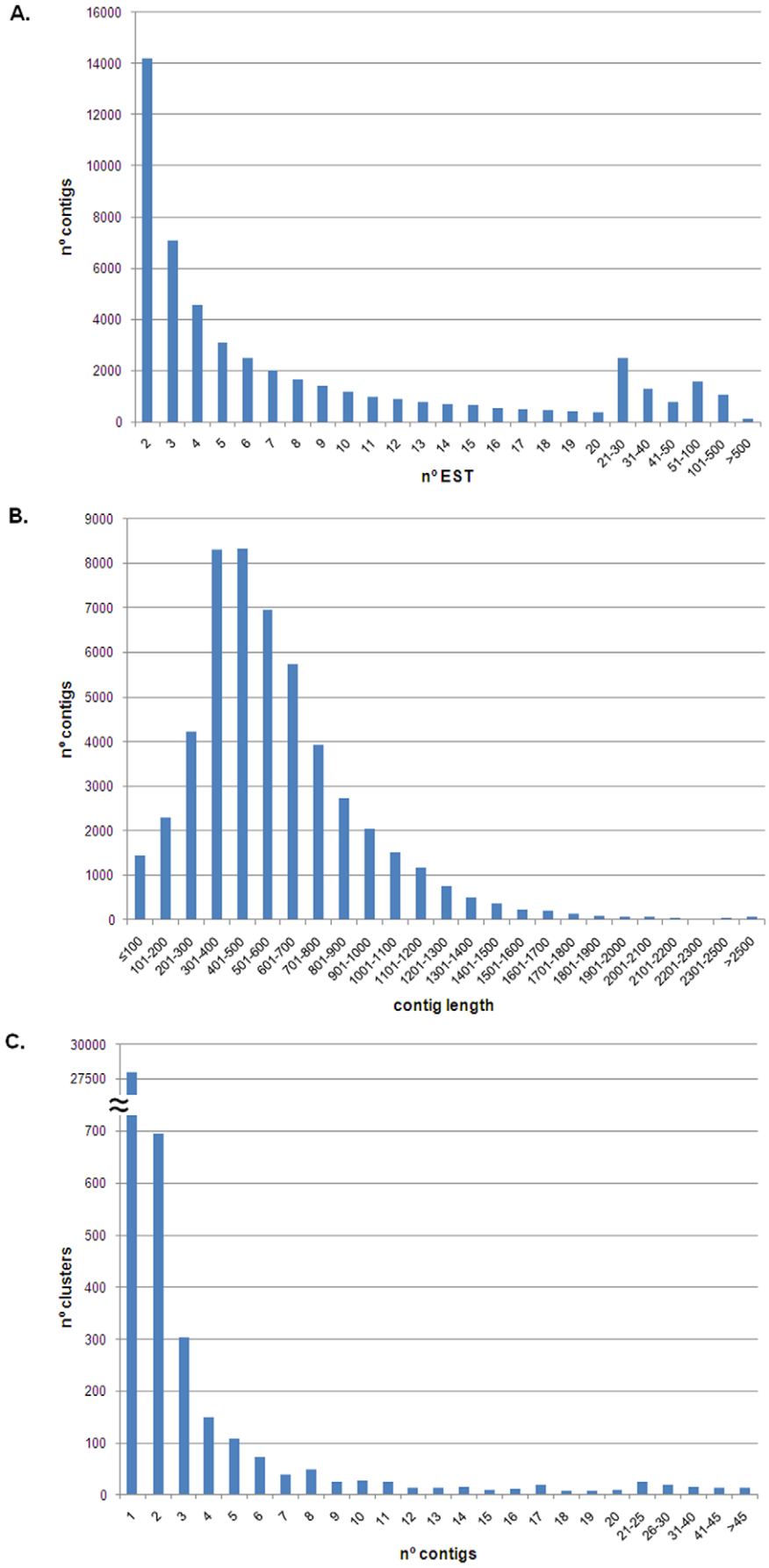


Figure 2. Transcriptome assembly statistics. A: Distribution of contig composition by EST. **B:** Distribution of contig length. **C:** Distribution of cluster composition by contigs.
doi:10.1371/journal.pone.0035009.g002

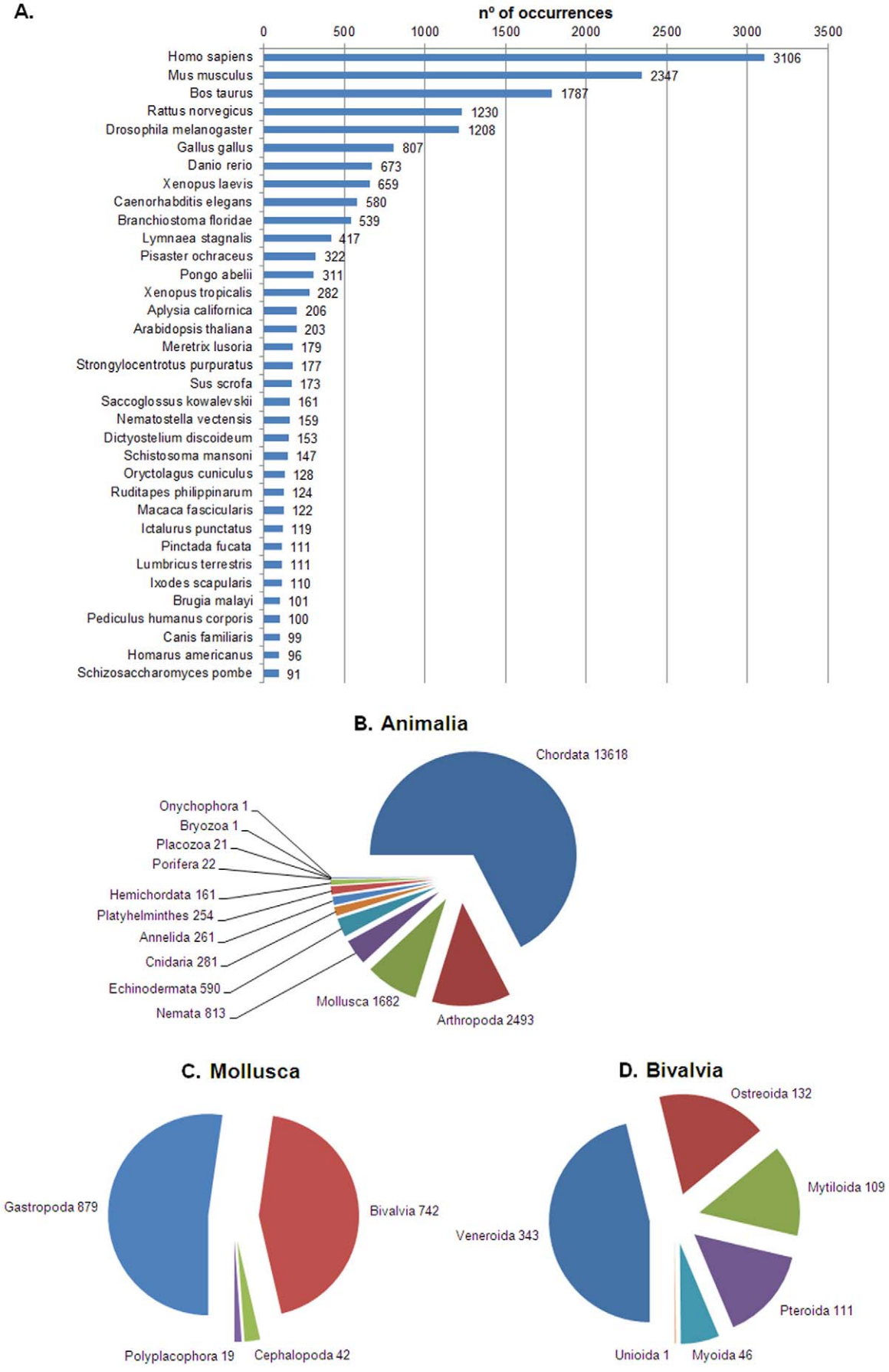

A.
nº of occurrences
0
500
1000
1500
2000
2500
3000
3500
Homo sapiens 3106
Mus musculus 2347
Bos taurus 1787
Rattus norvegicus 1230
Drosophila melanogaster 1208
Gallus gallus 807
Danio rerio 673
Xenopus laevis 659
Caenorhabditis elegans 580
Branchiostoma floridae 539
Lymnaea stagnalis 417
Pisaster ochraceus 322
Pongo abelii 311
Xenopus tropicalis 282
Aplysia californica 206
Arabidopsis thaliana 203
Meretrix lusoria 179
Strongylocentrotus purpuratus 177
Sus scrofa 173
Saccoglossus kowalevskii 161
Nematostella vectensis 159
Dictyostelium discoideum 153
Schistosoma mansoni 147
Oryctolagus cuniculus 128
Ruditapes philippinarum 124
Macaca fascicularis 122
Ictalurus punctatus 119
Pinctada fucata 111
Lumbricus terrestris 111
Ixodes scapularis 110
Brugia malayi 101
Pediculus humanus corporis 100
Canis familiaris 99
Homarus americanus 96
Schizosaccharomyces pombe 91
B. Animalia
Chordata 13618
Onychophora 1
Bryozoa 1
Placozoa 21
Porifera 22
Hemichordata 161
Platyhelminthes 254
Annelida 261
Cnidaria 281
Echinodermata 590
Nemata 813
Mollusca 1682
Arthropoda 2493
C. Mollusca
Gastropoda 879
Bivalvia 742
Polyplacophora 19
Cephalopoda 42
D. Bivalvia
Ostreoida 132
Mytiloida 109
Veneroida 343
Pteroida 111
Unioida 1
Myoida 46

Figure 3. Taxonomic classification and distribution of annotated sequences. Numbers on bars and pie charts refers to the n° of occurrences. **A:** BLASTx results for the top 35 species matching sequences. **B:** Kingdom Animalia distribution. **C:** *Phylum* Mollusca distribution. **D:** Class Bivalvia distribution.
doi:10.1371/journal.pone.0035009.g003

[33–34], in which the fraction of new transcripts was approximately 45%. One possible explanation for this discrepancy is the low number of nucleotide and EST sequences currently available in public databases for *R. philippinarum*, but these transcripts could also be regions in which homology is not reached, such as 5′ and 3′ untranslated regions or genes with a high mutation rate.

On the other hand, a comparison between our 454 results and the Milan *et al.* [16] transcriptome using a BLASTn approach is summarized in Table 2. It is worth noting that the number of hits for immune-related clusters is 1,120 (55.9%) when compared to Milan *et al.* transcriptome; consequently, there are 885 (44.1%) new clusters in our *R. philippinarum* immune transcriptome not represented in Milan *et al.* study. Moreover, 44% of our total results (13,059 clusters) were found in Milan *et al.* study; therefore, 56% of our total clusters (16,620) represent new information about Manila clam transcriptome compared to Milan *et al.* study.

Immune-related sequences

R. philippinarum hemocytes were subjected to immune stimulation using several different PAMPs to enrich the EST collection with immune-related sequences. The objective was to obtain a more complete view of clam responses to pathogens. A keyword list and GO immune-related terms were used to find proteins putatively involved in the immune system. After this selection step, we found that more than 10% of the proteins predicted from the contig sequences had a possible immune function. Some sequences were found to be clustered in common, well-recognized immune pathways, such as the complement, apoptosis and toll-like receptors pathways, indicating conserved ancient mechanisms in bivalves (Figures 6, 7, 8).

1. Complement pathway

The complement system is composed of over 30 plasma proteins that collaborate to distinguish and eliminate pathogens. C3 is the central component in this system. In vertebrates, it is proteolytically activated by a C3 convertase through both the classic, lectin-induced and alternative routes [35]. Although the complement pathway has not been extensively described in bivalves, there is evidence that supports the presence of this defense mechanism. ESTs with homology to the C1q domain have been detected in the American oyster, *C. virginica* [36], the tropical clam *Codakia orbicularis* [37], the Zhikong scallop *Chlamys farreri* [38] and the mussel *M. galloprovincialis* [39–40]. More recently, a novel C1q adiponectin-like, a C3 and a factor B-like proteins have been identified in the carpet shell clam *R. decussatus* [41–42]. These data support the putative presence of the complement system in bivalves.

Our pyrosequencing results, using the BLASTx similarity approach, showed that the complement pathway in *R. philippinarum* was almost complete as compared to the KEGG reference pathway (Figure 6). Only the complement components C1r, C1s, C6, C7 and C8 were not detected.

2. Pattern recognition receptors (PRRs)

i. Lectins. Lectins are a family of carbohydrate-recognition proteins that play crucial self- and non-self-recognition roles in innate immunity and can be found in soluble or membrane-associated forms. They may initiate effector mechanisms against pathogens, such as agglutination, immobilization and complement -mediated opsonization and lysis [43].

Several types of lectins have been cloned or purified from the Manila clam, *R. philippinarum* [44–46], and their function and expression were also studied [18,47]. Also, a Manila clam tandem-repeat galectin, which is induced upon infection with *Perkinsus olseni*, has been characterized [46].

Lectin sequences have been found in the stimulated hemocytes studied in our work: 23 of the contigs are homologous to C-type lectins (calcium-dependent carbohydrate-binding lectins that have characteristic carbohydrate-recognition domains), 115 are homologous to galectins (characterized by a conserved sequence motif in their carbohydrate recognition domain and a specific affinity for β-galactosides), 4 contigs have homology with ficolin A and B (a group of oligomeric lectins with subunits consisting of both collagen-like and fibrinogen-like domains) and 34 contigs have homology with other groups of lectins such as lactose-, mannose- or sialic acid-binding lectins.

ii. β-glucan recognition proteins. β-glucan recognition proteins are involved in the recognition of invading fungal organisms. They bind specifically to β-1,3-glucan stimulating short-term immune responses. Although these receptors have been partially sequenced in several bivalves, there is only one complete description of them in the scallop *Chlamys farreri* [48].

Two contigs with homology to the beta-1,3-glucan-binding protein were found in our study.

iii. Peptidoglycan recognition proteins. Peptidoglycan recognition proteins (PGRPs) specifically bind peptidoglycans, which is a major component of the bacterial cell wall. This family of proteins influences host-pathogen interactions through their pro- and anti-inflammatory properties that are independent of their hydrolytic and antibacterial activities. In bivalves, they were first identified in the scallops *C. farreri* and *A. irradians* [49,50] and the Pacific oyster *C. gigas*, and from the latter four different types of PGRPs were identified [51].

Peptidoglycan-recognition proteins and a peptidoglycan-binding domain containing protein have been found for the first time in *R. philippinarum* in our results and were present 4 and 1 times, respectively.

iv. Toll-like receptors. Toll-like receptors (TLRs) are an ancient family of pattern recognition receptors that play key roles in detecting non-self substances and activating the immune system. The unique bivalve TLR was identified and characterized in the Zhikong Scallop, *C. farreri* [52].

TLR 2, 6 and 13 were present among the pyrosequencing results. TLR2 and TLR6 form a heterodimer, which senses and recognizes various components from bacteria, mycoplasma, fungi and viruses [53]. TLR13 is a novel and poorly characterized member of the Toll-like receptor family. Although the exact role of TLR13 is currently unknown, phylogenic analysis indicates that TLR13 is a member of the TLR11 subfamily [54] suggesting that it could recognize urinary pathogenic *E. coli* [55]. It has been demonstrated that TLR13 colocalizes and interacts with UNC93B1, a molecule located in the endoplasmic reticulum, which strongly suggests that TLR13 might be found inside cells and might play a role in recognizing viral infections [56]. Figure 7 summarizes the TLR signaling pathway with the corresponding molecules found in the *R. philippinarum* transcriptome.

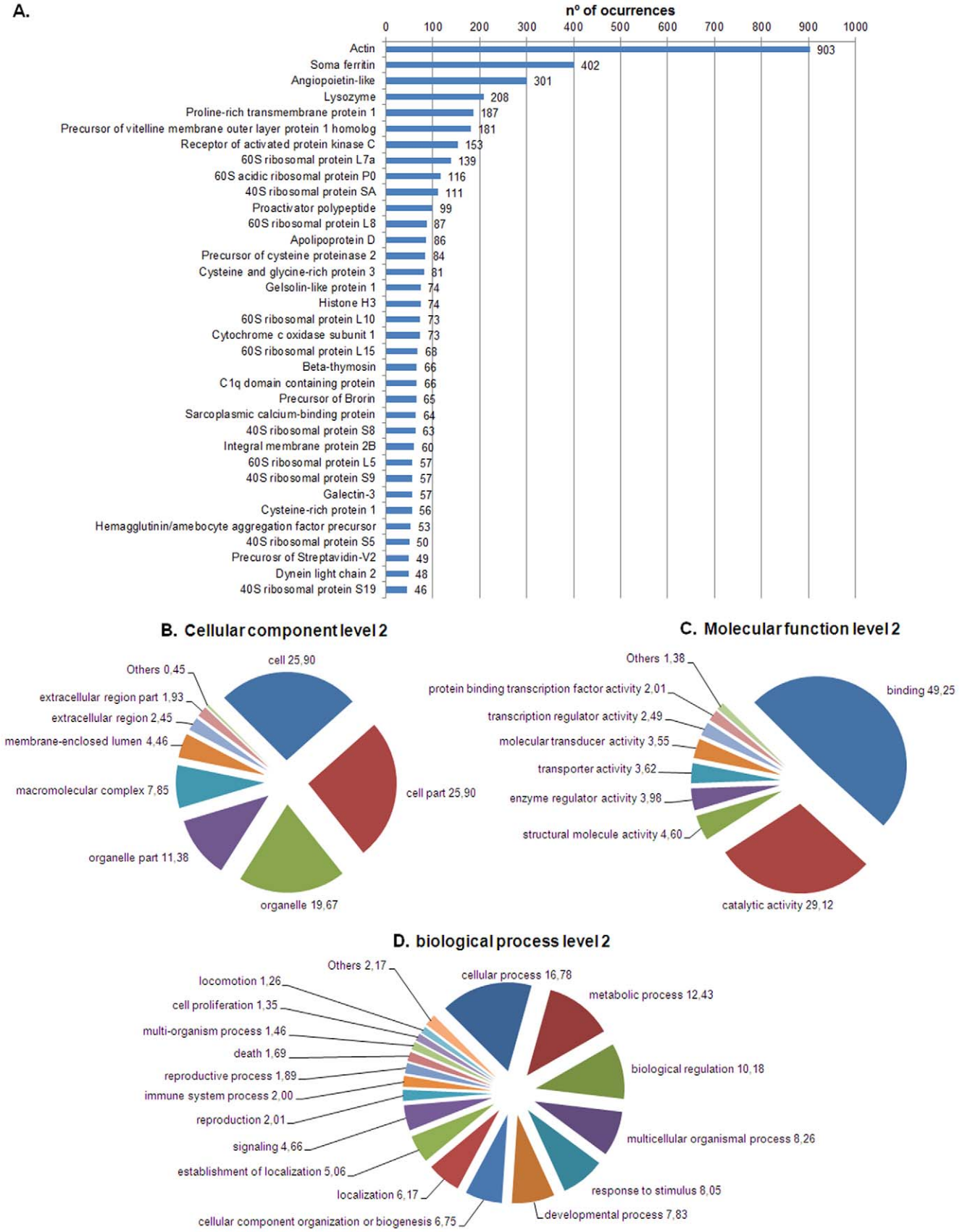


Figure 4. Classification of annotated sequences by BLASTx and GeneOntology Terms at level 2. A: The top 35 hit sequences by BLASTx. Numbers refer to the n° of occurrences **B:** Cellular component. **C:** Molecular function. **D:** Biological process. Numbers in B, C and D refer to the percentage of occurrences.
doi:10.1371/journal.pone.0035009.g004

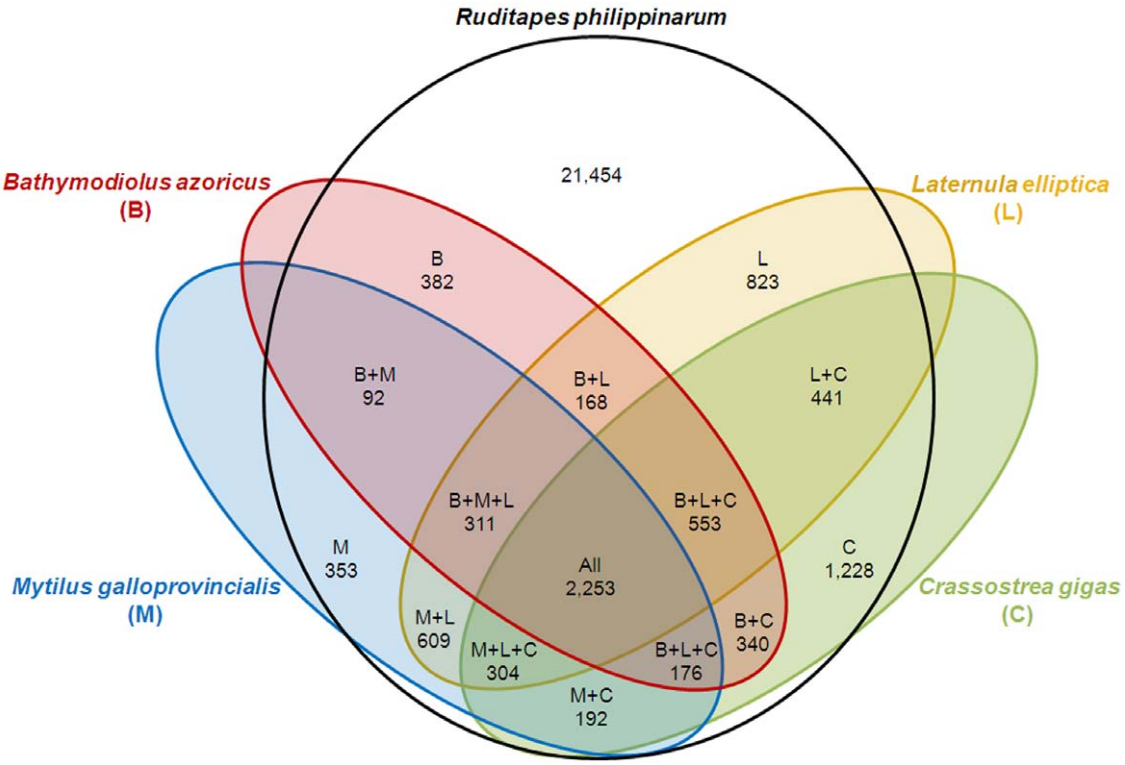


Figure 5. Venn diagram showing the comparison of *R. philippinarum* sequences with *B. azoricus* (B), *L.elliptica* (L), *M.galloprovincialis* (M) and *C.gigas* (C) known sequences. Numbers refer to the n° sequences belonging to each group.
doi:10.1371/journal.pone.0035009.g005

3. Protease inhibitors

Pathogen proteases are important virulence factors that facilitate infection, diminish the activity of lysozymes and quench the agglutination capacity of hemocytes. Because protease inhibitors play important roles in invertebrate immunity by protecting hosts through the direct inactivation of pathogen proteases, many bivalves have developed protease inhibitors to regulate the activities of pathogen proteases [1]. Some genes encoding protease inhibitors were identified in *C. gigas* [57], *A. irradians* [58], *C. farreri* [59] and *C. virginica*; in the latter a novel family of serine protease inhibitors was also characterized [60–62].

A total of 23 contigs with homology to Serine, Cystein, Kunitz- and Kazal- type protease inhibitors and metalloprotease inhibitors were found among our results.

4. Lysozyme

Lysozyme was one of the most represented groups of immune genes in this transcriptome study with 208 contigs present. It is an antibacterial molecule present in numerous animals including bivalves. Although lysozyme activity was first reported in molluscs over 30 years ago, complete sequences were published only recently including those of *R. philippinarum* [24].

Table 2. Comparison between Milan *et al.* and the current *R.philippinarum* transcriptome.

	Database (Milan et al. contigs)	
Query (current study)	**% hits**	**% coverage**
TOTAL Clusters (29,679)	44.0 (13,059)	66.6
INMUNE-RELATED Clusters (2,005)	55.9 (1,120)	67.6

doi:10.1371/journal.pone.0035009.t002

5. Antimicrobial peptides

Antimicrobial peptides (AMPs) are small, gene-encoded, cationic peptides that constitute important innate immune effectors from organisms spanning most of the phylogenetic spectrum. AMPs alter the permeability of the pathogen membrane and cause cellular lysis [63]. In bivalves, they were first purified from mussel hemocyte granules [64,65]. In mussels, the AMP myticin C was found to have a high polymorphic variability as well as chemotactic and immunoregulatory roles [66,67]. In clams, two AMPs with similarity to mussel myticin and mytilin [68] and a big defensin [69] are known.

We were able to detect 36 contigs with homology to different defensins: defensin-1 (American oyster defensin), defensin MGD-1 (Mediterranean mussel defensin) and the big defensin previously mentioned. Four contigs were similar to an unpublished defensin sequence from *Venerupis* (= *Ruditapes*) *philippinarum*.

6. Heat shock proteins

The primary role of heat shock proteins (HSPs) is to function as molecular chaperones. Their up-regulation also represents an important mechanism in the stress response [70], and their activity is closely linked to the innate immune system. HSPs mediate the mitochondrial apoptosis pathway and affect the regulation of NF-κB [71]. HSPs are well studied in bivalves. For *R. philippinarum*, several assays have been developed to better understand the HSPs profile in response to heavy metals and pathogen stresses [72–74].

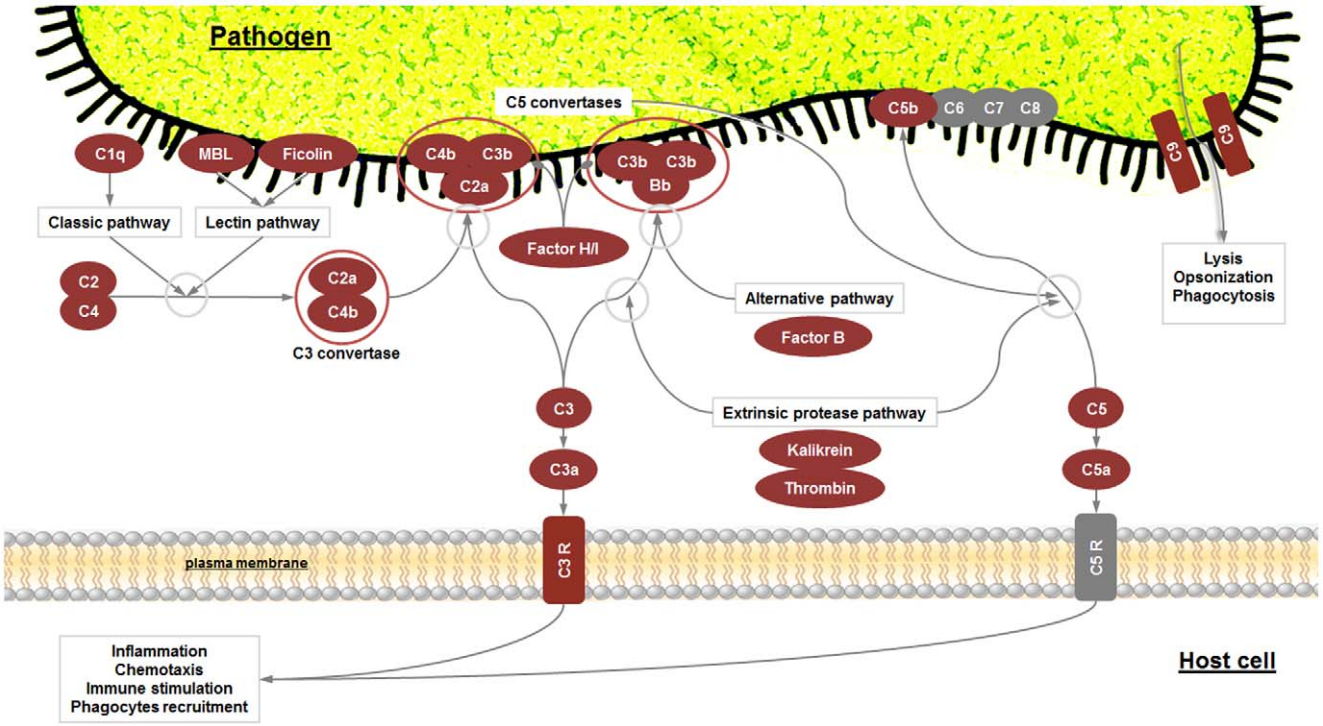


Figure 6. Complement pathway. Red boxes indicate proteins identified in our 454 results and grey boxes the absent ones. Connectors finishing in a circle indicate inhibition. C3R:C3 receptor; C5R: C5 receptor; MBL: Mannose-binding protein.
doi:10.1371/journal.pone.0035009.g006

The most important and well-studied groups of HSPs were present in our *R. philippinarum* transcriptome (HSP27, HSP40/DnaJ, HSP70 and HSP90), but other, less common HSPs were also represented (HSP10, HSP22, HSP83 and some members from the HSP90 family).

7. Other immune molecules

Recently, several genes related to the inflammatory response against LPS stimulation have been detected in bivalves. Such is the case of the LPS-induced TNF-α factor (LITAF), which is a novel transcription factor that critically regulates the expression of TNF-α and various inflammatory cytokines in response to LPS stimulation. It has been described in three bivalve species: *Pinctada fucata* [75], *C. gigas* [76] and *C. farreri* [77].

Other TNF-related genes have been identified in the Zhikong scallop, such as a TNFR homologue [78] and a tumor necrosis factor receptor-associated factor 6 (TRAF6), which is a key signaling adaptor molecule common to the TNFR superfamily and to the IL-1R/TLR family [79]. Figure 7 shows that several components of the TLR signaling pathway that are present in our transcriptomic sequences (MyD88, IRAK4, TRAF-3 and -6, TRAM, BTK, RAC-1, PI3K, AKT, BTK and TANK).

Pathogen sequences

A total of 1,918 contigs, 8.43% of those annotated, had homology with the main groups of putatively pathogenic organisms such as viruses (47 hits), bacteria (1,126 hits), protozoa (341 hits) and fungi (404 hits). Figure 9 displays the taxonomic classification of these sequences and Table 3 summarizes a list of the known bivalve pathogens found in our results.

Bacteria constitute the main group found among the sequences not belonging to the clam. As filter-feeding animals, bivalves can concentrate a large amount of bacteria and it could be one of their sources of food [24]. Because *Vibrio* spp. are ubiquitous in aquatic ecosystems, it was expected that the *Vibrionales* order, with 141 hits, would be the most predominant. Several species of the *Vibrio* genus are among the main causes of disease in bivalves specifically causing bacillary necrosis in larval stages [80]. Is noticeable that sequences belonging to the causative agent of Brown Ring Disease in adults of Manila Clam, *Vibrio tapetis*, have not been found.

Perkinsus marinus, with 2 matches, is the only bivalve pathogen found within the protozoa (*Alveolata*) group. Perkinsosis is produced by species from the genus *Perkinsus*. Both *P. marinus* and *P. olseni* have been associated with mortalities in populations of various groups of molluscs around the world and are catalogued as notifiable pathogens by the OIE.

Viruses were the least represented among pathogens. The *Baculoviridae* family was the most predominant with 21 matches, but the corresponding sequences were inhibitors of apoptosis (IAPs) [81] that could also be part of the clam's transcriptome. Five viral families were found in our transcriptome study: *Iridoviridae*, *Herpesviridae*, *Malacoherpesviridae*, *Picornaviridae* and *Retroviridae*. A well-known bivalve pathogen was also identified, the ostreid herpesvirus 1, which has been previously been found to infect clams [82].

Fungi had 404 matches in our results. It is known that bivalves are sensitive to fungal diseases, which can degrade the shell or affect the larval bivalve stages [83,84].

Conclusions

This study represents the first *R. philippinarum* transcriptome analysis focused on its immune system using a 454-pyrosequencing approach and complements the recent pyrosequencing assay carried out by Milan *et al.* [16]. The discovery of new immune sequences was effective, resulting in an enormous variety of contigs corresponding to molecules that could play a role in the defense

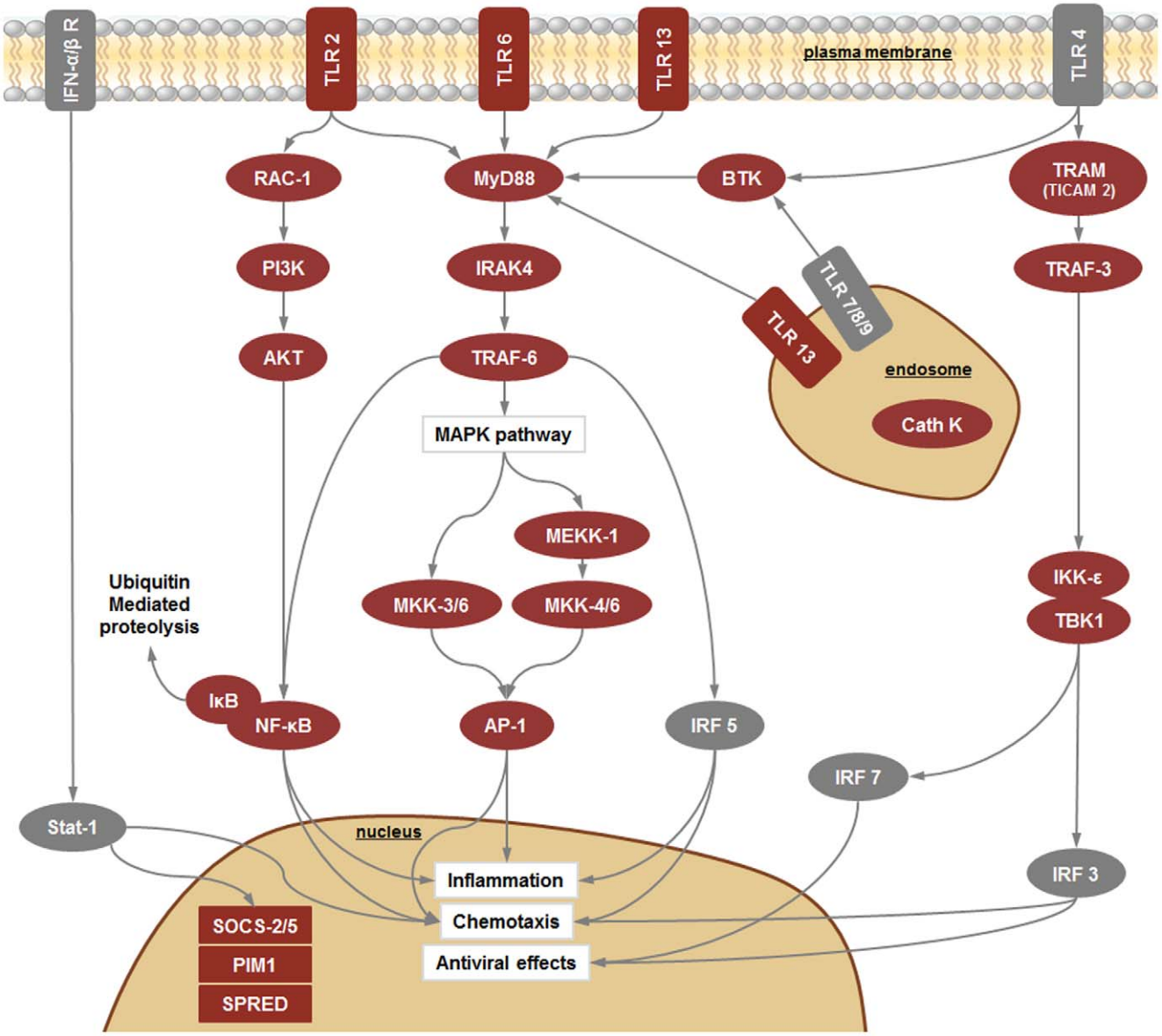


Figure 7. TLR signaling pathway. Red boxes indicate proteins identified in our 454 results and grey boxes the absent ones. Connectors finishing in a circle indicate inhibition. AKT: RAC-alpha serine/threonine-protein kinase = Protein kinase B; AP-1: Transcription factor AP-1 = Proto-oncogene c-Jun; BTK: Tyrosine-protein kinase BTK; Cath: Cathepsin; IFN-α/β R: Interferon alpha/beta receptor; IκB: Inhibitor of NF-κB; IKK-ε: Inhibitor of NF-κB kinase subunit epsilon; IRAK4: Interleukin-1 receptor-associated kinase 4; IRF: Interferon regulatory factor; MEKK: Mitogen-activated protein kinase kinase kinase; MKK: Mitogen-activated protein kinase kinase; MyD88: Myeloid differentiation primary response protein MyD88; NF-κB: Nuclear factor kappa B; PI3K: Phosphatidylinositol 3- kinase; PIM1: Proto-oncogene serine/threonine-protein kinase pim-1; Rac-1: Ras-related C3 botulinum toxin substrate 1; SOCS: Suppressor of cytokine signaling; SPRED: Sprouty-related, EVH1 domain-containing protein 2; Stat-1: signal transducer and activator of transcription 1; TBK1: TANK-binding kinase 1; TLR: Toll-like receptor; TRAF: TNF receptor-associated factor 3/6; TRAM: TIR domain-containing adapter molecule 2.
doi:10.1371/journal.pone.0035009.g007

mechanisms. More than 10% of our results had relationship with immunity. This new resource is now gathered in the NCBI Short Read Archive with the accession number: SRA046855.1.

Our results will provide a rich source of data to discover and identify new genes, which will serve as a basis for microarray construction and gene expression studies as well as for the identification of genetic markers for various applications including the selection of families in the aquaculture sector. We have found sequences from molecules never described in bivalves before like C2, C4, C5, C9, AIF, Bax, AKT, TLR6 and TLR13, among others. As a part of this work, three immune pathways in *R. philippinarum* have been characterized, the apoptosis, the toll like signaling pathway and the complement cascade, which could help us to better understand the resistance mechanisms of this economically important aquaculture clam species.

Materials and Methods

Animal sampling and in vitro stimulation of hemocytes

R. philippinarum clams were obtained from a commercial shellfish farm (Vigo, Galicia, Spain). Clams were maintained in open circuit filtered sea water tanks at 15°C with aeration and were fed

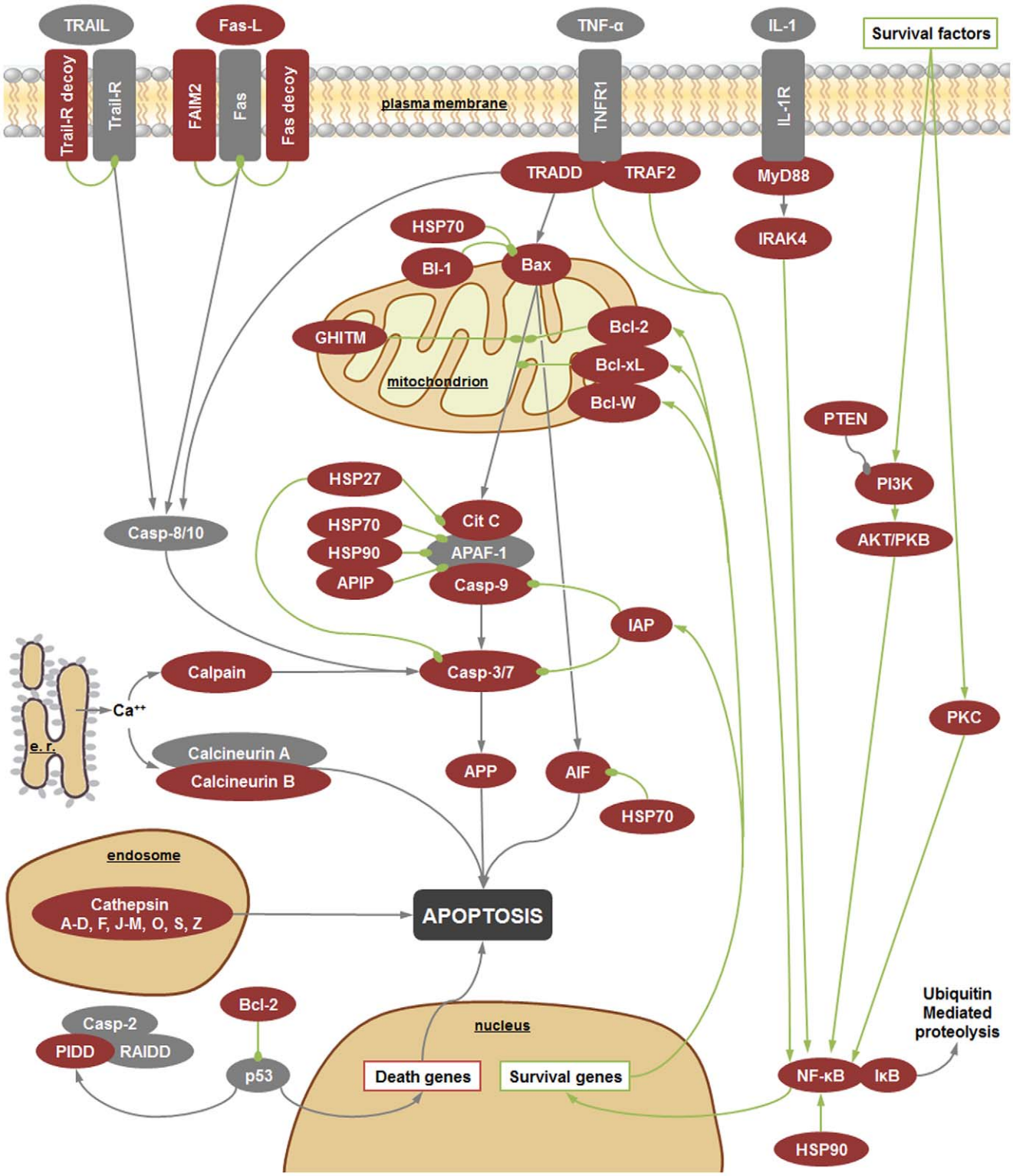


Figure 8. Apoptosis pathway. Red boxes indicate proteins identified in our 454 results and grey boxes the absent ones. Connectors finishing in a circle indicate inhibition. Green connectors highlight the survival pathways and grey ones indicate apoptosis. AIF: Apoptosis-inducing factor 1 mitochondrial; AKT/PKB: RAC-alpha serine/threonine-protein kinase = Protein kinase B; APAF-1: Apoptotic protease-activating factor 1; APIP: APAF1-interacting protein; APP: Amyloid beta A4 protein; Bax: Apoptosis regulator BAX; Bcl-2: Apoptosis regulator Bcl-2; Bcl-W: Bcl-2-like protein 2; Bcl-XL: Bcl-2-like protein 1; BI-1: Bax inhibitor 1; Casp: Caspase; Cit C: Cytochrome C; FAIM: Fas apoptotic inhibitory molecule 2; Fas: Apoptosis-mediating surface antigen FAS (CD95); Fas decoy: Decoy receptor for Fas ligand; Fas-L: Fas antigen ligand; GHITM: Growth hormone-inducible transmembrane protein = Transmembrane BAX inhibitor motif-containing protein 5; HSP: Heat shock protein; IAP: Inhibitor of apoptosis; IkB: Inhibitor of NF-kB; IL-1: Interleukin 1; IL-1 R: Interleukin 1 receptor; IRAK4: Interleukin-1 receptor-associated kinase 4; MyD88: Myeloid differentiation primary response protein MyD87; NF-kB: Nuclear factor kappa B; p53: Tumor suppressor p53; PI3K: Phosphatidylinositol 3- kinase; PIDD: p53-induced protein with a death

domain; PKC: Protein kinase C; PTEN: Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN; RAIDD: Caspase and RIP adapter with death domain; TNF R1: Tumor necrosis factor receptor 1; TNF-a: Tumor necrosis factor alpha; TRADD: TNF receptor type 1-associated DEATH domain protein; TRAF2: TNF receptor-associated factor 2; TRAIL: TNF-related apoptosis-inducing ligand; TRAIL decoy: Decoy TRAIL receptor without death domain; TRAIL-R: TRAIL receptor.
doi:10.1371/journal.pone.0035009.g008

daily with *Phaeodactylum tricornutum* and *Isochrysis galbana*. Prior to the experiments, clams were acclimatized to aquaria conditions for one week.

A total of 100 clams were notched in the shell in the area adjacent to the anterior adductor muscle. A sample of 500 ul of hemolymph was withdrawn from the adductor muscle of each clam with an insulin syringe, pooled and then distributed in 6-well plates, 7 ml per well, in a total of 7 wells, one for each treatment. Hemocytes were allowed to settle to the base of the wells for 30 min at 15°C in the darkness. Then, the hemocytes were stimulated with 50 µg/ml of Polyinosinic:polycytidylic acid (Poly I:C), Peptidoglycans, ß- Glucan, *Vibrio anguillarum* DNA (CpG), Lipopolysaccharide (LPS), Lipoteichoic acid (LTA) or 1×10^6 UFC/ml of heat-inactivated *Vibrio anguillarum* (one stimulus per well) for 3 h at 15°C. All stimuli were purchased from SIGMA.

RNA isolation and cDNA production

Pyrosequencing. After stimulation, hemolymph was centrifuged at 1700 g at 4°C for 5 minutes, the pellet was resuspended in 1 ml of Trizol (Invitrogen) and RNA was extracted following the manufacturer's protocol. After RNA extraction, samples were treated with *Turbo DNase free* (Ambion) to eliminate DNA. Next, the concentration and purity of the RNA samples were measured using a *NanoDrop ND1000* spectrophotometer. The RNA quality was assessed in a Bioanalyzer 2010 (Agilent Technologies). From each sample, 1 µg of RNA was pooled and used for the production of normalized cDNA for 454 sequencing in the Unitat de Genòmica (SCT-UB, Barcelona, Spain).

Full-length-enriched double stranded cDNA was synthesized from 1,5 µg of pooled total RNA using MINT cDNA synthesis kit (Evrogen, Moscow, Russia) according to manufacturer's protocol, and was subsequently purified using the QIAquick PCR Purification Kit (Qiagen USA, Valencia, CA). The amplified cDNA was normalized using Trimmer kit (Evrogen, Moscow, Russia) to minimize differences in representation of transcripts. The method involves denaturation-reassociation of cDNA, followed by a digestion with a Duplex-Specific Nuclease (DSN) enzyme [85,86]. The enzymatic degradation occurs primarily on the highly abundant cDNA fraction. The single-stranded cDNA fraction was then amplified twice by sequential PCR reactions according to the manufacturer's protocol. Normalized cDNA was purified using the QIAquick PCR Purification Kit (Qiagen USA, Valencia, CA).

To generate the 454 library, 500 ng of normalized cDNA were used. cDNA was fractionated into small, 300- to 800-basepair fragments and the specific A and B adaptors were ligated to both the 3′ and 5′ ends of the fragments. The A and B adaptors were

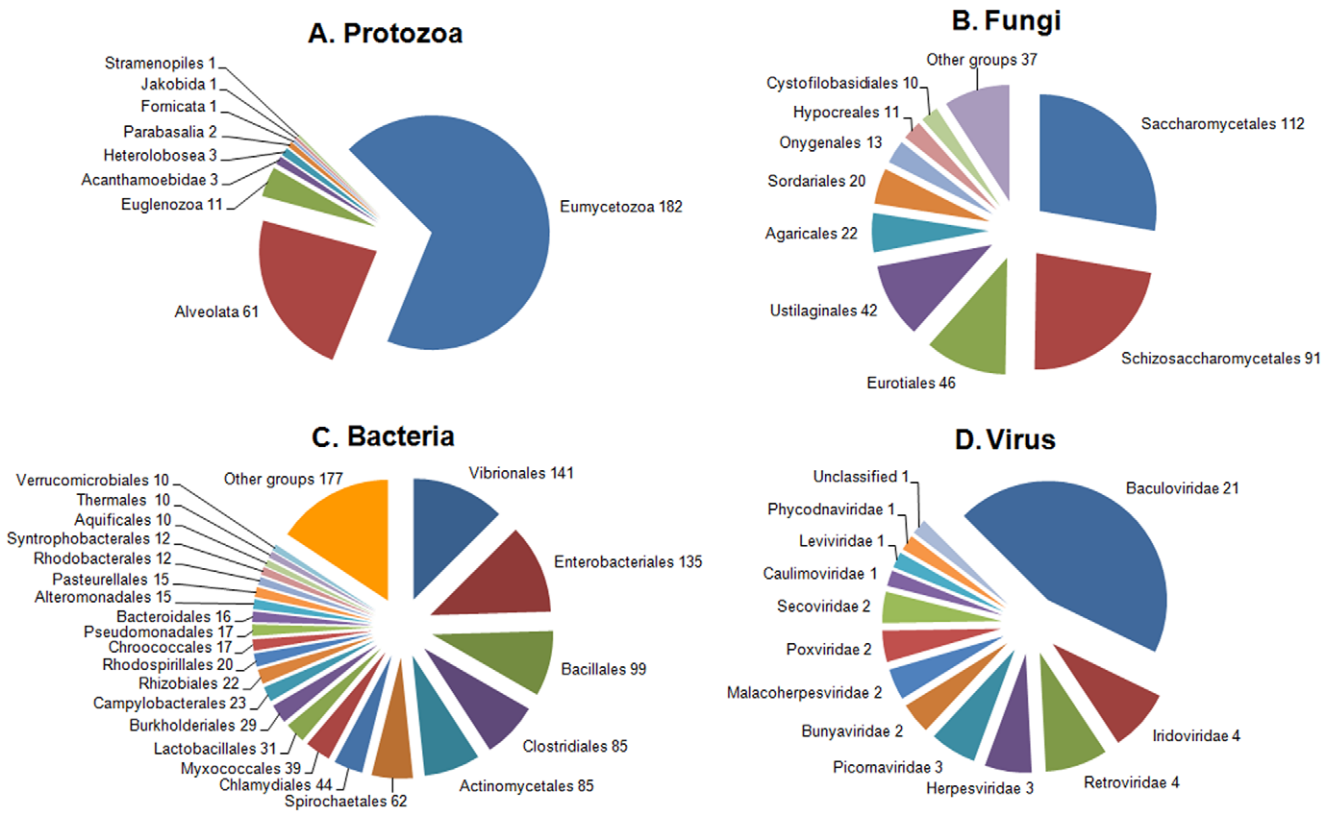


Figure 9. Classification of pathogen sequences. Numbers refer to the n° of occurrences. A: Protozoa. Results are observed at the phylum level. B: Fungi. Results are observed at the order level. C: Bacteria. Results are observed at the order level. D: Viruses. Results are observed at the family level.
doi:10.1371/journal.pone.0035009.g009

Table 3. Putative bivalve pathogen sequences.

Pathogen species	n° contigs	Effects in bivalves
Virus		
Ostreid herpesvirus 1	2	Early stages mortality
Bacteria		
Vibrio alginolyticus	1	Bacilar necrosis
Vibrio anguillarum	2	Bacilar necrosis; Inhibited capacity for filtering
Vibrio harveyi	4	Adult and larval mortality
Vibrio parahaemolyticus	13	Bacilar necrosis
Vibrio splendidus	11	Bacilar necrosis; Summer mortality
Other *Vibrio* species	108	Bacilar necrosis
Genus *Pseudomonas*	16	Bacilar necrosis
Genus *Aeromonas*	2	Bacilar necrosis
Genus *Alteromonas*	1	Bacilar necrosis
Order *Chlamydiales*	42	Early stages mortality
Order *Ricketsiales*	6	Early stages mortality; Gill hyperplasia; Withering Syndrome
Protozoa		
Perkinsus marinus	2	Perkinsosis
Fungi		
_	401	Fungal disease

doi:10.1371/journal.pone.0035009.t003

used for purification, amplification, and sequencing steps. One sequencing run was performed on the GS-FLX using Titanium chemistry. 454 Sequencing is based on sequencing-by-synthesis, addition of one nucleotide, or more, complementary to the template strand results in a chemiluminescent signal recorded by the CCD camera within the instrument. The signal strength is proportional to the number of nucleotides incorporated in a single nucleotide flow. All reagents and protocols used were from Roche 454 Life Sciences, USA.

Assembly and functional annotation

Pyrosequencing raw data, comprised of 975,190 reads, were processed with the Roche quality control pipeline using the default settings. *Seqclean* (http://compbio.dfci.harvard.edu/tgi/software/) software was used to screen for and remove normalization adaptor sequences, homopolymers and reads shorter than 40 bp prior to assembly. A total of 974,973 quality reads were subjected to MIRA, version 3.2.0 [87], to assemble the transcriptome. By default, MIRA takes into account only contigs with at least 2 reads. The other reads go into debris, which might include singletons, repeats, low complexity sequences and sequences shorter than 40 bp. NCBI *Blastclust* was used to group similar contigs into clusters (groups of transcripts from the same gene). Two sequences were grouped if at least 60% of the positions had at least 95% identity. The 51,265 contigs were grouped into a total of 29,679 clusters.

An iterative blast workflow was used to annotate the *R. philippinarum* contigs with at least 100 bp (49,847 contigs out of 51,265). Then, BLASTx [88] with a cut off value of 10e-3, was used to compare the *R. philippinarum* contigs with the NCBI Swissprot, the NCBI Metazoan Refseq, the NCBI nr and the UniprotKB/Trembl protein databases.

After annotation, Blast2GO software [89] was used to assign Gene Ontology terms [90] to the largest contig of a representative cluster (minimum of 100 bp). This strategy was used to avoid redundant results. Default values in Blast2GO were used to perform the analysis and ontology level 2 was selected to construct the level pie charts.

Comparative analysis

To make a comparison between *R. philippinarum* and other bivalve species, the nucleotide sequences and ESTs from *C. gigas*, *M. galloprovincialis*, *L. elliptica and B. azoricus* were obtained from GenBank and from dedicated databases, when available. http://metagenomics.anl.gov/?page = DownloadMetagenome&metagenome = 4442941.3, http://metagenomics.anl.gov/?page = DownloadMetagenome&metagenome = 4442947.3, http://metagenomics.anl.gov/?page = DownloadMetagenome&metagenome = 4442948.3, http://metagenomics.anl.gov/?page = DownloadMetagenome&metagenome = 4442949.3 for *M. galloprovincialis* [91], http://www.ncbi.nlm.nih.gov/sra?term = SRA011054 for *L. elliptica* [92] and http://transcriptomics.biocant.pt:8080/deepSeaVent/?rvn = 1 for *B. azoricus* [93]. Unique sequences from these databases (based on gi number) were used from each of the databases. These sequences were compared by BLASTn against the longest contig from each of 29,679 *R. philippinarum* clusters with a cut off e-value of 10e-05. Hits to *R. philippinarum* sequences were represented in a Venn diagram.

The comparison between our 454 results, the longest contig from each of 29,679 clusters, and the Milan *et al.* [16] transcriptome, contigs downloaded from RuphiBase (http://compgen.bio.unipd.it/ruphibase/query/), was made by BLASTn with a cut off e-value of 10e-05. Another analysis was carried out to compare just the longest contig from each of 2,005 clusters identified as immune-related and the Milan *et al.* contigs as well. The results were summarized in a table (Table 2). The percentage of coverage is the average % of query coverage by the best blast hit and the percentage of hits is the % of query with at least one hit in database, in parenthesis were added the total number of hits.

Identification of immune-related genes

All the contig annotations were revised based on an immunity and inflammation-related keyword list (i.e. apoptosis, bactericidal, C3, lectin, SOCS…) developed in our laboratory to select the candidate sequences putatively involved in immune response. The presence or absence of these words in the BLASTx hit descriptions was checked to identify putative immune-related contigs. The remaining non-selected contigs were revised using the GO terms at level 2, 3 and 4 assigned to each sequence after the annotation step that had a direct relationship with immunity. Selected contigs were checked again to eliminate non-immune ones and distributed into functional categories.

Immune-related genes were grouped in three reference immune pathways (complement cascade, TLR signaling pathway and apoptosis) to describe each route indicated by our pyrosequencing results.

Taxonomy analysis

To identify and classify the groups of organisms that had high similarity with our clam sequences, the Uniprot Taxonomy [94] was used except for the protozoa group. Because protozoa are a highly complex group, a specific taxonomy [95] was followed. Briefly, after the BLASTx annotation step all the hit descriptions included the species name (i.e. Homo sapiens) or a code (i.e. HUMAN) meaning that protein has been previously identified as belonging to that species. With such information sequences were classified in taxonomical groups and represented in pie charts.

Supporting Information

Table S1 List of contigs (e-value<10-3) of *Ruditapes philippinarum* including sequence, length, description (Hit description), accession number of description (Hit ACC), e-value obtained and database used for annotation (Blast).
(XLSX)

Author Contributions

Conceived and designed the experiments: BN AF. Performed the experiments: RM PB. Analyzed the data: BN AF RM PB. Contributed reagents/materials/analysis tools: SB BF JVP. Wrote the paper: RM.

References

1. Gestal C, Roch P, Renault T, Pallavicini A, Paillard C, et al. (2008) Study of diseases and the immune system of bivalves using molecular biology and genomics. Rev Fish Sci 16: 131–154.
2. Paillard C, Leroux F, Borrego JJ (2004) Bacterial disease in marine bivalves, Review of recent studies. Trends and evolution. Aquat Living Ress 17: 477–498.
3. Villalba A, Reece KS, Ordás MC, Casas SM, Figueras A (2004) Perkinsosis in molluscs: A review. Aquat Living Resour 17: 411–432.
4. Canesi L, Gallo G, Gavioli M, Pruzzo C (2002) Bacteria–hemocyte interactions and phagocytosis in bivalves. Microsc Res Technol 57: 469–476.
5. Olafsen JA (1995) Role of lectins (C-reactive protein) in defense of marine bivalves against bacteria. Adv Exp Med Biol 371A: 343–348.
6. Ordás MC, Novoa B, Figueras A (2000) Modulation of the chemiluminescence response of Mediterranean mussel (*Mytilus galloprovincialis*) haemocytes. Fish Shellfish Immunol 10: 611–622.
7. Ordás MC, Ordás A, Beloso C, Figueras A (2000) Immune parameters in carpet shell clams naturally infected with *Perkinsus atlanticus*. Fish Shellfish Immunol 10: 597–609.
8. Tafalla C, Gómez-León J, Novoa B, Figueras A (2003) Nitric oxide production by carpet shell clam (*Ruditapes decussatus*) hemocytes. Dev Comp Immunol 27: 197–205.
9. Fleury E, Huvet A, Lelong C, de Lorgeril J, Boulo V, et al. (2009) 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 10: 341.
10. Gueguen Y, Cadoret JP, Flament D, Barreau-Roumiguière C, Girardot AL, et al. (2003) Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*. Gene 303: 139–145.
11. Pallavicini A, Costa MM, Gestal C, Dreos R, Figueras A, et al. (2008) Sequence variability of myticins identified in haemocytes from mussels suggests ancient host-pathogen interactions. Dev Comp Immunol 32: 213–226.
12. Venier P, De Pittà C, Bernante F, Varotto L, De Nardi B, et al. (2009) MytiBase, a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics 10: 72.
13. Venier P, Varotto L, Rosani U, Millino C, Celegato B, et al. (2011) Insights into the innate immunity of the Mediterranean mussel *Mytilus galloprovincialis*. BMC Genomics 12: 69.
14. Wang A, Wang Y, Gu Z, Li S, Shi Y, et al. (2011) Development of Expressed Sequence Tags from the Pearl Oyster, *Pinctada martensii* Dunker. Mar Biotechnol 13: 275–283.
15. Moreira R, Balseiro P, Romero A, Dios S, Posada D, et al. (2012) Gene expression analysis of clams *Ruditapes philippinarum* and *Ruditapes decussatus* following bacterial infection yields molecular insights into pathogen resistance and immunity. Dev Comp Immunol 36: 140–149.
16. Milan M, Coppe A, Reinhardt R, Cancela LM, Leite RB, et al. (2011) Transcriptome sequencing and microarray development for the Manila clam, *Ruditapes philippinarum*: genomic tools for environmental monitoring. BMC Genomics 12: 234.
17. Ghiselli F, Milani L, Chang PL, Hedgecock D, Davis JP, et al. (2011) *De Novo* Assembly of the Manila Clam *Ruditapes philippinarum* Transcriptome Provides New Insights into Expression Bias, Mitochondrial Doubly Uniparental Inheritance and Sex Determination. Mol Biol EvolIn press.
18. Kang YS, Kim YM, Park KL, Cho SK, Choi KS, et al. (2006) Analysis of EST and lectin expression in hemocytes of Manila clams (*Ruditapes phylippinarum*) (Bivalvia, Mollusca) infected with *Perkinsus olseni*. Dev Comp Immunol 30: 1119–1131.
19. Tanguy A, Bierne N, Saavedra C, Pina B, Bachère E, et al. (2008) Increasing genomic information in bivalves through new EST collections in four species, development of new genetic markers for environmental studies and genome evolution. Gene 408: 27–36.
20. Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, et al. (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol Ecol 17: 1636–1647.
21. Metzker ML (2010) Sequencing technologies - the next generation. Nat Rev Genet 11: 31–46.
22. Huan P, Wang H, Liu B (2011) Transcriptomic Analysis of the Clam *Meretrix meretrix* on Different Larval Stages. Mar BiotechnolIn press.
23. Ruddell RG, Hoang-Le D, Barwood JM, Rutherford PS, Piva TJ, et al. (2009) Ferritin functions as a proinflammatory cytokine via iron-independent protein kinase C zeta/nuclear factor kappaB-regulated signaling in rat hepatic stellate cells. Hepatology 49: 887–900.
24. Zhao J, Qiu L, Ning X, Chen A, Wu H, et al. (2010) Cloning and characterization of an invertebrate type lysozyme from *Venerupis philippinarum*. Comp Biochem Physiol B Biochem Mol Biol 156: 56–60.
25. Kishore U, Gaboriaud C, Waters P, Shrive AK, Greenhough TJ, et al. (2004) C1q and tumor necrosis factor superfamily: modularity and versatility. Trends Immunol 25: 551–561.
26. Henderson NC, Sethi T (2009) The regulation of inflammation by galectin-3. Immunol Rev 230: 160–171.
27. Fujii N, Minetti CA, Nakhasi HL, Chen SW, Barbehenn E, et al. (1992) Isolation, cDNA cloning, and characterization of an 18-kDa hemagglutinin and amebocyte aggregation factor from *Limulus polyphemus*. J Biol Chem 267: 22452–22459.
28. Kadomatsu T, Tabata M, Oike Y (2011) Angiopoietin-like proteins: emerging targets for treatment of obesity and related metabolic diseases. FEBS J 278: 559–564.
29. Do Carmo S, Levros LC, Jr., Rassart E (2007) Modulation of apolipoprotein D expression and translocation under specific stress conditions. Biochim Biophys Acta 1773: 954–969.
30. Do Carmo S, Jacomy H, Talbot PJ, Rassart E (2008) Neuroprotective effect of apolipoprotein D against human coronavirus OC43-induced encephalitis in mice. J Neurosci 28: 10330–10338.
31. Rassart E, Bedirian A, Do Carmo S, Guinard O, Sirois J, et al. (2000) Apolipoprotein D. Biochim Biophys Acta 1482: 185–198.
32. Fleischer A, Ayllon V, Rebollo A (2002) ITM2BS regulates apoptosis by inducing loss of mitochondrial membrane potential. Eur J Immunol 32: 3498–3505.
33. Bai X, Rivera-Vega L, Mamidala P, Bonello P, Herms DA, et al. (2011) Transcriptomic signatures of ash (*Fraxinus spp.*) phloem. PLoS One 6: e16368.
34. Bai X, Mamidala P, Rajarapu SP, Jones SC, Mittapalli O (2011) Transcriptomics of the bed bug (*Cimex lectularius*). PLoS One 6: e16336.
35. Dunkelberger JR, Song WC (2010) Complement and its role in innate and adaptive immune responses. Cell Res 20: 34–50.
36. Jenny MJ, Ringwood AH, Lacy ER, Lewitus AJ, Kempton JW, et al. (2002) Potential indicators of stress response identified by expressed sequence tag

analysis of hemocytes and embryos from the American oyster, *Crassostrea virginica*. Mar Biotechnol 4: 81–93.
37. Gourdine JP, Smith-Ravin EJ (2007) Analysis of a cDNA-derived sequence of a novel mannose-binding lectin, codakine, from the tropical clam *Codakia orbicularis*. Fish Shellfish Immunol 22: 498–509.
38. Zhang H, Song L, Li C, Zhao J, Wang H, et al. (2008) A novel C1q-domain-containing protein from Zhikong scallop *Chlamys farreri* with lipopolysaccharide binding activity. Fish Shellfish Immunol 25: 281–289.
39. Gerdol M, Manfrin C, De Moro G, Figueras A, Novoa B, et al. (2011) The C1q domain containing proteins of the Mediterranean mussel *Mytilus galloprovincialis*: a widespread and diverse family of immune-related molecules. Dev Comp Immunol 35: 635–643.
40. Gestal C, Pallavicini A, Venier P, Novoa B, Figueras A (2010) MgC1q, a novel C1q-domain-containing protein involved in the immune response of *Mytilus galloprovincialis*. Dev Comp Immunol 34: 926–934.
41. Prado-Alvarez M, Gestal C, Novoa B, Figueras A (2009) Differentially expressed genes of the carpet shell clam *Ruditapes decussatus* against *Perkinsus olseni*. Fish Shellfish Immunol 26: 72–83.
42. Prado-Alvarez M, Rotllant J, Gestal C, Novoa B, Figueras A (2009) Characterization of a C3 and a factor B-like in the carpet-shell clam, *Ruditapes decussatus*. Fish Shellfish Immunol 26: 305–315.
43. Vasta GR, Ahmed H, Odom EW (2004) Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. Curr Opin Struct Biol 14: 617–630.
44. Bulgakov AA, Park KI, Choi KS, Lim HK, Cho M (2004) Purification and characterisation of a lectin isolated from the Manila clam *Ruditapes philippinarum* in Korea. Fish Shellfish Immunol 16: 487–499.
45. Kim JY, Adhya M, Cho SK, Choi KS, Cho M (2008) Characterization, tissue expression, and immunohistochemical localization of MCL3, a C-type lectin produced by *Perkinsus olseni*-infected Manila clams (*Ruditapes philippinarum*). Fish Shellfish Immunol 25: 598–603.
46. Kim JY, Kim YM, Cho SK, Choi KS, Cho M (2008) Noble tandem-repeat galectin of Manila clam *Ruditapes philippinarum* is induced upon infection with the protozoan parasite *Perkinsus olseni*. Dev Comp Immunol 32: 1131–1141.
47. Kim YM, Park KI, Choi KS, Alvarez RA, Cummings RD, et al. (2006) Lectin from the Manila clam *Ruditapes philippinarum* is induced upon infection with the protozoan parasite *Perkinsus olseni*. J Biol Chem 281: 26854–26864.
48. Su J, Song L, Xu W, Wu L, Lia H, et al. (2004) cDNA cloning and mRNA expression of the lipopolysaccharide- and beta-1,3-glucan-binding protein gene from scallop *Chlamys farreri*. Aquaculture 239: 69–80.
49. Su J, Ni D, Song L, Zhao J, Qiu L (2007) Molecular cloning and characterization of a short type peptidoglycan recognition protein (CfPGRP-S1) cDNA from Zhikong scallop *Chlamys farreri*. Fish Shellfish Immunol 23: 646–656.
50. Ni D, Song L, Wu L, Chang Y, Yu Y, et al. (2007) Molecular cloning and mRNA expression of peptidoglycan recognition protein (PGRP) gene in bay scallop (*Argopecten irradians*, Lamarck 1819). Dev Comp Immunol 31: 548–558.
51. Itoh N, Takahashi KG (2008) Distribution of multiple peptidoglycan recognition proteins in the tissues of Pacific oyster, *Crassostrea gigas*. Comp Biochem Physiol B Biochem Mol Biol 150: 409–417.
52. Qiu L, Song L, Xu W, Ni D, Yu Y (2007) Molecular cloning and expression of a Toll receptor gene homologue from Zhikong Scallop, *Chlamys farreri*. Fish Shellfish Immunol 22: 451–466.
53. Takeuchi O, Akira S (2010) Pattern recognition receptors and inflammation. Cell 140: 805–820.
54. Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, et al. (2005) The evolution of vertebrate Toll-like receptors. Proc Natl Acad Sci USA 102: 9577–9582.
55. Zhang D, Zhang G, Hayden MS, Greenblatt MB, Bussey C, et al. (2004) A toll-like receptor that prevents infection by uropathogenic bacteria. Science 303: 1522–1526.
56. Kim YM, Brinkmann MM, Paquet ME, Ploegh HL (2008) UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. Nature 452: 234–238.
57. Montagnani C, Le Roux F, Berthe F, Escoubas JM (2001) Cg-TIMP, an inducible tissue inhibitor of metalloproteinase from the Pacific oyster *Crassostrea gigas* with a potential role in wound healing and defense mechanisms. FEBS Lett 500: 64–70.
58. Zhu L, Song L, Chang Y, Xu W, Wu L (2006) Molecular cloning, characterization and expression of a novel serine proteinase inhibitor gene in bay scallops (*Argopecten irradians*, Lamarck 1819). Fish Shellfish Immunol 20: 320–331.
59. Wang B, Zhao J, Song L, Zhang H, Wang L, et al. (2008) Molecular cloning and expression of a novel Kazal-type serine proteinase inhibitor gene from Zhikong scallop *Chlamys farreri*, and the inhibitory activity of its recombinant domain. Fish Shellfish Immunol 24: 629–637.
60. Xue QG, Waldrop GL, Schey KL, Itoh N, Ogawa M, et al. (2006) A novel slow-tight binding serine protease inhibitor from eastern oyster (*Crassostrea virginica*) plasma inhibits perkinsin, the major extracellular protease of the oyster protozoan parasite *Perkinsus marinus*. Comp Biochem Physiol B Biochem Mol Biol 145: 16–26.
61. Xue Q, Itoh N, Schey KL, Cooper RK, La Peyre JF (2009) Evidence indicating the existence of a novel family of serine protease inhibitors that may be involved in marine invertebrate immunity. Fish Shellfish Immunol 27: 250–259.
62. La Peyre JF, Xue QG, Itoh N, Li Y, Cooper RK (2010) Serine protease inhibitor cvSI-1 potential role in the eastern oyster host defense against the protozoan parasite *Perkinsus marinus*. Dev Comp Immunol 34: 84–92.
63. Brogden KA (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat Rev Microbiol 3: 238–250.
64. Charlet M, Chernysh S, Philippe H, Hetru C, Hoffmann JA, et al. (1996) Innate immunity. Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, *Mytilus edulis*. J Biol Chem 271: 21808–21813.
65. Hubert F, Noel T, Roch P (1996) A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). Eur J Biochem 240: 302–306.
66. Costa MM, Dios S, Alonso-Gutierrez J, Romero A, Novoa B, et al. (2009) Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev Comp Immunol 33: 162–170.
67. Balseiro P, Falcó A, Romero A, Dios S, Martínez-López A, et al. (2011) *Mytilus galloprovincialis* Myticin C: A Chemotactic Molecule with Antiviral Activity and Immunoregulatory Properties. PLoS ONE 6: e23140.
68. Gestal C, Costa M, Figueras A, Novoa B (2007) 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 406: 134–143.
69. Zhao J, Li C, Chen A, Li L, Su X, et al. (2010) Molecular characterization of a novel big defensin from clam *Venerupis philippinarum*. PLoS One 5: e13480.
70. De Maio A (1999) Heat shock proteins: facts, thoughts, and dreams. Shock 11: 1–12.
71. Parcellier A, Gurbuxani S, Schmitt E, Solary E, Garrido C (2003) Heat shock proteins, cellular chaperones that modulate mitochondrial cell death pathways. Biochem Biophys Res Commun 304: 505–512.
72. Lakhal F, Bury-Moné S, Nomane Y, Le Goïc N, Paillard C, et al. (2008) DjlA, a membrane-anchored DnaJ-like protein, is required for cytotoxicity of clam pathogen *Vibrio tapetis* to hemocytes. Appl Environ Microbiol 74: 5750–5758.
73. Li C, Li L, Liu F, Ning X, Chen A, et al. (2011) Alternation of *Venerupis philippinarum* Hsp40 gene expression in response to pathogen challenge and heavy metal exposure. Fish Shellfish Immunol 30: 447–450.
74. Li C, Wang L, Ning X, Chen A, Zhang L, et al. (2011) Identification of two small heat shock proteins with different response profile to cadmium and pathogen stresses in *Venerupis philippinarum*. Cell Stress Chaperones 15: 897–904.
75. Zhang D, Jiang J, Jiang S, Ma J, Su T, et al. (2009) Molecular characterization and expression analysis of a putative LPS-induced TNF-alpha factor (LITAF) from pearl oyster *Pinctada fucata*. Fish Shellfish Immunol 27: 391–396.
76. Park EM, Kim YO, Nam BH, Kong HJ, Kim WJ, et al. (2008) Cloning, characterization and expression analysis of the gene for a putative lipopolysaccharide-induced TNF-alpha factor of the Pacific oyster, *Crassostrea gigas*. Fish Shellfish Immunol 24: 11–17.
77. Yu Y, Qiu L, Song L, Zhao J, Ni D, et al. (2007) Molecular cloning and characterization of a putative lipopolysaccharide-induced TNF-alpha factor (LITAF) gene homologue from Zhikong scallop *Chlamys farreri*. Fish Shellfish Immunol 23: 419–429.
78. Li L, Qiu L, Song L, Song X, Zhao J, et al. (2009) First molluscan TNFR homologue in Zhikong scallop: molecular characterization and expression analysis. Fish Shellfish Immunol 27: 625–632.
79. Qiu L, Song L, Yu Y, Zhao J, Wang L, et al. (2009) Identification and expression of TRAF6 (TNF receptor-associated factor 6) gene in Zhikong scallop *Chlamys farreri*. Fish Shellfish Immunol 26: 359–367.
80. Beaz-Hidalgo R, Balboa S, Romalde JL, Figueras MJ (2010) Diversity and pathogenecity of *Vibrio* species in cultured bivalve molluscs. Environ Microbiol Rep 2: 34–43.
81. Crook NE, Clem RJ, Miller LK (1993) An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. J Virol 67: 2168–2174.
82. Batista FM, Arzul I, Pepin JF, Ruano F, Friedman CS, et al. (2007) Detection of ostreid herpesvirus 1 DNA by PCR in bivalve molluscs: a critical review. J Virol Methods 139: 1–11.
83. Bower SM, McGladdery SE (2003) Synopsis of infectious diseases and parasites of commercially exploited shellfish. http://www-scipacdfo-mpogcca/shelldis/title_ehtm.
84. Davis HC, Loosanoff VL, Weston WH, Martin C (1954) A fungus disease in clam and oyster larvae. Science 120: 36–38.
85. Shagin DA, Rebrikov DV, Kozhemyako VB, Altshuler IM, Shcheglov AS, et al. (2002) A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. Genome Res 12: 1935–1942.
86. Zhulidov PA, Bogdanova EA, Shcheglov AS, Vagner LL, Khaspekov GL, et al. (2004) Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acids Res 32: e37.
87. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WE, et al. (2004) Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. Genome Res 14: 1147–1159.
88. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410.
89. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, et al. (2005) Blast2GO, a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.
90. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology, tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25–29.

91. Craft JA, Gilbert JA, Temperton B, Dempsey KE, Ashelford K, et al. (2010) Pyrosequencing of *Mytilus galloprovincialis* cDNAs: tissue-specific expression patterns. PLoS One 5: e8875.
92. Clark MS, Thorne MA, Vieira FA, Cardoso JC, Power DM, et al. (2010) Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyrosequencing. BMC Genomics 11: 362.
93. Bettencourt R, Pinheiro M, Egas C, Gomes P, Afonso M, et al. (2010) High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*. BMC Genomics 11: 559.
94. Phan IQ, Pilbout SF, Fleischmann W, Bairoch A (2003) NEWT, a new taxonomy portal. Nucleic Acids Res 31: 3822–3823.
95. Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, et al. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52: 399–451.

Supporting Information

Supporting information associated with this article can be found at:
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0035009

GENE EXPRESSION PROFILE ANALYSIS OF MANILA CLAM (*RUDITAPES PHILIPPINARUM*) HEMOCYTES AFTER A *VIBRIO ALGINOLYTICUS* CHALLENGE USING AN IMMUNE-ENRICHED OLIGO-MICROARRAY

Moreira et al. BMC Genomics 2014, 15:267
http://www.biomedcentral.com/1471-2164/15/267

BMC Genomics

RESEARCH ARTICLE Open Access

Gene expression profile analysis of Manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibrio alginolyticus* challenge using an immune-enriched oligo-microarray

Rebeca Moreira[1], Massimo Milan[2], Pablo Balseiro[1], Alejandro Romero[1], Massimiliano Babbucci[2], Antonio Figueras[1], Luca Bargelloni[2] and Beatriz Novoa[1*]

Abstract

Background: The Manila clam (*Ruditapes philippinarum*) is a cultured bivalve with worldwide commercial importance, and diseases cause high economic losses. For this reason, interest in the immune genes in this species has recently increased. The present work describes the construction of the first *R. philippinarum* microarray containing immune-related hemocyte sequences and its application to study the gene transcription profiles of hemocytes from clams infected with *V. alginolyticus* through a time course.

Results: The complete set of sequences from *R. philippinarum* available in the public databases and the hemocyte sequences enriched in immune transcripts were assembled successfully. A total of 12,156 annotated sequences were used to construct the 8 ×15 k oligo-microarray. The microarray experiments yielded a total of 579 differentially expressed transcripts. Using the gene expression results, the associated Gene Ontology terms and the enrichment analysis, we found different response mechanisms throughout the experiment. Genes related to signaling, transcription and apoptosis, such as IL-17D, NF-κB or calmodulin, were typically expressed as early as 3 hours post-challenge (hpc), while characteristic immune genes, such as PGRPs, FREPs and defense proteins appeared later at 8 hpc. This immune-triggering response could have affected a high number of processes that seemed to be activated 24 hpc to overcome the *Vibrio* challenge, including the expression of many cytoskeleton molecules, which is indicative of the active movement of hemocytes. In fact functional studies showed an increment in apoptosis, necrosis or cell migration after the infection. Finally, 72 hpc, activity returned to normal levels, and more than 50% of the genes were downregulated in a negative feedback of all of the previously active processes.

Conclusions: Using a new version of the *R. philippinarum* oligo-microarray, a putative timing for the response against a *Vibrio* infection was established. The key point to overcome the challenge seemed to be 8 hours after the challenge, when we detected immune functions that could lead to the destruction of the pathogen and the activation of a wide variety of processes related to homeostasis and defense. These results highlight the importance of a fast response in bivalves and the effectiveness of their innate immune system.

Keywords: *Ruditapes philippinarum*, *Vibrio alginolyticus*, Hemocytes, Oligo-microarray, Gene ontology, Blast2GO, Immune response

* Correspondence: beatriznovoa@iim.csic.es
[1]Instituto de Investigaciones Marinas (IIM), Consejo Superior de Investigaciones Científicas (CSIC), Eduardo Cabello 6, 36208 Vigo, Spain
Full list of author information is available at the end of the article

Background

The Manila clam (*Ruditapes philippinarum*) is a cultured bivalve species with high worldwide commercial importance and value. The culture of these clams has been increasing in recent years, especially in Europe and Asia. Therefore, diseases can result in significant economic losses. The majority of diseases in clams are associated with *Vibrio* [1,2] and *Perkinsus* [3,4] species. Diseases can affect not only the development and survivorship of these organisms [5,6] but also the quality and price of the product. Therefore, there has been a growing interest in the study of the molecular biology of the defense mechanisms in bivalves in the last decade. Although molluscs lack a specific immune system, the innate response involving circulating hemocytes and a large variety of molecular effectors seems to be an efficient defense method for responding to external aggressions by detecting the molecular signatures of an infection [7-11]. Hemocytes are the main line of cellular defense against invading pathogens in molluscs [12], and the immune genes expressed in bivalve hemocytes are of great interest to researchers [13-19]. In recent years, a serious effort has also been made to increase the number of *R. philippinarum* sequences in public databases. Before 2011, less than 6,000 nucleotide sequences were available for this species. Our two groups released 457,717 and 975,190 raw reads from adult/larval tissues and hemocytes, respectively [20,21]. Concurrently, studies by Ghiselli *et al.* on gonad tissue yielded approximately 90 million raw sequences [22].

Oligo-microarrays are a sensitive and reproducible high-throughput technology for analyzing the gene expression of thousands of genes simultaneously [23]. This platform has been used by scientists to study the expression profile in many species from yeast to human [24,25]. Microarrays have also been applied to bivalves for different purposes [26-28]. To our knowledge, a toxicogenomics study in the gill and the digestive gland [20] and a recent study of the response to brown ring disease, [29] are the only works to date to use an oligo-microarray in *R. philippinarum*.

The main objective of this work was to analyze the response of the Manila clam against strain TA15 of *V. alginolyticus* through a time course. *V. alginolyticus* was previously reported to produce important mortality in clam larvae [6]. For this study, a new version of Manila clam DNA microarray including the sequences of thousands of hemocyte-exclusive genes [21] was designed and developed.

Results and discussion

Assembly and annotation

A summary of the sequence origin, assembly and annotation results is shown in Table 1. From the total 1,438,665 sequences from *R. philippinarum,* the Newbler software package (GS *De Novo* Assembler v2.6, Roche) was able to assemble 88.24% of the raw sequences, and 11.76% of the sequences (169,223) remained as singletons. The assembly resulted in 26,708 isotigs grouped into 15,175 isogroups and 156 contigs.

Table 1 Description of the microarray design process

		n° sequences
Sequence origin		
Sanger (Milan *et al.* [20])		5,758
454 tissues (Milan *et al.* [20])		457,717
454 hemocytes (Moreira *et al.* [21])		975,190
	subTOTAL	1,438,665
NCBI		2,050
	TOTAL	1,440,715
Assembly		
Singletons		169,223
Singletons phred Q > 20		16,495
Isotigs		26,708
Isogroups		15,175
Contigs		156
Annotation		
Singletons phred Q > 20 and NCBI		5,914
Contigs and longest isotig of each isogroup		6,242
Frame +		7,827
Frame -		2,761
Ambiguous frame		1,568
TOTAL *R. philippinarum* successfully designed probes		13,671

The longest isotig of each isogroup, the contigs, the singletons with more than 200 bp of continuous sequence with a Phred Q > 20 (16,495) and the *R. philippinarum* ESTs in the NCBI database (2,050) were then considered for the annotation. The putative identities of these sequences were obtained by running BlastX and BlastN similarity searches in 48 different protein and nucleotide databases. Additional file 1: Table S1 shows the percentage value of annotation success of each database. The protein databases showed a higher average percentage of matches (18.84%) than the nucleotide ones (5.3%), presumably due to the degeneration of the genetic code. Furthermore, the species databases yielded much lesser annotation percentage than the general ones (NCBI) with the exception of certain databases, probably because of the huge amount of sequences available, in the case of *H. sapiens,* or the higher phylogenetic similarity, with *S. purpuratus* and *L. gigantea*. Thus, we found it is worth the use of general protein databases and the specific ones previously named since only 1.5% of the contigs/isotigs and the 4.1% of the singletons were annotated with the remaining specific databases. A total of 12,156 successfully annotated transcripts were considered for the *R. philippinarum* DNA

microarray design (see Methods). As most sequence reads were obtained from a non-directional cDNA library, sense strand orientation was inferred putatively from the homologous protein sequences of other species. One probe for annotated transcripts with known orientation was designed to construct a high-density oligo-DNA microarray, while two probes with both orientations were designed for contigs with ambiguous orientation (see Table 1). A total of 13,671 probes representing 12,108 unique transcripts were created using the Agilent eArray interface (https://earray.chem.agilent.com/earray/).

Microarray hybridization, robustness and validation

A total of 36 microarray experiments were performed. The upper and lower fluorescence values were erased from the raw data (20 - 90th percentile) in all of the experiments, and only robust fluorescence values were used to analyze the expression and function of the results.

Quantitative PCR (qPCR) is commonly used to confirm the results obtained from microarray analysis, and although microarray and qPCR data could disagree, it is known that qPCR using SYBR Green is useful to validate Agilent inkjet-printed 60-mer oligo-microarrays [30]. Specific primers were designed to perform a qPCR for four selected genes with cDNAs synthesized from the same RNAs used for the microarray hybridization. The expression patterns identified for these genes by the array and by qPCR showed similar profiles (Figure 1). Significant differences in expression compared with the controls (t-test, $p < 0.05$) were observed in the qPCR results and matched the microarray results in most of the cases (IL-17D 3 h, LTBP-4 8 h and 24 h) but showed some new significant results in one case (IL-17D 72 h). The direction of the regulation was always comparable in both techniques.

The four genes selected for qPCR were chosen based on their relevance in the immune system. The complement component C3 is the central component of the complement system, whose functions are to distinguish and eliminate pathogens and trigger an inflammatory response [31]. The complement system has been characterized in the carpet shell clam *R. decussatus* and was found to be inhibited early after a *Vibrio* challenge, which concurs with our results [32]. The interleukin (IL) 17D belongs to a particular family with no sequence similarity to any other known cytokines. In humans, IL-17D regulates cytokine production in endothelial cells and has an inhibitory effect on hematopoiesis [33]. Previous works in vertebrates and invertebrates have analyzed IL-17 expression after a bacterial challenge [34,35]. These studies showed a general increase in IL-17 expression until a maximum level was reached between 6 and 24 hours post-challenge in both vertebrates and invertebrates, with the exception of the turbot intestine. In this fish species, the expression profile of IL-17 is more similar to our results, with an initial increase and subsequent

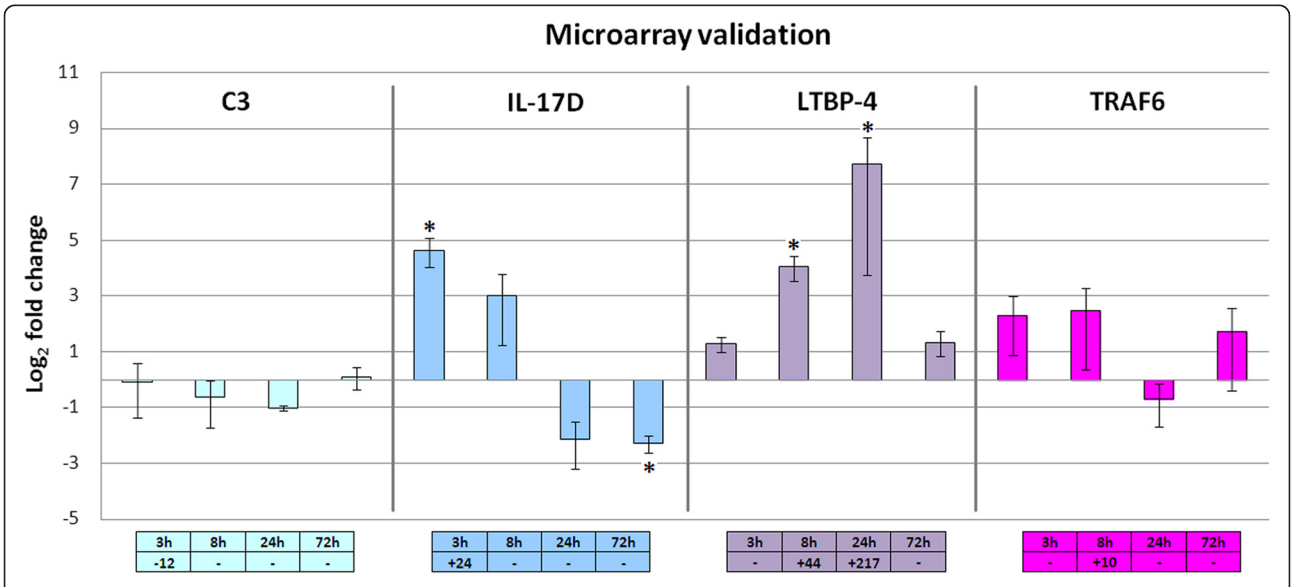


Figure 1 Validation of the microarray. qPCR expression of selected genes: C3 (Complement component C3), IL-17D (Interleukin 17 D), LTBP-4 (Latent Transforming Growth Factor-β Binding Protein 4) and TRAF6 (tumor necrosis factor receptor-associated factor 6). Data are log2 transformed to facilitate the illustration of downregulated genes. All qPCR reactions were performed as technical triplicates and the expression level of analyzed genes was normalized using the 18S rRNA. Fold change units were then calculated by dividing the normalized expression values of hemocytes from infected clams by the normalized expression values of the controls. Each bar represents the mean and standard error of the biological replicates. Asterisks highlight the significant fold changes (t-test, $p < 0.05$). Microarray results (t-test, $p < 0.01$) are indicated in the boxes below the graph.

decrease and inhibition. The latent-transforming growth factor beta-binding protein 4 (LTBP-4) assists in TGF-β folding, secretion and activation by mediating its matrix targeting [36] and is also involved in cell adhesion and migration [37]. The TNF receptor-associated factor 6 (TRAF6) is a key signaling adaptor molecule common to the TNFR superfamily and the IL-1R/TLR family that leads to the activation of the nuclear factor kappa-B (NF-κB) and AP-1 transcription factors. TRAF6 has been identified in other bivalves, such as the Zhikong scallop [38], and the expression of TRAF6 in hemocytes after a bacterial challenge (peptidoglycan) *in vitro* indicates an initial significant inhibition of TRAF6 transcription (3 h) followed by a subsequent recovery of the basal levels (6 h). This result does not coincide with our results in either expression or in timing, and this discrepancy could be a result of the different experimental designs. Nevertheless a late inhibition (24 h) and recuperation of the basal expression (72 h) was observed in *R. philippinarum*.

Gene expression profile after TA15 stimulation

To select genes that were significantly regulated by the *V. alginolyticus* challenge, we performed a t-test ($p < 0.01$) to find genes that were significantly different between the control and infected samples and an ANOVA ($p < 0.05$) to analyze the effect of the challenge in the whole dataset.

The t-test yielded the identification of 579 differentially expressed genes between the control groups and the *Vibrio*-infected groups. The highest number of differentially expressed genes (209) was found 8 hours after the challenge, while 125, 152 and 93 significantly differentially expressed genes were found 3, 24 and 72 hours after challenge, respectively. The expression pattern is an increase in the number of differentially expressed genes to 8 hours and a progressive decrease until reaching the minimum at 72 hours after the challenge. Figure 2 summarizes the results of the t-test and describes the number of regulated genes in fold change groups.

A Venn diagram illustrates the expression of the exclusive and common genes between the different sampling times in response to *V. alginolyticus* (Figure 3). This Venn diagram shows that the general transcription pattern of these genes is on/off, which implies that the majority of the genes were regulated at a single point (on) and only transcribed at basal levels the rest of the time (off). The genes expressed at several time points were scarce (21 out of 579, 3.63%), and only the ras-related c3 botulinum toxin substrate 3 (RAC3), a small GTPase of the Rho family implicated in cell differentiation, migration and apoptosis [39], was expressed in three consecutive sampling points (3, 8 and 24 hours post-challenge). In Figure 3, the coincident genes detected by the ANOVA and the t-test are also shown. The ANOVA analysis resulted in 15 regulated genes, and interestingly, 13 of these genes were in the group of 8 hours post-challenge and 8 genes were exclusive to this sampling point (for a description of these genes, see the ANOVA section).

The sequences of the regulated genes in each sampling time (t-test, $p < 0.01$) were analyzed in a Gene Ontology (GO) approach to cluster them into groups depending on the biological process in which the genes were involved and the time of maximum expression (Figure 4). Moreover, Table 2 shows the top 25 expressed genes in each sampling point. With this information, it is possible to infer the timing of the *R. philippinarum* response to a *V. alginolyticus* challenge. We have to take into account that the cumulative mortality rate at the end of the experiment, 72 hours, was 44%. The high mortality assures an infection but at the same time it is not too high allowing the clams to be able to fight against *Vibrio*.

In the first sampling point, 3 hours post-challenge, the results of the GO term analysis (Figure 4A) showed that most of the genes that were differentially expressed belonged to the cellular process category that involves many different functions. More specifically, one of the principal processes with a maximum expression at this sampling point was related to cellular component organization and biogenesis, leading to the assembly, disassembly or arrangement of the constituent macromolecules of the cell. Signaling and death processes were also present to trigger the response to the pathogen and eliminate compromised cells after an infection. Following the same line, the main functions of the top 25 hemocyte genes in the first sampling time were in accordance with the GO results. The most significantly upregulated gene 3 hpc (Table 2) is related to the cytoskeleton (dynein light chain) and could be involved in cellular spreading, chemotaxis and adhesion (RAC3, perlucin-like protein). However, the signaling and transcription processes (LTBP-4, phosphoinositide 3-kinase adapter protein 1, adp-ribosyl cyclase, HMG box transcription factor Sox17) could lead to a rapid triggering of several mechanisms, including the calcium-dependent processes that can lead to apoptosis via calmodulin activation [40]. Another upregulated gene was prothrombin, which is related to coagulation, but strikingly, there was an absence of the inflammatory response at the transcription level, as demonstrated by the strong inhibition of the complement factor C3.

In Figure 4B, 8 hours after challenge, the most represented functions are those related to response to stimulus and immune system. These processes initiate the response to the pathogen after the initial triggering of the signaling pathways. The GO term multi-organism process, with genes such as TRAF6, is also closely related because an interaction between clams and *V. alginolyticus* is implied. These results coincide with the information obtained in the top-list where genes involved in defense and wound

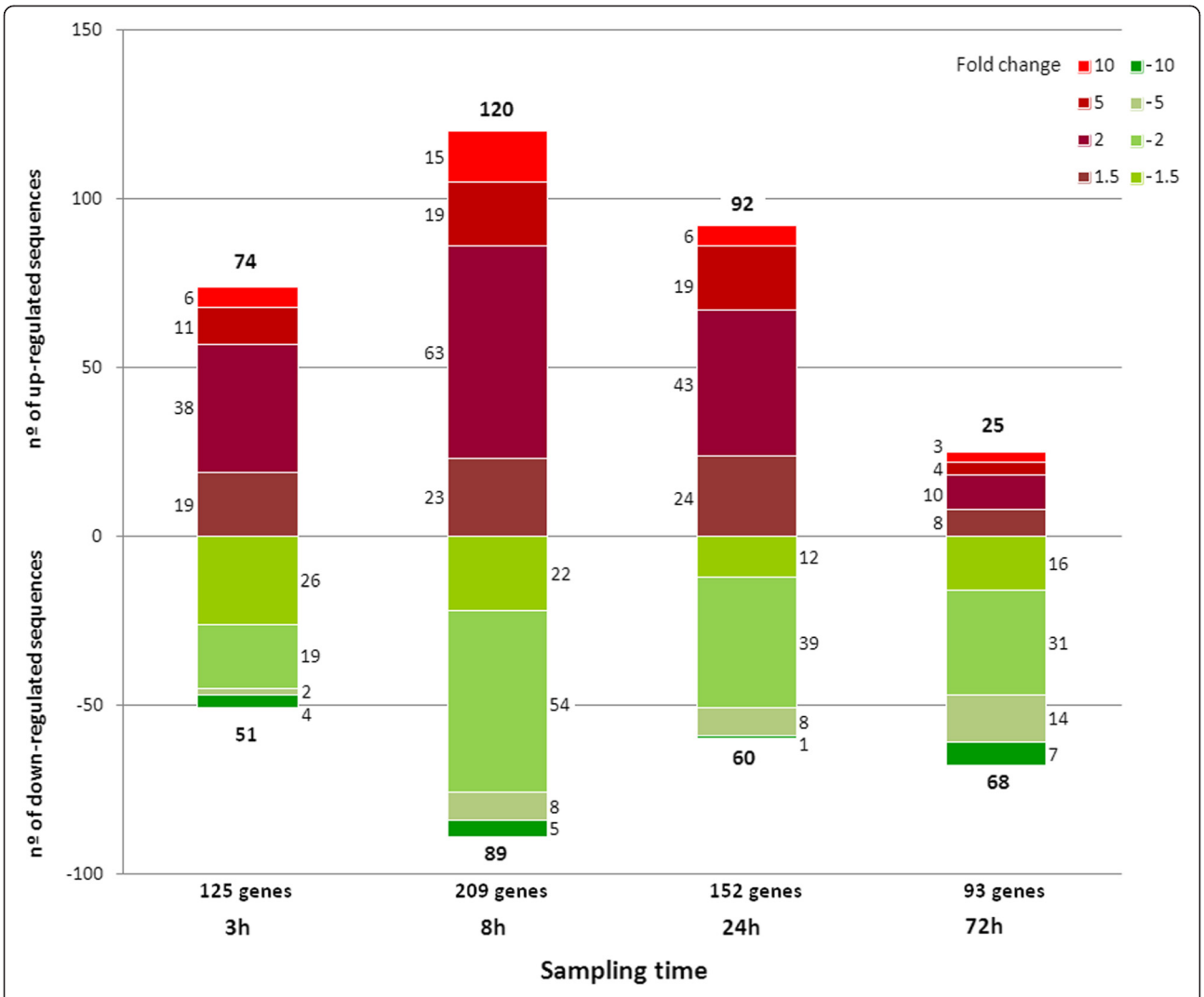


Figure 2 Distribution of regulated genes through time. Statistically significant gene modulation is subdivided according to intensity (fold change) and sense (up and down-regulation).

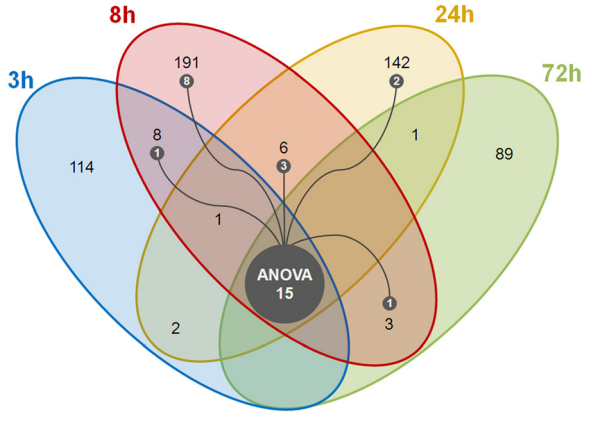


Figure 3 Venn diagram showing the common regulated genes among the four sampling times.

healing (pathogenesis-related protein 5, defense protein precursor and pro-neuregulin-2 membrane bound isoform) or metabolism (udp-glucose 4-epimerase) are included. Several processes were maintained from 3 hours to 8 hours, such as transcription (NF-κB p105 subunit) or recognition (peptidoglycan-recognition protein sb1 precursor (PGRP) and fibrinogen-related protein (FREP)). The expression of FREPs agrees with results reported in the mussel after a bacterial infection in which the maximum expression was obtained after 6 hours post-infection [19]. This result means that while the signaling is maintained because of the pathogen adhesion and recognition, the immune response seems to be initiated. The three immune-related molecules with C1q domains in the list were found to be downregulated. This effect could occur because C1q up-regulation is a very fast mechanism in molluscs as suggested by Gestal *et al.* [41].

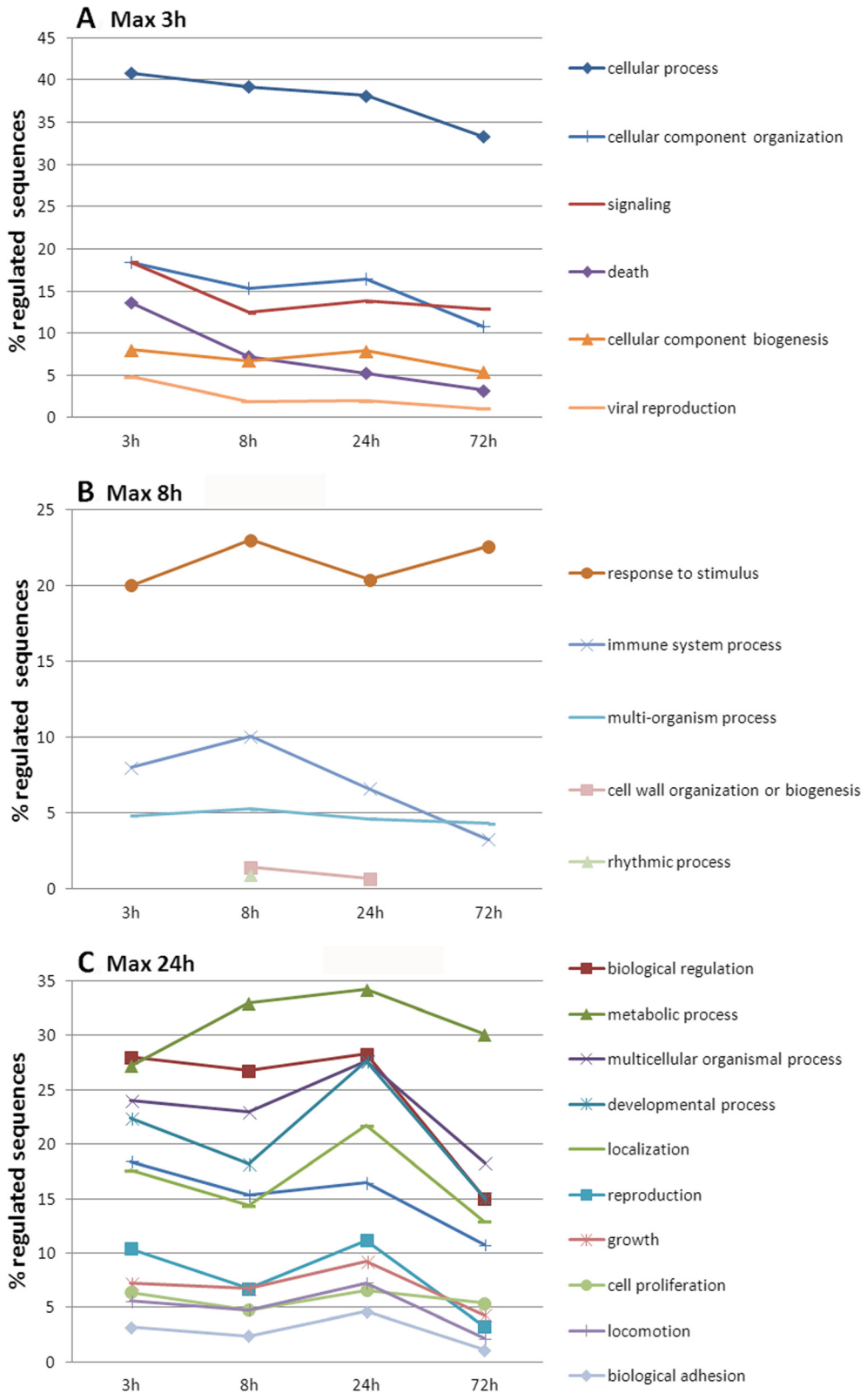


Figure 4 Biological process GO terms change in regulated genes through time. Numbers in abscissa refer to the percentage of GO term hits of the total of regulated sequences. GO terms with maximum representation in each sampling point are separately represented. **A**: 3 h, **B**: 8 h and **C**: 24 h. No GO processes exhibited a maximum representation 72 hours post-challenge.

Table 2 Top 25 expressed genes along the time course

Sequence description	3 hpc FC	Sequence description	24 hpc FC
Dynein light chain flagellar outer arm	48.91	Mariner mos1 transposase	13.2
Interleukin-17D	23.7	pdz and lim domain 5	6.31
Ras-related c3 botulinum toxin substrate 3	11.05	Zygote arrest protein 1	5.62
Barrier-to-autointegration factor	10.81	Monomeric sarcosine oxidase	5.57
Prothrombin	9.93	Cytochrome c-type heme lyase	5.45
Perlucin	7.94	Titin	5.27
Cubilin	6.23	Twitchin	4.77
Ets-related transcription factor elf-5	6.19	Pin2 terf1-interacting telomerase inhibitor 1	4.43
Pre-mrna-splicing factor cwc24	5.28	Leukocyte receptor cluster member 1 homolog	4.34
Adp-ribosyl cyclase	5	SART-3	4.19
Insoluble matrix shell protein 5	4.81	Myosin heavy striated muscle	4.03
MAPKKK1	3.92	Galaxin	3.45
Heat shock protein 70	3.91	Thap domain-containing protein 9	3.23
40s ribosomal protein s12	3.59	Calponin 3	3.09
HMG box transcription factor Sox17	3.57	Microtubule-associated protein futsch	2.94
LTBP-4	3.56	Nuclear speckle splicing regulatory protein 1	2.94
Histone acetyltransferase kat6a	3.54	Staphylococcal nuclease dom-cont protein 1	2.92
Calmodulin	3.46	Peroxisomal sarcosine oxidase	**−6.9**
Nucleolar complex protein 4 homolog	3.33	Mosc domain-containing protein mitochondrial	**−4.03**
Phosphoinositide 3-kinase adapter protein 1	3.27	Aldose 1-epimerase	**−3.88**
Cytochrome p450 2d6	3.21	Nuclease harbi1	**−3.8**
Electron transfer flavoprotein	**−67.28**	Lysocardiolipin acyltransferase 1	**−3.34**
Complement component c3	**−12.03**	Nuclear pore complex protein nup107	**−3.17**
Chaperonin	**−6.62**	Hamartin	**−3.09**
Collagen alpha-1 chain precursor	**−3.17**	Selectin	**−2.95**
Sequence description	**8 hpc FC**	**Sequence description**	**72 hpc FC**
Ganglioside gm2 activator	25.86	Mantle gene 6	52.66
Udp-glucose 4-epimerase	17.86	Serum amyloid a-3 protein precursor	19.53
Probable cubilin precursor	13.74	Elongation factor 1-gamma	7.95
Ets domain-containing protein elk-3	10.09	Biotin carboxylase	7.14
tnf receptor-associated factor 6	9.69	Kazal-type proteinase inhibitor protein	6.96
Nuclear factor nf-kappa-b p105 subunit	9.49	Abc transporter related protein	5.55
n-acylethanolamine-hydrolyzing acid amidase	9.44	Fizzy-related protein homolog	3.36
Pathogenesis-related protein 5	8.81	Nadh dehydrogenase 1 alpha subcomplex	2.47
Transmembrane protein 205	7.14	Sarcoplasmic calcium-binding protein	**−25.5**
Peptidyl-prolyl cis-trans isomerase b	6.64	High-affinity Na-dependent carnitine cotransporter	**−17.69**
Pro-neuregulin-2 membrane-bound isoform	6.42	Aragonite-binding protein	**−12.9**
Defense protein	5.56	Cytidine deaminase	**−12.29**
Peptidoglycan-recognition protein	4.94	Cat eye syndrome critical region protein 5	**−9.37**
Ficolin-2	4.85	Aldehyde cytosolic 2	**−8.3**
Cerebellin-3	**−46.98**	Lysozyme	**−7.67**
17-beta-hydroxysteroid dehydrogenase 14	**−18.9**	Glutamine synthetase/glutamate decarboxylase	**−6.71**
Complement c1q TNF-related protein 6	**−12.62**	Lambda-crystallin	**−5.72**
Ureidoglycolate hydrolase	**−12.39**	Organic cation transporter 1	**−5.17**

Table 2 Top 25 expressed genes along the time course *(Continued)*

Sodium-dependent serotonine	**−12.02**	Dna-directed rna polymerase ii subunit rpb3	**−4.52**
Cytochrome p450 2u1	**−9.19**	Monocarboxylate transporter 5	**−4.14**
Organic cation transporter protein	**−7.44**	Cholinesterase	**−3.96**
Chitotriosidase	**−7.41**	Muscle lim protein 1	**−3.92**
Protein acn9 mitochondrial precursor	**−6.89**	Probable 4-hydroxy-2-oxoglutarate mitochondrial	**−3.46**
C1q domain containing protein	**−6.84**	Protein nlrc3-like	**−3.19**
N-acetylneuraminate lyase	**−5.53**	Beta-microseminoprotein	**−3.03**

In bold, downregulated transcripts.

The major diversity of the processes found by the GO terms occurred one day after the *Vibrio* challenge (Figure 4C). General functions such as biological regulation or metabolism seemed to peak at this time point. Other processes such as development, localization (the process in which a cell, a macromolecule or an organelle is transported to and/or maintained in a specific location), reproduction, growth, cell proliferation, locomotion and biological adhesion were also important 24 hpc. These specific processes are closely related and could indicate the resolution of the challenge after the expression of the immune genes that were triggered at 8 hours. Locomotion and biological adhesion appeared together in the time course because these two processes are intimately related and lead to taxis and migration. In zebrafish, it is known that the chemotaxis of immune cells begins quickly and lasts for several days after an inflammatory stimulus (tailfin transection). Neutrophils are rapidly recruited with a maximum cell count at 6 h post injury, and the recruitment of macrophages progressively increases until at least 48 h after injury [42]. The temporal response of *R. philippinarum* hemocytes seems to be intermediate between these two vertebrate cell types. However, the cytoskeleton stays active, which suggests that locomotion is a very dynamic process with a high percentage of the top-regulated genes (Table 2, 24 hpc) such as titin, twitchin or calponin-3. In non muscular cells, as the hemocytes, titin has been shown to be related to chromosome condensation and segregation during mitosis [43]. This activity could be indicative of hemocyte proliferation and other upregulated processes, such as development. Furthermore, other functions are upregulated, such as transcription (staphylococcal nuclease domain-containing protein 1), or downregulated, such as metabolism or transport (aldose 1-epimerase, nuclear pore complex protein nup107).

Three days after the challenge, many genes were down-regulated, suggesting negative feedback of the majority of the activated genes (Table 2, 72 hpc). Figure 4 also showed that the GO analysis could not find any processes with a maximum representation 72 hours post-challenge. Metabolism, adhesion, cytoskeleton modulation, transcription and defense (lysozyme) returned to basal levels, which was also illustrated by the general expression pattern shown in Figure 2. Curiously, the most upregulated gene, mantle gene 6, is related to the biomineralization of the shell [44]. This result could suggest that after the infection is controlled, the clams refocus the genetic machinery to other matters, such as shell repair. This result agrees with the GO analysis, and all of the biological functions began to stabilize three days after the bacterial challenge.

Because the defense against pathogens in molluscs relies on the innate immune system the response is expected to be fast, to overcome the infections as soon as possible. The contrary often means that the organisms will die. In fact, the critical point of the experiment seemed to be 8 hours after the challenge, when many immune genes were upregulated. Regulation of the apoptotic processes in hemocytes was also triggered early after the challenge and barely no genes related to apoptosis were found 3 days after the challenge. On the other hand, some effector genes such as C3, C1q, defense protein, and lysozyme seemed to exert their function before 24 hpc; cathepsins, proteases and protease inhibitors were up or downregulated throughout all the challenge pointing to complex interactions between bacteria and clams. These findings suggest that the survival or death of the animals is decided very early after the challenge, suggesting the importance of a fast triggering of the defense mechanisms.

Enrichment analysis

A Fisher's exact test was performed to detect any significant enrichment of GO terms in each sampling point and regulation in comparison with the reference set of sequences present in the microarray. Only three analyses were enriched in specific GO terms compared with the entire microarray output: 3 hours upregulated (45 enriched GO terms), 8 hours downregulated (3 terms) and 72 hours downregulated (2 terms). After using the blast2GO tool to reduce the GO enriched terms, the most specific terms (the highest level GO terms of a parental line) were presented in Figure 5. It can be observed that the first significantly enriched processes of the upregulated genes were exclusively related to immune signaling and transcription, such as the toll-like receptor pathway or the regulation of NF-κB import to nucleus. An interesting result was the negative regulation of the IL-8

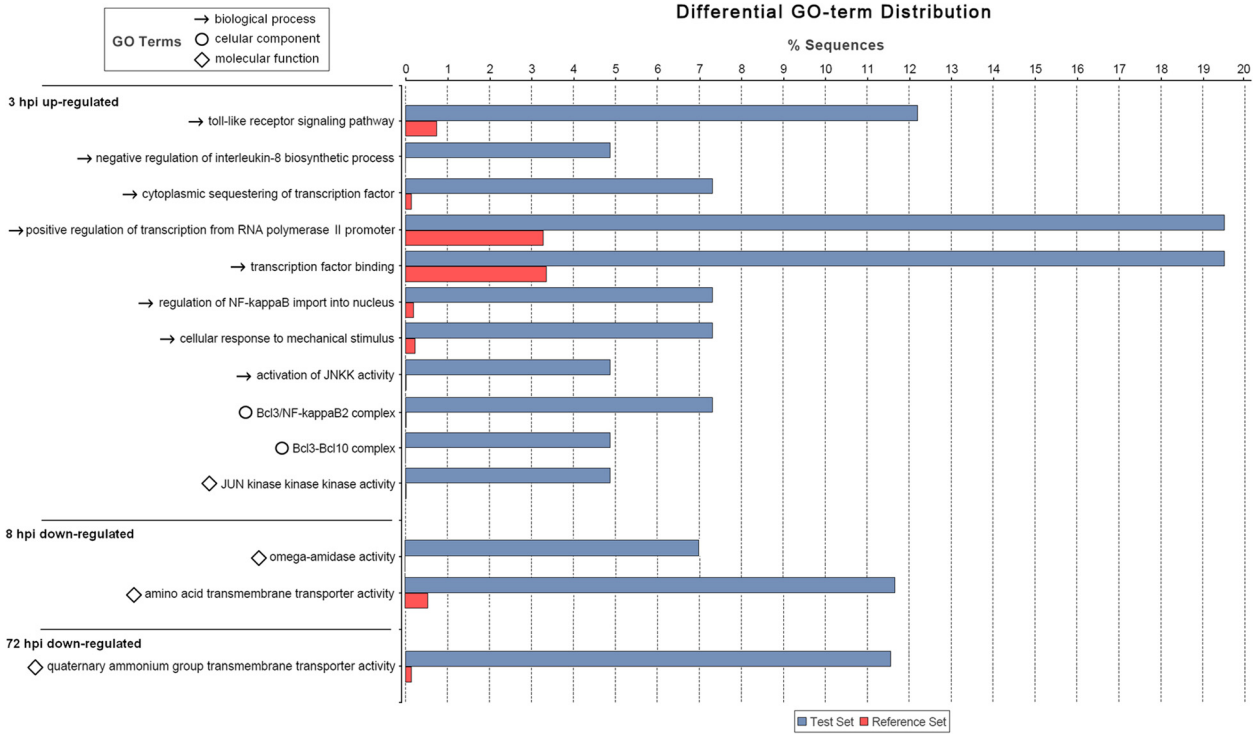


Figure 5 Enrichment analysis results. Differential GO term distribution between the test set (significantly expressed genes for each sampling point and regulation, up/down) and the reference set (all of the sequences present in the microarray). Only significant analyses were represented. hpc, hours post-challenge.

biosynthetic process, which is an indicator of an inhibition of the pro-inflammatory response. These results confirmed the previous findings of the GO annotation and the top-expressed genes, which showed an initial activation of the signaling and transcriptional processes and a strong inhibition of pro-inflammatory molecules such as C3, which seemed to exert their function much earlier than 3 hours after the stimulus, a trait that has been previously observed in bivalves [32].

The enrichment results at 8 hours and 72 hours post-challenge show that only downregulated processes related to metabolism and transport were found, but no specific enrichment of immunity genes was observed. These results imply the importance of an early response and an effective innate immune system in these bivalves.

ANOVA analysis: effect of the challenge in all of the experimental data

An ANOVA was performed to globally analyze the information of the microarray experiments after the challenge. Although a vast majority of the genes were regulated at a single sampling point, the ANOVA results showed a total of 15 genes that were significantly regulated throughout the time course. Table 3 summarizes a description of the 10 successfully annotated genes with their respective fold changes. The expression profile of the ten genes coincides with the observed pattern in Figure 1: the maximum fold change is achieved 8 hours after the bacterial challenge (except for LTBP4, whose maximum expression comes out at 24 hours). This finding suggests that the highest modulation of expression happens at 8 hpc. Thus, that time point after the challenge could represent the actual time point in which the fate of the surviving and dying animals is set.

These ten genes have important defense, protective and homeostasis functions, including cell death and apoptosis, that are key events to decide survivorship of the challenged clams: The phosphoinositide 3-kinase adapter protein 1 (PIK3AP1 or BCAP) links Toll-like receptor signaling and PI3K activation, preventing excessive inflammatory cytokine production and contributing to B-cell development in humans [45,46], this could mean that as the same time that defense cells produce an immune response, this is being controlled to avoid self-damage by an uncontrolled inflammatory reaction. The nuclear receptor dax-1 (NR0B1) is a co-regulatory inhibitor of the transcriptional activity of other nuclear receptors. NR0B1 is also related to the development and maintenance of stem cell pluripotency and lipogenesis and gluconeogenesis [47], important processes to maintain the energy homeostasis necessary to overcome an infection and regenerate damaged tissues. LTBP-4, described in the validation of the

Table 3 Differentially expressed genes detected by ANOVA and regulation after *V. alginolyticus* challenge

	3 h FC	8 h FC	24 h FC	72 h FC
Immune effectors				
Latent-transforming growth factor beta-binding protein 4	3.55	43.50	217.29	3.49
Heat shock 70 kda protein	1.31	5.30	2.18	−1.39
Phosphoinositide 3-kinase adapter protein 1	3.27	5.56	2.06	−1.13
Tnf receptor-associated factor 6	2.47	9.69	2.36	−1.27
Recognition/adhesion				
Perlucin-like protein	7.94	39.44	29.44	19.91
Protein lin-7 homolog c	−1.29	−2.37	−2.26	−1.25
Transcription				
Nuclear receptor dax-1	2.18	4.02	2.85	1.06
Ets-related transcription factor elf-5	6.19	26.84	10.00	1.35
Nuclear factor nf-kappa-b p50 subunit	1.74	9.48	1.89	−1.33
A20-binding inhibitor of nf-kappa-b activation 2	1.98	3.37	1.77	−1.13

microarray, modulates TGF-β activity and is related to cell adhesion and migration. The perlucin-like protein is a constitutive and regulating C-type lectin that promotes calcium carbonate precipitation and crystallization. Perlucin contains a carbohydrate recognition domain, and because this protein is expressed both in the mantle and in the gills and the digestive tract, it has been suggested that perlucin could play a role in non-self antigen recognition, similarly to other C-type lectins to trigger the immune response [48]. The upregulation of heat shock proteins (HSPs) represents an important mechanism in the stress response and their functions are closely linked to the innate immune system. Although the primary role of HSP70 is to function as a molecular chaperone, it inhibits the mitochondrial apoptosis pathway as well [49], a signal of the changes being produced after the *Vibrio* challenge and the survival efforts of hemocytes to overcome the infection. TRAF6, which has been described previously, is a signaling molecule that leads to the activation of the NF-κB and AP-1 transcription factors implicated in the regulation processes such as cell proliferation and survival, growth, differentiation, apoptosis, cell migration or metabolism; all of them closely related to immunity and response to a pathogen. The protein lin-7 homolog c establishes and maintains the distribution of channels and receptors at the plasma membrane of epithelial cells [50] and is likely to be involved in the formation of the cadherin-independent tight junction-like structure in epithelial cells and in the synapses in mammals [51], suggesting a role in establishment of the hemocytes adhesion to other cells and maybe intercellular signal transductions after the *Vibrio* challenge. The last three genes are a group related to transcription: the ets-related transcription factor elf-5, a transcriptional activator restricted to epithelial cells and involved in the keratinocyte differentiation in mammals [52], the NF-κB p50 subunit and the a20-binding inhibitor of NF-κB activation 2, two molecules that are implicated in the NF-κB signaling pathway, which controls many inflammatory and immune responses, cell proliferation, apoptosis and metabolism [53,54]. These molecules are closely involved in defense against pathogens, from recognition to effector functions in the immune system. Additionally, cell proliferation, survival or death were very represented processes, showing that apoptotic/antiapoptotic molecules are highly expressed after a *Vibrio* challenge, indicative of the active fight of the hemocytes.

Functional immune response of hemocytes after bacterial infection

To finally evaluate if the changes in gene expression corresponded to functional parameters, the immune response triggered by the *V. alginolyticus* infection was evaluated. In primary cultures of clam hemocytes (granulocytes and hyalinocytes) (Figure 6A), the *in vitro* infection with *V. alginolyticus* induced apoptotic and necrotic cell death 1 h after infection (Figure 1B). This functional result agrees with the GO analysis showing that most significantly upregulated genes at the beginning of the infection could be involved in chemotaxis, adhesion and therefore they lead to a rapid triggering of several mechanisms, such as apoptosis.

Results obtained after the intramuscular injection of *V. alginolyticus* also supported the previous findings of the GO annotation and the top-expressed genes. This bacterial infection induced significant changes in the cell concentration in hemolymph extracted from the site of infection. This parameter is an indirect indicator of cell migration to the site of infection. As soon as 6 h after infection the cell concentration registered in infected animals was significantly higher than that registered in

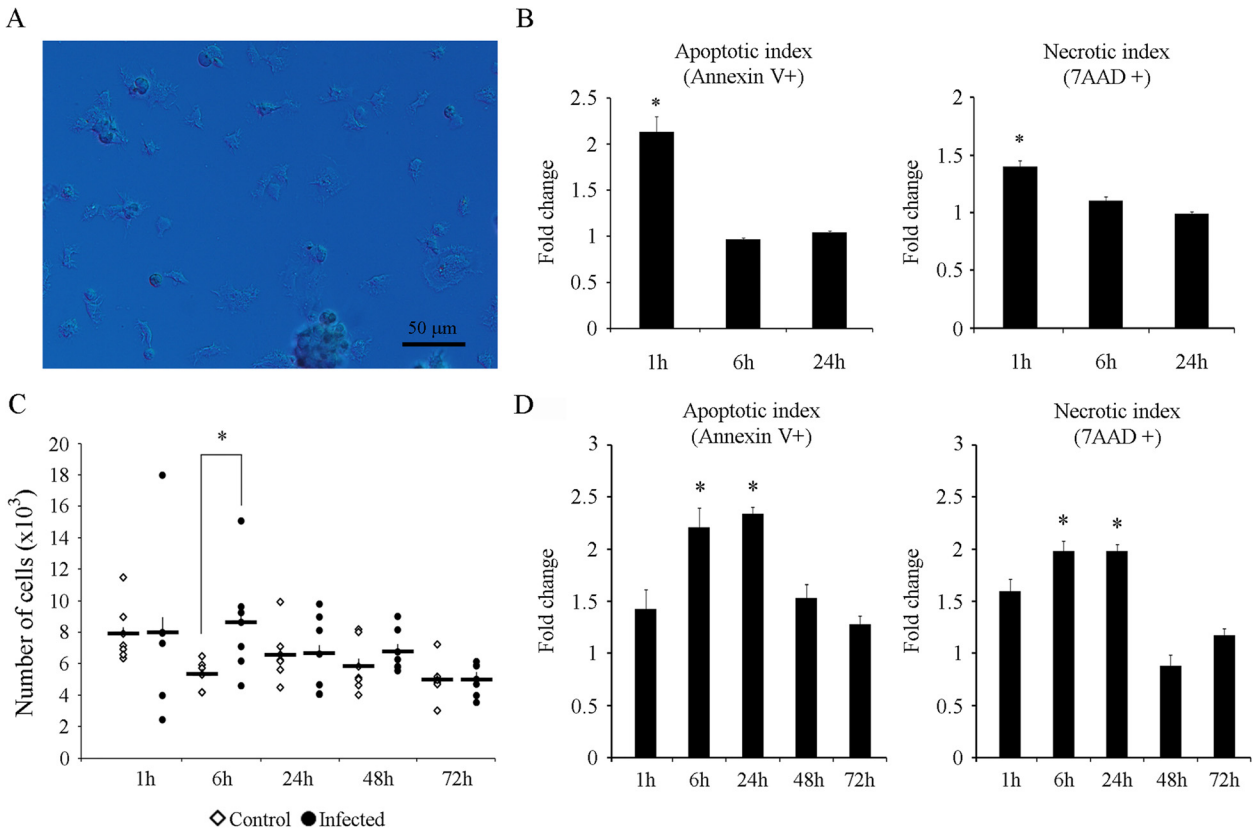


Figure 6 Evaluation of the immune response after bacterial infection. (A) The cellular components of the hemolymph are granulocytes and hyalinocytes. Photographs were taken using Nomarski DIC objective. **(B)** Apoptotic and necrotic levels induced by *V. alginolyticus* in primary cultures of hemocytes (n = 6). **(C)** *In vivo* modulation of the cellular concentration in hemolymph after bacterial infection. A significant increment in the number of cells was observed 6 h after infection. Graph represents the measurement of 6 independent animals as well as the mean + SEM. **(D)** Apoptotic and necrotic levels *in vivo* induced by *V. alginolyticus* at different times post infection. In all graphs, bars represent the mean + SEM of 6 independent measurements. Data were analyzed using the Student's t-test. (*) significant differences at $p < 0.05$.

controls. Cell migration seemed to be a quickly response since no significant variations were registered at 24, 48 and 72 h (Figure 6C). *V. alginolyticus* was also able to induce cell death after *in vivo* infections, although the timing for induction was different to that observed in cell cultures. A significant increment in the number of necrotic and apoptotic cells was registered at 6 and 24 h post-infection (Figure 6D).

Conclusions

The current work presents the second version of the *R. philippinarum* oligo-microarray enriched with immune sequences from hemocytes. We have analyzed the response of the Manila clam, *R. philippinarum*, after a *V. alginolyticus* challenge, strain TA15, and found almost 600 differentially expressed genes out of 13,671 probes. After the expression, GO term analyses and functional immune studies, we were able to establish a putative timing for a *Vibrio* infection in clams: there was an early response (3 hours) of hemocytes against the *Vibrio* challenge in which the main functions of hemocytes were mainly related to cellular component organization, biogenesis, cell migration, signaling and death. These functions were a fast response to the pathogen with related genes such as IL-17 or calmodulin. Eight hours seemed to be a key point in time to overcome the bacterial challenge, as the most important functions were related to response to stimulus, the immune system process and the multi-organism process to directly initiate the defense. Genes such as PGRPs, FREPs and defense proteins were present at this sampling point. After one day, there was a modulation of a great number of functions, from general processes such as metabolism to more specific functions such as development, localization and locomotion. Many cytoskeleton genes were found between the top-expressed genes, which is indicative of a possible chemotactic response in hemocytes. Finally, after 72 h, the number of regulated genes decreased substantially, indicating a stabilization of the status of the clam and the end of the cellular response to the challenge. We have to take into account that the modulation of the genes can be the result of an effective immune response but also a symptom of a

future death. In fact, several apoptotic and anti-apoptotic genes were regulated. This highlights that the obtained transcriptomic profile is a reflection of a complex interaction between the host immune system and the pathogen that deserves further studies.

In summary, bivalves are able to respond quickly against an infection with an orchestrated modulation of different genes. The immune-enriched oligo-microarray for *R. philippinarum* has proven to be useful in hemocyte expression analysis, and this tool has yielded interesting results.

Methods

Sequence assembly, annotation and microarray design

A total of 1,438,665 sequences from *R. philippinarum* were collected from different origins: 454 sequencing and Sanger sequencing [20,21] and available ESTs in the NCBI database. Table 1 details the number of sequences by origin, assembly and annotation criteria. After removing low-quality sequences and filtering for adaptors and primers, the cured sequences were assembled using the Newbler software package (GS *De Novo* Assembler v2.6, Roche). BlastX and BlastN algorithms [55] (http://blast.ncbi.nlm.nih.gov/Blast.cgi) were used to annotate the selected sequences. Alignments with an e-value threshold of 10^{-3} and 10^{-5} were considered significant for protein and nucleotide databases, respectively. The reference databases that were used are NCBI SwissProt, NCBI Metazoan RefSeq, NCBI nonredundant and ENSEMBL databases, and the dedicated databases for *B. glabrata* (http://www.snaildb.org/) and *L. gigantea* (http://genome.jgi-psf.org/Lotgi1/Lotgi1.download.html).

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.

Microarrays were synthesized *in situ* using the Agilent ink-jet technology with an 8 × 15 K format. Each array included default positive and negative controls. Probe sequences and other details on the microarray platform can be found in the GEO database (http://www.ncbi.nlm.nih.gov/geo/) [56] under accession number GPL16450. The GEO accession for the series data (microarray gene expression data) is GSE43274.

Animal sampling and RNA isolation

R. philippinarum clams were obtained from a commercial shellfish farm (Vigo, Galicia, Spain). Clams were maintained in open circuit filtered sea water tanks at 15°C with aeration and were fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana.* Prior to the experiments, the clams were acclimatized to the aquarium conditions for one week.

Clams (n = 100) were notched in the shell and injected in the muscle with 100 μl of 10^8 CFU/ml of *Vibrio alginolyticus*, strain TA15. The inoculation dose was chosen according to previous *Vibrio* challenges in *R. philippinarum* [5,18,57]. Controls (n = 100) were injected with 100 μl of PBS. After challenge, the clams were returned to the tanks and maintained at 15°C until sampling at 3, 8, 24, and 72 hours after challenge. The cumulative mortality rate at the end of the experiment, 72 hours, was 44%. The controls showed an end point mortality of 10%.

Hemolymph (1 ml) was withdrawn from the adductor muscle of the clams with a 0.5 mm diameter (25 G) disposable needle. Hemolymph from four individuals was pooled, and five biological replicates were taken at each sampling point. The hemolymph was centrifuged at 3000 g for 10 minutes at 4°C. The pellet was resuspended in 250 μl of TRIzol (Invitrogen). Total RNA isolation was conducted following the manufacturer's specifications in combination with the RNeasy Mini kit (Qiagen) for RNA purification after DNase I treatment. Next, the concentration and purity of the RNA were measured using a *NanoDrop ND1000* spectrophotometer. Finally, RNA integrity was tested on an Agilent 2100 Bioanalyzer (Agilent Technologies). Only the samples with high RNA quality were used for labeling and hybridization, 3 to 5 biological replicates were used for each condition and time.

Cy3 labeling

Sample labeling and hybridization were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol. Briefly, 100 ng of RNA from each RNA sample was amplified and labeled with Cy3 using the Low Input Quick Amp labeling kit (Agilent Technologies), according to the manufacturer's instructions. Each sample included a mixture of 10 different viral poly-acetylated RNAs (Agilent Spike-In Mix) before amplification and labeling so the microarray analysis workflow could be assessed. Qiagen RNeasy mini spin columns were used to purify amplified RNA. Finally, amplification and dye incorporation rates were verified using a *NanoDrop ND1000* spectrophotometer. These values should be between 200 and 500 ng/μL (RNA concentration) and between 20 and 50 pmol/μg aRNA (dye incorporation).

Microarray hybridization and analysis

Cy3 labeled RNA (600 ng) was fragmented with 5 μl of 10× Blocking Agent and 1 μl of 25× Fragmentation Buffer at 60°C for 30 min. Finally, 55 μl of 2× GE Hybridization buffer was added to dilute the fragmented RNA. The eight spaces of the gasket slide were filled with 40 μl of the correspondent hybridization solution and then assembled on the microarray slide (each slide contained 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 disassemble the slides and then washed in GE Wash Buffer 1 for 1 minute followed by one additional wash in pre-warmed (37°C) GE Wash Buffer 2.

Hybridized slides were scanned at 5 μm resolution using an Agilent G2565BA DNA microarray scanner. Default settings were modified to scan the same slide at two different sensitivity levels (XDR Hi 100% and XDR Lo 10%). The two linked images generated were analyzed together, the data were extracted, and the background was subtracted using the standard procedures in the Agilent Feature Extraction Software version 9.5.1. The software returned a series of spot quality measures to evaluate the goodness and the reliability of the spot intensity estimates. After ensuring that all of the microarrays passed the quality tests (spatial distribution of median signals for row and column, distribution of the spot signal, spatial distribution of all outliers of the array, coefficient of variation of the technical replicates of the SpikeIn and the regression correlation between signal and concentration) (Additional file 2: 1–36), control features (positive, negative, etc.), except for SpikeIn Viral RNAs, were excluded from subsequent analyses. SpikeIn control intensities were used as internal controls and were expected to be uniform across the experiments of a given dataset.

The GeneSpring software (Agilent) was used to normalize and analyze the microarray fluorescence data. To identify regulated genes, two statistical analyses in filtered raw data (20 - 90th Percentile) were carried out: a t-test ($p < 0.01$) and a two-way ANOVA ($p < 0.05$) with a Benjamini-Hochberg multiple testing correction. The t-test was used to find the genes that were significantly different between the controls and the infected samples at each sampling point. Genes with a fold change between ±1.5 were not used for further investigation. ANOVA analysis was performed to analyze the whole dataset taking into account the challenge along all the sampling times.

GO terms and enrichment analysis

After statistical analysis, blast2GO software [58] was used to assign GO terms [59] to the significantly expressed genes (t-test, $p < 0.01$) through the time course. Default values (annotation cutoff = 55, GOweight = 5) in blast2GO were used to perform the analysis and biological process ontology level 2 was selected.

The enrichment analyses were made with the total microarray information as the reference set and each sampling time (3 h, 8 h, 24 h, 72 h) and regulation (up or down) as the test sets. Then, Fisher's exact test was run with default values: a one-tailed test without removing double IDs and 0.05 false discovery rate (FDR) cut-off. To construct Figure 5, the blast2GO option to show only the most specific terms (0.05 FDR cut-off) was performed.

Validation of the microarray

Specific PCR primers were designed from the sequences of the selected probes (Table 4) using the *Primer3* program [60] according to qPCR restrictions. *Oligo Analyzer 1.0.2* was used to check for dimer and hairpin formation. The efficiency of each primer pair was then analyzed with seven serial five-fold dilutions of cDNA of *R. philippinarum* and was calculated from the slope of the regression line of the quantification cycle versus the relative concentration of cDNA [61]. A melting curve analysis was also performed to verify that only specific amplification occurred and that no primer dimers were amplified. If these conditions were not satisfied, new primer pairs were designed.

The cDNA synthesis was performed on 1 μg of total RNA using SuperScript™ III Reverse Transcriptase (Invitrogen) following the manufacturer's protocol. The same RNAs used for hybridization of the microarrays were used for the retrotranscription.

Real-time quantitative PCR was performed in the *7300 Real Time PCR System* (Applied Biosystems).

Table 4 Primers designed for the microarray validation

Primer name	Sequence 5′ → 3′	Tm	Product size	Ct slope	Acc. n°/Probe name
Clam 18S F	CCGAACATCTAAGGGCATCA	60.12	169 bp	−3.0	EF426293.1
Clam 18S R	AGTTGGTGGAGCGATTTGTC	60.99			
C3 F	CCCAGGTGCCAAAGAACA	55.56	162 bp	−3.5	S_isotig17547_isogroup06014
C3 R	GCGGGGTACACATACTCGTC	60.00			
IL-17D F	CTCAAAAAGACTCACAGGGAATG	60.15	186 bp	−3.8	P_isotig09595_isogroup01099
IL-17D R	CTGGCAATGATGTACTGTCGTAA	60.07			
LTBP-4 F	TAATCATTGCCGCCTTATCG	60.92	188 bp	−3.6	S_isotig21547_isogroup10014
LTBP-4 R	GCGACCTGAATCAAATTCGT	60.08			
TRAF6 F	GCCAACATAGTAGCTCAGGAACA	60.68	148 bp	−3.4	P_isotig17929_isogroup06396
TRAF6 R	TTCCAATATAGCTTACAACATCAACA	59.08			

One microlitre of fivefold-diluted cDNA template was mixed with 0.5 μl of each primer (10 μM) and 12.5 μl of SYBR Green PCR master mix (Applied Biosystems) in a final volume of 25 μl. The standard cycling conditions were 95°C for 10 min, followed by 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. All reactions were performed as technical triplicates, and an analysis of melting curves was performed in each reaction. The relative expression levels of the genes were normalized using 18S as a reference gene, which was constitutively expressed and not affected by the Vibrio challenge, following the Pfaffl method [61].

Functional immune response of hemocytes after bacterial infection

The immune response triggered after *V. alginolyticus* infection was evaluated through *in vitro* as well as *in vivo* experiments. Flow cytometry was used to evaluate different immune parameters such as cell migration and bacterial-induced apoptosis.

For *in vitro* experiments hemolymph from 12 animals were diluted 1:1 in ice-cold filtered sea water to prevent aggregation and dispensed (200 μl) into 24 wells plates. After 30 min of incubation at 15°C in the dark for cell adhesion, cells were infected with 200 μl of the TA15 strain *V. alginolyticus* (10^8 CFUs/ml) and maintained at 15°C. Induction of apoptosis was evaluated 1 h, 6 h and 24 h after infection in six different cell cultures. Supernatants were removed and 500 μl of ice-cold binding buffer (10 mM Hepes/NaOH, pH 7.4, 140 mM NaCl, 2.5 mM $CaCl_2$) containing 5 μl Annexin V-FITC (BD Biosciences) and 10 μl of 7-amino-actinomycin D (7-AAD, BD Biosciences) were added. Cells were incubated for 15 min in the dark and analyzed in a FACSCalibur flow cytometer using Cell Quest software (BD Biosciences). The apoptotic process was evaluated using the FITC and 7-AAD content detected in the FL-1 (530 nm) and FL-3 (650 nm) channels, respectively. Fold units were calculated by dividing the percentage of FL-1 positive hemocytes obtained after infection by the values recorded in the control group. In all experiments data were analyzed using the Student's t-test ($p < 0.05$). Fresh hemocytes were visualized using an Eclipse 80i light microscopy (Nikon) with Nomarski DIC prism to enhance the contrast.

For *in vivo* experiments 40 clams were intramuscular injected with 100 μl of the TA15 strain *V. alginolyticus* (4×10^{10} CFUs/ml). Control animals were injected with the same volume of filtered sea water. Animals were maintained in closed-circuit aquarium at 15°C. Hemolymph was extracted from 6 infected and 6 control animals at 1, 6, 24, 48 and 72 h after infection. Samples were immediately diluted 1:1 in ice-cold binding buffer and stained with Annexin V-FITC (BD Biosciences) and 7-amino-actinomycin D (7-AAD, BD Biosciences). Cells were incubated for 15 min in the dark and analyzed by flow cytometry as previously described. Cell migration to the site of infection was estimated by measuring the cell concentration in hemolymph extracted from the infected muscle.

Availability of supporting data

Microarray data are deposited in the public functional genomics data repository GEO (Gene Expression Omnibus):

- Microarray platform: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL16450
- Microarray gene expression data: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE43274

Additional files

Additional file 1: Description of the reference databases used to annotate *R. philippinarum* sequences.

Additional file 2: Quality control of each microarray experiment provided by the Agilent Feature Extraction Software.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

BN and AF conceived and designed the experiments. RM prepared the samples. LB, MM, MB and RM made the assembly and annotation of the sequences and designed the microarray platform. RM and MM hybridized the microarrays. AR performed functional immune assays. BN, AF, PB, AR and RM analyzed the data. RM wrote the paper. All authors read and approved the manuscript.

Acknowledgements

This work has been partially funded by the EU Project REPROSEED (245119). RM wishes to acknowledge the Spanish MICINN for her FPI Spanish research grant (BES-2009-029765).

Author details

[1]Instituto de Investigaciones Marinas (IIM), Consejo Superior de Investigaciones Científicas (CSIC), Eduardo Cabello 6, 36208 Vigo, Spain. [2]Department of Comparative Biomedicine and Food Science (BCA), University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy.

Received: 13 March 2013 Accepted: 31 March 2014
Published: 7 April 2014

References

1. Gestal C, Roch P, Renault T, Pallavicini A, Paillard C, Novoa B, Oubella R, Venier P, Figueras A: **Study of diseases and the immune system of bivalves using molecular biology and genomics.** *Rev Fish Sci* 2008, **16:**131–154.
2. Paillard C, Leroux F, Borrego JJ: **Bacterial disease in marine bivalves, Review of recent studies. Trends and evolution.** *Aquat Living Resour* 2004, **17:**477–498.
3. Villalba A, Reece KS, Ordás MC, Casas SM, Figueras A: **Perkinsosis in molluscs: A review.** *Aquat Living Resour* 2004, **17:**411–432.
4. Waki T, Shimokawa J, Watanabe S, Yoshinaga T, Ogawa K: **Experimental challenges of wild Manila clams with Perkinsus species isolated from naturally infected wild Manila clams.** *J Invertebr Pathol* 2012, **111:**50–55.
5. Allam B, Paillard C, Auffret M: **Alterations in hemolymph and extrapallial fluid parameters in the Manila clam, *Ruditapes philippinarum*, challenged with the pathogen *Vibrio tapetis*.** *J Invertebr Pathol* 2000, **76:**63–69.

6. Gómez-León J, Villamil L, Lemos ML, Novoa B, Figueras A: **Isolation of *V. alginolyticus* and *Vibrio splendidus* from aquacultured carpet shell clam (Ruditapes decussatus) larvae associated with mass mortalities.** *Appl Environ Microb* 2005, **71:**98–104.
7. Canesi L, Gallo G, Gavioli M, Pruzzo C: **Bacteria–hemocyte interactions and phagocytosis in bivalves.** *Microsc Res Technol* 2002, **57:**469–476.
8. Olafsen JA: **Role of lectins (C-reactive protein) in defense of marine bivalves against bacteria.** *Adv Exp Med Biol* 1995, **371A:**343–348.
9. Ordás MC, Novoa B, Figueras A: **Modulation of the chemiluminescence response of Mediterranean mussel (*Mytilus galloprovincialis*) haemocytes.** *Fish Shellfish Immunol* 2000, **10:**611–622.
10. Ordás MC, Ordás A, Beloso C, Figueras A: **Immune parameters in carpet shell clams naturally infected with *Perkinsus atlanticus*.** *Fish Shellfish Immunol* 2000, **10:**597–609.
11. Tafalla C, Gómez-León J, Novoa B, Figueras A: **Nitric oxide production by carpet shell clam (*Ruditapes decussatus*) hemocytes.** *Dev Comp Immunol* 2003, **27:**197–205.
12. Bayne CJ, Sminia T, Van der Knaap WPW: **Immunological Memory: Status of Molluscan Studies.** In *Phylogeny of Immunological Memory*, Developments in immunology, Volume 10. Edited by Manning MJ. Amsterdam: Elsevier; 1980:57–64.
13. Gueguen Y, Cadoret JP, Flament D, Barreau-Roumiguiere C, Girardot AL, Garnier J, Hoareau A, Bachere E, Escoubas JM: **Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*.** *Gene* 2003, **303:**139–145.
14. Costa MM, Novoa B, Figueras A: **Influence of β-glucans on the immune responses of carpet shell clam (*Ruditapes decussatus*) and Mediterranean mussel (*Mytilus galloprovincialis*).** *Fish Shellfish Immunol* 2008, **24:**498–505.
15. Costa MM, Prado-Alvarez M, Gestal C, Roch P, Li H, Novoa B, Figueras A: **Functional and molecular immune response of Mediterranean mussel (*Mytilus galloprovincialis*) hemocytes against pathogen associated molecular patterns and bacteria.** *Fish Shellfish Immunol* 2009, **26:**515–523.
16. Li H, Venier P, Prado-Alvarez M, Gestal C, Toubiana M, Quartesan R, Borghesan F, Novoa B, Figueras A, Roch P: **Expression of *Mytilus* immune genes in response to experimental challenges varied according to the site of collection.** *Fish Shellfish Immunol* 2010, **28:**640–648.
17. Araya MT, Markham F, Mateo DR, McKenna P, Johnson GR, Berthe FCJ, Siah A: **Identification and expression of immune-related genes in hemocytes of soft-shell clams, *Mya arenaria*, challenged with *Vibrio splendidus*.** *Fish Shellfish Immunol* 2010, **29:**557–564.
18. Moreira R, Balseiro P, Romero A, Dios S, Posada D, Novoa B, Figueras A: **Gene expression analysis of clams *Ruditapes philippinarum* and *Ruditapes decussatus* following bacterial infection yields molecular insights into pathogen resistance and immunity.** *Dev Comp Immunol* 2012, **36:**140–149.
19. Romero A, Dios S, Poisa-Beiro L, Costa MM, Posada D, Figueras A, Novoa B: **Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates.** *Dev Comp Immunol* 2011, **35:**334–344.
20. Milan M, Coppe A, Reinhardt R, Cancela LM, Leite RB, Saavedra C, Ciofi C, Chelazzi G, Patarnello T, Bortoluzzi S, Bargelloni L: **Transcriptome sequencing and microarray development for the Manila clam, *Ruditapes philippinarum*: genomic tools for environmental monitoring.** *BMC Genomics* 2011, **12:**234.
21. 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:**e35009.
22. Ghiselli F, Milani L, Chang PL, Hedgecock D, Davis JP, Nuzhdin SV, Passamonti M: ***De Novo* Assembly of the Manila Clam *Ruditapes philippinarum* Transcriptome Provides New Insights into Expression Bias, Mitochondrial Doubly Uniparental Inheritance and Sex Determination.** *Mol Biol Evol* 2012, **29:**771–786.
23. Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, Kobayashi S, Davis C, Dai H, He YD, Stephaniants SB, Cavet G, Walker WL, West A, Coffey E, Shoemaker DD, Stoughton R, Blanchard AP, Friend SH, Linsley PS: **Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer.** *Nat Biotechnol* 2001, **19:**342–347.
24. Iwahashi H, Kitagawa E, Suzuki Y, Ueda Y, Ishizawa YH, Nobumasa H, Kuboki Y, Hosoda H, Iwahashi Y: **Evaluation of toxicity of the mycotoxin citrinin using yeast ORF DNA microarray and Oligo DNA microarray.** *BMC Genomics* 2007, **8:**95.
25. Wang YZ, Han YS, Ma YS, Jiang JJ, Chen ZX, Wang YC, Che W, Zhang F, Xia Q, Wang XF: **Differential gene expression of Wnt signaling pathway in benign, premalignant, and malignant human breast epithelial cells.** *Tumour Biol* 2012, **33:**2317–2327.
26. Dheilly NM, Lelong C, Huvet A, Favrel P: **Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: identification of reference genes and tissue-enriched expression patterns.** *BMC Genomics* 2011, **12:**468.
27. Wang S, Peatman E, Liu H, Bushek D, Ford SE, Kucuktas H, Quilang J, Li P, Wallace R, Wang Y, Guo X, Liu Z: **Microarray analysis of gene expression in eastern oyster (*Crassostrea virginica*) reveals a novel combination of antimicrobial and oxidative stress host responses after dermo (*Perkinsus marinus*) challenge.** *Fish Shellfish Immunol* 2010, **29:**921–929.
28. Xu W, Faisal M: **Development of a cDNA microarray of zebra mussel (*Dreissena polymorpha*) foot and its use in understanding the early stage of underwater adhesion.** *Gene* 2009, **436:**71–80.
29. Allam B, Tanguy A, Jeffroy F, Le Bris C, Espinosa EP, Paillard C: **Transcriptional changes in Manila clam (*Ruditapes philippinarum*) in response to brown ring disease.** *Fish Shellfish Immunol* 2013, **34:**1636.
30. Morey JS, Ryan JC, Van Dolah FM: **Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR.** *Biol Proced Online* 2006, **8:**175–193.
31. Dunkelberger JR, Song WC: **Complement and its role in innate and adaptive immune responses.** *Cell Res* 2010, **20:**34–50.
32. Prado-Alvarez M, Rotllant J, Gestal C, Novoa B, Figueras A: **Characterization of a C3 and a factor B-like in the carpet-shell clam, *Ruditapes decussatus*.** *Fish Shellfish Immunol* 2009, **26:**305–315.
33. Starnes T, Broxmeyer HE, Robertson MJ, Hromas R: **Cutting edge: IL-17D, a novel member of the IL-17 family, stimulates cytokine production and inhibits hemopoiesis.** *J Immunol* 2002, **169:**642–646.
34. Costa MM, Pereiro P, Wang T, Secombes CJ, Figueras A, Novoa B: **Characterization and gene expression analysis of the two main Th17 cytokines (IL-17A/F and IL-22) in turbot, *Scophthalmus maximus*.** *Dev Comp Immunol* 2012, **38:**505–516.
35. Roberts S, Gueguen Y, de Lorgeril J, Goetz F: **Rapid accumulation of an interleukin 17 homolog transcript in *Crassostrea gigas* hemocytes following bacterial exposure.** *Dev Comp Immunol* 2008, **32:**1099–1104.
36. Zhou Y, Koli K, Hagood JS, Miao M, Mavalli M, Rifkin DB, Murphy-Ullrich JE: **Latent transforming growth factor-beta-binding protein-4 regulates transforming growth factor-beta1 bioavailability for activation by fibrogenic lung fibroblasts in response to bleomycin.** *Am J Pathol* 2009, **174:**21–33.
37. Kantola AK, Keski-Oja J, Koli K: **Fibronectin and heparin binding domains of latent TGF-beta binding protein (LTBP)-4 mediate matrix targeting and cell adhesion.** *Exp Cell Res* 2008, **314:**2488–2500.
38. Qiu L, Song L, Yu Y, Zhao J, Wang L, Zhang Q: **Identification and expression of TRAF6 (TNF receptor-associated factor 6) gene in Zhikong scallop *Chlamys farreri*.** *Fish Shellfish Immunol* 2009, **26:**359–367.
39. Walker MP, Zhang M, Le TP, Wu P, Lainé M, Greene GL: **RAC3 is a pro-migratory co-activator of ERα.** *Oncogene* 2011, **30:**1984–1994.
40. Can G, Akpinar B, Baran Y, Zhivotovsky B, Olsson M: **5-Fluorouracil signaling through a calcium-calmodulin-dependent pathway is required for p53 activation and apoptosis in colon carcinoma cells.** *Oncogene* 2013, **32:**4529–4538.
41. Gestal C, Pallavicini A, Venier P, Novoa B, Figueras A: **MgC1q, a novel C1q-domain-containing protein involved in the immune response of *Mytilus galloprovincialis*.** *Dev Comp Immunol* 2010, **34:**926–934.
42. Gray C, Loynes CA, Whyte MK, Crossman DC, Renshaw SA, Chico TJ: **Simultaneous intravital imaging of macrophage and neutrophil behaviour during inflammation using a novel transgenic zebrafish.** *Thromb Haemost* 2011, **105:**811–819.
43. Machado C, Andrew DJ: **D-Titin:a giant protein with dual roles in chromosomes and muscles.** *J Cell Biol* 2000, **151:**639–652.
44. Liu HL, Liu SF, Ge YJ, Liu J, Wang XY, Xie LP, Zhang RQ, Wang Z: **Identification and characterization of a biomineralization related gene PFMG1 highly expressed in the mantle of *Pinctada fucata*.** *Biochemistry* 2007, **46:**844–851.
45. Ni M, MacFarlane AW IV, Toft M, Lowell CA, Campbell KS, Hamerman JA: **B-cell adaptor for PI3K (BCAP) negatively regulates Toll-like receptor signaling through activation of PI3K.** *Proc Natl Acad Sci U S A* 2012, **109:**267–272.

46. Maruoka M, Suzuki J, Kawata S, Yoshida K, Hirao N, Sato S, Goff SP, Takeya T, Tani K, Shishido T: **Identification of B cell adaptor for PI3-kinase (BCAP) as an Abl interactor 1-regulated substrate of Abl kinases.** *FEBS Lett* 2005, **579:**2986–2990.
47. Laurenzana EM, Chen T, Kannuswamy M, Sell BE, Strom SC, Li Y, Omiecinski CJ: **The Orphan Nuclear Receptor, DAX-1, Functions as a Potent Co-repressor of the Constitutive Androstane Receptor (CAR, NR1I3).** *Mol Pharmacol* 2012, **82:**918–928.
48. Wang N, Lee YH, Lee J: **Recombinant perlucin nucleates the growth of calcium carbonate crystals: molecular cloning and characterization of perlucin from disk abalone, *Haliotis discus discus*.** *Comp Biochem Physiol B Biochem Mol Biol* 2008, **149:**354–361.
49. Parcellier A, Gurbuxani S, Schmitt E, Solary E, Garrido C: **Heat shock proteins, cellular chaperones that modulate mitochondrial cell death pathways.** *Biochem Biophys Res Commun* 2003, **304:**505–512.
50. Shelly M, Mosesson Y, Citri A, Lavi S, Zwang Y, Melamed-Book N, Aroeti B, Yarden Y: **Polar expression of ErbB-2/HER2 in epithelia. Bimodal regulation by Lin-7.** *Dev Cell* 2003, **5:**475–486.
51. Irie M, Hata Y, Deguchi M, Ide N, Hirao K, Yao I, Nishioka H, Takai Y: **Isolation and characterization of mammalian homologues of *Caenorhabditis elegans* lin-7: localization at cell-cell junctions.** *Oncogene* 1999, **18:**2811–2817.
52. Oettgen P, Kas K, Dube A, Gu X, Grall F, Thamrongsak U, Akbarali Y, Finger E, Boltax J, Endress G, Munger K, Kunsch C, Libermann TA: **Characterization of ESE-2, a novel ESE-1-related Ets transcription factor that is restricted to glandular epithelium and differentiated keratinocytes.** *J Biol Chem* 1999, **274:**29439–29452.
53. Leotoing L, Chereau F, Baron S, Hube F, Valencia HJ, Bordereaux D, Demmers JA, Strouboulis J, Baud V: **A20-binding inhibitor of nuclear factor-kappaB (NF-kappaB)-2 (ABIN-2) is an activator of inhibitor of NF-kappaB (IkappaB) kinase alpha (IKKalpha)-mediated NF-kappaB transcriptional activity.** *J Biol Chem* 2011, **286:**32277–32288.
54. Tornatore L, Thotakura AK, Bennett J, Moretti M, Franzoso G: **The nuclear factor kappa B signaling pathway: integrating metabolism with inflammation.** *Trends Cell Biol* 2012, **22:**557–566.
55. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215:**403–410.
56. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Holko M, Ayanbule O, Yefanov A, Soboleva A: **NCBI GEO:archive for functional genomics data sets–10 years on.** *Nucleic Acids Res* 2011, **39**(Database issue):D1005–D1010.
57. Allam B, Paillard C, Ford SE: **Pathogenicity of *Vibrio tapetis*, the etiological agent of brown ring disease in clams.** *Dis Aquat Organ* 2002, **48:**221–231.
58. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M: **Blast2GO, a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21:**3674–3676.
59. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene ontology, tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25:**25–29.
60. Rozen S, Skaletsky HJ: **Primer3 on the WWW for general users and for biologist programmers.** *Methods Mol Biol* 2000, **132:**365–386.
61. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Res* 2001, **29:**e45.

doi:10.1186/1471-2164-15-267
Cite this article as: Moreira *et al.*: Gene expression profile analysis of Manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibrio alginolyticus* challenge using an immune-enriched oligo-microarray. *BMC Genomics* 2014 **15**:267.

Additional files

Additional files associated with this article can be found at:
http://www.biomedcentral.com/1471-2164/15/267

IMMUNE RESPONSES DURING THE LARVAL STAGES OF *MYTILUS GALLOPROVINCIALIS*: METAMORPHOSIS ALTERS IMMUNOCOMPETENCE, BODY SHAPE AND BEHAVIOR

Fish & Shellfish Immunology 35 (2013) 438–447

Contents lists available at SciVerse ScienceDirect

Fish & Shellfish Immunology

journal homepage: www.elsevier.com/locate/fsi

Immune responses during the larval stages of *Mytilus galloprovincialis*: Metamorphosis alters immunocompetence, body shape and behavior

Pablo Balseiro, Rebeca Moreira, Rubén Chamorro, Antonio Figueras, Beatriz Novoa*

Instituto de Investigaciones Marinas (IIM), CSIC, C/Eduardo Cabello 6, C.P.36208 Vigo, Pontevedra, Spain

ARTICLE INFO

Article history:
Received 8 January 2013
Received in revised form
16 April 2013
Accepted 29 April 2013
Available online 17 May 2013

Keywords:
Larvae
Ontogeny
Antimicrobial
Metamorphosis
Mussel

ABSTRACT

We investigated the development of the immune system during the larval stages of the mussel *Mytilus galloprovincialis*. The ability of trochophore and veliger larvae to phagocytose foreign particles (*Escherichia coli* and zymosan) was measured. Phagocytosis was detected as early as 24 h post-fertilization (hpf) using flow cytometry and fluorescence microscopy. However, although there was a high basal production of reactive oxygen and nitrogen species (ROS and NRS), the phagocytosis of zymosan did not trigger an associated increase in radical production. In addition, a panel of immune-related mussel genes (Myticin B, Myticin C, Mytilin B, Mytimycin precursor 1, Macrophage migration inhibition factor, lysozyme, C1q, membrane attack complex protein and fibrinogen-related protein) was selected for expression profile analysis throughout the different developmental stages (trochophore, veliger, metamorphosis, post-settlement and spat). The expression of these genes increased during the transition from trochophore to spat, and the level of expression was higher in oocytes than in trochophores, suggesting that gene expression during the first larval stages might be maternal in origin. Metamorphosis was identified as a crucial stage when larvae increased the expression of immune-related genes and responded to environmental signals. Whole-mount *in situ* hybridization studies showed the mantle edge as an important area in the development of immunocompetence in bivalve larvae. Larvae responded to both live and heat-inactivated bacteria by modulating expression of immune-related genes. Altogether, our results support that during the early stages of *M. galloprovincialis* development, immune mechanisms emerge to aid larvae in managing infections.

1. Introduction

Marine invertebrate reproduction generally occurs by the release of a huge number of eggs into the water column and the formation of larvae after egg fertilization. After a period of free living, larvae undergo metamorphosis, transitioning from free-swimming larvae to benthonic postlarvae or juveniles [1]. This mechanism ensures the spread of the species throughout nature and can partially explain the rapid colonization of new habitats by marine animals. Embryonic and larval development are important phylogenetic events because bilateral animals are divided into deuterostomes or protostomes based on the embryonic origin of the mouth and anus. Mollusks are lophotrochozoans, a group characterized by the presence of trochophore larvae [2]. The ontogeny of mollusks is a complex process that is only completed by a small percentage of the original larvae produced after fecundation. The main threats to the completion of the larval life cycle are predation and poor environmental conditions, although larvae also have to manage pathogenic infections [3,4].

Among Protostomia, research into larval immune capacity as well as development, genetics and physiology has focused on model species such as the arthropod *Drosophila* or the nematode *Caenorhabditis*, which belong to the *superphylum* Ecdysozoa [5–7], paying less attention to the *superphylum* Lophotrochozoa, which includes mollusks. Mollusks are an interesting group, not only in terms of aquaculture production [8] but also because they have an important ecological role in the depuration of waters and as environmental sentinels [9]. Additionally, this group includes intermediate hosts for serious parasitic human diseases [10].

Larval settlement and metamorphosis is a vital transition period that is associated with the evolution of metazoans, as well as differentiation and speciation [11]. This stage is also very important because settlement can modify the larval dispersal capacity, leading individuals to settle near their source or to be dispersed with a

* Corresponding author. Tel.: +34 986214463; fax: +34 986292762.
E-mail addresses: pablobal@iim.csic.es (P. Balseiro), rebecamoreira@iim.csic.es (R. Moreira), ruchaval@iim.csic.es (R. Chamorro), antoniofigueras@iim.csic.es (A. Figueras), beatriznovoa@iim.csic.es, virus@iim.csic.es (B. Novoa).

1050-4648/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsi.2013.04.044

concomitant mortality risk. Therefore, competent and metamorphosing larvae are important for the ecological and evolutionary success of natural populations of marine invertebrates [12]. Even in cultured bivalves, metamorphosis is considered a crucial step for the overall success of the aquaculture facility because animals that have correctly settled usually show increased survival. However, it is unknown if competence is related to the expression of particular genes or if the immune system plays a role in the protection of larval stages. In fact, studies about the gene expression of mollusk larvae during development have been focused mainly on anatomical structural ontogenesis [13,14], and little is known about bivalve larval immune defense or bivalve response to external stimuli [15–17]. Phagocytosis, antimicrobial peptide production and chemical defenses that serve both to control pathogens and to protect larvae from predation have been proposed as immune mechanisms in veliger larvae [18]. It has been reported that larvae initiate metamorphosis when they are developmentally competent to respond to environmental signals [19–21]; however, the factors that regulate this process have been difficult to identify. In this work, we have focused on the immune competence of several developmental stages of *Mytilus galloprovincialis* because immune-related genes are essential for responding to external stimuli. We have identified the immune functions occurring at larval stages and profiled gene expression during each developmental stage to understand this complex process.

2. Materials and methods

2.1. Larval rearing

Mature Mediterranean mussels (*M. galloprovincialis*) were obtained from a commercial shellfish farm (Vigo, Northwestern Spain) during the spawning season. Mussels were subjected to mechanical (brush cleaning) and environmental stress (left at least 1 h without water), and spawning was induced with 1 μm filtered seawater (FSW) at 30 °C containing *Isochrysis galbana*. The mature animals responding to spawning induction were placed in individual containers and oocyte and sperm quality was observed (bright orange color, cell integrity and sperm motility) under a Nikon Optiphot light microscope (Nikon Instruments Inc., NY, USA).

Fertilization was performed by mixing high quality oocytes and sperm at a ratio of 1:10 in a sterile glass container containing 2 L of FSW. After trochophore development for 24 h at 21 °C, larvae were placed into 150 L tanks. Larvae were fed with *I. galbana* during the first week post-fertilization and a combination of *I. galbana* and *Phaeodactylum tricornutum* after the first week. Samples were taken and observed under a Nikon Eclipse E600 light microscope or a Nikon SMZ800 stereomicroscope. Measures were performed on pictures taken with a DXM1200 operated with ACT-1 v2.70 software (Nikon instruments Inc., NY, USA).

2.2. Immune stimulation

Bacterial challenge of larvae was performed with both heat-inactivated and live *Vibrio anguillarum* at all developmental stages. Larvae were stimulated with 10^6 CFU mL^{-1} of either live or inactivated bacterial challenges for 3 h at 15 °C. Three biological replicates of stimulated larvae or controls were used for each stage. Larvae were concentrated, counted and then distributed into 6-well plates at a final volume of 7 mL per well.

2.3. Functional immunology

The capacity of larvae to phagocytose foreign particles and to produce radical oxygen and nitrogen species was studied from 24 to 72 h post-fertilization (hpf). Briefly, *M. galloprovincialis* larvae were mechanically disaggregated in calcium–magnesium free saline (CMFS) buffer (137 mM sodium chloride, 4 mM potassium chloride, 0.4 mM sodium dihydrogen phosphate, 0.2 mM potassium hydrogen phosphate, 12 mM sodium bicarbonate, and 10 mM D-glucose), and the cells were resuspended in FSW prior to flow cytometry experiments.

Phagocytosis assays were performed using the larval cell suspensions and two different fluorescein-labeled (FITC) particles: *Escherichia coli* or zymosan (Molecular Probes, Life Technologies, Carlsbad, CA, USA). Cells and FITC particles were incubated for 90 min at 17 °C in the dark and counterstained with trypan blue to quench non-phagocytosed particles. The level of phagocytosis was determined using a FACSCalibur flow cytometer (Becton and Dickinson, San Jose, CA, USA). In addition, phagocytosis was also studied with fluorescence microscopy using pHrodo-labeled *E. coli* and Texas Red-labeled zymosan (Molecular Probes). Whole larvae were incubated with fluorescent particles as explained above and fixed overnight at 4 °C with 4% paraformaldehyde (Sigma Chem. Co, St. Louis, MO, USA) in FSW. After fixation, larvae were stained with DAPI to visualize nuclei and mounted with ProLong Gold Antifade Reagent (Life Technologies Carlsbad, CA, USA). Larvae were observed with a Leica TCS SPE confocal microscope (Leica Microsystems GmbH, Wetzlar, Germany).

The production of reactive oxygen species (ROS) was measured with flow cytometry using a 5,6-chloromethyl-2′,7′-dichlorodihydrofluorescein diacetate probe (CM-H_2DCFDA, Molecular Probes, Life Technologies, Carlsbad, CA, USA). After the probe penetrates the cells, it is oxidized by ROS, producing green fluorescence. Similarly, the production of reactive nitrogen species (RNS) was measured with flow cytometry using the 4-amino-5-methylamino-2′,7′-difluorofluorescein diacetate (DAF-FM diacetate) probe. In both cases, fluorescence in the FL-1 channel was measured after trypan blue counterstaining in a FACSCalibur flow cytometer (Becton and Dickinson, San Jose, CA, USA). Zymosan at 0.1 μg mL^{-1} was used to stimulate the immune response. The DNA-specific stain 7-amino-actinomycin D (7AAD) was used to determine the integrity of cell membranes.

2.4. Gene expression analysis using quantitative PCR

Larval gene expression was measured both at basal conditions and in challenged larvae. Each experiment was conducted using three different biological replicates, and each replicate belonged to a different family. For ontogeny studies at basal conditions, larvae were collected, concentrated using nylon mesh and centrifuged. The pellet was resuspended in 500 μL of Trizol (Life Technologies Carlsbad, CA, USA), and total RNA extraction and cDNA synthesis were performed using the previously described standard protocols [22]. In the studies of challenged larvae, RNA isolation was performed using the Maxwell 16 LEV Simply RNA Tissue Kit (Promega, Madison, WI, USA) following manufacturer's instructions. cDNA was synthesized as previously described.

We examined the expression of several immune-related genes previously described in mussel [23]. The following genes were included in the analysis: the antimicrobial peptides Mytimycin precursor 1, Mytilin B, Myticin B and Myticin C; the putative genes implicated in pathogen recognition, C1q and Fibrinogen-related protein (FREP); the pore-forming molecule MacP; the inflammatory regulator macrophage inhibition factor and lysozyme. The expression of these immune-related genes was also performed in hemocytes from adult mussels.

Quantitative PCR was performed using the PCR primers summarized in Table 1, which were selected according to qPCR restrictions. Oligo Analyzer 1.0.2 was used to assess dimer and

Table 1
Immune-related genes. Immune-related genes analyzed in the present work and primer sequences used for qPCR. The main function of each gene is also summarized.

Gene	Putative function	Primer name	Sequence (5′ > 3′)	Reference
Myticin C	Antimicrobial peptide	*Myt C fw*	ATTTGCTACTGCCTTCATTG	[23]
		Myt C rv	TCCATCTCGTTGTTCTTGTC	
Mytimicin Precursor 1	Antimicrobial peptide	*MMG1 Mytimycin Prec qPCR1 S*	ACGGATGACGCTTTTGTTTG	[23]
		MMG1 Mytimycin Prec qPCR1 As	GCAGTCCCAGCAATGTTTC	
Macrophage inhibition factor	Inflammatory regulator	*MIF qPCR 1 S*	TACACCCAGACCAAATGATG	[23]
		MIF qPCR 1 As	TTCTCCTAATGCTCCAATACTG	
Myticin B	Antimicrobial peptide	*Myticin B qPCR1 S*	AATGTCTTCGTTGTTCCAG	[23]
		Myticin B qPCR1 As	AATGCCAGTTTCACCTTG	
C1q	Antigen recognition	*MgC1q qPCR1 S*	ATTTATGCGTTCACTTGGAC	[23]
		MgC1q qPCR1 As	ACACCGATTTTTGTGCTG	
Lysozyme	Bacterial wall hydrolysis	*Mg Lysozyme qPCR1 S*	TGTCTGTCGCACTATTCTTC	[23]
		Mg Lysozyme qPCR1 As	AGTCCGCAACAAACATTC	
Mytilin B	Antimicrobial peptide	*Myl 4*	TGAAGGCAGGAGTTATTCTGGC	[56]
		Myl 3	ACAACGAAGACATTTGCAGTAGC	
Membrane attack complex and perforin-domain-containing protein (MacP)	Pore-forming molecule	*Perforin PCRq-F*	AAGGTGGATGTTGGTTATGGAGAA	[23]
		Perforin PCRq-R	GCCCAATCAGGCATCATGTTA	
Fribinogen-related protein (FREP)	Antigen recognition	*qPCR FREP G1 S*	GCAAATGTGCTACAAACTACCG	Present work
		qPCR FREP G1 As	TCCAAGTACGCCAGTAAACG	
18S	Ribosomal constituent	*Mussel-18S-F*	GTACAAAGGGCAGGGACGTA	[55]
		Mussel-18S-R	CTCCTTCGTGCTAGGGATTG	

hairpin formation. The efficiency of each primer pair was analyzed with seven serial fivefold dilutions of *M. galloprovincialis* cDNA and calculated from the slope of the regression line during the quantification cycle and compared to the relative concentration of cDNA [24]. A melting curve analysis was also performed to verify that primer dimers were not amplified. Real-time quantitative PCR was performed in the 7300 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). Reactions were performed using SYBR green PCR master mix (Applied Biosystems, Foster City, CA, USA) at a final volume of 25 μL (0.2 μM each primer). The standard cycling conditions were 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. All reactions were performed as technical triplicates, and melting curves were analyzed for each reaction. The relative expression levels of the genes were normalized using 18S RNA as a reference gene following the Pfaffl method [24] because it is constitutively expressed and not affected by the challenge.

2.5. In situ hybridization

Larvae were fixed overnight at 4 °C in 4% PFA, dehydrated through successive methanol dilutions and stored at −20 °C until whole-mount *in situ* hybridization (WISH) was performed. Sense and antisense probes were produced by PCR amplification using standard PCR conditions (35 cycles, 60 °C annealing temperature). Primers were designed including Sp6 (sense probe) or T7 (antisense probe) promoter sequences for labeling purposes (Table 2). Probes were purified with Amicon Ultra columns (Millipore Bedford, MA, USA), labeled with the Roche DIG RNA Labeling Kit (SP6/T7) (Roche Diagnostics, Mannheim, Germany), purified with Sigmaspin columns (Sigma Chem. Co, St. Louis, MO, USA) and stored at −20 °C until use.

WISH was set up for *M. galloprovincialis* larvae and performed in 24-well plates with 0.5–1 mL of the appropriate solution in each well. All steps were performed in 41-μm-diameter mesh hand-made

Table 2
Primer sequences used to synthesize probes and corresponding gene sequences. Genes and primers uses for probe synthesis. Highlighted bold nucleotides are the Sp6 and T7 promoters.

Gene	Probe	Primer sequence
Mytimycin	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**ACGGATGACGCTTTTGTTTG
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**GCAGTCCCAGCAATGTTTC
MIF	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**TACACCCAGACCAAATGATG
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**TTCTCCTAATGCTCCAATACTG
Myticin C	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**ATATTCCTCAAAACTCAAAACATTCA
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**TTCAAGCTGAAAACGTCGAA
Myticin B	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**AATGTCTTCGTTGTTCCAG
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**AATGCCAGTTTCACCTTG
Mytilin B	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**TGAAGGCAGGAGTTATTCTGGC
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**ACAACGAAGACATTTGCAGTAGC
C1q	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**ATTTATGCGTTCACTTGGAC
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**ACACCGATTTTTGTGCTG
Lysozyme	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**TGTCTGTCGCACTATTCTTC
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**AGTCCGCAACAAACATTC
MacP	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**CTATTTGGCGGGAGTTTGAC
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**TCTGAATCCACAAAGGAGCA
FREP	Sense-sp6	**ACGATTTAGGTGACACTATAGAA**GCAGGTAACAGCATAGAACATC
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA**CCAAGTACGCCAGTAAACG
Actin	Sense-sp6	**ACGATTTAGGTGACACTATAGAA** AACCGCCGCTTCTTCATCTTC
	Antisense-T7	**AGTTAATACGACTCACTATAGGGA** TACCACCAGACAAGACGG

baskets. Larvae stored in methanol were re-hydrated in a methanol/PBS series and permeabilized using Proteinase K (Roche Diagnostics, Mannheim, Germany) at a final concentration of 10 ng mL^{-1} at 37 °C for 10 min (trochophores and veliger larvae) or 20 min (metamorphosis larvae). The hybridization was carried out overnight at 70 °C with 250–500 ng of the corresponding probe in 1 mL of hybridization mix (50% formamide, 5× SSC, 0.1% Tween-20, 50 mg mL^{-1} heparin and 0.5 mg mL^{-1} torula yeast tRNA). After washing and preincubation with blocking solution (2% sheep serum and 2 mg mL^{-1} BSA in PBT), larvae were incubated with alkaline phosphatase-conjugated anti-DIG secondary antibody (1:5000) (Roche Diagnostics, Mannheim, Germany) overnight at 4 °C. After washing and presoaking, larvae were stained with NBT/BCIP (Roche Diagnostics, Mannheim, Germany) at room temperature. Larvae were observed and photographed using a Nikon Eclipse TS100 light microscope with a Nikon DS-Fi1 camera (Nikon Instruments Inc., NY, USA).

2.6. Statistics

Phagocytosis was compared using Student's *t*-test with unequal variances to identify statistically significant differences between the control and treatment groups. The results were expressed as the mean ± SEM and differences were considered significant at $p < 0.01$. All treatments were assayed in triplicate. Production of radical species was considered significant at $p < 0.05$.

Gene expression levels were transformed to a log2 scale before comparison with Student's *t*-test with equal variances to look for statistical significance between oocytes and different developmental stages (larval basal gene expression) or between stimulated and non-stimulated samples. The results were expressed as the mean ± SEM and differences were considered significant at $p < 0.05$.

3. Results and discussion

3.1. Larval production and development

M. galloprovincialis larvae were obtained in the laboratory, and the typical larval stages were observed as previously described (Fig. 1A–C) [25] with similar growth rates (Fig. 1D) as those obtained in closed systems with small volumes [26]. This is in agreement with previous data from laboratory culture [27,28], where growth is slower than in hatcheries, that are more appropriate for obtaining larvae with higher survival and growth rates [25].

After 24 h at 21 °C, the majority of fertilized oocytes develop into trochophore larvae (Fig. 1A) that are motile and characterized by the presence of two ciliar rings and an apical flagellum (Fig. 1A). At 48 h, trochophore larvae transformed into the veliger stage (Fig. 1B) in which motile larvae develop a *velum* that is enclosed in the primordial shell when resting. After 19 days, the majority of larvae swelled and took an umbonate shape. The pediveliger stage is characterized by the formation of a foot and is the larval stage that fixes to the substratum (Fig. 1C). Fig. 1D summarizes the length of larvae from five different *M. galloprovincialis* families. Larval length is altered by metamorphosis events and changes in form, particularly when larvae become umbonate in shape prior to settlement.

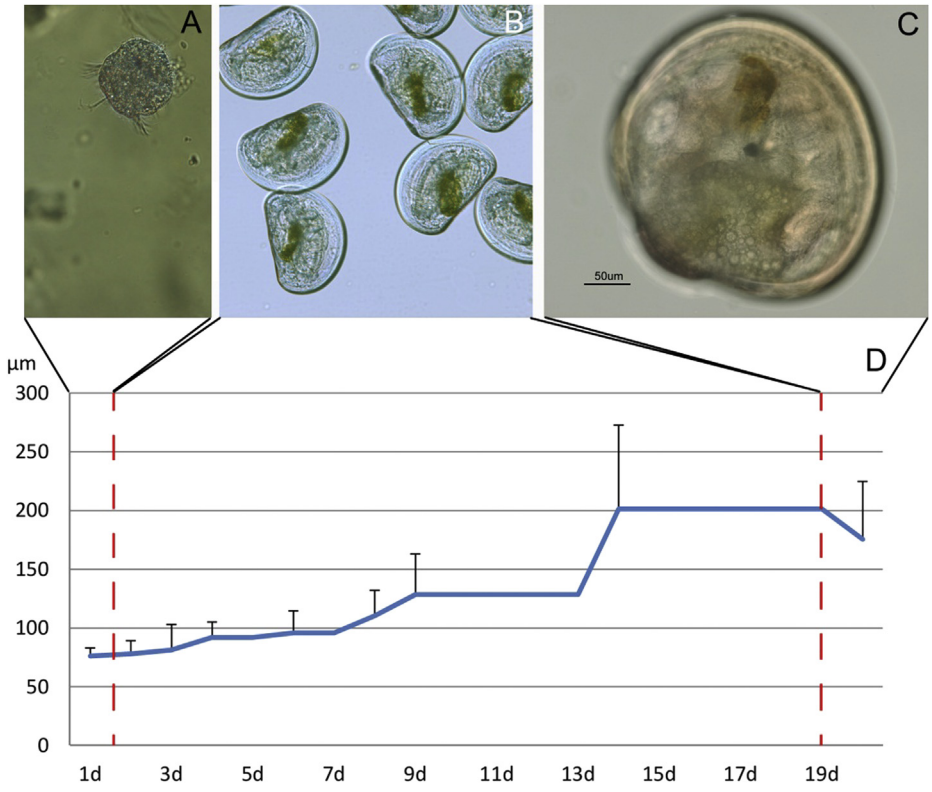


Fig. 1. A. Photomicrographs of *Mytilus galloprovincialis*. A. Trochophore stage. B. Veliger stage. C. Pediveliger stage. The scale bar for all three micrographs is 50 µm. D. *M. galloprovincialis* larval growth throughout development was measured as the mean shell length ± SD. Dashed bars indicate approximate transition from trochophores to veliger and veliger to pediveliger.

3.2. Immune capacity of early larval stages

The phagocytic capacity of mussel larvae during their first hours of development is summarized in Fig. 2, where the distribution of phagocytosing populations after treatment with the two FITC-labeled particles is represented (Fig. 2A). Phagocytic activity was observed as early as 24 hpf, and zymosan was the most phagocytosed particle (Fig. 2B).

Results obtained using flow cytometry were confirmed with fluorescence microscopy at 24 and 48 hpf. At 24 hpf (Fig. 3), phagocytosis of both *E. coli* and zymosan labeled with pHrodo and Texas Red, respectively, was observed. Phagocytosis occurred at a higher rate at 48 hpf than at 24 hpf. The early appearance of phagocytosis is in agreement with what has been previously observed in *Crassostrea virginica* larvae, where two different types of blood cells, SER cells (non-phagocytic cells) and phagocytic cells with intracytoplasmic granules and ameboid processes, were capable of internalizing India ink and observed as early as 24 hpf. Phagocytes from oyster larvae were also able to engulf cellular debris and foreign organisms [29,30]. Phagocytosis could be related to tissue reorganization during development [31], supporting the early appearance of this ability in the larval ontogeny.

Although the production of ROS and NO is considered to be associated with phagocytic processes and has been described in bivalve hemocytes [32–34], we were not able to detect triggering of these activities in larvae after stimulation with zymosan (Fig. 4). Basal levels of both reactive intermediates appeared high (a high percentage of fluorescence larvae were detected after treatment with the CM-H_2DCFDA or DAF-FM probes) but did not increase after stimulation (data not shown). The functions of ROS and NO are not restricted to immunity [35]. ROS have been implicated in cellular signaling and hematopoiesis in mice [36] and *Drosophila* [37], and nitric oxide is associated with metamorphosis in several different invertebrates [38,39]. High NO concentrations seem to block metamorphosis [1,40], and inhibitors of NO production such as L-NAME can induce metamorphosis of the sea urchin *Lytechinus pictus* [1]. The lack of modulation of ROS and NO by zymosan in our experiments could be in agreement with these observations and suggests that reactive radical production is tightly regulated during these early developmental stages.

3.3. Expression throughout development

Analysis of the constitutive expression of the selected immune-related genes throughout *M. galloprovincialis* larval development is summarized in Figs. 5 and 6. In general, the expression of these genes increases during the transition from trochophore larvae to later developmental stages, reaching the maximum in adult hemocytes, specialized cells known to be involved in immune defense [41] with high expression levels of immune-related genes [22,23,42,43]. This increase throughout development might suggest the onset of immune competence, in parallel with larval

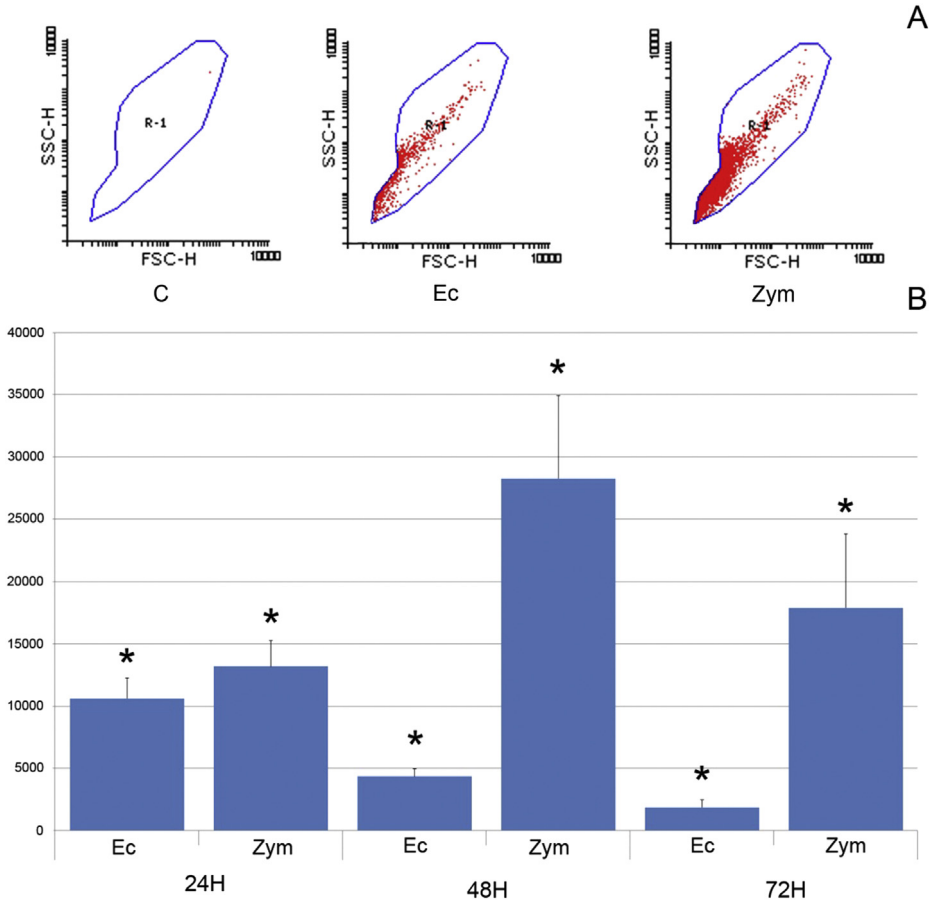


Fig. 2. Phagocytosis in *Mytilus galloprovincialis* larvae. A. Representative dot plot illustration of phagocytosing cells (FL-1 fluorescence events) at 24 hpf according to the cell size (FSC-H) and the cell granularity (SSC-H). The treatments are FSW (C), *Escherichia coli* (Ec) and zymosan (Zym). B. Phagocytosis is represented as the percentage of fluorescence events (FL-1) for each time and treatment normalized to the control (100%). The results are presented as the mean of at least two different families and three different technical replicates ± square error of the mean. Asterisks denote significant differences ($p < 0.01$).

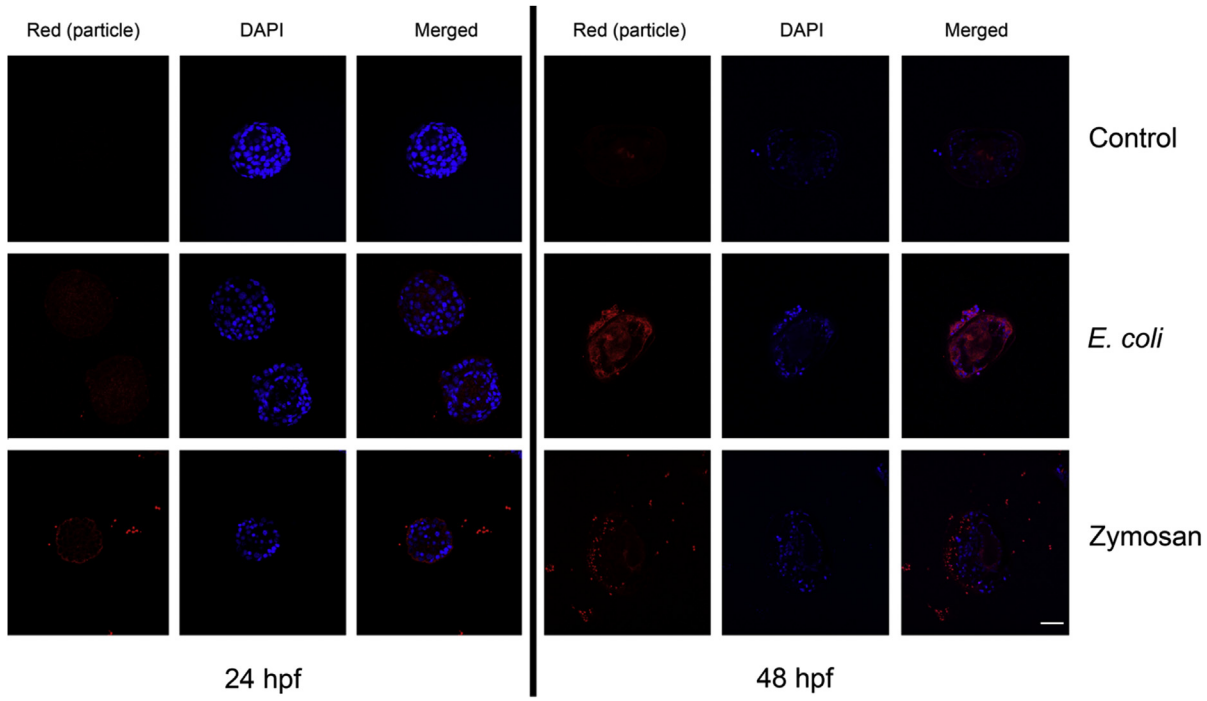


Fig. 3. Microscopic detection of larval phagocytosis. Mussel larvae phagocytosis was examined in complete larvae using pHrodo-labeled *Escherichia coli* and Texas Red-labeled zymosan. Red fluorescence of labeled particles is shown in columns 1 and 4 for 24 hpf and 48 hpf larvae, respectively. Nuclear DAPI staining and merged images of red and blue fluorescence are also shown in columns 2 and 3 for 24 hpf larvae and columns 5 and 6 for 48 hpf, respectively. Control larvae were treated with FSW. Scale bar is 25 μm in all pictures.

competence, and the preparation of the mussel planktonic larvae for an adult benthic life. Interestingly, there was a clear shift at the metamorphosis stage: most of the genes had low expression levels related to oocytes and were upregulated as development progressed, a change that persisted in adult hemocytes. This finding suggests that at this stage, animals not only modify their body structure and behavior but also acquire a clear immunocompetence by expressing a complete repertoire of immune-related genes. There were two genes, lysozyme and mytimicin, that showed a different expression pattern, with lower expression in all developmental stages compared with oocytes or hemocytes. This result might indicate maternal inheritance (high expression values in oocytes) probably related to immunoprotection of the progeny, although this phenomenon requires further research. Parental investment in the immunoprotection of offspring has been observed in other mollusks [44], and after development, planktotrophic larvae can begin using a fraction of the energy obtained by active feeding for active transcription. Finally, MacP showed high expression levels in all samples compared to oocyte expression, with an onset of expression observed in mussel veliger larvae, which is in agreement with previous studies [43]. This finding suggests possible involvement of this gene in development, as has been reported in other bivalves for genes with a similar expression pattern, such as the metalloprotease inhibitor from *Crassostrea gigas*, *Cg-timp* [16]. The basal expression of immune-related genes was confirmed by the WISH studies (Fig. 6). The trochophore larvae show a conspicuous staining pattern in which some of the larvae are completely stained, whereas others are not. This discrepancy could be due to individual variability at the onset of larval gene expression. During development, the variability among larvae is reduced, with the veliger larvae showing specific staining in the mantle edge. This pattern of expression is very different when compared to the positive control actin, which is expressed in all the mussel cells but was mainly observed in the muscles on the interior of the shell, specially in the veliger stage. However, lysozyme and mytimicin were exceptions to this expression pattern; lysozyme was not detected during the veliger stage, and mytimicin showed very low expression in the velum, further supporting the qPCR results. The majority of gene expression observed in pediveliger larvae stained in a patchy pattern along the mantle edge (Myticin B and C, Mytilin B, MgC1q and MacP), although expression of lysozyme and MIF was observed in the digestive gland, and FREP and mytimycin were expressed through the entire larval body in a pattern similar to the positive control actin.

Mussel larval development is rapid and is driven by energy stored by the oocyte, as well as protein and mRNA of maternal origin. After planktotrophic larvae begin actively feeding, the larvae synthesize their own mRNA and proteins, acquiring competence and preparing the animals to settle down and begin benthic life. This anticipatory pathway allows larvae to complete metamorphosis in a very short time [45]. Larvae in a metamorphic state are weak in terms of defense against predators and likely also against pathogens [46]. At this metamorphic stage, expression levels of immune-related genes were higher than in oocytes, suggesting active immune gene expression in mussel larvae. Previous mollusk ontogeny studies showed that this pattern of gene expression is followed by the expression of several proteins, both immune- and non-immune-related [16,47,48]. The upregulation of innate immune-related genes during metamorphosis has been previously reported in *Boltenia villosa*, highlighting the role of innate immunity during ascidian metamorphosis [49]. The upregulation observed during metamorphosis might simply reflect the maturation of the innate immune system but might also be related to the resorption and reorganization of larval tissues, as well as the larval ability to detect and respond to bacterial settlement cues.

WISH revealed the individual pattern of immune-related gene expression and provided a more realistic representation of gene expression than qPCR. Although several studies have been

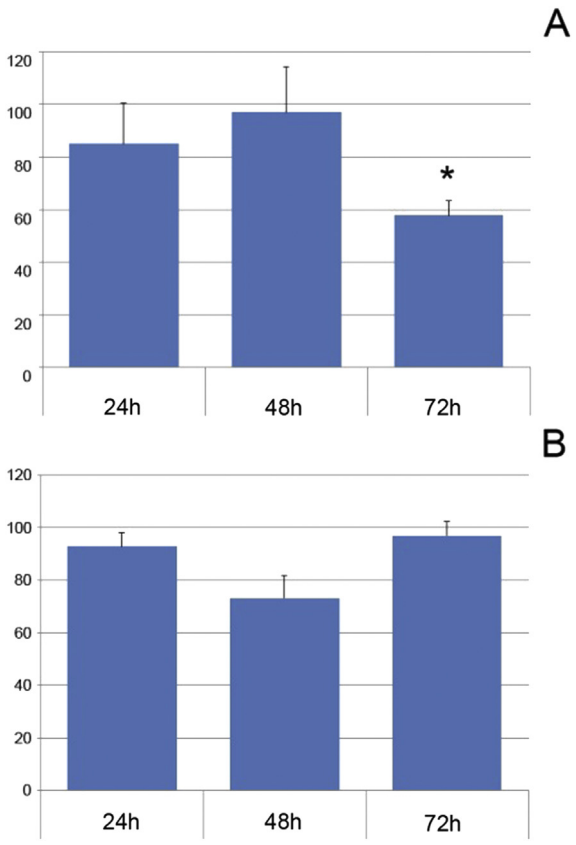


Fig. 4. Radical production by *Mytilus galloprovincialis* larvae. A. ROS production after zymosan stimulation in 24 hpf, 48 hpf or 72 hpf larvae. The results are presented as the percentage of ROS production in the absence of stimulus (100%) of at least two different mussel families. B. RNS production after zymosan stimulation in 24 hpf, 48 hpf or 72 hpf larvae. The results are presented as the percentage of RNS production in the absence of stimulus (100%). The results are presented as the mean ± square error of the mean for at least two different mussel families, and asterisks denote significantly different treatments ($p < 0.05$).

published concerning different WISH protocols in bivalves [50–52], to our knowledge, the present work depicts the largest array of gene expression examined by WISH in any mollusk species. Strikingly, several of the immune-related genes analyzed were expressed in the mantle edge (Fig. 6). In *C. virginica*, the accumulation of non-phagocytosing cells has been observed near the thickened mantle region of the hinge, as well as the presence of mitotic figures [30]. Strong labeling of the veliger mantle edge seen by WISH of the hemocyte-expressed mRNAs Myticin B and C and FREP [22,23,42] suggests that accumulation of hemocytes is occurring in this region during larval stages because these genes are expressed mainly in hemocytes.

3.4. Modulation of immune expression in challenged larvae

Immune challenge with heat-inactivated (Fig. 7A) or live *V. anguillarum* (Fig. 7B) induced differential expression of the selected immune-related genes. Stimulation with heat-inactivated *V. anguillarum* induced downregulation of gene expression compared to control larvae for most of the immune-related genes during the planktotrophic stages except for MIF, lysozyme, MacP and FREP in trochophores. In veliger larvae, the expression of immune genes was similar for heat-inactivated and live *V. anguillarum* stimulation, where we observed the downregulation of all genes examined. It is likely that these larvae invest their energetic resources in fast metabolic development to reach settlement and to develop competence [46]. The modulation observed in *C. gigas* during early stages [16] might be related to differences in habitat (sandy shallow waters for oysters and rough rocky shores for mussels). However, larvae in metamorphosis showed important differences in response to heat-inactivated and live *V. anguillarum* stimulation. Heat-inactivated bacteria induced a mild upregulation of immune-related genes, whereas live *V. anguillarum* induced downregulation of all genes except MMG1 and FREP. This effect was also observed in spat, where heat-inactivated *V. anguillarum* induced upregulation of all of the studied genes, whereas live *V. anguillarum* induced a decrease in gene expression. This result

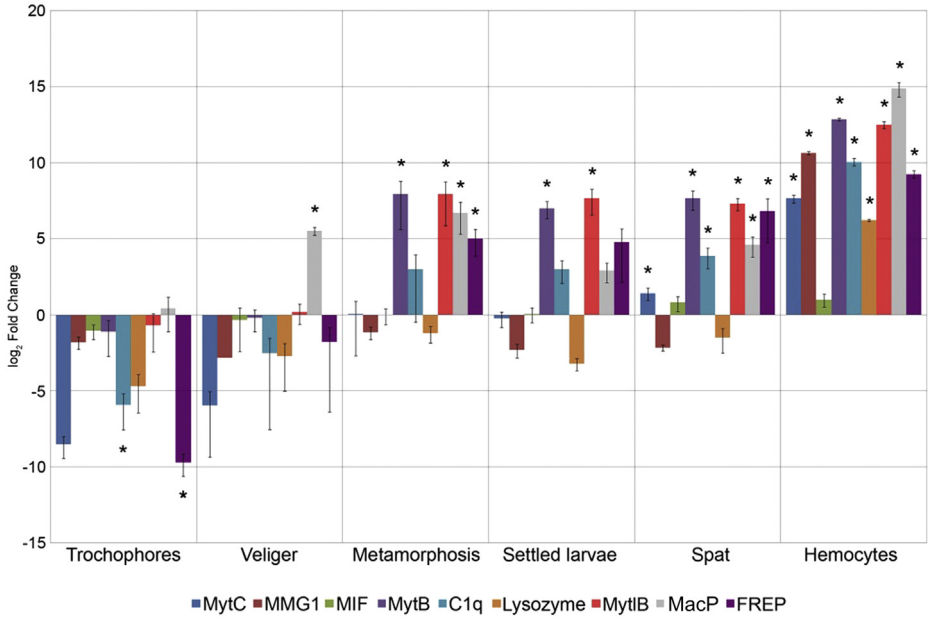


Fig. 5. Constitutive expression levels of selected immune-related genes throughout different larval stages and in adult hemocytes. Fold change data for each developmental stage related to oocyte expression have been log2 transformed to facilitate the visual depiction of downregulation. The results are presented as the mean ± square error of the mean of three different biological replicates. Asterisks indicate significant differences ($p < 0.05$) in gene expression. Myt C: Myticin C, MMG1: Mytimycin Precursor 1, MIF: Macrophage migration inhibition factor, Myt B: Myticin B, C1q: Complement C1q, Lysozyme, Mytl B: Mytilin B, MacP: Membrane Attack Complex/perforin, FREP: Fibrinogen-related protein.

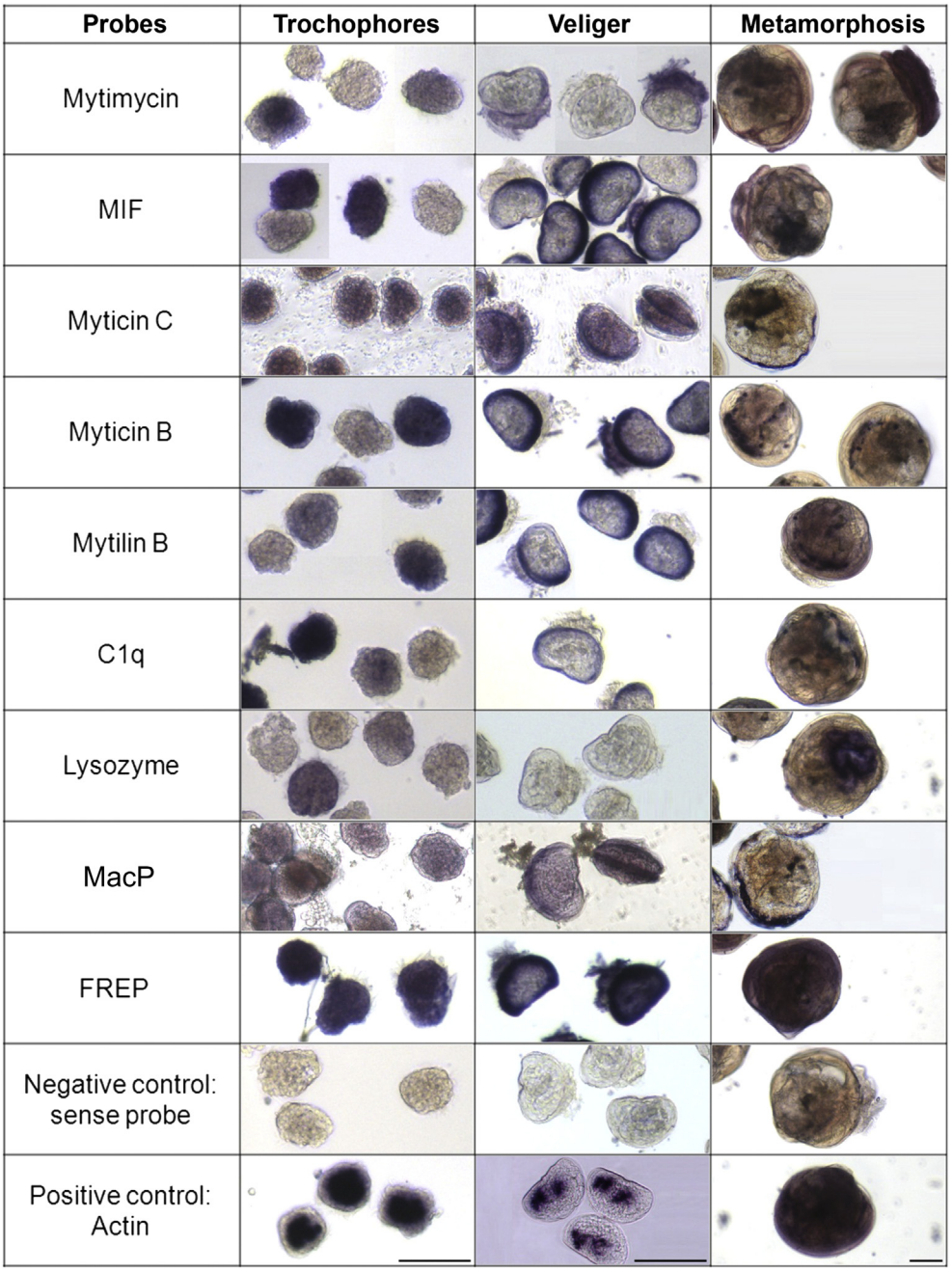


Fig. 6. Whole-mount *in situ* hybridization of selected immune-related genes throughout different larval stages. Antisense probes for each gene are shown for trochophore, veliger and metamorphosed larvae. The corresponding *in situ* hybridization using a sense probe for each gene was performed as the negative control, and a representative figure is shown. Actin antisense probe was used as the positive control. Scale bar: 100 μm.

suggests a mechanism for antimicrobial evasion by live but not heat-inactivated *V. anguillarum*, as it has been described for other virulent bacteria such as *Vibrio parahaemolyticus* [53]. This effect seemed more drastic in larvae, perhaps because of the stress associated with metamorphosis. Our data support that bacterial stimulus can trigger an appropriate immune response in larvae, as previously described for the adult mussels [54,55].

In summary, *M. galloprovincialis* immune capacities arise during mussel development as early as the trochophore stage. At this developmental stage, gene expression has contributions of maternal origin, but stimulation induces the expression of immune-related genes. Gene expression starts to rise after the veliger stage, preparing larvae for settlement. At this stage, the expression of immune-related genes characteristic of hemocytes is localized to the mantle edge. Further research is necessary in order to clarify if the accumulation of hemocytes in the mantle edge arises from hemocyte migration or if this organ could have a role in hematopoiesis. At metamorphosis, there is an important switch in the gene expression pattern, which is more similar to adults when examining both constitutive expression and the response to bacterial stimuli. The upregulation of immune-related genes observed during metamorphosis might reflect the maturation of the innate immune system or the reorganization of larval tissues but also plays an active role in immunity at this stage.

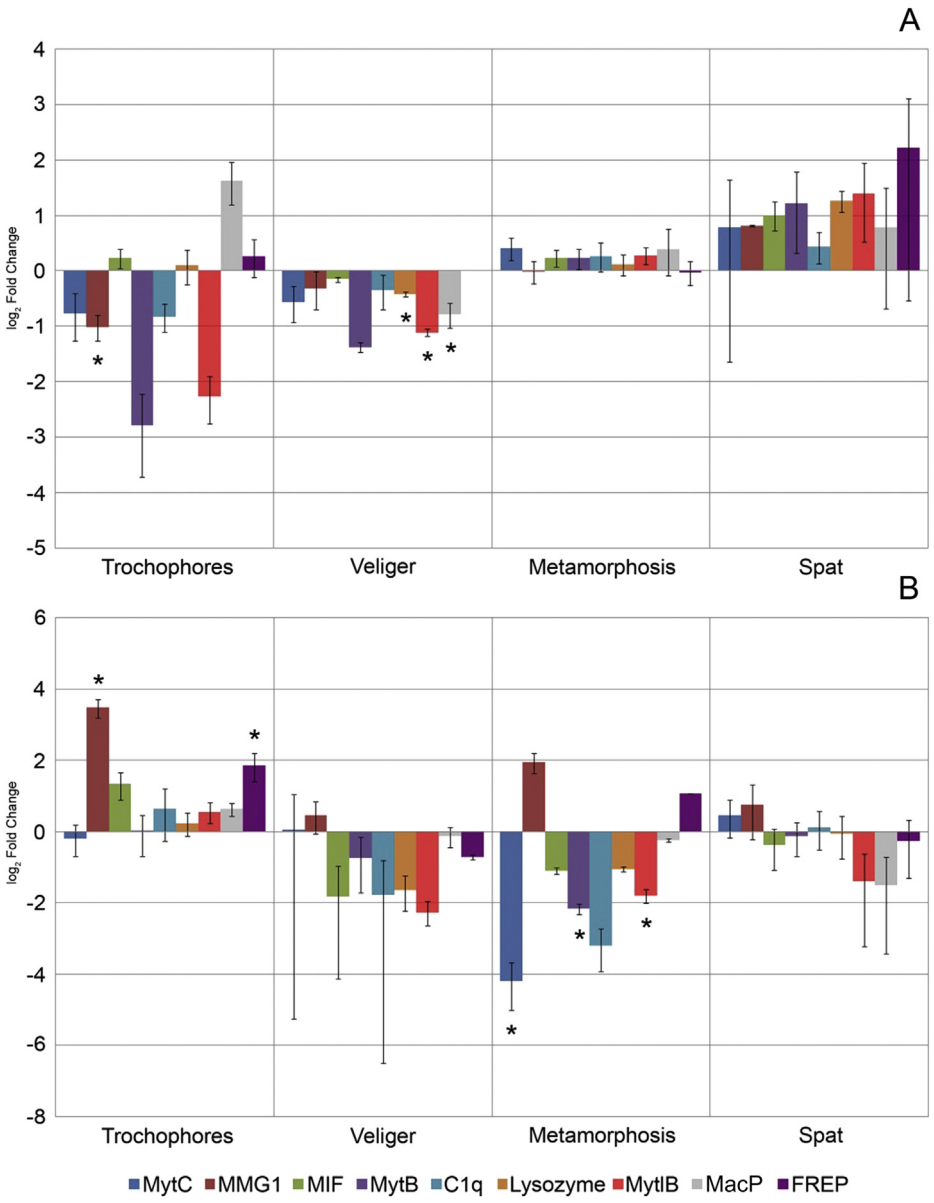


Fig. 7. Modulation of the expression of selected immune-related genes in larval stages challenged with *Vibrio anguillarum*. Gene expression fold change data for stimulated larvae compared to non-stimulated larvae were log2 transformed. A. Heat-inactivated *V. anguillarum*. B. Live *V. anguillarum*. The results are presented as the mean ± square error of the mean for three different mussel families performed in triplicate. Asterisks indicate significant differences ($p < 0.05$) in gene expression. Myt C: Myticin C, MMG1: Mytimycin Precursor 1, MIF: Macrophage migration inhibition factor, Myt B: myticin B, C1q: Complement C1q, Lysozyme, Mytl B: Mytilin B, MacP: Membrane Attack Complex/perforin, FREP: Fibrinogen-related protein.

Acknowledgements

This work was funded by the EU Project REPROSEED (FP7 245119) and European structural funds (FEDER) / Ministerio de Ciencia e Innovación (CSIC08-1E-102). R.M. wishes to acknowledge additional funding from the Spanish MICINN through an FPI Spanish research grant (BES-2009-029765). We want to acknowledge Dr. Alejandro Romero for flow cytometry assistance and Noelia Estévez-Calvar for technical assistance.

References

[1] Bishop CD, Huggett MJ, Heyland A, Hodin J, Brandhorst BP. Interspecific variation in metamorphic competence in marine invertebrates: the significance for comparative investigations into the timing of metamorphosis. Integr Comp Biol 2006;46:662–82.

[2] Halanych KM. The new view of animal phylogeny. Annu Rev Ecol Evol Syst 2004;35:229–56.

[3] Eckman JE. Closing the larval loop: linking larval ecology to the population dynamics of marine benthic invertebrates. J Exp Mar Biol Ecol 1996;200:207–37.

[4] Lambert C, Nicolas JL, Cilia V, Corre S. *Vibrio pectenicida* sp. nov., a pathogen of scallop (*Pecten maximus*) larvae. Int J Syst Bacteriol 1998;48:481–7.

[5] Shia AK, Glittenberg M, Thompson G, Weber AN, Reichhart JM, Ligoxygakis P. Toll-dependent antimicrobial responses in *Drosophila* larval fat body require Spatzle secreted by haemocytes. J Cell Sci 2009;122:4505–15.

[6] Makki R, Meister M, Pennetier D, Ubeda J-M, Braun A, Daburon V, et al. A short receptor downregulates JAK/STAT signalling to control the *Drosophila* cellular immune response. PLoS Biol 2010;8:e1000441.

[7] Ogawa A, Streit A, Antebi A, Sommer RJ. A conserved endocrine mechanism controls the formation of Dauer and infective larvae in nematodes. Curr Biol 2009;19:67–71.

[8] FAO. Fisheries and aquaculture topics. The State of World Fisheries and Aquaculture (SOFIA). In: Pulvenis J-F, editor. Topics fact sheets. Rome, Italy: FAO Fisheries and Aquaculture Department, Food and Agricultural Organization (FAO); 2012. p. 150.

[9] Wootton EC, Dyrynda EA, Pipe RK, Ratcliffe NA. Comparisons of PAH-induced immunomodulation in three bivalve molluscs. Aquat Toxicol 2003;65:13–25.
[10] Morgan JAT, Dejong RJ, Snyder SD, Mkogi GM, Loker ES. *Schistosoma mansoni* and *Biomphalaria*: past history and future trends. Parasitology 2001;123:211–28.
[11] Degnan SM, Degnan BM. The initiation of metamorphosis as an ancient polyphenic trait and its role in metazoan life-cycle evolution. Phil Trans R Soc B 2010;365:641–51.
[12] Williams EA, Degnan SM. Carry-over effect of larval settlement cue on post-larval gene expression in the marine gastropod *Haliotis asinina*. Mol Ecol 2009;18:4434–49.
[13] Dyachuk VA, Plotnikov SV, Odintsova NA. Appearance of muscle proteins in ontogenesis of the mussel *Mytilus trossulus* (Bivalvia). Russ J Mar Biol/Biol Morya 2005;31:327–30.
[14] Ellis I, Kempf SC. Characterization of the central nervous system and various peripheral innervations during larval development of the oyster *Crassostrea virginica*. Invertebr Biol 2011;130:236–50.
[15] Jenny MJ, Warr GW, Ringwood AH, Baltzegar DA, Chapman RW. Regulation of metallothionein genes in the American oyster (*Crassostrea virginica*): ontogeny and differential expression in response to different stressors. Gene 2006;379:156–65.
[16] Tirapé A, Bacque C, Brizard R, Vandenbulcke F, Boulo V. Expression of immune-related genes in the oyster *Crassostrea gigas* during ontogenesis. Dev Comp Immunol 2007;31:859–73.
[17] Montagnani C, Tirape A, Boulo V, Escoubas JM. The two Cg-timp mRNAs expressed in oyster hemocytes are generated by two gene families and differentially expressed during ontogenesis. Dev Comp Immunol 2005;29:831–9.
[18] Dyrynda EA, Pipe RK, Ratcliffe NA. Host defence mechanisms in marine invertebrate larvae. Fish Shellfish Immunol 1995;5:569–80.
[19] Hadfield MG, Carpizo-Ituarte EJ, del Carmen K, Nedved BT. Metamorphic competence, a major adaptive convergence in marine invertebrate larvae. Am Zool 2001;41:1123–31.
[20] Pawlik JR. Chemical ecology of the settlement of benthic marine invertebrates. Oceanogr Mar Biol Annu Rev 1992;30:273–335.
[21] Degnan BM, Morse DE. Developmental and morphogenetic gene regulation in *Haliotis rufescens* larvae at metamorphosis. Am Zool 1995;35:391–8.
[22] Romero A, Dios S, Poisa-Beiro L, Costa MM, Posada D, Figueras A, et al. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. Dev Comp Immunol 2011;35:334–44.
[23] Balseiro P, Falcó A, Romero A, Dios S, Martínez-López A, Figueras A, et al. *Mytilus galloprovincialis* myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS ONE 2011;6:e23140.
[24] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.
[25] Ruiz M, Tarifeno E, Llanos-Rivera A, Padget C, Campos B. Temperature effect in the embryonic and larval development of the mussel, *Mytilus galloprovincialis* (Lamarck, 1819). Rev Biol Mar Oceanogr 2008;43:51–61.
[26] Sánchez-Lazo C, Martínez-Pita I. Effect of temperature on survival, growth and development of *Mytilus galloprovincialis* larvae. Aquacult Res 2012;43:1127–33.
[27] His E, Robert R, Dinet A. Combined effects of temperature and salinity on fed and starved larvae of the Mediterranean mussel *Mytilus galloprovincialis* and the Japanese oyster *Crassostrea gigas*. Mar Biol 1989;100:455–63.
[28] Satuito CG, Natoyama K, Yamazaki M, Fusetani N. Larval development of the mussel *Mytilus edulis galloprovincialis* cultured under laboratory conditions. Fish Sci 1994;60:65–8.
[29] Elston R. Functional anatomy, histology and ultrastructure of the soft tissues of the larval American oyster, *Crassostrea virginica*. Proc Natl Shellfish Ass 1980;70:65–93.
[30] Elston R. Functional morphology of the coelomocytes of the larval oysters (*Crassostrea virginica* and *Crassostrea gigas*). J Mar Biol Assoc UK 1980;60:947–57.
[31] Vernon PJ, Tang D. Eat-me: autophagy, phagocytosis, and reactive oxygen species signaling. Antioxid Redox Signal 2012.
[32] Prado-Álvarez M, Romero A, Balseiro P, Dios S, Novoa B, Figueras A. Morphological characterization and functional immune response of the carpet shell clam (*Ruditapes decussatus*) haemocytes after bacterial stimulation. Fish Shellfish Immunol 2012;32:69–78.
[33] García-García E, Prado-Álvarez M, Novoa B, Figueras A, Rosales C. Immune responses of mussel hemocyte subpopulations are differentially regulated by enzymes of the PI 3-K, PKC, and ERK kinase families. Dev Comp Immunol 2008;32:637–53.
[34] Torreilles J, Guerin MC, Roch P. Peroxidase-release associated with phagocytosis in *Mytilus galloprovincialis* haemocytes. Dev Comp Immunol 1997;21:267–75.
[35] Sardina JL, López-Ruano G, Sánchez-Sánchez B, Llanillo M, Hernández-Hernández A. Reactive oxygen species: are they important for haematopoiesis? Crit Rev Oncol/Hematol 2012;81:257–74.
[36] O'Brien JJ, Spinelli SL, Tober J, Blumberg N, Francis CW, Taubman MB, et al. 15-deoxy-Δ12,14-PGJ2 enhances platelet production from megakaryocytes. Blood 2008;112:4051–60.
[37] Owusu-Ansah E, Banerjee U. Reactive oxygen species prime *Drosophila* haematopoietic progenitors for differentiation. Nature 2009;461:537–41.
[38] Kuzin B, Roberts I, Peunova N, Enikolopov G. Nitric oxide regulates cell proliferation during *Drosophila* development. Cell 1996;87:639–49.
[39] Enikolopov G, Banerji J, Kuzin B. Nitric oxide and *Drosophila* development. Cell Death Differ 1999;6:956–63.
[40] Hodin J. Expanding networks: signaling components in and a hypothesis for the evolution of metamorphosis. Integr Comp Biol 2006;46:719–42.
[41] Cheng TC. Bivalves. In: Ratcliffe NA, Rowley AF, editors. Invertebrate blood cells. London: Academic Press; 1981. p. 233–300.
[42] Mitta G, Vandenbulcke F, Noel T, Romestand B, Beauvillain JC, Salzet M, et al. Differential distribution and defence involvement of antimicrobial peptides in mussel. J Cell Sci 2000;113(Pt 15):2759–69.
[43] Estévez-Calvar N, Romero A, Figueras A, Novoa B. Involvement of pore-forming molecules in immune defense and development of the Mediterranean mussel (*Mytilus galloprovincialis*). Dev Comp Immunol 2011;35:1017–31.
[44] Hathaway JJM, Adema CM, Stout BA, Mobarak CD, Loker ES. Identification of protein components of egg masses indicates parental investment in immunoprotection of offspring by *Biomphalaria glabrata* (Gastropoda, Mollusca). Dev Comp Immunol 2010;34:425–35.
[45] Jackson D, Leys SP, Hinman VF, Woods R, Lavin MF, Degnan BM. Ecological regulation of development: induction of marine invertebrate metamorphosis. Int J Dev Biol 2002;46:679–86.
[46] Hadfield MG. Why and how marine-invertebrate larvae metamorphose so fast. Semin Cell Dev Biol 2000;11:437–43.
[47] Zhou Z, Wang L, Shi X, Yue F, Wang M, Zhang H, et al. The expression of dopa decarboxylase and dopamine beta hydroxylase and their responding to bacterial challenge during the ontogenesis of scallop *Chlamys farreri*. Fish Shellfish Immunol 2012;33:67–74.
[48] Ge H, Wang G, Zhang L, Zhang Z, Wang S, Zou Z, et al. Molecular cloning and expression of interleukin-1 receptor-associated kinase 4, an important mediator of toll-like receptor signal pathway, from small abalone *Haliotis diversicolor*. Fish Shellfish Immunol 2011;30:1138–46.
[49] Davidson B, Swalla BJ. A molecular analysis of ascidian metamorphosis reveals activation of an innate immune response. Development 2002;129:4739–51.
[50] Andersen Ø, Torgersen J, Pagander H, Magnesen T, Johnston I. Gene expression analyses of essential catch factors in the smooth and striated adductor muscles of larval, juvenile and adult great scallop (*Pecten maximus*). J Muscle Res Cell Motil 2009;30:233–42.
[51] Fabioux C, Huvet A, Lelong C, Robert R, 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.
[52] Wang X, Liu B, Xiang J. Cloning, characterization and expression of ferritin subunit from clam *Meretrix meretrix* in different larval stages. Comp Biochem Physiol B Biochem Mol Biol 2009;154:12–6.
[53] Rahman MM, McFadden G. Modulation of NF-κB signalling by microbial pathogens. Nat Rev Micro 2011;9:291–306.
[54] Li H, Venier P, Prado-Álvarez M, Gestal C, Toubiana M, Quartesan R, et al. Expression of *Mytilus* immune genes in response to experimental challenges varied according to the site of collection. Fish Shellfish Immunol 2010;28:640–8.
[55] Costa MM, Prado-Álvarez 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 Shellfish Immunol 2009;26:515–23.
[56] Mitta G, Hubert F, Dyrynda EA, Boudry P, Roch P. Mytilin B and MGD2, two antimicrobial peptides of marine mussels: gene structure and expression analysis. Dev Comp Immunol 2000;24:381–93.

RNA-SEQ IN *MYTILUS GALLOPROVINCIALIS*: COMPARATIVE TRANSCRIPTOMICS AND EXPRESSION PROFILES AMONG DIFFERENT TISSUES

Moreira et al. BMC Genomics (2015) 16:728
DOI 10.1186/s12864-015-1817-5

BMC Genomics

RESEARCH ARTICLE Open Access

RNA-Seq in *Mytilus galloprovincialis*: comparative transcriptomics and expression profiles among different tissues

Rebeca Moreira[1], Patricia Pereiro[1], Carlos Canchaya[2], David Posada[2], Antonio Figueras[1] and Beatriz Novoa[1*]

Abstract

Background: The Mediterranean mussel (*Mytilus galloprovincialis*) is a cosmopolitan, cultured bivalve with worldwide commercial and ecological importance. However, there is a qualitative and quantitative lack of knowledge of the molecular mechanisms involved in the physiology and immune response of this mollusc. In order to start filling this gap, we have studied the transcriptome of mantle, muscle and gills from naïve Mediterranean mussels and hemocytes exposed to distinct stimuli.

Results: A total of 393,316 million raw RNA-Seq reads were obtained and assembled into 151,320 non-redundant transcripts with an average length of 570 bp. Only 55 % of the transcripts were shared across all tissues. Hemocyte and gill transcriptomes shared 60 % of the transcripts while mantle and muscle transcriptomes were most similar, with 77 % shared transcripts. Stimulated hemocytes showed abundant defense and immune-related proteins, in particular, an extremely high amount of antimicrobial peptides. Gills expressed many transcripts assigned to both structure and recognition of non-self patterns, while in mantle many transcripts were related to reproduction and shell formation. Moreover, this tissue presented additional and interesting hematopoietic, antifungal and sensorial functions. Finally, muscle expressed many myofibril and calcium-related proteins and was found to be unexpectedly associated with defense functions. In addition, many metabolic routes related to cancer were represented.

Conclusions: Our analyses indicate that whereas the transcriptomes of these four tissues have characteristic expression profiles in agreement with their biological structures and expected functions, tissue-specific transcriptomes reveal a complex and specialized functions.

Keywords: *Mytilus galloprovincialis*, Transcriptome, NGS, RNA-Seq, NOISeq, KEGG, Gene Ontology, Blast2GO

Background

The Mediterranean mussel (*Mytilus galloprovincialis*) is a cultured bivalve species with an important commercial and ecological value worldwide [1, 2]. In contrast to other cultured bivalves such as clams and oysters, where different pathogens may result in massive mortalities and therefore, substantial economic losses [3–5], *M. galloprovincialis* displays an extraordinary resistance to a variety of pathogens [6]. Although molluscs lack a specific immune response, their innate response, which involves circulating hemocytes and a large variety of molecular effectors, constitutes an efficient defense mechanism [7–9]. While a wide range of molecules involved in the bivalve immune system have been described [10–13], particularly for mussels and oysters [14–18], the information is very limited compared, for example, to vertebrates.

Unfortunately, most bivalve genomic resources are not annotated or well described, with the exception of the Pacific oyster, *Crassostrea gigas*, whose genome has been recently published [19] or the pearl oyster, *Pinctada fucata,* in which genome annotation is still at the draft level [20]. Several bivalve transcriptomes are publicly available for *M. galloprovincialis* [21, 22], *Bathymodiolus azoricus* [23], *Patinopecten yessoensis* [24], *Ruditapes philippinarum* [25, 26] *Corbicula fluminea* [27] and *Crassotrea gigas* [19, 28, 29]. There are also 666 entries from the class Bivalvia deposited in the NCBI Short

* Correspondence: beatriznovoa@iim.csic.es
[1]Instituto de Investigaciones Marinas (IIM), Consejo Superior de Investigaciones Científicas (CSIC), Eduardo Cabello, 6, 36208 Vigo, Spain
Full list of author information is available at the end of the article

Read Archive (SRA) (25/03/2015). The number of available sequences for *M. galloprovincialis* is constantly increasing [30–33]. As an example, 23 *M. galloprovincialis* entries are publicly available in the SRA database, including whole-body, digestive gland and hemocytes transcriptomes, a *M. galloprovincialis* EST database called Mytibase [16].

In this study we show the results of the first comparative RNA-Seq analysis of gene expression in different *M. galloprovincialis* tissues, including gills, muscle, mantle and hemocytes. The raw data are accessible from the NCBI Short Read Archive (SRA: SRP033481). Additional files 1, 2, 3, 4 and 5 include all the transcripts obtained, together with their expression values, annotation and sequences in FASTA format.

Results and discussion

Sequence analysis and functional annotation

Mussel samples were processed as depicted in Fig. 1. The 7 cDNA libraries obtained (2 from stimulated hemocytes, 2 from mantle, 2 from muscle and 1 from gills) were sequenced on the Illumina HiSeq™ 2000 platform.

The sequencing and assembly statistics are summarized in Table 1. Briefly, we obtained a total of 393.3 million raw reads (with an average of 56.2 million reads per run). Of these, more than 95 % passed the quality standards and were subjected to further analyses. The filtered high-quality reads were assembled in a three-step approach with the Trinity software [34] into 1,242,475 contigs, which after clustering resulted in 479,806 unigenes. Until this point, the assembly protocol was individual for each sample, but the third and last clustering step was performed in common for the 7 samples. A total of 151,320 non-redundant unigenes ("transcripts" hereafter) were obtained, which could represent the *M. galloprovincialis* global transcriptome for these 4 tissues.

The length of the transcripts ranged from 200 to 17,690 bp, with an average length of 570 bp, a similar size to that obtained with Roche 454 technology in other bivalves, e.g., 582 bp in the Manila clam [26]. Furthermore,

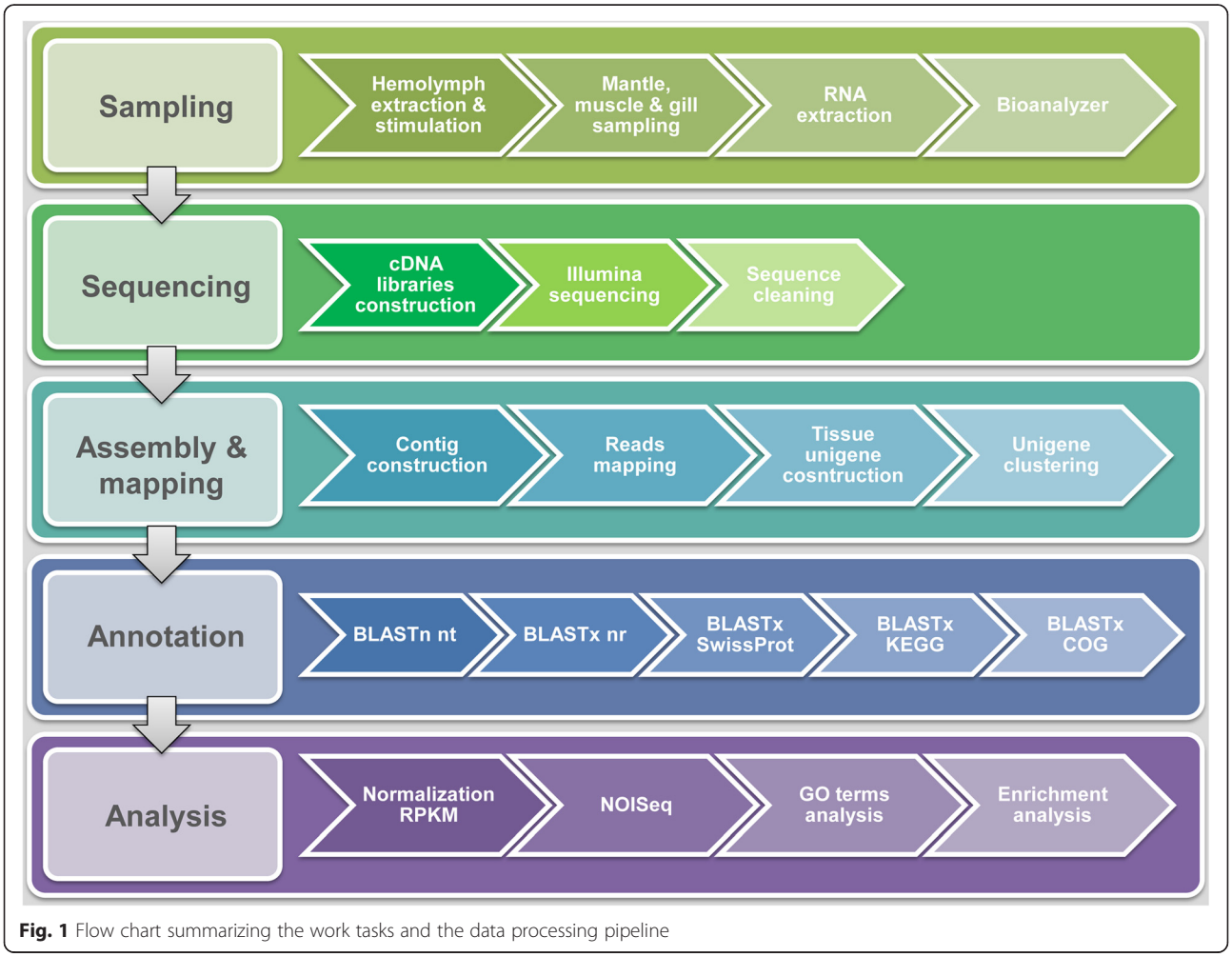


Fig. 1 Flow chart summarizing the work tasks and the data processing pipeline

Table 1 Summary of sequencing and assembly data

Sequencing statistics	Hemocytes	Mantle	Muscle	Gill
Millions of raw reads	112.706	111.322	113.045	56.244
Millions of clean reads	107.386	106.060	107.127	53.335
Total Megabases	9,665	9,545	9,641	4,800
% GC content	38.99 %	38.32 %	38.28 %	37.54 %
Assembly statistics				
Number of contigs	261,332	428,939	313,554	238,650
Tissue unigenes	107,045	131,935	120,572	120,254
	All			
Total number of transcripts	151,320			
Average transcript length	570			
N50 transcript length	774			
Range of transcript lengths	200 – 17,690			
Number of transcripts < 500pb	104,757			
Number of transcripts > 500pb	46,563			
Annotation statistics				
Annotated transcripts by nt	14,207 (9.4 %)			
Annotated transcripts by nr	45,182 (29.8 %)			
Annotated transcripts by SwissProt	36,656 (24.2 %)			
Annotated transcripts by KEGG	31,144 (20.6 %)			
Annotated transcripts by COG	14,503 (9.6 %)			
TOTAL annotated transcripts	50,998 (33.7 %)			
Transcripts with GO terms	18,899 (12.5 %)			

when we compared our results with SOLiD and Illumina RNA-Seq analyses conducted in oyster, we obtained larger transcripts than those reported by Gavery and Roberts [28] or Zhao et al. [29], averaging 276 bp (554 using GigasDatabase v8 as a reference for mapping) and 322 bp, respectively.

The NCBI's nucleotide and non-redundant, SwissProt, KEGG [35] and COG [36] databases were chosen to annotate the transcripts. The percentage of transcripts annotated with an e-value threshold of $1x10e^{-5}$ was 33.7 %. The annotations and expression values are included in Additional file 1. Our annotation percentage was similar to previous transcriptome studies conducted in bivalves using 454 technology [26], with 45 % of hemocyte transcripts being annotated with an e-value threshold of $1x10e^{-3}$. Similar approaches applied in oyster using the SOLiD [28] or Illumina [29] sequencing platform resulted in an annotation success of 41 % or 16 %, respectively, while in the *M. galloprovincialis* digestive gland transcriptome, about half (48.1 %) of the transcripts were successfully annotated [22].

The coverage of the whole transcript for each specific tissue sample (calculated as the percentage of base pairs in a transcript covered by reads of a specific sample per transcript length) is summarized in Table 2. The mean coverage was 69.87 %, with an average of 256.55 reads

Table 2 Coverage, mapping and new discoveries using a 2nd replicate

	Average coverage of transcripts	Average unique mapped reads	New transcripts in the 2nd replicate
Hemocytes1	67.58 %	279.50	8,890 (8.31 %)
Hemocytes2	62.94 %	334.05	
Mantle1	75.76 %	256.79	12,437 (9.43 %)
Mantle2	72.60 %	214.44	
Muscle1	69.48 %	248.45	13,726 (11.38 %)
Muscle2	66.73 %	221.61	
Gill1	73.99 %	241.03	-

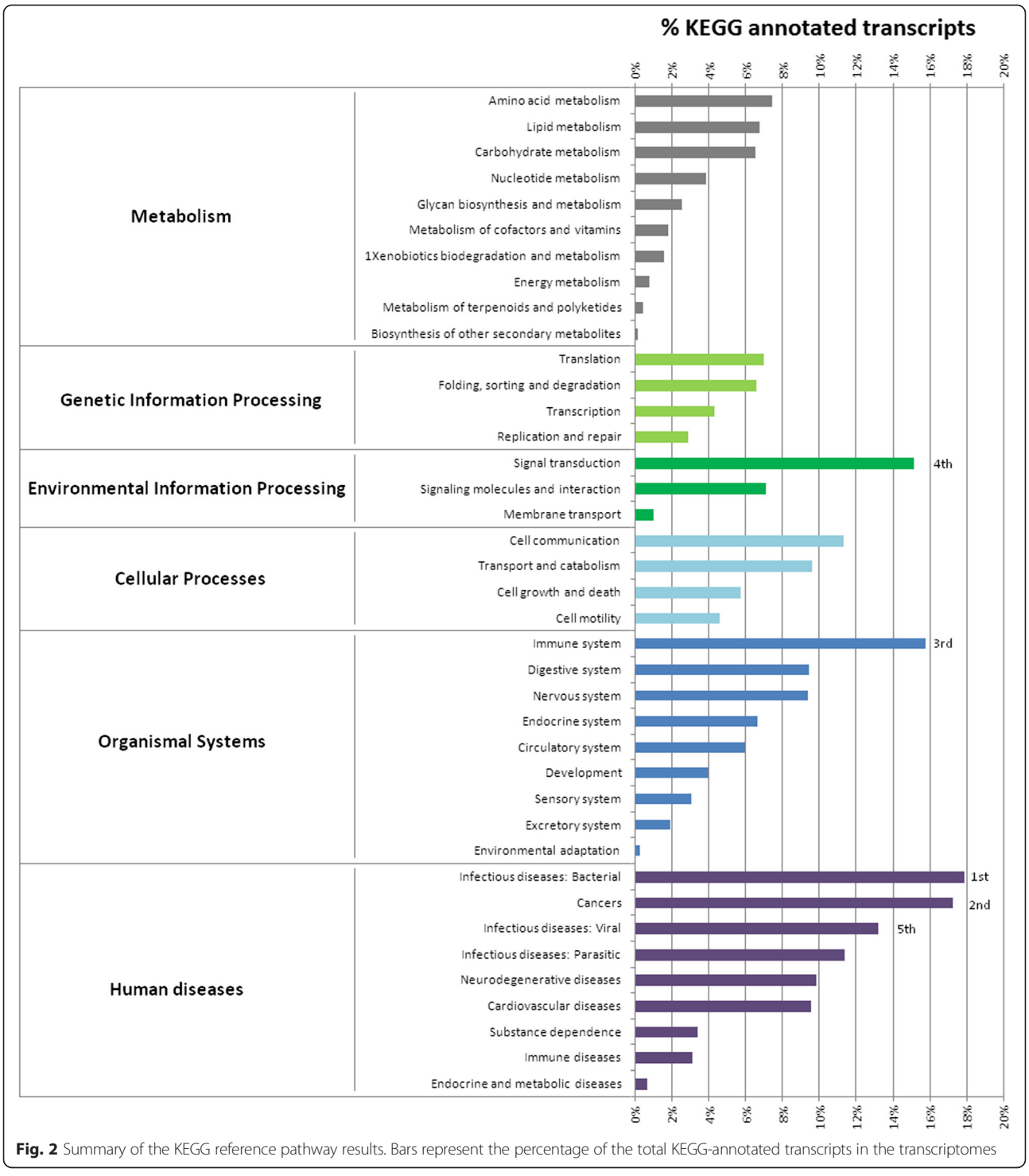


Fig. 2 Summary of the KEGG reference pathway results. Bars represent the percentage of the total KEGG-annotated transcripts in the transcriptomes

being mapped to each transcript, which is lower than the values reported in the oyster gills study by Gavery and Roberts [28], in which 454 reads per transcript were mapped (376 using GigasDatabase v8 as a reference). Hemocytes were the sample with the lowest coverage, but also showed the highest number of mapped reads per transcript. If this fact is related with the specific immune function of hemocytes, as was previously reported in other bivalves [26], is something that deserves further investigation.

We sequenced 2 samples of hemocytes, mantle and muscle, in order to understand whether a second biological

replicate would effectively increase the sequencing depth (Table 2). Although this second replicate resulted in an average increase of 9.7 % transcripts for each tissue, it did not result in a significantly higher number of transcripts when all tissues were considered (Additional file 2). However, the use of pools of individuals, different tissues and biological replicates increased the reliability and robustness of the results, as previously reported [37]. For example, we achieved a transcriptome completeness of 88.71 % or 95.16 % (considering whole or partial sequence comparisons, respectively) using the CEGMA package (http://korflab.ucdavis.edu/datasets/cegma/).

Qualitative description of the *M. galloprovincialis* transcriptome

Using KEGG, we annotated 31,144 transcripts (20.6 %). This annotation served as a basis for analyzing not only the role of individual transcripts, but also the interaction with other genes. Figure 2 provides a representation of the global functionality of the transcripts and summarizes the 256 molecular pathways found in the transcriptomes. It was interesting that a high number of these transcripts had annotations related to the immune system, signal transduction and infectious diseases (bacterial, viral and parasitic). A possible explanation for this could be that, as a filter feeding animal, *M. galloprovincialis* is permanently in contact with microorganisms and with toxic/pollutant substances in their marine environment [38], and has adapted to become very resistant to these impacts [39, 40]. Another group of disease-associated pathways were those related to cancer, which ranked second among the most represented pathways, like in other studies in oysters [29]. Interestingly, although mussels and oysters shared less than 10 % of proteins with a sequence identity over 80 % (Fig. 3), the response to infectious disease and cancer were highly represented in both transcriptomes [29]. Although these cancer-related genes may have other functions, this subject obviously requires further attention specially taking into account that some bivalves are affected by a disease of the circulatory system closely resembling leukemia [41, 42].

The information about the molecules that were present and absent in each pathway is available in Additional file 3. The specialization and diversification observed throughout the phylogeny of the immune system [43] suggests that the absence of some key molecules in the pathways can be an artifact. It is possible that they were not annotated or that other molecules could play a similar function.

A comparative analysis among the *M. galloprovincialis* transcriptomes was conducted to identify transcripts conserved in the 4 tissues and those unique to each tissue (Fig. 4). Among the total 151,320 transcripts, 54.57 % were shared by all the tissues. The most related pair of tissues, muscle and mantle, had 76.63 % transcripts in common, whereas hemocytes and gills shared only 59.56 % of the transcripts.

The tissue with the fewest private transcripts was muscle, with only 769 unique transcripts (0.51 %), while gills presented the highest number of non-shared transcripts, 9178 (6.07 %). This might be due to the filter feeding behavior of bivalves, where the gills are in constant contact with the surrounding habitat and exposed to more stress factors such as microorganisms, pollutants, pH or salinity changes.

Tissue-specific transcriptome portions are presented in Table 3. A high number of lectins, C1q domain-

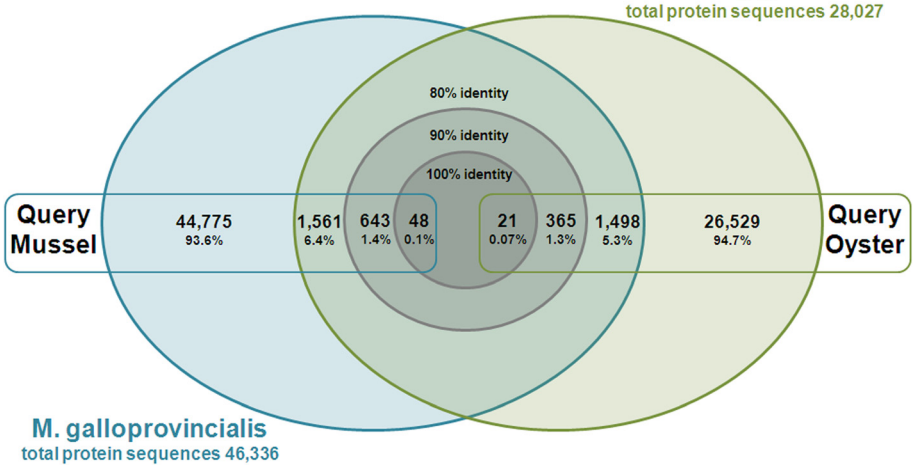


Fig. 3 Comparison of the translated *Mytilus galloprovincialis* sequences with the *Crassostrea gigas* proteome downloaded from http://www.oysterdb.com/FrontDownloadAction.do?method=download

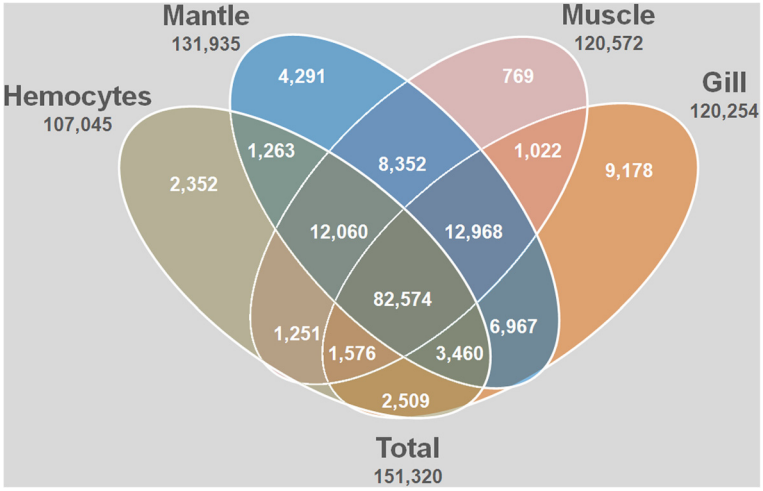


Fig. 4 Venn diagram showing a comparison of the *R. philippinarum* tissue transcriptomes: hemocytes, mantle, muscle and gills. Numbers refer to the transcripts that belong to each group

containing proteins and fibrinogen-related proteins were detected in gills. Their direct contact with the environment could explain the high presence of these putative recognition and immune-triggering molecules [11, 12, 44]. Hemocytes, as key players in the invertebrate immune response [45], showed a high percentage of antimicrobial peptides (AMPs), such as defensin, mytilins, and myticins, as well as other immune-related proteins, such as FREPs, serine protease inhibitors, complement component C4, HSP90 and C1q domain-containing proteins. Hemocyte hematopoiesis is a poorly described process in bivalves, but some studies suggest that heart and mantle could be possible hematopoietic tissues [46, 47]. In our case, peroxidasin transcripts, an early hematopoietic differentiation marker in *Drosophila* [48], were found only in the mantle, reinforcing the hypothesis that mantle could be the main hematopoietic tissue of bivalves.

The lowest number of tissue-specific transcripts was observed in muscle, which could be due to its limited functional diversity (Fig. 4). In this case, the most abundant transcripts corresponded to heat shock proteins (HSP90, HSP70, HSP40, HSP20, HSP24 and GprE) and ribosome-related sequences. Interestingly, some immune-related molecules, such as myticin C, ficolin-2, C1q domain-containing protein and the scavenger receptor MARCO, were also present in the muscle. In mammals, MARCO is a pattern recognition receptor for Gram-positive and negative bacteria expressed mainly in macrophages [49] and it has not been described in invertebrates to date.

Some tissue-specific transcripts presented different variants depending on the tissue. This could mean that the gene is not tissue-exclusive but instead tissue-exclusive variants may exist. This was clearly observed for the C1q domain-containing proteins, a group of molecules that show high variability in *M. galloprovincialis* [12]. The C1q annotation for all the non-shared transcripts did not coincide between tissues (Additional file 4), which could suggest a high specialization of this molecule in each tissue and, possibly, specialization for non-self recognition, as might be inferred from the high abundance of the C1q sequences in gills. Without the *M. galloprovincialis* genome we can not know if the C1q variants are different isoforms or belong to different loci, further research is needed to clarify this issue.

Quantitative analysis between tissues: RNA-Seq

The transcriptomes were also quantitatively analyzed. We first normalized the number of reads that were mapped to each transcript into RPKM units (Reads Per Kilobase of exon model per Million mapped reads). To evaluate differentially expressed genes (d.e.g.) among tissues we used NOISeq [50], a nonparametric statistical approach that presents a low false discovery rate. The expression distribution of all the transcripts is showed in Fig. 5. As it is reflected by the red color intensity in each chart, the most similar tissues are mantle and muscle (d) while the most different are hemocytes and gills (c). The pairwise comparisons between the four tissues are summarized in Table 4. In the tissues that exhibited more transcripts in common (mantle and muscle) only 256 d.e.g. were found. In contrast, the comparison between hemocytes and gill, the tissues with the most dissimilar transcriptomes, revealed almost 2000 significantly different transcripts.

The heatmap provided in Fig. 6 illustrates, as an example, the quantitative expression of 5 among the top-

Table 3 Top 25 non-shared transcripts

Reads	Hemocyte top 25 non-shared	Reads	Mantle top 25 non-shared
905.5	Apolipoprotein L	305	Von Willebrand factor D and EGF dom-contprot
501	PugilistDominant	293	Fibroin heavy chain
466	C1q dom-cont prot MgC1q28	283.5	C1q domain containing protein MgC1q95
300	Defensin	245	Nacrein B3
277.5	Toxin CrTX-A	215	ADAM family mig-17
268	Mytilin B	165.5	Fibrocystin L
255	Conodipine-M alpha chain	160.5	Gigasin-6
238.5	DNA ligase 1	152.5	C1q domain containing protein MgC1q69
159	Fibrinogen-related protein	149	ATP-dependent RNA helicase A
146	Mytilin-6	132	GTPase IMAP family member 8-like
145.5	Transcription antiterminator	120.5	Lactase-phlorizin hydrolase
141	Serine protease inhibitor Cvs.i-2	120.5	Processed variable antigen (Fragment)
125	Rossmann fold nucleotide-binding protein	116	TPR repeat
121	Reverse transcriptase-like protein	115.5	Peroxidasin homolog (Drosophila)-like
112.5	Complement component 4	114	Beta-hexosaminidase
107	ATP synthase subunit a	105	Nicotinic acetylcholine receptor alpha subunit
92	Heat shock protein 90 (HSP90-2)	103	LDL receptor-related protein 8 (LRP8)
81.5	Pol-like protein	101	Electrogenic NBC-like protein
73.5	C1q dom-cont prot MgC1q56	99.5	Basic proline-rich protein
72.5	Ribosome-associated protein Y (PSrp-1)	99.5	Inter-alpha-trypsin inhibitor heavy chain H5
70	ATP-dependent RNA helicase ddx41	94.5	Golgi-associated plant pathogenesis-related protein 1
69.5	Myticin C	94	Fatty acid synthase
63	Nephrin	93.5	Myb-related transcription factor, partner of profilin
61.5	Cytosolic phospholipase A2	91.5	C1q domain containing protein MgC1q48
56.5	Fibrinogen-related protein (FREP_G1)	89	RING finger protein 13
Reads	Muscle top 25 non-shared	Reads	Gill top 25 non-shared
107	Ribosomal RNA	6644	Perlucin-like protein
62	Heat shock protein 90 (HSP90-2)	3223	C1q domain containing protein MgC1q71
48	Mammaglobin-A precursor	2935	Yolk ferritin
44.5	Gill symbiont ribosomal RNA	1655	WSC domain-containing protein 2
36.5	28S ribosomal RNA gene, partial sequence	1423	Short-chain collagen C4 (Fragment)
31	Myticin C	1407	Apextrin-like protein
30.5	Angiopoietin-4	1342	Fibroin heavy chain
30	Basal body protein NBP-2	1248	GTPase IMAP family member 4
27	Stress-70 protein, mitochondrial-like	1030	Nicotinic acetylcholine receptor alpha subunit
25.5	Ficolin-2-like, partial	997	Fucolectin
23.5	Collagen alpha-2(I) chain	989	Collagen alpha-1(XII) chain
17.5	16S ribosomal RNA	969	Multiple EGF-like domains protein 6-like
17.5	Rps19	916	C1q domain containing protein MgC1q17
17	ABC protein, subfamily ABCC	897	C1q domain containing protein MgC1q52
17	Catecholamine binding protein	889	C1q domain containing protein MgC1q36
15.5	Oxidoreductase, FAD/FMN-binding family protein	879	Eggshell protein
14.5	Large exoprotein involved in heme utilization or adhesion	875	GTPase IMAP family member 7
14.5	Ribulose-phosphate 3-epimerase,	858	C1q domain containing protein MgC1q81

Table 3 Top 25 non-shared transcripts *(Continued)*

12.5	C1q domain containing protein MgC1q22	853	Codakine
11.5	Ribosomal protein L32	825	Cathepsin L
11.5	Small heat shock protein hspl, mitochondrial	799	C-type lectin
11	DnaJ homolog dnj-10	795	Fibrinogen-related protein
10.5	GrpE-like protein	772	Fibrinogen C domain-containing protein 1
10.5	Zn-finger domain associated with topoisomerase type I	742	Tetraspanin-CD63 receptor
9.5	Macrophage receptor MARCO	728	Calmodulin

Reads are the averaged read number per tissue library

expressed genes in each tissue, showing also the high reproducibility of the two biological replicates.

Table 5 shows the 25 most highly expressed genes in each tissue compared to the other 3 transcriptomes. The top d.e.g. in hemocytes were immune-related, including AMPs, such as myticin A, mytilin B, mytilin 6 and 7 or defensin 2b; pore-forming molecules, such as apextrin and MAC/perforin; lectins (C-type, nacre protein, macrophage mannose receptor) and many other genes that are directly or indirectly related to the immune system, such as ADAMTS16, a metalloprotease required for remodeling the basement membrane during cell migration [51] (Table 5). This list also included C1q, a serine protease inhibitor that modulates host-pathogen interactions [52] and HSP70, ISG12 and IAP, which play important roles in apoptosis and immunity [53–55]. The expression fold change was relatively high, varying approximately between 200 and 2000. Nevertheless, we have to consider that hemocytes were stimulated with different treatments whereas the other tissues were sampled from unstimulated mussels.

In mantle the most highly expressed protein was usherin, showing a fold change of 5008 (Table 5). This protein is involved in visual and auditory transduction in mammals [56]. Other bivalves, such as scallops, possess 'eyes' at the mantle edge that influence their relationship

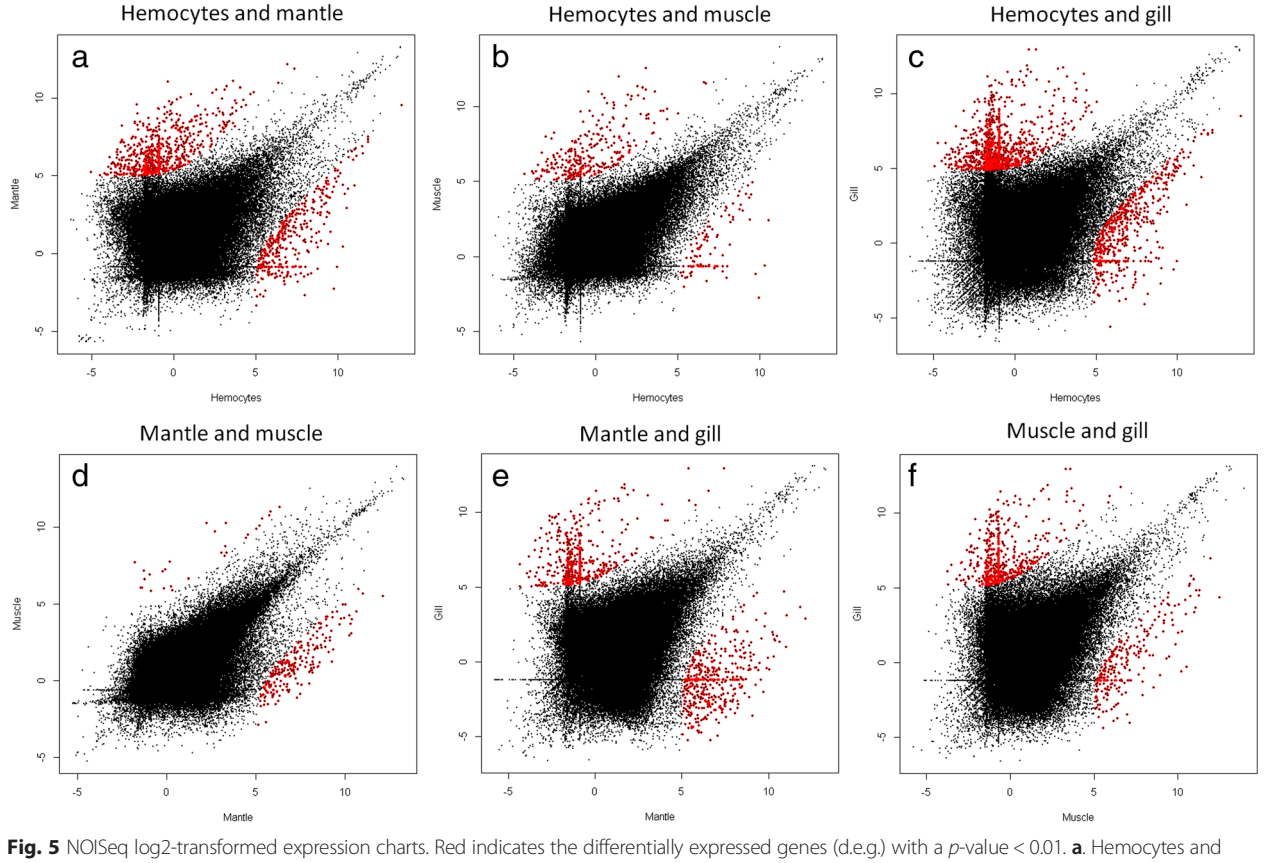


Fig. 5 NOISeq log2-transformed expression charts. Red indicates the differentially expressed genes (d.e.g.) with a *p*-value < 0.01. **a**. Hemocytes and Mantle; **b**. Hemocytes and Muscle; **c**. Hemocytes and Gills; **d**. Mantle and Muscle; **e**. Mantle and Gills; **f**. Muscle and Gills

Table 4 Number of differentially expressed genes between tissues

Analysis	*p*-value0.01	
	Total	Annotated
Hemocytes vs. Mantle	1,086	707 (238 h + 469 m)
Hemocytes vs. Muscle	399	264 (55 h + 209mu)
Hemocytes vs. Gills	1,928	1,040 (357 h + 683 g)
Mantle vs. Muscle	256	169 (149 m + 20mu)
Mantle vs. Gills	1,016	566 (285 m + 281 g)
Muscle vs. Gills	905	496 (182mu + 314 g)

h hemocytes; *m* mantle; *mu* muscle; *g* gills

with the environment [57]. The presence of this highly expressed gene in the mantle suggests that it might play a sensory role, in addition to its shell-forming and reproduction functions, which are also represented by genes such as vitellogenin 6, which is a precursor to egg-yolk proteins during embryonic development [58], or fibronectin 3, which is involved in shell formation in bivalves but also in mammal spermatogenesis [59, 60]. Interestingly, 3 of the 25 most highly expressed genes in mantle were related to antifungal functions or chitin metabolism: the PIF protein, mytimycin and chitinase 3, which showed fold changes of over 1000 compared with the other tissues. The shell of bivalves is a substrate for epibiotic communities, including fungi. Some fungi possess the ability to penetrate into the internal organs of animals and cause mycoses if the host-pathogen relationship is altered [61]. Therefore, the shell and the mantle could represent the first antifungal barrier, which would explain the presence of these d.e.g.

The muscle showed many typical myofibril molecules presenting fold changes of over 400, such as myosin light and heavy chain; paramyosin, which is typical of invertebrates; obscurin, which is involved in myofibrillogenesis [62]; calcium-related proteins, such as calmodulin or calpain, which are linked to muscle remodeling and contraction [63, 64]; and angiogenesis- and migration-related genes, such as angiopoietin-4 [65] and viral response molecules (e.g., TRIM56) [66], which exhibited expression increases of 465 and 776 fold, respectively (Table 5). These results suggest other possible functions of muscle in bivalves, as mentioned above.

The expression profile observed in gills confirmed previous studies showing that collagen is a major compound of this tissue [67]. Collagen was represented at more than 1000 fold the levels found in the other tissues, and showed higher levels than other extracellular matrix-related genes, such as dermatopontin, ST14 and TLL1. Dermatopontin accelerates and stabilizes collagen fibril formation, but this protein also presents other functions that are closely related to immune defense, such as cell adhesion via integrin binding, enhancing Transforming Growth Factor β1 activity or inhibiting cell proliferation [68]. ST14 degrades the extracellular matrix [69] and TLL1 processes procollagen C-propeptides [70]. However, as previously noted, the gills showed a significant expression of some immune-related molecules, such as the PIF protein, perlucin, LPS binding protein, big defensin or apextrin, which displayed fold changes ranging from 1243, in the case of apextrin, to 2837, in the case of PIF.

Enrichment analyses to compare qualitative and quantitative results

Gene Ontology (GO) terms were assigned to the non-redundant transcripts. A total of 18,899 (12.5 %) transcripts matched at least one GO term, which is twice as much as what was obtained in other reported Illumina

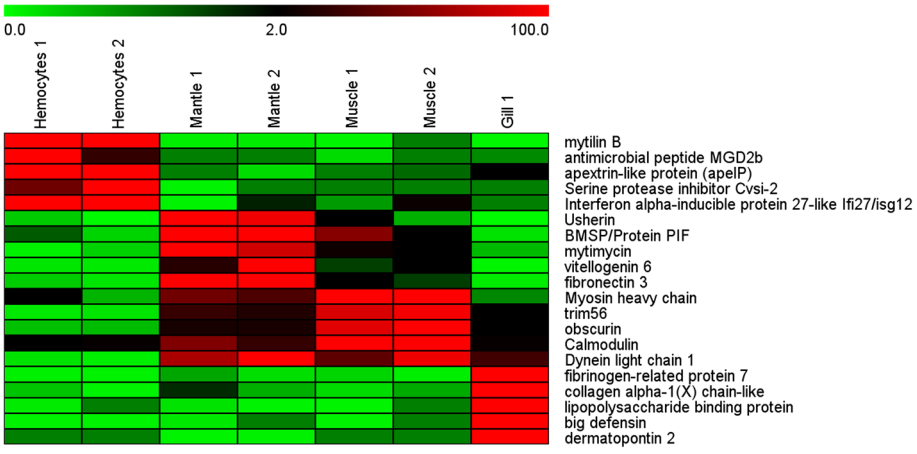


Fig. 6 Heatmap of a selection of 5 of the most highly expressed genes by tissue, which shows the expression level of all the biological replicates used in this study. The scale bar is a non-linear representation of the normalized expression: Saturated green: no expression, 0 %. Black: 2 % of maximum expression. Saturated red: maximum expression, 100 %

Table 5 Top 25 differentially expressed transcripts

FC	Hemocyte top 25 expressed	FC	Mantle top 25 expressed
1,820	Mytilin B	5,008	Usherin
750	Procollagen type VI alpha 4	2,702	Mitochondrial glycine cleavage system H prot
744	Metallothionein MT-20	2,521	BMSP / Protein PIF
709	C-type lectin	1,992	MAM dom-cont glycosylphosphatidylinositol anchor protein 2
685	Disintegrin and metalloproteinase with thrombospondin motifs 16 (ADAMTS16)	1,951	Protocatechuate 3,4-dioxygenase beta subunit
588	Cystatin-A-like	1,795	Collagen alpha-1, IV/III
530	Melatonin receptor-like (1A/1B)	1,758	Keratin, type II cytoskeletal 2 epidermal
508	Nacre protein	1,618	Fibroin heavy chain
478	Defensin 2b (MGD2b)	1,552	ATP-dependent RNA helicase A
465	C1q domain containing protein	1,541	Endo-1,4-mannanase
428	Mytilin-6	1,489	Heterogeneous nuclear ribonucleoprotein A3
407	Spermine oxidase	1,489	Mytimycin
383	Gly, Ala and Asn-rich protein	1,489	Splicing factor 3A subunit 2
357	Apextrin-like protein (apelP)	1,409	L-rhamnose-binding lectin CSL3
320	Mucin-2	1,389	Sarcoplasmic calcium-binding protein
296	Mytilin-7	1,314	Protein diaphanous
269	Macrophage mannose receptor 1-like	1,314	Vitellogenin 6
260	Serine protease inhibitor Cvs.i-2	1,261	Fibronectin 3
258	MAC/perforin- and kringle-dom-cont prot	1,218	Heat shock protein 70
252	Peptide O-xylosyltransferase	1,193	Whey acidic protein-like
251	MAM and LDL-receptor class A dom-cont prot	1,184	Porin-like
246	Heat shock protein 70	1,184	Hornerin / filaggrin
242	Myticin-A	1,168	Chitinase 3
241	Interferon alpha-inducible protein 27 2B (IFI27/ISG12)	1,144	Protein unc-93 homolog A
237	Inhibitor of apoptosis 7B/2/3	1,136	TPA: SCO-spondin protein
FC	Muscle top 25 expressed	FC	Gill top 25 expressed
2,702	Collagen alpha-3/6(VI) chain	3,083	Inner centromere protein A
1,226	Myosin heavy chain	2,837	BMSP / protein PIF
1,128	Heat shock protein 70	2,817	Perlucin
861	Collagen pro alpha-chain	2,592	Caveolin-1/3
787	C1q domain containing protein	2,120	Peptide O-xylosyltransferase
776	Tripartite motif-containing protein 2/56 (TRIM2/56)	2,048	Endonuclease domain-containing 1 protein
695	Proteoglycan 4	1,872	Insulin-like growth factor binding protein 2b
690	Protein LEA-1	1,833	Fibrinogen-related protein 7
662	Sushi, VWF type A, EGF and pentraxin dom-cont prot 1 (SVEP1)	1,771	Synaptotagmin
657	Beta-glucanase/Beta-glucan synthetase	1,734	Collagen alpha-1(X) chain-like
644	Enzymatic glycosylation-regulating-like	1,722	LDL receptor-related protein 8 (LRP8)
553	Fatty acid-binding protein homolog 9/7	1,652	Suppressor of tumorigenicity 14 protein (ST14)
530	Nucleolar protein 12	1,552	Antistasin
488	Forkhead box L2	1,479	Collagen triple helix repeat protein
481	Obscurin	1,458	LPS binding protein / Bactericidal permeability-increasing protein
471	Calmodulin	1,305	Viral A-type inclusion protein repeat
471	Dynein light chain 1	1,269	Big defensin

Table 5 Top 25 differentially expressed transcripts *(Continued)*

465	Angiopoietin-4	1,269	Notch gene homolog 3-like
465	GTP-binding protein REM 1	1,261	Alpha 1 type V collagen
461	Synaptopodin 2	1,243	Golgin subfamily A member 4
458	Myosin light chain	1,243	Apextrin-like protein
452	Calpain-5	1,235	Short-chain collagen C4
443	Plasminogen	1,209	Dermatopontin 2
440	Paramyosin	1,184	Stanniocalcin
428	BTB/POZ domain-containing protein KCTD7	1,160	Tolloid-like 1 precursor (TLL1)

FC fold change

transcriptomes (6 % GO annotation) [29]. This GO information was used to identify overrepresented biological processes in each transcriptome and in each group of d.e.g. by tissue. The results are summarized in Fig. 7 for the different transcriptomes and in Fig. 8 for the d.e.g. Figure 8 also shows overrepresented cellular components and molecular functions of hemocytes, mantle and muscle transcriptomes. First, it is important to note the large differences in enriched terms when the whole transcriptomes are compared with those of the d.e.g., which do not present a single term in common. The complete transcriptomes appeared to show more general functions, such as metabolism, transport or transcription, whereas the differentially expressed transcriptomes presented more detailed terms and functions, such as defense or regulation of specific signaling processes.

Mussel hemocytes exhibited the most divergent transcriptome due to the large number of significantly enriched processes found (Fig. 7). These processes included immune-related functions, regulation of the apoptotic process, cellular response to chemical stimulus, intracellular protein kinase cascade and hemopoiesis. Other functions regarding cell proliferation (regulation of cell cycle, DNA-dependent DNA replication, DNA repair) or migration (cell junction assembly) were also found to be overrepresented. In contrast, the hemocyte d.e.g. (Fig. 8) showed a quite different profile, with a high representation of categories related to the immune response, such as innate immune response (12 % of d.e.g.), defense response to bacterium (21 %), regulation of immune defense to virus (12 %) or defense response to fungus (7 %). Figure 8 provides other interesting results as well; for example, almost 50 % of the hemocyte d.e.g. had functions involved in the response to stimulus, and over 20 % of them had ontologies for extracellular proteins and functions related to receptor binding.

The mantle transcriptome showed some remarkable enriched processes (Fig. 7), such as the evolutionarily conserved Wnt receptor signaling pathway, which plays a key role in development, including stem cell proliferation and cancer [71]. This finding is also congruent with the mantle being the hematopoietic tissue in mussels. Moreover, the differentially expressed transcriptome of the mantle confirmed the functional similarity between the mantle and muscle, as it included three GO terms related to muscle contraction: myofibril, myosin filament and calmodulin binding (Fig. 8). The results of the muscle d.e.g. enrichment analysis fully coincide with the mantle analysis, adding more terms related to contraction, such as motor activity, sarcomere, actin binding or calcium binding (Fig. 8). The complete muscle transcriptome, as well as the transcriptomes of the other tissues, presented general processes such as metabolism or transcription, but also two specific processes: response to DNA damage stimulus and hemopoietic or lymphoid organ development, including differentiation of resident and migratory cell types (Fig. 7). The GO terms related to immune response may be due to the normal presence of some hemocytes in the muscle.

The enriched functions of the gill transcriptome showed a similar, but reduced profile compared with that from hemocytes (Fig. 7). The gill transcriptome was not as closely related to immunity as it was to signaling (intracellular protein kinase cascade) and cell proliferation (cell cycle phase, chromosome organization and cytokinesis), which were most likely overrepresented due to the direct contact of this tissue with the environment, as such contact could lead to a regular renewal of the tissue. The enrichment analysis of the gill d.e.g. (Fig. 8) produced the highest number of results among all the analyzed data, with 37 GO biological process categories being overrepresented. All of these processes can be grouped into three main categories: calcium homeostasis, coagulation and defense, which are intimately related to each other. The identified coagulation processes (blood coagulation, platelet activation and degranulation) could also be included in the defense group because coagulation triggers the complement cascade [72]

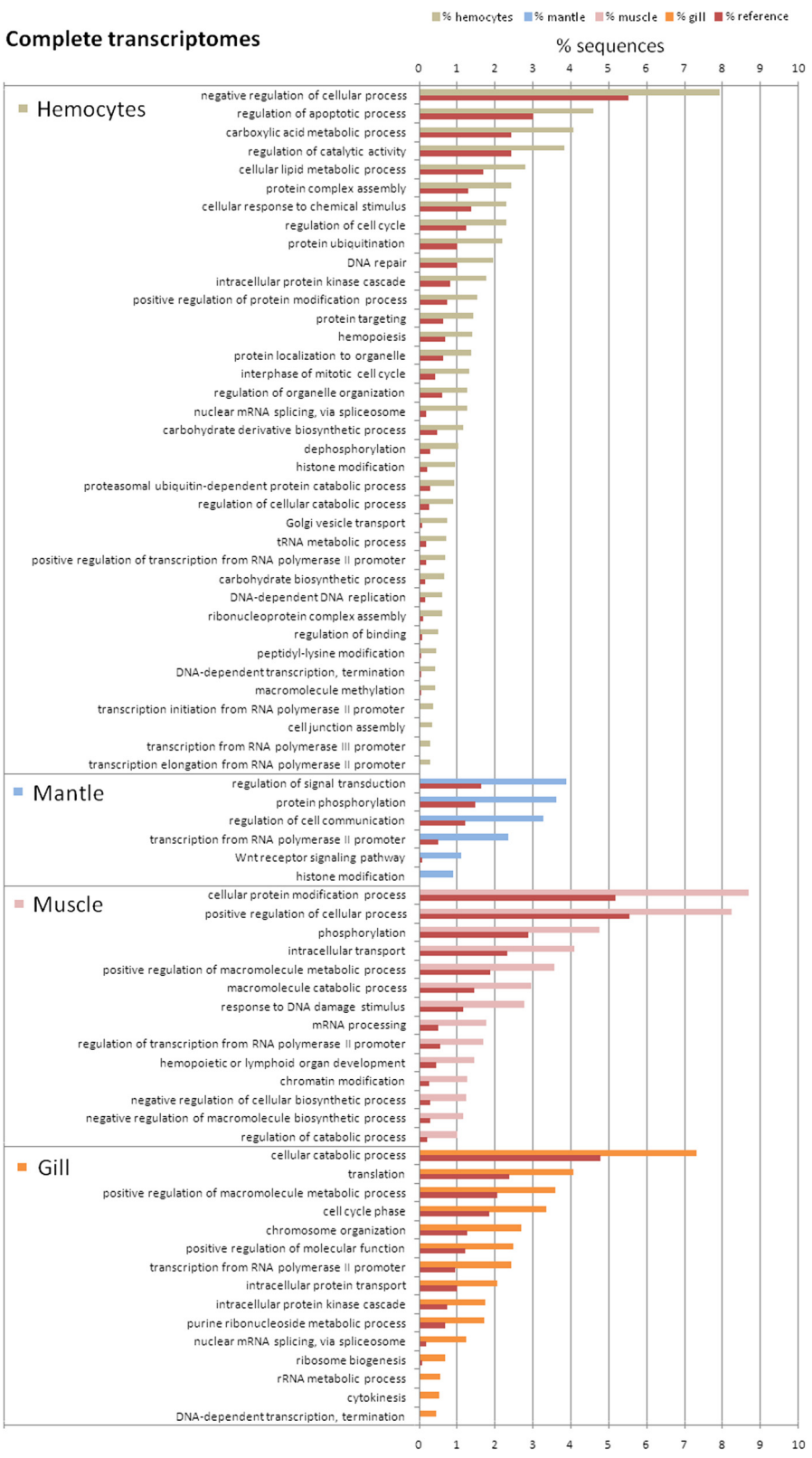


Fig. 7 Classification of the complete transcriptomes by tissue type after a Blast2GO Enrichment Analysis. Only overrepresented biological process GO terms are shown

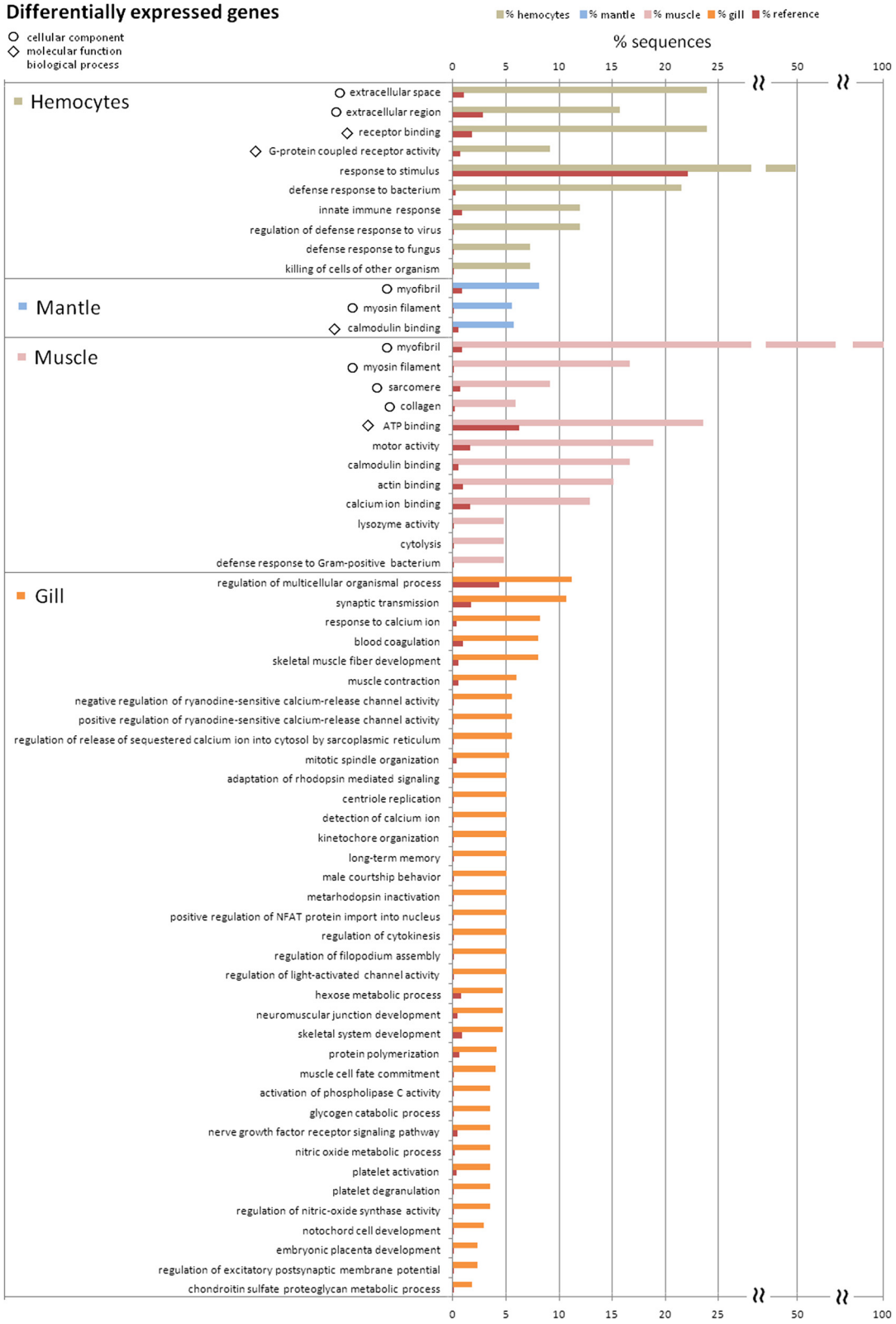


Fig. 8 Classification of the d.e.g. by tissue after a Blast2GO Enrichment Analysis. Only overrepresented GO terms are shown and for the gills, only the biological process enriched GO terms are shown

and is critical in immune defense, as well as the production of toxic radicals such as nitric oxide (NO) (represented with the categories nitric oxide metabolic process and regulation of nitric oxide synthase activity), which has been shown to occur in the gills. NO production is known to be up-modulated in bivalves stimulated with bacteria and parasites [73, 74]. Calcium homeostasis processes were clearly represented in the gill d.e.g. (Fig. 8), such as the detection and response to calcium ions, regulation of the release of sequestered calcium and activation of phospholipase C activity. In addition to their role in gas exchange, gills exhibit osmoregulatory, ion transport and homeostasis functions in crustaceans and fish [75, 76]; however, these functions have not been studied in bivalves. The cells involved in these processes in fish are ionocytes, a mitochondria-rich cell (MRC) type. In bivalves, three types of MRCs are present in the gills [67]. These factors suggest that there is calcium homeostasis activity in mussel gills.

Conclusions

We have shown the value of whole-transcriptome analysis generated via RNA-Seq for accurate quantification of gene expression. Using almost 400 million reads, we described the transcriptome and expression profiles of *M. galloprovincialis* tissues and the generated data has enriched the genomic resources available for this organism.

This study represents the first RNA-Seq approach applied in bivalves to describe and analyze tissue-specific transcriptomes. We identified a high number of transcripts related to the immune system, signal transduction and infectious diseases that highlight immune functions in all the tissues studied, probably as a result of mussel's open circulatory system. Another group of disease-associated pathways were those related to cancer, which ranked second among the most represented pathways. Moreover, we also found specific and unexpected functions in specific tissues: mussel hemocytes showed the greatest number of antimicrobial and defense proteins; mantle appeared to exhibit a more specific antifungal function and even to be a firm candidate of the hematopoietic tissue; gills presented a large number of putative recognition molecules; and muscle expressed stress- and defense-related proteins.

Our results shed light into the transcriptomics and physiology of the Mediterranean mussel. This species has a great economical and ecological importance, it has been extensively used as pollution sentinel and the present findings related to immunity, hematopoiesis and cancer confirm that *M. galloprovincialis* is a very interesting candidate to be the model species for bivalves and even molluscs. The mussel genome project, that will come soon, will further support this candidature.

Methods

Tissue sampling, *in vitro* stimulation of hemocytes and RNA isolation

M. galloprovincialis mussels were obtained from a commercial shellfish farm (Vigo, Galicia, Spain) after depuration. The animals were maintained in open-circuit filtered sea water tanks at 15 °C with aeration and were fed daily with *Phaeodactylum tricornutum* and *Isochrysis galbana* until 2 days before sampling. Prior to the experiments, the mussels were acclimatized to aquarium conditions for one week.

The mantle, muscle and gill tissues from 5 mussels were sampled, pooled and conserved in 1 ml of TRIzol (Invitrogen). All samplings were performed as 2 biological replicates from all the tissues, except for the gills (which included only 1 biological replicate).

For hemolymph collection, approximately 50 mussels were notched in the shell and hemolymph (1–3 ml) was withdrawn from the adductor muscle of each mussel with a 0.5-mm-diameter (25G) disposable needle. The hemolymph was pooled and distributed in 6-well plates, with 7 ml per well, in a total of 9 wells, one for each treatment. The hemocytes were allowed to settle to the base of the wells for 30 min at 15 °C in the dark. Then, the hemocytes were stimulated for 3 h at 15 °C with 50 μg/ml polyinosinic:polycytidylic acid (Poly I:C), peptidoglycans (PG), zymosan, *Vibrio anguillarum* DNA (CpG), lipopolysaccharide (LPS), lipoteichoic acid (LTA), 100 ng/ml flagellin and 1 x 10^6 CFU/ml of heat-inactivated *Vibrio anguillarum* (one stimulus per well). The last group of hemocytes remained unstimulated. All the stimuli were purchased from SIGMA, except for CpG and *V. anguillarum*, which were produced in our laboratory. This procedure was performed twice to obtain 2 biological replicates. Hemolymph was centrifuged at 4 °C at 3000 g for 10 min and the pellet was resuspended in 500 μl of TRIzol (Invitrogen).

From this step onwards the methodology used was the same for all the tissues. Total RNA isolation was conducted following the manufacturer's protocol using the RNeasy Mini kit (Qiagen) for RNA purification after DNase I treatment. Next, the concentration and purity of the RNA were measured using a *NanoDrop ND1000* spectrophotometer. Finally, RNA integrity was tested on an Agilent 2100 Bioanalyzer (Agilent Technologies) to produce cDNA libraries for Illumina sequencing.

cDNA production and Illumina sequencing

The mRNA-Seq sample preparation kit from Illumina was used according to the manufacturer's instructions. Briefly, eukaryotic mRNA was extracted from total RNA using oligo (dT) magnetic beads and cleaved into short fragments using fragmentation buffer. A cDNA library compatible with the Illumina NGS technology was then

prepared from the fragmented mRNA via reverse transcription, second-strand synthesis and ligation of specific adapters (paired-ends) after cDNA purification using the QIAquick PCR Purification Kit (Qiagen). The amount of cDNA in each library was quantified through spectrofluorometric analysis using the Qbit system. Next-generation sequencing was performed using Illumina HiSeq™ 2000 technology at the Beijing Genomics Institute (BGI-HongKong Co., Ltd., Tai Po, Hong Kong).

Bioinformatics workflow

Assembly and functional annotation

The image data output from the sequencing apparatus was transformed via base calling into raw data and stored in FASTQ format. The raw data were cleaned with filter_fq software to discard low-quality reads, reads with regions with greater than 5 % unknown bases or reads with adapters.

De novo transcriptome assembly was conducted with the short reads assembly program Trinity [34, 77] (minimal contig_length: 100; group_pairs distance: 250; minimal kmer_cov: 2). Trinity first combined overlapping reads to form contigs with at least a 100-bp length and a minimum of 2 reads to be assembled. Then, the contigs were assembled again to obtain longer sequences that could not be further extended, which are unigenes. During this process and before obtaining the final unigenes, the reads were mapped against the contigs to confirm the assembly procedure. When multiple samples from the same species are sequenced (biological replicates or different tissues), unigenes from each sample can be applied together to perform another assembly step. This process detects sequence splicing and redundancy to acquire the longest sequences and group them into clusters. Each cluster is formed by several unigenes with more than 70 % similarity. To simplify the terminology employed in this study, all the non-redundant sequences will be called "transcripts", regardless of whether they are unique unigenes or belong to a cluster. The completeness of the mussel transcriptome was confirmed with the CEGMA package (http://korflab.ucdavis.edu/datasets/cegma/).

A total of 151,320 transcripts were obtained following this protocol. This number represents all the detectable variability in the mRNAs from the four studied tissues, including splicing variants, non-overlapping fragments of the same mRNA, UTRs or mRNAs in different splicing stages.

The transcripts were first annotated using BLASTx and BLASTn (with an e-value threshold of $10e^{-5}$) against the NCBI nr, Swiss-Prot, KEGG and COG protein databases and the NCBI nt nucleotide database. The annotation step provided the identity of the transcript with the species harboring the matching sequence, which is useful for detecting possible contaminants in our samples. Using the KEGG database information, the metabolic pathways and functions of the annotated transcripts could be obtained and presented.

The oyster proteome was downloaded from http://www.oysterdb.com/FrontDownloadAction.do?method=download and compared with the translated mussel transcripts.

RNA-Seq with NOISeq: Quantitative analysis between tissues

RNA-Seq compares the number of reads that align to a specific transcript in different samples or cDNA libraries. The calculation of expression uses the RPKM (Reads Per Kilobase of exon model per Million mapped reads) normalization, while accounting for the length of the transcript that they belong to, its number of base pairs and the total number of reads in the transcriptome [78]. This normalization can eliminate the influence of different gene lengths and sequencing levels on the calculation of the gene expression. Therefore, the calculated gene expression can be directly used for comparison of the differences in gene expression between tissues in pairwise comparisons. The chosen method for evaluating the d.e.g. between tissues was NOISeq (http://bioinfo.cipf.es/noiseq) [50]. NOISeq is a nonparametric statistical approach that creates an empirical distribution of count changes that are adapted to the available data. This method has been proven to be the most effective in controlling the false discovery rate. The p-value threshold used to detect d.e.g. was 0.01.

To present the quantitative results and to facilitate their visualization, the pairwise comparisons (three per tissue) were fused, calculating the average of the three fold change values of the transcripts with the same annotation. Only one table/figure per tissue is presented, rather than all the possible comparisons.

The heatmap shown in Fig. 6 was designed with the software TMeV [79]. The normalized values (RPKM) for each gene by tissue and biological replicate were used to represent their expression in a green/0 – red/100 scale, with green representing the lower expression values and red the higher expression values.

GO classification and enrichment analysis

The nr annotation was used to obtain the GO term assignments of the transcripts with the Blast2GO program [80]. Then, enrichment analyses were conducted with the total information from all the tissues, including the reference set and each tissue and expression analysis test set. Next, Fisher's exact test was run with default values (a two-tailed test that removes double IDs, with a false discovery rate (FDR) cut-off of 0.01). The Blast2GO option to show only the most specific terms (0.01 FDR cut-off) was used once. To reduce the dimensions of Fig. 8,

the enrichment analyses of the expression results were combined according to the tissue types. Thus, only one graph per tissue is represented, instead of all the possible comparisons. Non-redundant categories were aggregated. For the coincident categories, the average of the percent representation was calculated.

Additional files

Additional file 1: Table listing Mytilus galloprovincialis transcripts, including the sequence, length, RPKM, description, accession number of the description (Hit ACC) and e-value obtained in each database used for annotation and the GO terms ascribed to each sequence. (XLSX 38412 kb)

Additional file 2: Figure representing the number of transcripts obtained after the assembly of increasing number of mapped reads. (XLSX 743 kb)

Additional file 3: Pathways found in the annotated portion of the transcriptomes. (ZIP 4950 kb)

Additional file 4: Tables and figures showing the non-shared C1q variants found in the hemocytes, mantle, muscle and gills. (XLSX 17 kb)

Additional file 5: Mytilus galloprovincialis transcriptome in FASTA format. (ZIP 25884 kb)

Abbreviations

ACC Nº: Accession number; ADAMTS16: A disintegrin and metalloproteinase with thrombospondin motifs 16; AMPs: Antimicrobial peptides; BLAST: Basic local alignment search tool; bp: Base pairs; C1q: C1q domain-containing proteins; CEGMA: Core Eukaryotic Genes Mapping Approach; COG: Clusters of orthologous groups of proteins; d.e.g.: Differentially expressed genes; EST: Expressed sequence tags; FC: Fold change; FDR: False discovery rate; FREPs: Fibrinogen-related proteins; GO: Gene Ontology; GrpE: GroP-like gene E; HSP: Heat shock protein; IAP: Inhibitor of apoptosis protein; ID: Identity; IFI27: Interferon alpha-inducible protein 27-like protein; ISG12: Interferon-stimulated gene 12 protein; KEGG: Kyoto encyclopedia of genes and genomes; LPS: Lipopolysaccharide; LTA: Lipoteichoic acid; *M. galloprovincialis*: *Mytilus galloprovincialis*; MAC: Membrane attack complex; MARCO: Macrophage receptor with collagenous structure; MRC: Mitochondria-rich cell; NCBI: National Center for Biotechnology Information; NGS: Next generation sequencing; NO: Nitric oxide; PG: Peptidoglycans; Poly I:C: Polyinosinic:polycytidylic acid; RNA-Seq: RNA sequencing; RPKM: Reads Per Kilobase of exon model per Million mapped reads; SRA: Short read archive; ST14: Suppressor of tumorigenicity 14 protein; TLL1: Tolloid like-1 precursor; TMeV: TIGR Multiexperiment Viewer; TRIM56: Tripartite motif-containing protein 56; *V. anguillarum*: *Vibrio anguillarum*.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

BN, AF, DP and CC conceived and designed the experiments. RM, PP prepared the samples. RM and BN analyzed the data. RM wrote the paper. BN, AF, DP, CC and PP provided corrections to the manuscript. All authors read and approved the manuscript.

Acknowledgments

This work has been funded by the projects BIVALIFE, FP7-KBBE-2010-4/266157 and 10 PXIB 402 096 PR from Xunta de Galicia. RM wishes to acknowledge the Spanish MICINN for her FPI Spanish research grant (BES-2009-029765) and PP gratefully acknowledges the Spanish Ministerio de Educación for her FPU fellowship (AP2010-2408). We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

Author details

[1]Instituto de Investigaciones Marinas (IIM), Consejo Superior de Investigaciones Científicas (CSIC), Eduardo Cabello, 6, 36208 Vigo, Spain. [2]Departamento de Bioquímica, Genética e Inmunología, Facultad de Biología, Unidad Asociada CSIC, Universidade de Vigo, 36310 Vigo, Spain.

Received: 4 May 2015 Accepted: 5 August 2015
Published online: 24 September 2015

References

1. Figueras A. Biología y cultivo de mejillón (*Mytilus galloprovincialis*) en Galicia. Madrid: Consejo Superior de Investigaciones Científicas; 2007.
2. Gosling E. Bivalve Molluscs: Biology, Ecology and Culture. New York: John Wiley & Sons; 2008.
3. 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.
4. Paillard C, Leroux F, Borrego JJ. Bacterial disease in marine bivalves, Review of recent studies. Trends and evolution. Aquat Living Resour. 2004;17:477–98.
5. Villalba A, Reece KS, Ordás MC, Casas SM, Figueras A. Perkinsosis in molluscs: A review. Aquat Living Resour. 2004;17:411–32.
6. Figueras A. Cultured Aquatic Species Information Programme. *Mytilus galloprovincialis*. In: FAO Fisheries and Aquaculture Department. 2004. http://www.fao.org/fishery/culturedspecies/Mytilus_galloprovincialis/en. Accessed 10 February 2015.
7. Canesi L, Gallo G, Gavioli M, Pruzzo C. Bacteria–hemocyte interactions and phagocytosis in bivalves. Microsc Res Technol. 2002;57:469–76.
8. Olafsen JA. Role of lectins (C-reactive protein) in defense of marine bivalves against bacteria. Adv Exp Med Biol. 1995;371A:343–8.
9. Ordás MC, Novoa B, Figueras A. Modulation of the chemiluminescence response of Mediterranean mussel (*Mytilus galloprovincialis*) haemocytes. Fish Shellfish Immunol. 2000;10:611–22.
10. Pallavicini A, Costa MM, Gestal C, Dreos R, Figueras A, Venier P, et al. Sequence variability of myticins identified in haemocytes from mussels suggests ancient host-pathogen interactions. Dev Comp Immunol. 2008;32:213–26.
11. Romero A, Dios S, Poisa-Beiro L, Costa MM, Posada D, Figueras A, et al. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. Dev Comp Immunol. 2011;35:334–44.
12. Gerdol M, Manfrin C, De Moro G, Figueras A, Novoa B, Venier P, et al. The C1q domain containing proteins of the Mediterranean mussel *Mytilus galloprovincialis*: a widespread and diverse family of immune-related molecules. Dev Comp Immunol. 2011;35:635–43.
13. Criscitiello MF, de Figueiredo P. Fifty shades of immune defense. PLoS Pathog. 2013;9:e1003110.
14. 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.
15. Gueguen Y, Cadoret JP, Flament D, Barreau-Roumiguière C, Girardot AL, Garnier J, et al. Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*. Gene. 2003;303:139–45.
16. Venier P, De Pittà C, Bernante F, Varotto L, De Nardi B, Bovo G, et al. MytiBase: a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics. 2009;10:72.
17. 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;13:275–83.
18. Tanguy A, Bierne N, Saavedra C, Pina B, Bachère E, Kube M, et al. Increasing genomic information in bivalves through new EST collections in four species, development of new genetic markers for environmental studies and genome evolution. Gene. 2008;408:27–36.
19. 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.
20. Takeuchi T, Kawashima T, Koyanagi R, Gyoja F, Tanaka M, Ikuta T, et al. Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. DNA Res. 2012;19:117–30.

21. Rosani U, Varotto L, Rossi A, Roch P, Novoa B, Figueras A, et al. Massively parallel amplicon sequencing reveals isotype-specific variability of antimicrobial peptide transcripts in *Mytilus galloprovincialis*. PLoS One. 2011;6:e26680.
22. Gerdol M, De Moro G, Manfrin C, Milandri A, Riccardi E, Beran A, et al. RNA sequencing and de novo assembly of the digestive gland transcriptome in *Mytilus galloprovincialis* fed with toxinogenic and non-toxic strains of *Alexandrium minutum*. BMC Res Notes. 2014;7:722.
23. Bettencourt R, Pinheiro M, Egas C, Gomes P, Afonso M, Shank T, et al. High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*. BMC Genomics. 2010;11:559.
24. 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:e21560.
25. 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.
26. Moreira R, Balseiro P, Planas JV, Fuste B, Beltran S, Novoa B, et al. Transcriptomics of in vitro immune-stimulated hemocytes from the Manila clam *Ruditapes philippinarum* using high-throughput sequencing. PLoS One. 2012;7:e35009.
27. Chen H, Zha J, Liang X, Bu J, Wang M, Wang Z. Sequencing and De Novo Assembly of the Asian Clam (*Corbicula fluminea*) Transcriptome Using the Illumina GAIIx Method. PLoS One. 2013;8:e79516.
28. Gavery MR, Roberts SB. Characterizing short read sequencing for gene discovery and RNA-Seq analysis in *Crassostrea gigas*. Comp Biochem Physiol Part D Genomics Proteomics. 2012;7:94–9.
29. Zhao X, Yu H, Kong L, Li Q. Transcriptomic responses to salinity stress in the Pacific oyster *Crassostrea gigas*. PLoS One. 2012;7:e46244.
30. Apeti DA, Lauenstein GG, Christensen JD, Kimbrough K, Johnson WE, Kennedy M, et al. A historical assessment of coastal contamination in Birch Harbor, Maine based on the analysis of mussels collected in the 1940s and the Mussel Watch Program. Mar Pollut Bull. 2010;60:732–42.
31. Montes MO, Hanna SK, Lenihan HS, Keller AA. Uptake, accumulation, and biotransformation of metal oxide nanoparticles by a marine suspension-feeder. J Hazard Mater. 2012;225–6:139–45.
32. 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.
33. Balseiro P, Falcó A, Romero A, Dios S, Martínez-López A, Figueras A, et al. *Mytilus galloprovincialis* myticin C: a chemotactic molecule with antiviral activity and immunoregulatory properties. PLoS One. 2011;6:e23140.
34. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644–52.
35. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008;36(Database issue):D480–4.
36. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 2000;28:33–6.
37. St Laurent G, Shtokalo D, Tackett MR, Yang Z, Vyatkin Y, Milos PM, et al. On the importance of small changes in RNA expression. Methods. 2013;63:18–24.
38. Gonzalez-Rey M, Lau TC, Gomes T, Maria VL, Bebianno MJ, Wu R. Comparison of metal accumulation between 'Artificial Mussel' and natural mussels (*Mytilus galloprovincialis*) in marine environments. Mar Pollut Bull. 2011;63:149–53.
39. Romero A, Costa M, Forn-Cuni G, Balseiro P, Chamorro R, Dios S, et al. Occurrence, seasonality and infectivity of *Vibrio* strains in natural populations of mussels *Mytilus galloprovincialis*. Dis Aquat Organ. 2014;108:149–63.
40. Kurelec B, Pivčević B. Evidence for a multixenobiotic resistance mechanism in the mussel *Mytilus galloprovincialis*. Aquat Toxicol. 1991;19:291–301.
41. Carella F, Figueras A, Novoa B, De Vico G. Comparative cytomorphological features and PCNA expression pattern in Haemic Neoplasia from mediterranean mussels (*Mytilus galloprovincialis*) and Galician common cockles (*Cerastoderma edule*). Dis Aquat Organ. 2013;105:81–7.
42. Arriagada G, Metzger MJ, Muttray AF, Sherry J, Reinisch C, Street C, et al. Activation of transcription and retrotransposition of a novel retroelement, Steamer, in neoplastic hemocytes of the mollusk *Mya arenaria*. Proc Natl Acad Sci U S A. 2014;111:14175–80.
43. Boon E, Faure MF, Bierne N. The flow of antimicrobial peptide genes through a genetic barrier between *Mytilus edulis* and *M. galloprovincialis*. J Mol Evol. 2009;68:461–74.
44. Wang N, Lee YH, Lee J. Recombinant perlucin nucleates the growth of calcium carbonate crystals: molecular cloning and characterization of perlucin from disk abalone, *Haliotis discus discus*. Comp Biochem Physiol B Biochem Mol Biol. 2008;149:354–61.
45. Bayne CJ, Sminia T, Van der Knaap WPW. Immunological memory: status of molluscan studies. In: Manning MJ, editor. Phylogeny of immunological memory. Developments in immunology, vol. 10. Amsterdam: Elsevier; 1980. p. 57–64.
46. 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.
47. Balseiro P, Moreira R, Chamorro R, Figueras A, Novoa B. Immune responses during the larval stages of *Mytilus galloprovincialis*: metamorphosis alters immunocompetence, body shape and behavior. Dev Comp Immunol. 2013;35:438–47.
48. Grigorian M, Mandal L, Hartenstein V. Hematopoiesis at the onset of metamorphosis: terminal differentiation and dissociation of the Drosophila lymph gland. Dev Genes Evol. 2011;221:121–31.
49. Ojala JR, Pikkarainen T, Tuuttila A, Sandalova T, Tryggvason K. Crystal structure of the cysteine-rich domain of scavenger receptor MARCO reveals the presence of a basic and an acidic cluster that both contribute to ligand recognition. J Biol Chem. 2007;282:16654–66.
50. Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-Seq: a matter of depth. Genome Res. 2011;21:2213–23.
51. Surridge AK, Rodgers UR, Swingler TE, Davidson RK, Kevorkian L, Norton R, et al. Characterization and regulation of ADAMTS-16. Matrix Biol. 2009;28:416–24.
52. Xue Q, Itoh N, Schey KL, Cooper RK, La Peyre JF. Evidence indicating the existence of a novel family of serine protease inhibitors that may be involved in marine invertebrate immunity. Fish Shellfish Immunol. 2009;27:250–9.
53. Parcellier A, Gurbuxani S, Schmitt E, Solary E, Garrido C. Heat shock proteins, cellular chaperones that modulate mitochondrial cell death pathways. Biochem Biophys Res Commun. 2003;304:505–12.
54. Rosebeck S, Leaman DW. Mitochondrial localization and pro-apoptotic effects of the interferon-inducible protein ISG12a. Apoptosis. 2008;13:562–72.
55. Silke J, Meier P. Inhibitor of apoptosis (IAP) proteins-modulators of cell death and inflammation. Cold Spring Harb Perspect Biol. 2013;5(2):a008730.
56. Yang J, Liu X, Zhao Y, Adamian M, Pawlyk B, Sun X, et al. Ablation of whirlin long isoform disrupts the USH2 protein complex and causes vision and hearing loss. PLoS Genet. 2010;6:e1000955.
57. Speiser DI, Loew ER, Johnsen S. Spectral sensitivity of the concave mirror eyes of scallops: potential influences of habitat, self-screening and longitudinal chromatic aberration. J Exp Biol. 2011;214:422–31.
58. Nakamura A, Yasuda K, Adachi H, Sakurai Y, Ishii N, Goto S. Vitellogenin-6 is a major carbonylated protein in aged nematode, *Caenorhabditis elegans*. Biochem Biophys Res Commun. 1999;264:580–3.
59. Joubert C, Piquemal D, Marie B, Manchon L, Pierrat F, Zanella-Cléon I, et al. Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell:focus on biomineralization. BMC Genomics. 2010;11:613.
60. Obholz KL, Akopyan A, Waymire KG, MacGregor GR. FNDC3A is required for adhesion between spermatids and Sertoli cells. Dev Biol. 2006;298:498–513.
61. Morley NJ. Interactive effects of infectious diseases and pollution in aquatic molluscs. Aquat Toxicol. 2010;96:27–36.
62. Young P, Ehler E, Gautel M. Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly. J Cell Biol. 2001;154:123–36.
63. Tavi P, Westerblad H. The role of in vivo Ca^{2+} signals acting on Ca^{2+}-calmodulin-dependent proteins for skeletal muscle plasticity. J Physiol. 2011;589:5021–31.
64. Murphy RM. Calpains, skeletal muscle function and exercise. Clin Exp Pharmacol Physiol. 2010;37:385–91.

65. Lee HJ, Cho CH, Hwang SJ, Choi HH, Kim KT, Ahn SY, et al. Biological characterization of angiopoietin-3 and angiopoietin-4. FASEB J. 2004;18:1200–8.
66. Shen Y, Li NL, Wang J, Liu B, Lester S, Li K. TRIM56 is an essential component of the TLR3 antiviral signaling pathway. J Biol Chem. 2012;287:36404–13.
67. de Oliveira David JA, Salaroli RB, Fontanetti CS. Fine structure of *Mytella falcata* (Bivalvia) gill filaments. Micron. 2008;39:329–36.
68. Okamoto O, Fujiwara S. Dermatopontin, a novel player in the biology of the extracellular matrix. Connect Tissue Res. 2006;47:177–89.
69. Ge W, Hu H, Ding K, Sun L, Zheng S. Protein interaction analysis of ST14 domains and their point and deletion mutants. J Biol Chem. 2006;281:7406–12.
70. Berry R, Jowitt TA, Garrigue-Antar L, Kadler KE, Baldock C. Structural and functional evidence for a substrate exclusion mechanism in mammalian tolloid like-1 (TLL-1) proteinase. FEBS Lett. 2010;584:657–61.
71. Niehrs C. The complex world of WNT receptor signalling. Nat Rev Mol Cell Biol. 2012;13:767–79.
72. Dunkelberger JR, Song WC. Complement and its role in innate and adaptive immune responses. Cell Res. 2010;20:34–50.
73. 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 Shellfish Immunol. 2009;26:515–23.
74. Villamil L, Gomez-Leon J, Gomez-Chiarri M. Role of nitric oxide in the defenses of *Crassostrea virginica* to experimental infection with the protozoan parasite *Perkinsus marinus*. Dev Comp Immunol. 2007;31:968–77.
75. Freire CA, Onken H, McNamara JC. A structure-function analysis of ion transport in crustacean gills and excretory organs. Comp Biochem Physiol A Mol Integr Physiol. 2008;151:272–304.
76. Conte FP. Origin and differentiation of ionocytes in gill epithelium of teleost fish. Int Rev Cell Mol Biol. 2012;299:1–25.
77. Yang Y, Smith SA. Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. BMC Genomics. 2013;14:328.
78. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5:621–8.
79. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, et al. TM4 microarray software suite. Methods Enzymol. 2006;411:134–93.
80. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO, a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005;21:3674–6.

Additional files

Additional files associated with this article can be found at:
http://www.biomedcentral.com/1471-2164/16/728

FINAL DISCUSSION

Physiological studies are basing more and more on molecular biology and genomic tools. From 1982 to the present, the number of bases in GenBank has doubled approximately every 18 months (GenBank Release Notes, 2015). Regarding bivalves, in 2006: A search of the GenBank database for the keywords "Mollusca" + "Bivalvia" returned about 32,000 entries. At that time, the most of the sequences were the result of single-gene research (Saavedra and Bachere, 2006). In the moment of deposit of this PhD thesis (September 2015) the search for bivalve nucleotide sequences returned more than 720,000 results, and the Short Read Archive (SRA) repository has 771 entries. That is the result of almost 10 years of research in bivalve genomics. The direct contribution for clams and mussels presented in this PhD thesis is the following: from the 30 SRA projects for *R. philippinarum,* 1 was uploaded after this research (SRX100159), and 7 out of the 30 SRA projects for *M. galloprovincialis* were presented in this thesis (SRP033481). This means tens of thousands and millions of new reads of new identified nucleotide sequences, for clams and mussels respectively, which enriched the genomic background for bivalves.

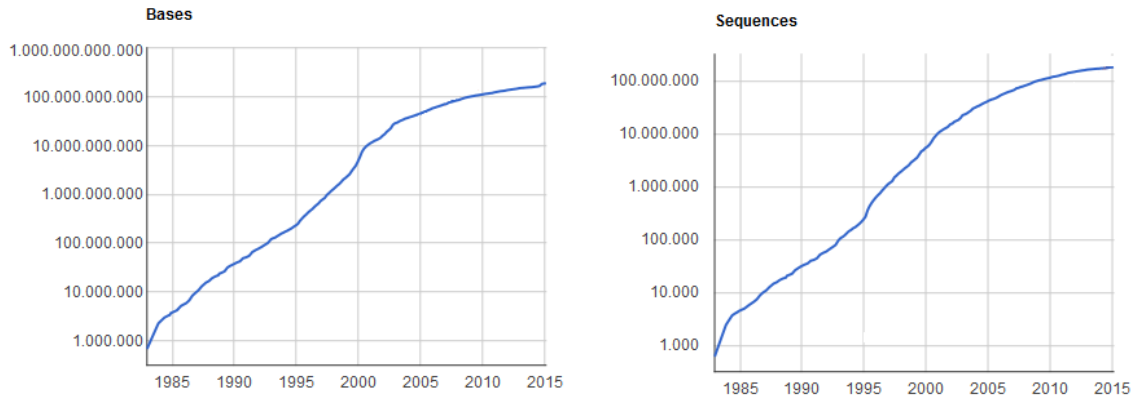


Figure 16. Evolution of the total number of bases and sequences in the GenBank database for all organisms. Font: GenBank Release Notes, 2015.

Indeed, in 2012, when the first paper of this PhD thesis was published, "Gene expression analysis of clams *Ruditapes philippinarum* and *Ruditapes decussatus* following bacterial infection yields molecular insights into pathogen resistance and immunity", only 3,784 ESTs of R. *decussatus* and 4,607 of *R. philippinarum* were present in GenBank. This amount was reduced to 424 and 464 for each species respectively if we only take into account the annotated sequences. In the era of the genetic studies those are unacceptable numbers. Even so, we were able to make the first comparative study of the expression of 13 immune-related genes between *R. decussatus* and *R. philippinarum*. It was also the first time that several genes were described in mollusks (IAP, APIP, thrombin and HMG1). We were able to conclude that *R. decussatus* deficiency in expression of transcription factors and genes for non-self recognition may explain its low resistance to pathogens and stress and that *R. philippinarum* is able to develop rapid pathogen detection, promoting an effective and competent defense response.

As the number of ESTs was very low at that time, we decided to vastly increase the number of sequences for some bivalve species in the framework of the ReProSeed project. At the moment only *R. philippinarum* transcriptome, obtained by 454 pyrosequencing, was published: "Transcriptomics of *in vitro* immune-stimulated hemocytes from the Manila clam *Ruditapes philippinarum* using high-throughput sequencing". We discovered thousands of new immune sequences for *R. philippinarum*, many of them were never described before in bivalves (C2, C4, C5, C9, AIF, Bax, AKT, TLR6 or TLR13). As a central part of this work, we gathered the sequences that belonged to three key immune pathways: apoptosis, toll like signaling pathway and the complement cascade. This helped to better understand the immune mechanisms of bivalves.

All this sequencing work was the basis to construct the microarray presented in "Gene expression profile analysis of Manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibrio alginolyticus* challenge using an immune-enriched oligo-microarray". This way we could study the timing of the immune response of hemocytes to a bacterial challenge. We found almost 600 differentially expressed genes out of the 13,671 probes included in the microarray. The early response revealed signaling and apoptosis as the initial steps of the defense, the most characteristic genes of the immune system processes (PGRPs, FREPs or NF-κB) were expressed later and at the end of the defense course chemotaxis revealed its importance. All these processes occurred within the first 24 hours after the challenge. Eventually the decrease of the regulated genes showed the stabilization of the status of the clam and the end of the cellular response. This immune-enriched microarray has demonstrated to be useful in hemocyte expression analysis and proved the effective and quick response of *R. philippinarum* against an infection.

After these efforts we decided to explore the ontogeny of the immune system in bivalves, as a high percentage of mortality occurs during early development. *M. galloprovincialis* was the chosen species to continue our research because, compared to other bivalves, it is relatively easy to obtain offspring in aquaria. As in larvae it is virtually impossible to obtain separated tissue samples, we decided to explore other techniques that complemented qPCR and microarray data: *in situ* hybridization and flow cytometry. These three techniques combined provided us the means to study the ontogeny of the immune system in *M. galloprovincialis* from a functional, physical and genetic point of view as presented in "Immune responses during the larval stages of *Mytilus galloprovincialis*: Metamorphosis alters immunocompetence, body shape and behavior". We were able to assess with flow cytometry that the *M. galloprovincialis* larvae are able to phagocyte in the trochophore stage already, 24 hours after the fecundation. With qPCR we observed a maternal contribution for gene expression during the first hours of development. Larvae induce own expresion especially after the veliger stage, preparing larvae for settlement. qPCR results suggest that the expression pattern after metamorphosis is more similar to adults, which reflects the maturation of the immune system. The *in situ* hybridization revealed that the expression of these immune-related genes, characteristic of hemocytes, is localized in the mantle edge, suggesting its possible role in hematopoiesis.

To complete the study in adult *M. galloprovincialis* we decided to perform an RNAseq study to gain insight into the transcriptomes of several tissues and published the last

paper presented in this PhD thesis, "RNA-Seq in *Mytilus galloprovincialis*: Comparative transcriptomics and expression profiles among different tissues". *M. galloprovincialis* could become the model species for the study of immunity and host-pathogen interactions in bivalve molluscs, but the genomic information for this bivalve could be improved. We generated almost 400 million reads to describe the transcriptome and expression profiles of *M. galloprovincialis* hemocytes, mantle, muscle and gill. These transcriptomes showed a high number of transcripts related to the immune system. Specific new tissue functions were also described: hemocytes showed the greatest number of antimicrobial and defense proteins, mantle exhibited antifungal and hematopoietic functions, gill presented a large number of recognition molecules as well as osmotic functions and muscle expressed stress- and defense-related proteins. This way we have contributed to increase the genomic background for bivalves and to set the basis for future research.

The key contributions to this thesis were the incredible amount of new sequences in two bivalve species belonging to different families. These results are now freely available in the NCBI databases, as well as the new knowledge in bivalve immunology regarding to response to pathogens, ontology of immune system and comparative transcriptomics.

All the work produced for this PhD thesis has generated knowledge to initiate and support many other academic findings, not only to our own department but to many collaborators and researchers from other institutions all over the world. We have published other research articles in bivalve immunology regarding reference genes to be used for qPCR experiments in *R. philippinarum* and *M. galloprovincialis* challenged with different bacterial strains (Moreira *et al.,* 2014) or the comparison of the immune response to diverse *Vibrios* of two mussel species from distant habitats, the intermareal *M. galloprovincilais* and the abyssal *B. azoricus* (Martins *et al.,* 2014). About the enrichment of the public databases in bivalve sequences the ReProSeed project has released, and is still releasing, a lot of information about bivalve transcriptomes, like the recently published *Pecten maximum* transcriptome (Pauletto *el al*., 2014). A more direct outcome of this thesis is the recent publication of the results of a *Perkinsus* infection in *R. philippinarum* using the microarray presented in the third manuscript (Romero *et al.,* 2015).

Many works related to molluscs have been published after the results of this thesis, indicative of the high importance of these animals in diverse scientific fields like gene characterization (Saranya *et al.,* 2012; Maldonado-Aguayo *et al.,* 2014; Niu *et al.,* 2014), genomics, transcriptomics and proteomics (Romero *et al.,* 2012; Castellanos-Martínez *et al.,* 2014; Rocher *et al.,* 2015), toxicity and stress (Figueira *et al.,* 2012; Suárez-Ulloa *et al.,* 2013; Li *et al.,* 2013; Menike *et al.,* 2014), ontology (Bassim *et al.,* 2014a; Bassim *et al.,* 2014b) and, of course immunology (Estévez-Calvar *et al.,* 2013; Soudant *et al.,* 2013; Bjelland *et al.,* 2013; Leite *et al.,* 2013; Tavares *et al.,* 2013; Toubiana *et al.,* 2014; Allam *et al.,* 2014; Sartim et al, 2014; Gerdol and Venier, 2015). All the research in bivalves is of great interest for scientists and for the related industries. This thesis provides many basic results that can be object of research in the following years.

Currently, the knowledge about the molecular and cellular mechanisms involved in the physiological processes of interest in aquaculture (reproduction, growth, immunity) is

rapidly increasing thanks to the genomic research. The next challenge is probably to find the specific functional roles of all these new genes and proteins, and to curate the possible bioinformatic issues in sequencing, mapping and assembly or annotation. The 'big data' obtained by NGS tends to generate large numbers of errors, requiring monitoring for methodologic biases and strategies for replication and validation (Goldman *et al.,* 2014). The future of all the expression and sequencing techniques used in this thesis is exciting and promising. These tools have never stopped their technical improvement and science is always finding new applications and updates to solve the infinite questions that the new discoveries set out.

In the near future, other new and interesting possibility for bivalve research is epigenetics, an almost unexplored technique regarding bivalve biology. This branch of science only has been applied to bivalves in the case of the Pacific oyster (*Crassostrea gigas*) (Gavery and Roberts, 2014; Olson and Roberts, 2014). It seems that DNA methylation has regulatory functions in *C. gigas*, particularly in gene families that have inducible expression, including those involved in stress, environmental responses and male gamete tissue gene regulatory activity.

Another possibility is the genome sequencing, limited to two species in bivalves: the Pacific oyster genome (Zhang *et al.*, 2012) and the draft genome of the Japanese pearl oyster (*Pinctada fucata*) (Takeuchi *et al.,* 2012).

One more very new method, not yet used in bivalves, is ribosome profiling (Ingolia *et al.,* 2009). This method reflects the translation process of specific tissues and developmental stages or conditions. As a result, it was suggested to estimate the translation efficiency of genes as well as the interaction between translation initiation, elongation, and termination. Expression and sequencing tools alone cannot explain critical aspects of gene expression regulation. This technique fills the gap between gene expression and protein quantification. In combination with methods such as RNA immunoprecipitation, miRNA profiling or proteomics, it is possible to get a new point of view of post-transcriptional and translational gene regulation (Kuersten *et al.,* 2013). In addition, these techniques also provide new insight into new regulatory elements, such as alternative open reading frames, and translation regulation under different conditions (Jackson and Standard, 2015).

As a consequence of these years of research, the enrichment of the bivalve sequences in the databases and after the release of the two bivalve genomes, five functional categories specific to molluscan biology, such as "Byssus Formation" and "Shell Formation", have been added to the Gene Ontology database (Kawashima *et al.,* 2013). The resulting data will serve as a useful reference for future genomic analyses in related species. But this is only the beginning and in the years to come new bivalve and mollusk genomes, as well as new GO categories will further enrich the databases, as the researchers demand them.

It is challenging the speed at which all these techniques appear, are updated as their cost decrease and get obsolete. This fact highlights the importance of everything related to gene expression and regulation in natural sciences and the need of all the scientists to keep into the stream and find collaborators of other fields, especially informatics, to continue moving on.

CONCLUSIONS

1. Based on the gene expression analysis, *R. philippinarum* response to a *Vibrio* infection appeared to be more effective and faster than that from *R. decussatus*. Additionally, *R. decussatus* did not seem to express transcription factors that could initiate an inflammatory response. Indeed, the expression pattern of IκB in *R. philippinarum* suggested a negative regulation of this inflammatory response 24 h after infection.
2. The first *R. philippinarum* transcriptome analysis focused on its immune system was completed. Near 30,000 new sequences have been discovered and more than 10% of the results had relationship with immunity. Immune-ralated sequences from molecules never described before in bivalves were found: C2, C4, C5, C9, AIF, Bax, AKT, TLR6 and TLR13. Three immune pathways have been characterized: the apoptosis, the toll like signaling pathway and the complement cascade.
3. It has been designed the second version of the *R. philippinarum* oligo-microarray enriched with immune sequences from hemocytes, and it has proven to be useful in hemocytes expression analysis, proved by the establishment of the timing for the immune response against a *Vibrio* infection. This research highlights the importance of a fast and effective response in bivalves:
    - Genes related to signaling, transcription and apoptosis were typically expressed as early as 3 hpi.
    - The key point to overcome the infection seemed to be 8 hpi, with characteristic immune and defense genes.
    - A high number of processes were activated 24 hpi to overcome the infection, including chemotaxis.
    - 72 hpi a negative feedback of all of the previously activated processes was observed.
4. *M. galloprovincialis* immune capacities arise during mussel development as early as the trochophore stage, in which the larvae are capable to phagocyte. Gene expression starts to rise after the veliger stage, preparing larvae for settlement. At the veliger and metamorphosis stages, the expression of immune-related genes, characteristic of hemocytes, is localized to the mantle edge, suggesting a possible hematopoietic role.
5. Almost 400 million reads belonging to transcriptomic and expression profiles of *M. galloprovincialis* tissues were established using RNA-Seq and Illumina for the first time in bivalves. Immune sequences are present in every tissue, indicative of the extremely high resistance of mussel to stress and pathogens.
    - Hemocytes: Typically defensive profile, with a huge amount of AMPs.
    - Mantle: Reproductive, sensorial and antifungal profile with novel hematopoyesis-related transcripts.
    - Muscle: Classic contractile sequences were found and some immune-related functions.
    - Gills: Many transcripts related to structure and non-self recognition, but also new osmotic and homeostatic functions were discovered.

QUALITY CRITERIA OF THE PUBLICATIONS

Developmental and Comparative Immunology

Developmental and Comparative Immunology (DCI) is an international journal that publishes articles describing original research in all areas of immunology, including comparative aspects of immunity and the evolution and development of the immune system. Manuscripts describing studies of immune systems in both vertebrates and invertebrates are welcome. All levels of immunological investigations are appropriate: organismal, cellular, biochemical and molecular genetics, extending to such fields as aging of the immune system, interaction between the immune and neuroendocrine system and intestinal immunity.

DCI 2014 Impact Factor: 2.815.

DCI 2012 Impact Factor: 3.238. 2012 was the year of publication of "Gene expression analysis of clams *Ruditapes philippinarum* and *Ruditapes decussatus* following bacterial infection yields molecular insights into pathogen resistance and immunity"

The ranking of DCI in Zoology category is 13th out of 153: Q1.

PLoS ONE

PLoS ONE is an international, peer-reviewed, open-access, online publication. *PLOS ONE* welcomes reports on primary research from any scientific discipline. It provides: open-access-freely accessible online; authors retain copyright; fast publication times; peer review by expert, practicing researchers; post-publication tools to indicate quality and impact; community-based dialogue on articles; worldwide media coverage.

PLoS ONE 2014 Impact Factor: 3.234.

PLos ONE 2012 Impact Factor: 3.702. 2012 was the year of publication of "Transcriptomics of *In Vitro* Immune-Stimulated Hemocytes from the Manila Clam Ruditapes philippinarum Using High-Throughput Sequencing".

The ranking of PLoS ONE in Multidisciplinary Sciences category is 8th out of 56: Q1.

Fish & Shellfish immunology

Fish and Shellfish Immunology rapidly publishes high-quality, peer-refereed contributions in the expanding fields of fish and shellfish immunology. It presents studies on the basic mechanisms of both specific and non-specific defense systems, the cells, tissues, and humoral factors involved, their dependence on environmental and intrinsic factors, response to pathogens, response to vaccination, and applied studies on the development of specific vaccines for use in the aquaculture industry.

Fish and Shellfish Immunology 2014 Impact Factor: 2.674.

Fish and Shellfish Immunology 2013 Impact Factor: 3.034. 2013 was the year of publication of "Immune responses during the larval stages of *Mytilus galloprovincialis*: Metamorphosis alters immunocompetence, body shape and behavior".

The ranking of Fish and Shellfish Immunology in Fisheries category is 4th out of 52: Q1; and in Marine and Freshwater Biology is 16th out of 102: Q1.

BMC genomics

BMC Genomics is an open access, peer-reviewed journal that considers articles on all aspects of genome-scale analysis, functional genomics, and proteomics.

BMC Genomics is part of the BMC series which publishes subject-specific journals focused on the needs of individual research communities across all areas of biology and medicine. We offer an efficient, fair and friendly peer review service, and are committed to publishing all sound science, provided that there is some advance in knowledge presented by the work.

BMC Genomics 2014 Impact Factor: 3.986.

2014 was the year of publication of "Gene expression profile analysis of Manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibrio alginolyticus* challenge using an immune-enriched oligo-microarray".

It is nor still available the BMC Genomics Impact Factor for 2015, the year of publication of "RNA-Seq in *Mytilus galloprovincialis*: Comparative transcriptomics and expression profiles among different tissues".

The ranking of BMC genomics in Biotehnology & Applied Microbiology category is 26th out of 162: Q1; and in Genetics & Heredity is 40th out of 167: Q1.

	IF 2014	5-year IF	ISSN
Developmental & Comparative Immunology	2.815	3.339	0145-305X
PLoS One	3.234	3.702	1932-6203
Fish & Shellfish Immunology	2.674	2.996	1050-4648
BMC Genomics	3.986	4.360	1471-2164

AUTHORS' CONTRIBUTION

Gene expression analysis of clams *Ruditapes philippinarum* and *Ruditapes decussatus* following bacterial infection yields molecular insights into pathogen resistance and immunity.

Rebeca Moreira, Pablo Balseiro, Alejandro Romero, Sonia Dios, David Posada, Beatriz Novoa, Antonio Figueras.

BN and AF conceived and designed the experiments. RM prepared the biological samples, designed the primers and performed the qPCR assay. DP and RM constructed the phylognetic trees. BN, AF, PB, AR, SD and RM analyzed the data. RM wrote the paper. All authors read and approved the manuscript.

Transcriptomics of *in vitro* immune-stimulated hemocytes from the Manila clam *Ruditapes philippinarum* using high-throughput sequencing.

Rebeca Moreira, Pablo Balseiro, Josep V. Planas, Berta Fuste, Sergi Beltran, Beatriz Novoa, Antonio Figueras.

BN and AF conceived and designed the experiments. RM and PB performed the experiments. SB, BF and JVP contributed reagents/materials/analysis tools. BN, AF, PB and RM analyzed the data. RM wrote the paper. All authors read and approved the manuscript.

Gene expression profile analysis of Manila clam (*Ruditapes philippinarum*) hemocytes after a *Vibrio alginolyticus* challenge using an immune-enriched oligo-microarray.

Rebeca Moreira, Massimo Milan, Pablo Balseiro, Alejandro Romero, Massimiliano Babbucci, Antonio Figueras, Luca Bargelloni and Beatriz Novoa.

BN and AF conceived and designed the experiments. RM prepared the samples. LB, MM, MB and RM made the assembly and annotation of the sequences and designed the microarray platform. RM and MM hybridized the microarrays. AR performed functional immune assays. BN, AF, PB, AR and RM analyzed the data. RM wrote the paper. All authors read and approved the manuscript.

Immune responses during the larval stages of *Mytilus galloprovincialis*: Metamorphosis alters immunocompetence, body shape and behavior.

Pablo Balseiro, Rebeca Moreira, Rubén Chamorro, Antonio Figueras, Beatriz Novoa.

BN and AF conceived and designed the experiments. PB and RC reared the larvae. PB performed the functional immune assays. RM prepared the samples and performed the expression analyses. BN, AF, PB and RM analyzed the data. PB wrote the paper. All authors read and approved the manuscript.

RNA-Seq in *Mytilus galloprovincialis*: Comparative transcriptomics and expression profiles between different tissues.

Rebeca Moreira, Patricia Pereiro, Carlos Canchaya, David Posada, Antonio Figueras, Beatriz Novoa.

BN, AF, DP and CC conceived and designed the experiments. RM and PP prepared the samples. RM and BN analyzed the data. RM wrote the paper. All authors read and approved the manuscript.

ADDITIONAL REFERENCES

1. 454 Life Sciences, a Roche Company. 13/04/2015. http: //454.com/
2. Allam B, Pales Espinosa E, Tanguy A, Jeffroy F, Le Bris C, Paillard C. Transcriptional changes in Manila clam (*Ruditapes philippinarum*) in response to Brown Ring Disease. Fish Shellfish Immunol. 2014; 41: 2-11.
3. Bardach JE, Ryther JH, McLarney O. Aquaculture – the farming and husbandry of freshwater and marine organisms. John Wiley & Sons, Inc. (New York). 1972.
4. Barnabé G. Acuicultura. Volumen I. Omega. 1991.
5. Bassim S, Genard B, Gauthier-Clerc S, Moraga D, Tremblay R. Ontogeny of bivalve immunity: assessing the potential of next-generation sequencing techniques. Reviews in Aquaculture 2014a; 6: 1-21.
6. Bassim S, Tanguy A, Genard B, Moraga D, Tremblay R. Identification of *Mytilus edulis* genetic regulators during early development. Gene. 2014b; 551: 65-78.
7. Bayne CJ, Sminia T, Van der Knaap WPW: Immunological Memory: Status of Molluscan Studies. In Phylogeny of Immunological Memory, Developments in immunology, Volume 10. Edited by Manning MJ. Amsterdam: Elsevier; 1980: 57-64.
8. Benjamini E, Coico R, Sunshine G. Immunology A Short Course. Fourth Edition. Wiley (New York). 2000.
9. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara E Catenazzi M, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman

E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008; 456: 53-59.

10. Bjelland AM, Fauske AK, Nguyen A, Orlien IE, Ostgaard IM, Sørum H. Expression of *Vibrio salmonicida* virulence genes and immune response parameters in experimentally challenged Atlantic salmon (*Salmo salar* L.). Front Microbiol. 2013; 4: 401.

11. Bricelj VM, Shumway SE. Paralytic shellfish toxins in bivalve molluscs: occurrence, transfer kinetics and biotransformation. Rev. Fish. Sci. 6: 315-383, 1998.

12. British Museum - Room 61 Tomb-chapel of Nebamun. 20/04/2015. http: //www.britishmuseum.org/explore/galleries/ancient_egypt/room_61_tomb-chapel_nebamun.aspx

13. Carballal MJ, López MC, Azevedo C, Villalba A. Hemolymph cell types of the mussel *Mytilus galloprovincialis*. Dis Aquat Organ. 1997; 29: 127-135.

14. Castellanos-Martínez S, Arteta D, Catarino S, Gestal C. De novo transcriptome sequencing of the *Octopus vulgaris* hemocytes using Illumina RNA-Seq technology: response to the infection by the gastrointestinal parasite *Aggregata octopiana*. PLoS One. 2014; 9: e107873.

15. Cerenius L, Lee BL, Söderhäll K.The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol. 2008; 29: 263-271.

16. Chu FE. Humoral defense factors in marine bivalves. Am Fish Soc Spec Pub. 1988; 18: 178-188.

17. Chu Y, Corey DR. RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic Acid Ther. 2012; 22: 271-274.

18. Criscitiello MF, de Figueiredo P. Fifty shades of immune defense. PLoS Pathog. 2013; 9: e1003110.

19. de Sousa JT, Milan M, Bargelloni L, Pauletto M, Matias D, Joaquim S, Matias AM, Quillien V, Leitão A, Huvet A. A microarray-based analysis of gametogenesis in two Portuguese populations of the European clam *Ruditapes decussatus*. PLoS One. 2014; 9: e92202.

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

21. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, *et al.* Real-time DNA sequencing from single polymerase molecules. Science. 2009; 323: 133-138.

22. Estévez-Calvar N, Romero A, Figueras A, Novoa B.Genes of the mitochondrial apoptotic pathway in *Mytilus galloprovincialis*. PLoS One. 2013; 8: e61502.

23. FAO Fisheries & Aquaculture – Acuicultura. 10/04/2015. http: //www.fao.org/fishery/aquaculture/es

24. FAO. 2014. The State of World Fisheries and Aquaculture 2014. Rome.

25. Farber JM. Ancient Hawaiian fishponds: can restoration succeed on Moloka'i? Cornell University. First Edition. Neptune House Publications (California). 1997.

26. Feng SY. Cellular defense mechanisms of oysters and mussels. Am Fish Soc Spec Publ. 1988; 18: 153-168.

27. Figueira E, Cardoso P, Freitas R.Ruditapes decussatus and *Ruditapes philippinarum* exposed to cadmium: toxicological effects and bioaccumulation patterns. Comp Biochem Physiol C Toxicol Pharmacol. 2012; 156: 80-86.

28. Gall JG, Pardue ML. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A. 1969; 63: 378-383.

29. Gavery MR, Roberts SB. A context dependent role for DNA methylation in bivalves. Brief Funct Genomics. 2014; 13: 217-222.

30. GenBank Release Notes. 20/04/2015. http: //www.ncbi.nlm.nih.gov/genbank/statistics

31. 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: 390-399.

32. Gerdol M, Venier P. An updated molecular basis for mussel immunity. Fish Shellfish Immunol. 2015; 46: 17-38.

33. Gestal C, Pallavicini A, Venier P, Novoa B, Figueras A. MgC1q, a novel C1q-domain-containing protein involved in the immune response of *Mytilus galloprovincialis*. Dev Comp Immunol. 2010; 34: 926-934.

34. Goldman D, Domschke K. Making sense of deep sequencing. Int J Neuropsychopharmacol. 2014; 17: 1717-1725.

35. Gómez-León 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

36. Guo K, Kang NX, Li Y, Sun L, Gan L, Cui FJ, Gao MD, Liu KY. Regulation of HSP27 on NF-kappaB pathway activation may be involved in metastatic hepatocellular carcinoma cells apoptosis. BMC Cancer. 2009; 9: 100.

37. Hedgecock D, Gaffney,P, Goulletquer P, Guo X, Reece K, Warr GW. The case for sequencing the oyster genome. J. Shellfish Res. 2005; 24: 429-442.

38. Huber-Lang M, Sarma JV, Zetoune FS, Rittirsch D, Neff TA, McGuire SR, Lambris JD, Warner RL, Flierl MA, Hoesel LM, Gebhard F, Younger JG, Drouin SM, Wetsel RA, Ward PA. Generation of C5a in the absence of C3: a new complement activation pathway. Nat Med. 2006; 12: 682-687.

39. Humphries JE, Yoshino TP. Cellular receptors and signal transduction in molluscan hemocytes: connections with the innate immune system of vertebrates. Integr Comp Biol. 2003; 43: 305-312.

40. Illumina - History of Illumina Sequencing. 20/04/2015. http: //www.illumina.com/technology/next-generation-sequencing/solexa-technology.html

41. Illumina - Next-Generation Sequencing (NGS). 20/04/2015. http: //www.illumina.com/technology/next-generation-sequencing.html

42. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. 2009; 324: 218-223.

43. Jackson R, Standart N. The awesome power of ribosome profiling. RNA. 2015; 21: 652-654.

44. Kawashima T, Takeuchi T, Koyanagi R, Kinoshita S, Endo H, Endo K. Initiating the mollusk genomics annotation community: toward creating the complete curated gene-set of the Japanese Pearl Oyster, *Pinctada fucata*. Zoolog Sci. 2013; 30: 794-796.

45. Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonák J, Lind K, Sindelka R, Sjöback R, Sjögreen B, Strömbom L, Ståhlberg A, Zoric N. The real-time polymerase chain reaction. Mol Aspects Med. 2006; 27: 95-125.

46. Kuersten S, Radek A, Vogel C, Penalva LO. Translation regulation gets its 'omics' moment. Wiley Interdiscip Rev RNA. 2013; 4: 617-630.

47. Kurelec B, Pivčević B: Evidence for a multixenobiotic resistance mechanism in the mussel *Mytilus galloprovincialis*. Aquat Toxicol 1991; 19: 291–301

48. Leite RB, Milan M, Coppe A, Bortoluzzi S, dos Anjos A, Reinhardt R, Saavedra C, Patarnello T, Cancela ML, Bargelloni L. mRNA-Seq and microarray development for the Grooved Carpet shell clam, *Ruditapes decussatus*: a functional approach to unravel host-parasite interaction. BMC Genomics. 2013; 14: 741.

49. Li H, Parisi MG, Parrinello N, Cammarata M, Roch P. Molluscan antimicrobial peptides, a review from activity-based evidences to computer-assisted sequences. Invertebr Surviv J. 2011; 8: 85-97.

50. Li Q, Zhao X, Kong L, Yu H. Transcriptomic response to stress in marine bivalves. Invertebr Surviv J. 2013; 10: 84.

51. Liao Z, Wang XC, Liu HH, Fan MH, Sun JJ, Shen W. Molecular characterization of a novel antimicrobial peptide from *Mytilus coruscus*. Fish Shellfish Immunol. 2013; 34: 610-616.

52. Maldonado-Aguayo W, Teneb J, Gallardo-Escárate C. A galectin with quadruple-domain from red abalone *Haliotis rufescens* involved in the immune innate response against to *Vibrio anguillarum*. Fish Shellfish Immunol. 2014; 40: 1-8.

53. Manfrin C, De Moro G, Torboli V, Venier P, Pallavicini A, Gerdol M. Physiological and molecular responses of bivalves to toxic dinoflagellates. Invertev Surviv J. 2012; 9: 184-199.

54. Martins E, Figueras A, Novoa B, Santos RS, Moreira R, Bettencourt R. Comparative study of immune responses in the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus* and the shallow-water mussel *Mytilus galloprovincialis* challenged with *Vibrio* bacteria. Fish Shellfish Immunol. 2014; 40: 485-499.

55. Maskos U, Southern E M. Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ. Nucleic Acids Res. 1992; 20: 1679-1684.

56. Menike U, Lee Y, Oh C, Wickramaarachchi WD, Premachandra HK, Park SC, Lee J, De Zoysa M.Oligo-microarray analysis and identification of stress-immune response genes from manila clam (*Ruditapes philippinarum*) exposure to heat and cold stresses. Mol Biol Rep. 2014; 41: 6457-6473.

57. Metzker ML. Emerging technologies in DNA sequencing. Genome Res. 2005; 15: 1767-1776.

58. Moreira R, Pereiro P, Costa MM, Figueras A, Novoa B. Evaluation of reference genes of *Mytilus galloprovincialis* and *Ruditapes philippinarum* infected with three bacteria strains for gene expression analysis. Aquat Living Res. 2014; 27: 147-152.

59. Morgan JAT, Dejong RJ, Snyder SD, Mkogi GM, Loker ES. *Schistosoma mansoni* and *Biomphalaria*: past history and future trends. Parasitology. 2001; 123: 211e28.

60. Mullis KB, Erlich HA, Arnheim N, Horn GT, Saiki RK, Scharf J. Process for amplifying, detecting, and/or-cloning nucleic acid sequences. US 4683195 A; July 28, 1987.

61. Niu D, Xie S, Bai Z, Wang L, Jin K, Li J. Identification, expression, and responses to bacterial challenge of the cathepsin C gene from the razor clam *Sinonovacula constricta*. Dev Comp Immunol. 2014; 46: 241-245.

62. Novoa B, Figueras A. Virus-like particles associated with mortalities of the carpet-shell clam *Ruditapes decussatus*. Dis Aquat Organ. 2000; 39: 147-149.

63. Nyren P. Method of sequencing DNA based on the detection of the release of pyrophosphate and enzymatic nucleotide degradation. US 6 258 568BI; July 10, 2001.

64. Olson CE, Roberts SB. Genome-wide profiling of DNA methylation and gene expression in *Crassostrea gigas* male gametes. Front Physiol. 2014; 5: 224.

65. Ordás MC, Novoa B, Faisal M, McLaughlin S, Figueras A. Proteolytic activity of cultured *Pseudoperkinsus tapetis* extracellular products. Comp Biochem Physiol B Biochem Mol Biol. 2001; 130: 199-206.

66. Parcellier A, Gurbuxani S, Schmitt E, Solary E, Garrido C. Heat shock proteins, cellular chaperones that modulate mitochondrial cell death pathways. Biochem Biophys Res Commun. 2003; 304: 505-512.

67. Pauletto M, Milan M, Moreira R, Novoa B, Figueras A, Babbucci M, Patarnello T, Bargelloni L. Deep transcriptome sequencing of *Pecten maximus* hemocytes: a genomic resource for bivalve immunology. Fish Shellfish Immunol. 2014, 37: 154-165.

68. Pfaffl MW. A new mathematical model for relative quantification in realtime RT-PCR. Nucleic Acids Res. 2001; 29: 2002–2007

69. Pillay TVR. Economic and social dimensions of aquaculture management. Aquacult Econ Manag 1997; 1: 3-11.

70. Pompilii J. "How Pyrosequencing Works" Polytechnic University of Milan, DensityDesign Research Lab. Image is released under CC-BY-SA licence. Licensed under CC BY-SA 4.0 via Wikimedia Commons. http: //commons.wikimedia.org/wiki/File: How_Pyrosequencing_Works.svg#mediaviewer/File: How_Pyrosequencing_Works.svg

71. Prado-Alvarez M, Romero A, Balseiro P, Dios S, Novoa B, Figueras A.Morphological characterization and functional immune response of the carpet shell clam (*Ruditapes decussatus*) haemocytes after bacterial stimulation. Fish Shellfish Immunol. 2012; 32: 69-78.

72. Prado-Alvarez M, Rotllant J, Gestal C, Novoa B, Figueras A.Characterization of a C3 and a factor B-like in the carpet-shell clam, *Ruditapes decussatus*. Fish Shellfish Immunol. 2009; 26: 305-315.

73. Qin CL, Huang W, Zhou SQ, Wang XC, Liu HH, Fan MH, Wang RX, Gao P, Liao Z. Characterization of a novel antimicrobial peptide with chitin-biding domain from *Mytilus coruscus*. Fish Shellfish Immunol. 2014; 41: 362-370.

74. Ratcliffe NA, Rowley AF, Fitzgerald SW, Rhodes CP. Invertebrate immunity: basic concepts and recent advances. Int Rev Cytol. 1985; 97: 183-350.

75. Relieve - Foro Egipto Viajar e Historia de Egipto. 20/04/2015. http: //www.egiptoforo.com/forums/showthread.php?t=17759

76. Ricklin D, Lambris JD. Complement-targeted therapeutics. Nat Biotechnol. 2007; 25: 1265-1275.

77. Roberts S, Gueguen Y, de Lorgeril J, Goetz F. Rapid accumulation of an interleukin 17 homolog transcript in *Crassostrea gigas* hemocytes following bacterial exposure. Dev Comp Immunol. 2008; 32: 1099-1104.

78. Rocher B, Bultelle F, Chan P, Le Foll F, Letendre J, Monsinjon T, Olivier S, Péden R, Poret A, Vaudry D, Knigge T. 2-DE Mapping of the Blue Mussel Gill Proteome: The Usual Suspects Revisited. Proteomes 2015; 3: 3-41.

79. Romero A, Novoa B, Figueras A. Genomics, immune studies and diseases in bivalve aquaculture. Invertebr Surviv J. 2012; 9: 110-121.

80. Romero A, Costa M, Forn-Cuni G, Balseiro P, Chamorro R, Dios S, Figueras A, Novoa B: Occurrence, seasonality and infectivity of *Vibrio* strains in natural populations of mussels *Mytilus galloprovincialis*. Dis Aquat Organ. 2014; 108: 149-163.

81. Romero A, Forn-Cuní G, Moreira R, Milan M, Bargelloni L, Figueras A, Novoa B. An immune-enriched oligo-microarray analysis of gene expression in Manila clam (*Venerupis philippinarum*) haemocytes after a *Perkinsus olseni* challenge. Fish Shellfish Immunol. 2015; 43: 275-286.

82. Ronaghi M, Uhlén M, Nyrén P. A sequencing method based on real-time pyrophosphate. Science. 1998; 281: 363-365.

83. Saavedra C, Bachère E. Bivalve genomics. Aquaculture. 2006; 256: 1-14.

84. Saranya Revathy K, Umasuthan N, Lee Y, Choi CY, Whang I, Lee J. First molluscan theta-class Glutathione S-Transferase: identification, cloning, characterization and transcriptional analysis post immune challenges. Comp Biochem Physiol B Biochem Mol Biol. 2012; 162: 10-23.

85. Sartim MA, Riul TB, Del Cistia-Andrade C, Stowell SR, Arthur CM, Sorgi CA, Faccioli LH, Cummings RD, Dias-Baruffi M, Sampaio SV. Galatrox is a C-type lectin in *Bothrops atrox* snake venom that selectively binds LacNAc-terminated glycans and can induce acute inflammation. Glycobiology. 2014; 24: 1010-1021.

86. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995; 270: 467-470.

87. Soudant P, E Chu FL, Volety A. Host-parasite interactions: Marine bivalve molluscs and protozoan parasites, *Perkinsus* species.J Invertebr Pathol. 2013; 114: 196-216.

88. Squidonius (talk) "Microarray exp horizontal". Licensed under Public Domain via Wikipedia. - http: //en.wikipedia.org/wiki/File: Microarray_exp_horizontal.svg#mediaviewer/File: Microarray_exp_horizontal.svg

89. Squidonius (talk) "NA hybrid". Licensed under Public Domain via Wikipedia. http: //en.wikipedia.org/wiki/File: NA_hybrid.svg#mediaviewer/File: NA_hybrid.svg

90. 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.

91. Takeuchi T, Kawashima T, Koyanagi R, Gyoja F, Tanaka M, Ikuta T, Shoguchi E, Fujiwara M, Shinzato C, Hisata K, Fujie M, Usami T, Nagai K, Maeyama K, Okamoto K, Aoki H, Ishikawa T, Masaoka T, Fujiwara A, Endo K, Endo H, Nagasawa H, Kinoshita S, Asakawa S, Watabe S, Satoh N. Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. DNA Res. 2012; 19: 117-130.

92. Tavares LS, Silva CS, de Souza VC, da Silva VL, Diniz CG, Santos MO. Strategies and molecular tools to fight antimicrobial resistance: resistome, transcriptome, and antimicrobial peptides. Front Microbiol. 2013; 4: 412.

93. Tomkinson AE, Vijayakumar S, Pascal JM, Ellenberger T. DNA ligases: structure, reaction mechanism, and function. Chem. Rev. 2006; 106: 687-699.

94. Toubiana M, Rosani U, Giambelluca S, Cammarata M, Gerdol M, Pallavicini A, Venier P, Roch P. Toll signal transduction pathway in bivalves: complete cds of intermediate elements and related gene transcription levels in hemocytes of immune stimulated *Mytilus galloprovincialis*. Dev Comp Immunol. 2014; 45: 300-312.

95. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002; 3, Research0034.

96. Vasta, G.R. Roles of galectins in infection. Nat Rev Microbiol. 2009; 7: 424-438.

97. Wootton EC, Dyrynda EA, Pipe RK, Ratcliffe NA. Comparisons of PAH-induced immunomodulation in three bivalve molluscs. Aquat Toxicol. 2003; 65: 13e25.

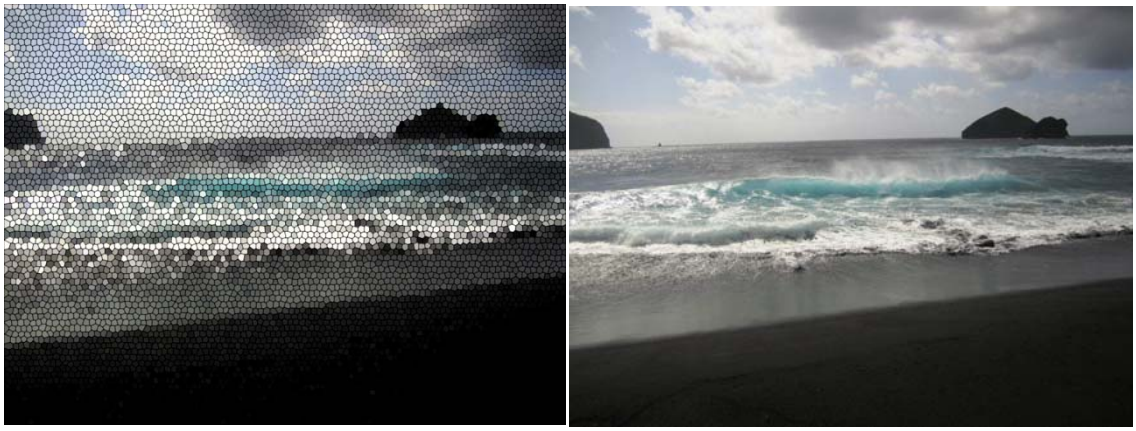
98. Xue QG, Waldrop GL, Schey KL, Itoh N, Ogawa M, Cooper RK, Losso JN, La Peyre JF. A novel slow-tight binding serine protease inhibitor from eastern oyster (*Crassostrea virginica*) plasma inhibits perkinsin, the major extracellular protease of the oyster protozoan parasite *Perkinsus marinus*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 2006; 145: 16-26.

99. Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PW, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Lošo T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Wang J, Chen M, He Y, Tan F, Song X, Zheng Q, Huang R, Yang H, Du X, Chen L, Yang M, Gaffney PM, Wang S, Luo L, She Z, Ming Y, Huang W, Zhang S, Huang B, Zhang Y, Qu T, Ni P, Miao G, Wang J, Wang Q, Steinberg CE, Wang H, Li N, Qian L, Zhang G, Li Y, Yang H, Liu X, Wang J, Yin Y, Wang J. The oyster genome reveals stress adaptation and complexity of shell formation. Nature. 2012; 490: m49-54.

100. Zhao J, Qiu L, Ning X, Chen A, Wu H, Li C. Cloning and characterization of an invertebrate type lysozyme from *Venerupis philippinarum*. Comp Biochem Physiol B Biochem Mol Biol. 2010; 156: 56-60.

La foto de portada la saqué en la Praia de Mosteiros, Ilha de São Miguel, Açores, Portugal.

I took the cover photo in Praia de Mosteiros, Ilha de São Miguel, Açores, Portugal.



La capa en forma de vidriera, en la foto de la izquierda, representa la visión que tenemos los científicos de la situación fisiológica de un ser vivo al estudiar la expresión genética: Nosotros estudiamos un momento concreto, una foto, en su vida. Con esta foto indefinida, que contiene la información de los genes de los que disponemos en ese momento, tenemos que interpretar cual es su realidad y cómo responde a ella. A medida que se avanza en el descubrimiento de nuevos genes y se estudia su expresión, la foto irá adquiriendo cada vez más nitidez, pareciéndose más a la de la derecha, a la realidad.

The stained glass layout, left photo, represents the scientific vision of the physiological status of a living being when we study its gene expression: We study a specific moment, a photograph, in their lives. With this undefined photo, containing the information about the available genes in this moment, we have to figure out what their reality is and how they respond to it. With the discovery of new genes and the study of their expression the photo will gain on definition, resembling more and more the right photograph, the reality.